UNIVERSITY OF JYVÄSKYLÄ
JYVÄSKYLÄN YLIOPISTO

Proceedings of the 12th European Conference on Information Warfare and Security

University of Jyväskylä, Finland
11-12 July 2013

Edited by
Prof Rauno Kuusisto, Finnish Defence Forces Technical Research Centre and
Dr Erkki Kurkinen, University of Jyväskylä

A conference managed by ACI, UK

acpi

Proceedings of
The 12th European Conference on Information Warfare and Security

University of Jyväskylä
Finland

11-12 July 2013

Edited by
Rauno Kuusisto
Finnish Defence Forces Technical Research Centre (PVTT), Finland
and
Erkki Kurkinen
University of Jyväskylä, Finland

Papers have been double-blind peer reviewed before final submission to the conference. Initially, paper abstracts were read and selected by the conference panel for submission as possible papers for the conference.

Many thanks to the reviewers who helped ensure the quality of the full papers.

These Conference Proceedings have been submitted to Thomson ISI for indexing. Please note that the process of indexing can take up to a year to complete.

Further copies of this book and previous year's proceedings can be purchased from http://academic-bookshop.com

E-Book ISBN: 978-1-909507-36-4
E-Book ISSN: 2048-8610
Book version ISBN: 978-1-909507-34-0
Book Version ISSN: 2048-8602
CD Version ISBN: 978-1-909507-37-1
CD Version ISSN: 2048-8629

The Electronic version of the Proceedings is available to download at ISSUU.com. You will need to sign up to become an ISSUU user (no cost involved) and follow the link to http://issuu.com

Published by Academic Conferences and Publishing International Limited
Reading
UK
44-118-972-4148
www.academic-publishing.org

Contents

Preface

This year sees the 12th European Conference on Information Warfare and Security (ECIW 2013), which is hosted by the University of Jyväskylä, Finland.

The Conference Co-Chairs are Prof Mikko Siponen and Researcher Martti Lehto from the University of Jyväskylä, Finland. The Programme Co-Chairs are Prof Rauno Kuusisto, from the Finnish Defence Forces Technical Research Centre (PVTT), Finland and Researcher Dr Erkki Kurkinen, from the University of Jyväskylä, Finland

The Conference continues to bring together individuals working in the area of cyberwar and cyber security in order to share knowledge and develop new ideas with their peers. The range of papers presented at the Conference will ensure two days of interesting discussions. The topics covered this year illustrate the depth of the information operations' research area, with the subject matter ranging from the highly technical to the more strategic visions of the use and influence of information.

The opening keynote is given by Dr Jarno Limnéll, from Stonesoft Helsinki, Finland on the topic of "The changing reality of security" and the second day keynote will be presented by Aapo Cederberg, from the Ministry of Defence, Finland, on the topic of "Finland´s Cyber Security Strategy."

With an initial submission of 99 abstracts, after the double blind, peer review process there are 39 Research papers, 8 PHD Papers, 1 non Academic paper, 3 Work in Progress Papers published in these Conference Proceedings. These papers come from all parts of the globe including Australia, Czech Republic, Estonia, Finland, France, Hungary, Israel, Italy, Portugal, Russia, South Africa, Turkey, United Kingdom and the United States of America.

I wish you a most interesting conference and an enjoyable stay in Finland.

Rauno Kuusisto, Finnish Defence Forces Technical Research Centre (PVTT), Finland
and
Erkki Kurkinen, from the University of Jyväskylä, Finland
Programme Co-Chairs
July 2013

Conference Committee

Conference Executive
Prof Mikko Siponen, University of Jyväskylä, Finland
Martti Lehto, University of Jyväskylä, Finland
Prof Rauno Kuusisto, Finnish Defence Forces Technical Research Centre (PVTT), Finland
Erkki Kurkinen, University of Jyväskylä, Finland

Mini track Chairs:
Dr. Nasser Abouzakhar, University of Hertfordshire, UK
Dr Rain Ottis, University of Jyväskylä, Jyväskylä, Finland
Dr Aki Huhtinen Finnish National Defence University, Finland
Dr Jari Rantapelkonen, Finnish National Defence University, Finland

Committee Members

The conference programme committee consists of key individuals from countries around the world working and researching in the Information Warfare and Security community. The following have confirmed their participation:

Dr Mohd Faizal Abdollah (University Technical Malaysia Melaka, Melaka); Dr Nasser Abouzakhar (University of Hertfordshire, UK); Dr Kari Alenius (University of Oulu, Finland); Dr Olga Angelopoulou (University of Derby, UK); Dr Leigh Armistead (Edith Cowan University, Australia); Colin Armstrong (Curtin University, Australia, Australia); Debi Ashenden (Cranfield University, Shrivenham, UK); Laurent Beaudoin (ESIEA, Laval, France); Maumita Bhattacharya (Charles Sturt University, Australia); Prof Matt Bishop (University of California at Davis, USA); Andrew Blyth (University of Glamorgan, UK); Martin Botha (South African Police, South Africa); Colonel (ret) Colin Brand(Graduate School of Business Leadership, South Africa); Dr Svet Braynov (University of Illinois at Springfield, USA); Bill Buchanen (Napier University, UK); Dr Joobin Choobineh (Texas A&M University, USA); Dr Maura Conway (Dublin City University, Ireland); Michael Corcoran(DSTL, UK); Dr Paul Crocker (Universidade de Beira Interior, Portugal); Christian Czosseck (NATO Cooperative Cyber Defence Centre of Excellence, Estonia); Geoffrey Darnton (Bournemouth University, UK); Josef Demergis (University of Macedonia, Greece); Moses Dlamini(SAP Research Pretoria, South Africa); Paul Dowland (University of Plymouth, UK); Marios Efthymiopoulos (Political Science Department University of Cyprus, Cyprus); Dr Ramzi El-Haddadeh (Brunel University, UK); Daniel Eng (C-PISA/HTCIA, China); Prof. Dr. Alptekin Erkollar(ETCOP, Austria); Professor of CS and Security Robert Erra (ESIEA PARIS, France); John Fawcett (University of Cambridge, UK);Professor and Lieutenant-colonel Eric Filiol (Ecole Supérieure en Informatique, Electronique et Automatique, France); Dr Chris Flaherty(University of New South Wales, Australia); Steve Furnell (University of Plymouth, UK); Javier Garci'a Villalba (Universidad Complutense de Madrid, Spain); Kevin Gleason (KMG Consulting, MA, USA); Dr Michael Grimaila (Air Force Institute of Technology, USA); Professor Stefanos Gritzalis (University of the Aegean, Greece); Marja Harmanmaa (University of Helsinki, Finland); Dr Julio Cesar Hernandez Castro(Portsmouth University, UK); Ulrike Hugl (University of Innsbruck, Austria); Aki Huhtinen (National Defence College, Finland); Bill Hutchinson(Edith Cowan University, Australia); Dr Berg Hyacinthe (Assas School of Law-CERSA-CNRS, La Sorbonne, France); Dr Abhaya Induruwa(Canterbury Christ Church University, UK); Hamid Jahankhani (University of East London, UK); Dr Amit Jain (BenefitFocus Inc, USA); Dr Helge Janicke (De Montfort University, UK); Joey Jansen van Vuuren (CSIR, South Africa); Saara Jantunen (University of Helsinki, Finland); Andy Jones (BT, UK); James Joshi (University of Pittsburgh, USA); Nor Badrul Anuar Jumaat (University of Malaya, Malaysia); Maria Karyda (University of the Aegean, Greece); Vasilis Katos (Democritus University of Thrace, Greece); Dr Anthony Keane (Institute of Technology Blanchardstown, Dublin, Ireland); Jyri Kivimaa (Cooperative Cyber Defence and Centre of Excellence, Tallinn, Estonia); Spyros Kokolakis (University of the Aegean, Greece); Prof. Ahmet Koltuksuz (Yasar University, Dept. of Comp. Eng., Turkey); Theodoros Kostis(Hellenic Army Academy, Greece); Prashant Krishnamurthy (University of Pittsburgh, USA); Dan Kuehl (National Defense University, Washington DC, USA); Peter Kunz (DiamlerChysler, Germany); Pertti Kuokkanen (Finnish Defence Forces, Finland); Erikk Kurkinen(University of Jyväskylä, Finland); Takakazu Kurokawa (National Defence Academy, Japan); Rauno Kuusisto (Finish Defence Force, Finland); Tuija Kuusisto (Internal Security ICT Agency HALTIK, Finland); Dr Laouamer Lamri (Al Qassim University and European University of Brittany, Saudi Arabia); Michael Lavine (John Hopkins University's Information Security Institute, USA); Martti Lehto (National Defence University, Finland); Tara Leweling (Naval Postgraduate School, Pacific Grove, USA); Paul Lewis (technology strategy board, UK); Peeter Lorents(CCD COE, Tallinn, Estonia); Hossein Malekinezhad (Islamic Azad University,Naragh Branch, Iran); Mario Marques Freire (University of Beira Interior, Covilhã, Portugal); Ioannis Mavridis (University of Macedonia, Greece); Rob McCusker (Teeside University, Middlesborough, UK); Jean-Pierre Molton Michel (Ministry of Agriculture, Haiti); Durgesh Mishra (Acropolis Institute of Technology and Research, India); Dr. Yonathan Mizrachi (University of Haifa, Israel, Israel); Edmundo Monteiro (University of Coimbra, Portugal); Evangelos Moustakas(Middlesex University, London, UK); Dr Kara Nance (University of Alaska Fairbanks, USA); Muhammad Naveed (IQRA University Peshawar, Pakistan, Pakistan); Mzukisi Njotini (University of South Africa, South Africa); Rain Ottis (Cooperative Cyber Defence Centre of Excellence, Estonia); Tim Parsons (Selex Communications, UK); Dr Andrea Perego (European Commission - Joint Research Centre, Ispra, , Italy); Michael Pilgermann (University of Glamorgan, UK); Engur Pisirici (govermental - independent, Turkey); Dr Muttukrishnan Rajarajan (City University London, UK); LtCol Jari

Rantapelkonen (National defense University, Finland); Andrea Rigoni (Booz & Company , USA); Dr Neil Rowe (US Naval Postgraduate School, Monterey, USA); Raphael Rues (DigiComp Academy, Switzerland); Filipe Sa Soares (University of Minho, Portugal); Professor Henrique Santos (University of Minho, Portugal); Prof Chaudhary Imran Sarwar (Mixed Reality University, Pakistan); Dr Damien Sauveron (Mathematics and Computer Sciences, University of Limoges, France); Sameer Saxena (IAHS Academy, Mahindra Special Services Group , India); Prof. Dr. Richard Sethmann (University of Applied Sciences Bremen, Germany); Dr Yilun Shang(University of Texas at San Antonio, USA); Paulo Simoes (University of Coimbra, Portugal); Prof Mikko Siponen (University of Jyväskylä, Finland); Professor Jill Slay (University of South Australia, Australia); Anna Squicciarini (University of Milano, Italy); Iain Sutherland(University of Glamorgan, Wales, UK); Jonas Svava Iversen (Danish Broadcast Corporation, Denmark); Anna-Maria Talihärm (Tartu University, Estonia); Dr Selma Tekir (Izmir Institute of Technology, Turkey); Prof Sérgio Tenreiro de Magalhães (Universidade Católica Portuguesa, Portugal); Prof. Dr. Peter Trommler (Georg Simon Ohm University Nuremberg, Germany); Dr Theodore Tryfonas (University of Bristol, UK); Craig Valli (Edith Cowan UniversitY, Australia); Rudi Vansnick (Internet Society, Belgium); Richard Vaughan (General Dynamics UK Ltd, UK); Stilianos Vidalis (Newport Business School, Newport, UK); Dr Natarajan Vijayarangan (Tata Consultancy Services Ltd, India); Marja Vuorinen (University of Helsinki, Finland); Michael Walker (Vodafone, UK); Mat Warren (Deakin University, Australia, Australia); Dr Kenneth Webb (Edith Cowan University , Australia); Dr Stuart Weinstein (University of Hertfordshire School of Law, UK); Dr Santoso Wibowo (CQUniversity, Australia); Dr Trish Williams (Edith Cowan University, Australia); Simos Xenitellis (Royal Holloway University, London, UK); Dr Omar Zakaria (National Defence University of Malaysia, Malaysia);

Biographies

Conference Co-Chairs

Prof Mikko Siponen is a Professor at the University of Jyväskylä in the Department of Mathematical Information Technology, Finland. Before that he was a professor and Director of the IS Security Research Centre in the Department of Information Processing Science at the University of Oulu, Finland. He holds a Ph.D. in philosophy from the University of Joensuu, Finland, and a Ph.D. in Information Systems from the University of Oulu, Finland. His research interests include IS security, IS development, computer ethics, and philosophical aspects of IS. In addition to more than 70 conference articles, he has more than 40 papers in journals such as MIS Quarterly, Journal of the Association for Information Systems, European Journal of Information Systems, Information & Organization, Information Systems Journal, Information & Management, ACM Database, and Communications of the ACM, IEEE Computer, IEEE IT Professional and ACM Computers & Society. He has received over 5.4million USD of research funding from corporations and numerous funding bodies. 20011 he was ranked as the most productive Information Systems scholar in Europe. The Finnish Funding Agency for Technology and Innovation ranked IS security research center that Dr. Siponen established and led as the top ICT research group in Finland.

Dr Martti Lehto Researcher, PhD Martti graduated as doctor of military science at the Finnish National Defence University on 2012. His major subject was Finnish Air Force C4ISR system evolution. He has graduated as general staff officer at the Finnish National Defence University on 1987. Lehto has over 30 years of experience mainly as developer and leader of C4ISR Systems in Finnish Defence Force and Air Force. He is a researcher in the University of Jyväskylä in the Department of Mathematical Information Technology and his research areas are cyber security, cyber defence and security clusters. He has about 20 publications and research reports on areas of C4ISR systems, cyber security and defence, information warfare, defence policy, leadership and management. Since 2001 he has been also the Editor-in-Chief of the Military Magazine.

Programme Co-Chairs

Prof Rauno Kuusisto graduated as doctor of philosophy at Helsinki University of Technology on 2004. His major subject was corporate security and minor subjects were knowledge management and futures studies. Kuusisto has graduated as general staff officer at the Finnish National Defence University on 1993. Kuusisto has contributed as an expert, consultant and manager roles particularly on the areas of information availability in strategic decision-making, strategy thinking, product development, research purchasing, project portfolio management, network enabled management and leadership in innovative environment, systems thinking, as well as modeling comprehensive challenges. Kuusisto works at the moment as a head of the Electronics and Information Technology Division of the Finnish Defence Forces Technical Research Centre (PVTT). Kuusisto is adjunct professor of network enabled defense at the Finnish National Defence University. He is also a visiting professor in the University of Jyväskylä in the Department of Mathematical Technology.

Dr Erkki Kurkinen Researcher has over 30 years history in the industry of mobile information technology, both in public safety and security (PSS) markets, and in consumer markets. His PhD dissertation is on mobile technology acceptance among law enforcement. He possesses experience in utilization of the TETRA technology in PSS-market. He has an excellent track record in developing products and services for the consumer market based on 2G/3G GSM technology as well. At present he is a researcher in the University of Jyväskylä in the Department of Mathematical Information Technology. His research areas are technology acceptance, use of social media in authority operations and cyber security.

Keynote Speakers

Dr Jarno Limnéll is the Director of Cyber Security at Stonesoft. He is an expert in international security politics and reality of today´s threats and security environment. He has profound understanding of the global threat landscape, combined with the courage to address even the most complex issues and the ability to lead discussions. Mr. Limnéll holds a Doctor of Military Science degree of Strategy from the National Defense University, a Master of Social Science degree from the Helsinki University and an Officer´s degree from the National Defense University. Mr. Limnéll holds a comprehensive list of publications on security issues, and his latest book is called “The World and Finland after 9/11.” Prior to Stonesoft, Mr. Jarno Limnéll acted as Manager, Defense & Public Safety at Accenture. Previously, he worked as Director of Administration and Development at Evli Bank Plc and Lecturer of Strategy at the Department of Strategic and Defence Studies, National Defence University in Helsinki.

Aapo Cederberg has a long career in the Finnish Armed Forces, lastly as Senior Military Adviser at the Permanent Mission of Finland to the OSCE in 1999- 2003, Commander of the Häme GBAD Battalion in 2003 – 05, Head of Strategic Planning at the Ministry of Defence in 2005 – 2007 and at present Secretary General for the Security Committee at the Ministry of Defence since 2007. The Committee supports the Government in comprehensive security matters and has provided the Security Strategy for the Society as well as Finland´s first Cyber security Strategy

Mini Track Chairs

Dr Nasser S. Abouzakhar is a senior lecturer at the University of Hertfordshire, UK. Currently, his research area is mainly focused on critical infrastructure security and applying machine learning solutions to various Internet and Web security and forensics related problems. He received MSc (Eng) in Data Communications in 2000 followed by PhD in Computer Sci Engg in 2004 from the University of Sheffield, UK. Nasser worked as a lecturer at the University Of Hull, UK in 2004-06 and a research associate at the University of Sheffield in 2006-08. He is a technical studio guest to various BBC World Service Programmes such as Arabic 4Tech show, Newshour programme and Breakfast radio programme. Nasser is a BCS assessor for the accreditation of Higher Education Institutions (HEIs) in the UK, BCS chartered IT professional (CITP), CEng and CSci. His research papers were published in various international journals and conferences.

Dr Rain Ottis is a scientist at the NATO Cooperative Cyber Defence Centre of E xcellence, in Tallinn, Estonia. He previously served as a communications officer in the Estonian Defence Forces, focusing primarily on cyber defence training and awareness. He is a graduate of the United States Military Academy (BS, Computer Science) and Tallinn University of Technology (PhD, Computer Science; MSc, Informatics). His research interests include cyber conflict, national cyber security, politically motivated cyber attacks and the role of volunteers in cyber security. In addition to his current assignment, he is teaching cyber security in Tallinn University of Technology (EST) and University of Jyväskylä (FIN).

Dr Aki Huhtinen is a professor at Finnish National Defence University. His expertise areas are military leadership, and philosophy of war.

Dr Jari Rantapelkonen is a professor at Finnish National Defence University. His expertise areas are strategic communication and operational art and tactics.

Biographies of Presenting Authors

Nasser Abouzakhar is a senior lecturer at the University of Hertfordshire, UK. Currently, his research area is mainly focused on applying machine learning solutions to critical infrastructure protection. He received PhD in 2004, the University of Sheffield. Nasser is a BCS assessor for the accreditation of Higher Education Institutions in the UK, CEng and CSci.

Kari Alenius is Associate Professor in the Department of History at the University of Oulu, Finland, since 1998. He also has Adjunct Professorship at the University of Oulu (1997). His research interests include the history of propaganda and mental images, the history of Eastern Europe between the World Wars, and the history of ethnic minorities.

Olga Angelopoulou, BSc, MSc, PhD is a lecturer and the programme leader for the MSc Computer Forensic Investigation at the University of Derby. She obtained a doctorate in Computing with the title: 'Analysis of Digital Evidence in Identity Theft Investigations' from the University of Glamorgan. Her research interests include Digital Forensics, Identity Theft, Online Fraud, Digital Investigation Methodologies and Online Social Networking.

Dr Edwin "Leigh" Armistead is the President of Peregrine Technical Solutions, which focuses on IO and Cyber Security. Leigh received his PhD from Edith Cowan University with an emphasis on IO, and serves as Co-Editor for the Journal of International Warfare, plus the Editorial Review Board for ECIW and is the Programme Director for ICIW.

Amir Averbuch holds B.Sc, M.Sc in Mathematics, Hebrew University, Jerusalem and Ph.D, Columbia University. During 1966-1970, 1973-1976 he served in the Israeli Defense Forces. Research Staff Member at IBM, Watson Research Centre. In

1987 he joined the School of Computer Science, Tel Aviv University, where he is professor. His research interests include applied harmonic analysis.

Mohd Rizuan Baharon is currently a second year PhD student at Liverpool John Moores University, Liverpool, UK. Mohd completed his master degree (MSc in Mathematics) in 2006 and his undergraduate studies in 2004 at University Technology Malaysia, Malaysia. His research interests lie in the area of Cryptography and Network Security.

Daria Yu. Bazarkina PhD is the associate professor of the chair of journalism and media education at the Sholokhov Moscow State Humanitarian University. Author of more than thirty articles on the communication aspect of the terrorist activity and counter-terrorist struggle.

Muhammad Bilal University Salford Manchester, UK. Well-motivated, logical minded individual with good interpersonal skills and will always drive to succeed in a competitive environment. When presented with a difficult task I can adapt to the working environment and where necessary, work as part of a team to achieve the objectives of the organization

Prof. Robin Bloomfield is Professor of Software and System Dependability at the City University, London. His research interests are in the dependability (reliability, safety, security) of computer-based systems. His work in safety in the past 20 yrs has combined policy formulation, technical consulting and underpinning research. He is currently the Industrial Liaison Director for a UK interdisciplinary research project on dependability (DIRC).

Ms. Larisa Breton is a Strategic Communication practitioner and theoretician with COCOM, NATO, and commercial experience. Larisa is published in The Small Wars Journal and is forthcoming in the Journal of Information Warfare. She has Guest Lectured at the JFK School for Special Operations. She is adjunct faculty for University of the District of Columbia.

Teodora Ciocoiu University of Alcalá, System Informatics, Spain, is a 4th year Informatics Systems student at the Alcalá University in Spain. I think IP E-Discovery Program was a great opportunity to learn more about forensic science and all the procedures that must be done in order to respect all the rules and laws applied.

Natalia Chaparro is from MidSweden University, Archival- and Information Science and Informatics, Sweden. She is studying at Mid-Sweden University in Sweden. The primary studies are in the field of Archival- and Information Science, and the secondary studies are in informatics. The IP has helped her. She has not only learned about digital forensics, she has also learned about how it is to work in an international team.

Grégory Commin is a a student at ESIEA (Ecole Supérieur Informatique Electronique ET Automatique) and I am participated in the (C+V) ° Laboratory. I speak French, English and Spanish. In 2007, I passed my baccalaureate Scientist as an option Engineer Science.

Paul Crocker has a PhD in Mathematics from the University of Leeds, UK. After working in software development he joined the Computer Science Department at the University of Beira Interior, Portugal. His research and teaching interest include Parallel Computing, Security and Operating systems. He is a member of the Portuguese research Institute of Telecommunications.

Asaf David is currently a M.Sc candidate in computer science at the Tel Aviv Jaffa Academic collage. He has received B.Sc in Computer Science from Tel Aviv University; Asaf is currently with the israeli diffence forces

Austin Davies is currently residing in the United Kingdom in the small town of Bolton. Austin currently studies Software Engineering, in his second year at the University of Salford.

Denis Edgar-Nevill holds the post of Head of Department of Computing at Canterbury Christ Church University in the UK as well as Chair of the British Computer Society Cybercrime Forensics Specialist Group which has 1,600 members in 44 countries. He is Principal Researcher/Project Manager for the EU funded ECENTRE project.

Karin Ehnberg id Sweden University, Archive and Information Science. Earlier degree in Archeology at University of Lund. Education combined with professional experience in participating and leading projects have led to experience in organizing, searching and interpreting information, archives and writing scientific reports.

Yu-Chun Lu is a researcher of Electrical Engineering at the Institute of Computer and Communication Engineering at National Cheng Kung University, Tainan, Taiwan. He received a B.Sc. degree in Computer Information and Science from Tung-Hai University, Taiwan in 1997, and both a M.Sc. degree and a Ph.D. (Computer Science and Engineering) from National Sun Yat-sen University, Taiwan in 1999 and 2010, respectively. His research interests include Internet security, wireless network, home network system, IoT, and learning cloud services.

Eric Filiol is the head of the Operational Cryptology and Virology at ESIEA. He has spent 21 years in the French Army. He holds an Engineer diploma in Cryptology, a PhD and a Habilitation Thesis in applied mathematics and computer science. He is also the Scientific Director of EICAR and the Editor-in-chief of the Journal in Computer Virology.

Jason Flood, MSc is currently an Ethical Hacking Architect at IBM in Dublin. He is also a PhD student at the Institute of Technology Blanchardstown where he investigates better ways of training Network Administrators. Jason is the co-founder Irish Chapter of the Honeynet Project and works with OWASP and Facebook in running CTF competitions.

Grigorios Fragkos, BSc, MSc, PhD, Certified TigerScheme, AST and QSTM. He has a number of publications in Computer Security and Computer Forensics. He has been part of the CyberDefense dept. of the Hellenic Army acting as Information Security consultant and Penetration tester. Currently, works for Sysnet Global **Solutions as Sr. Consultant and Penetration tester.**

Wendy Goucher is about to enter the final phase of her PhD research. She is a part time student at University of Glasgow and is also an information security consultant with Idrach Ltd. where she specialises in assisting in the design and communication of operationally effective security policy.

Dijana Grd comes from Croatia. She is a second year student of graduate study programme Information and Software Engineering at Faculty of Organization and Informatics in Varazdin. Her studies have provided her an insight into area of identification, collection, processing, analysis and production of electronically stored information. She has some work experience as a student assistant in Informatics and as a project manager in student organization AIESEC. She enjoys participating in all kind of international conferences and projects. She also likes to travel and meet new people.

Clement Guitton is a PhD candidate in War Studies at King's College London focusing on cyber security. He holds a master degree both in international relations and in electrical engineering. Fluent in English, French, and German, he previously worked at the International Telecommunication Union, the United Nation agency specialised on information and communication technologies.

Håkan Gunneriusson has a PhD in History 2002, Uppsala University. Hakan is interested in sociological and historical perspectives on current and coming issues regarding military tactical and cultural issues. Hakan is currently head of research ground operative and tactical areas, Swedish National Defence College.

Samuli Haataja is a PhD candidate in the Griffith Law School at Griffith University on the Gold Coast, Australia. He holds a Bachelor of Laws (Hons) and Bachelor of International Relations from Griffith University. His research focuses on cyber attacks and international law – specifically on the relationship of technology, violence and law in this context.

Mikko Hakuli is currently employed as security specialist at JyvSecTec-project in Jyväskylä University of Applied Sciences (JAMK), where his main responsible are technical security testing and development of various situational awareness "best practices" in cyber-security area. Formerly he worked as Head of information security on Finnish Airforces. Currently he also make studies in University of Jyväskylä and Jyväskylä University of Applied Sciences.

Juhani Hämäläinen received his PhD degree in theoretical physics from the University of Jyväskylä in 2004. He is currently in the position of principal scientist at Finnish Defense Forces Technical Research Centre (PVTT). His research interests include mathematical model development and operational analysis.

Major Arto Hirvelä is an instructor (leadership) in a research group at the Finnish National Defence University. He is preparing a doctoral dissertation in Military Science (leadership). His research interests are information environment, strategic communication, and information operations.

Ilona Ilvonen is a doctoral student at Tampere University of Technology, department of Information Management and Logistics. Her doctoral thesis topic is the management of knowledge security, and the thesis is due in 2013. She has published conference papers on information security management, knowledge management and relating topics since the year 2003.

Margarita Jaitner is a research intern at the Finnish National Defence University. She received a Bachelor's degree in Political Science at the Swedish Defence College and is currently pursuing a Master's degree in Societal Risk Management at the Karlstad University in Sweden.

Saara Jantunen has studied English language and culture in the University of Groningen in the Netherlands and the University of Helsinki in Finland. She has completed her PhD in the Finnish National Defence University, where she researched military communication concepts and doctrine.

Roman Jašek is the head of Department of Informatics and Artificial Intelligence, Faculty of Applied Informatics, Tomas Bata University in Zlín. His habilitation thesis focused on implementing information security paradigm into commercial organizations as well as tertiary education institutions. Professional interests include computer security auditing, knowledge protection, information systems, and informatics.

Georgios A. Kallos received Dipl.–Eng. degree in Electrical and Computer Engineering from the University of Patras, Greece, in 2006. In 2010, he obtained M.Sc. degrees in Wireless Systems and Information and Communication Technologies from the Royal Institute of Technology (KTH), Stockholm, Sweden and the Technical University of Catalonia (UPC), Barcelona, Spain. He is current a research scientist at the British Telecommunications, UK.

Producer, LT (SG), Tommi Kangasmaa is Producer of the Finnish Defence Forces at Defence Command, Public Information Division. He is also leader of the Combat Camera capability project and information operations planner.

Harry Kantola Major, (General Staff) Kantola has studied (Masters of arts) in Swedish National Defense College (SNDC) and to General Staff Officer at Finnish National Defense University (FNDU) 2011. Currently he is carying out doctoral studies and teaching at the department of tactics and op-erational arts at the same university.

Captain, M.Sc. (Eng.) Anssi Kärkkäinen graduated from the Finnish National Defence University in 2000. He also graduated a Master of Science (Engineering) degree from Helsinki University of Technology (currently Aalto University) in 2005. Currently he is carrying out his doctoral studies at the same university. His current assignment is a Staff Engineer for Defence Command Finland.

Anthony Keane, MSc, PhD is currently a lecturer in the Institute of Technology Blanchardstown, Dublin and principal investigator of the research group in Information Security. Anthony is a board member of the Irish Reporting & Information Security Service (IRISS) and the Irish Chapters of the Cloud Security Alliance and the Irish Honeynet Project.

Michael Kiperberg is currently a Ph.D candidate in mathmatical information technology at the university of Jyväskylä . Michael received B.Sc (Cum Laude) and M.Sc (Cum Laude) in Computer Science from Tel Aviv University,Ramat Aviv, Israel in 2009 and 2012 respectively.Michael is currently with the israeli diffence forces

Boris Kišić lives in Croatia, and he is a second year student of graduate study of Databases and Knowledge Bases at Faculty of Organization and Informatics in Varaždin. Currently he has a title of Bachelor of Informatics. Apart from the databases, he has good skills and interest in programming and application development.

Dr. Koltuksuz received his Ph.D. with a dissertation thesis of "Cryptanalytical Measures of Turkey Turkish for Symmetrical Cryptosystems" in 1995. Currently, affiliated with Computer Engineering of Yaşar University of İzmir Turkey. His research interests are Cryptology, Theory of Numbers, Information Theory, Theory of Computation, Operating Systems, Multicore Architectures, Cyberspace Defense & Security, Open Sources Intelligence Analysis and of Computer Forensics.

Dr. Erkki Kurkinen has a long history in the industry of mobile information technology. He has an excellent track record in developing products and services both professional and consumer use. He is a researcher in the University of Jyväskylä in the Department of Mathematical Information Technology. His research area covers technology acceptance and cyber security.

Jaana Kuula works as a project manager at the Department of Mathematical Information Technology of the University of Jyväskylä and has a 30 years' experience in the field. Currently she works in the fields of crisis management, mobile emergency communication and forensics with the Police and other security authorities in Finland.

Professor Rauno Kuusisto has contributed as an expert, consultant and manager roles particularly on the areas of information availability in strategic decision-making, strategy thinking, product development, research purchasing, project portfolio management, network enabled management and leadership in innovative environment, systems thinking, as well as modeling comprehensive challenges. He has published about fifty academic papers, edited books and research reports. At the moment he is working at the Finnish Defence Forces.

Tuija Kuusisto holds a PhD degree in Geoinformatics. She has over 12 years work experience in local and national IT administration and over 10 years work experience in national and global software business and telecommunications industry. She is a special expert at Ministry of Finance and an adjunct professor of information management for decision-making at National Defence University in Finland. Her current research interests lay around information management and security. She has about 60 scientific publications in international and national journals, conference proceedings and books.

Michael Kyobe is A/Professor of Information System. He holds a PhD in Computer Information Systems and an MBA. Michael worked as a project manager and IT manager for several years and has consulted extensively with the public, and SMEs. His research interests include business-IT alignment, governance, computer security, ethics, knowledge management and SMEs.

Aubrey Labuschagne started his career in tertiary education in the fields of programming, networking and security. In 2010 he moved to the CSIR and since then contributed on various projects within Cyber Defence. Currently he is a technology researcher in the fields of social media, social engineering and cyber awareness. He is currently completing his masters on how to improve the effectiveness of security awareness programs.

Dr. (Military Sciences) Col (ret.) Martti Lehto has over 30 years of experience in the Finnish Air Force. He is a researcher in the University of Jyväskylä (the Department of Mathematical Information Technology). He has about 25 publications and research reports on areas of C4ISR systems, cyber security and defence, IW, defence policy. Since 2001 he has been the Editor-in-Chief of the Military Magazine

Áine MacDermott is a PhD research student studying at Liverpool John Moores University in the School of Computing and Mathematical Sciences. She achieved a 1st class BSc. in the field of Computer Forensics. Her current research towards her PhD is focusing on protecting critical infrastructure services in the cloud environment.

José Carlos Lourenço Martins Lieutenant-Colonel and Professor at Military Academy in the Area of Information Warfare and Information Security Management. Computer Science Engineering Degree, Military Science Degree (five years courses), Masters Degree in Information Systems and PhD Student in Technology and Information Systems at the University of Minho.

Dr. Montgomery 'Mitzy' McFate is the Minerva Chair at the US Naval War College in Newport, Rhode Island. Formerly, she was the Senior Social Scientist for the US Army's Human Terrain System. Dr. McFate received a B.A. from UC Berkeley, a PhD in Anthropology from Yale University, and a J.D. from Harvard Law School.

Karie Nickson received the BSc degree with honors in Information Technology from Kurukshetra University (India) in 2005 and MSc degree in 2007 from Sikkim Manipal University of Health and Technological Sciences (India). Currently I am a PhD student at the University of Pretoria South Africa focusing on Digital Forensics.

Paweł Niziński is an R&D scientist at NATO CCDCOE in Tallinn, Estonia. Pawel's research intrests include network intrusions, network and computer forensics, intrusion detection and prevention systems. As specialist he spent past few years implementing, maintaining IT security solutions, specifically working with Microsoft-based directory services, e-mail, terminal services. Currently his main area of responsibility is Computer Forensics.

Riku Nykänen obtained his masters degree in computer science at University of Jyväskylä. Currently he is carrying out doctoral studies at the same university. He is consultant and partner in ICT consulting company and working on information security domain.

Abimbola A. Olabelurin is a PhD researcher at City University London and his research works his supported by British Telecommunications. He obtained B.Sc. in Electronic Engineering from Obafemi Awolowo University, Ile-Ife, Nigeria in 2006 and M.Sc. in Telecommunications and Networks from City University London in 2010. He is currently with Centre for Cyber and Security Science at City University London. His research interests include data mining and machine learning.

Grant Oosterwyk is a part time Masters student studying Information Systems at the University of Cape Town. He received his BTech (Honours: Information Technology) from Cape Peninsula University of Technology with a major in Communication Networks. His currently employed as a full time Network Engineer. His research focuses on Social Media, Mobile Technology and Mobile Security.

Professor Evgeny N. Pashentsev is a research supervisor of the specialization “Communication management” at Lomonosov Moscow State University. He is also a head of the Communication Management Centre at the Russian-German Graduate School of Management – the Faculty at the Russian Presidential Academy of National Economy and Public Administration.

Jyri Rajamaki received his D.Sc. degree in electrical and communications engineering from Helsinki University of Technology in 2002. Since 2006 he has headed Laurea’s NETWORKS lab. His research interests include ICT systems for safety and security services. He has authored over 70 scientific publications. He has acted as a scientific supervisor for several research projects.

Dr Muttukrishnan Rajarajan leads the Information Security Group at City University London. He has published more than 100 journal and conference papers. His research focus has been in the areas of mobile security, identity management, intrusion detection and cloud security. He is a Senior Member of IEEE and also acts as an external advisor to Goverment of India Cyber Security Research labs.

Dr. Jari Rantapelkonen is currently serving at the Finnish National Defence University’s Department of Operational Art and Tactics. In 2006 he authored an award winning PhD thesis on The Narrative Leadership of War. Professor Rantapelkonen’s current research focuses on theoretical and practical issues of future warfare.

Amit Resh is an MSc. candidate in Information Technology from the University of Jyväskylä and an MBA and B.Sc. in Computer Science from Technion University in Israel. Amit worked as a Program Manager for Apple Israel. He has extensive experience in business development and software R&D specializing in Internet Communications and Security Protocols.

Eng. Francisco Ribeiro, 2009: Eng. degree in Informatics at the University of Minho, Portugal, 2011: MSc degree in Informatics at the University.

Simon Robin is from ESIEA, IT and electronics engineering school, France. His passion is to study, read information and news, learn and find different ways how something works about one particular subject - the information security. There are as many things to discover as possibilities. He also likes to be exposed to a problem to solve it, to have objectives to complete a good work in a team.

Char Sample has over 18 years of experience in the information security industry, and presently works for CERT at Carnegie Mellon University where she supports various cyber efforts. Presently, Ms. Sample is a doctoral candidate at Capitol College in Laurel, Maryland, and her research area is the “Applicability of Cultural Dimensions in Computer Network Attack Attribution”. Other areas of research interest include: Cloud Computing, Anomaly Detection methods, and DNS.

Prof. Dr. Henrique Santos 1984: Eng. degree in Electronic Engineering, (Computer Systems option) at the University of Coimbra, Coimbra, Portugal.1988: Academic degree equivalent to the MSc., at the University of Minho, Portugal.1996: Ph.D, on Computer Engineering: "Specification Methodologies and Analysis of Digital Systems: development of am APA controller (GLiTCH)", also at the University of Minhorsity of Minho, Portugal

Libor Sarga is a doctoral worker at the Department of Statistics and Quantitative Methods, Faculty of Management and Economics, Tomas Bata University in Zlín. His dissertation is focused on computer and data security in the presence of unreliable human element as an exploitable attack vector. His personal interests include technology and practical security applications.

Mr. Alexander Semenov holds a M.Eng. degree from Saint-Petersburg State Electrotechnical University, Saint Petersburg, Russia. He is pursuing his Ph.D. degrees at the National Research University ITMO, Saint-Petersburg, Russia (Dept. of Computer Technologies), and at the University of Jyväskylä, Finland (Dept. of Computer Science and Information Systems), since 2009. His research interests are in databases, social network analysis, and software tools development for it.

Nathan Shone received a BSc (Hons) in Computer Forensics from Liverpool John Moores University. He is currently a PhD student of Network Security and a member of the PROTECT research centre at Liverpool John Moores University. His current research focuses on security monitoring in System-of-Systems and his research interests include network, computer and mobile security.

Paulo Simões is a Professor at the University of Coimbra. His main research interests are Security, Network Management and Critical Infrastructures. He has over 100 journal and conference publications in these areas. He has been involved in several European research projects. He also participated in several industry-funded research projects in these areas.

Lior Tabansky is a Researcher at the Yuval Ne'eman Workshop for Science, Technology and Security, exploring political aspects of information technologies. He is a Ph.D. candidate at the Department of Political Science at Tel Aviv University, examining Cyberspace in comparative national security perspectives.

Risto Vaarandi received his PhD degree in Computer Engineering from Tallinn University of Technology, in June 2005. Since May 2006, he has been holding a position of a scientist at NATO CCDCOE. Risto's research interests include event correlation, data mining for event logs, network security, and system monitoring.

Namosha Veerasamy has obtained a BSc:IT Computer Science Degree, and both a BSc: Computer Science (Honours Degree) and MSc: Computer Science with distinction from the University of Pretoria. She is currently employed as a researcher at the Council for Scientific and Industrial Research (CSIR) in Pretoria. Namosha is also qualified as a Certified Information System Security Professional (CISSP) and Certified Information Security Manager (CISM).

Dr. Suresh Veluru is a Research Fellow in the School of Engineering and Mathematical Sciences at City University London, United Kingdom. Prior to this, he worked as a Research Fellow in Computer Science at University of York, United Kingdom and University of New Brunswick, Canada. He received his PhD in Computer Science from Indian Institute of Technology Guwahati, India in 2009. His research interests include data mining, natural language processing, and privacy preserving data mining.

Prof. H.S. Venter is an Associate Professor and research group leader at the Information and Computer Security Architectures research group at the University of Pretoria's Department of Computer Science. He holds a PhD in Computer Science from Rand Afrikaans University (now University of Johannesburg). His research interests include network security, intrusion detection, information privacy, and digital forensics.

Critical Infrastructure Cybersecurity: A Review of Recent Threats and Violations

Nasser Abouzakhar
School of Computer Science, College Lane, University of Hertfordshire, Hatfield, UK
N.Abouzakhar@herts.ac.uk

Abstract: Most of current industries and their critical infrastructure rely heavily on the Internet for everything. The increase in the online services and operations for various industries has led to an increase in different security threats and malicious activities. In US, the department of homeland security reported recently that there have been 200 attacks on core critical infrastructures in the transportation, energy, and communication industries (Erwin et al., 2012). This paper is concerned with the growing dependence of modern society on the Internet, which has become an ideal channel and vital source of malicious activities and various security threats. These threats could have an impact on different distributed systems within and across all the critical infrastructures, such as industrial networks, financial online systems and services, nuclear power generation and control systems, airlines and railway traffic controllers, satellite communication networks, national healthcare information systems … etc. The major problem is that the existing Internet mechanisms and protocols are not appropriately designed to deal with such recently developed problems. Therefore, a rigorous research is required to develop security approaches and technologies that are capable of responding to this new evolving context. This paper presents various security threats and incidents over the past recent years on different critical infrastructure domains. It introduces some security measures including vulnerability assessment and penetration testing approaches for critical infrastructure.

Keywords: critical infrastructure protection, cyber security, security threats and violations

1. Introduction

Critical infrastructure cyber security is concerned with the protection and response to malicious activities that involve the critical infrastructure of a particular country. It is about the protection of electronic systems from malicious electronic attack and the means of dealing with such attacks. Critical infrastructure cyber security comprises technical, operational and managerial activities, and relates to the application processes, electronic systems and to the information stored and processed by such systems. During recent years the context of cyber security threats to critical infrastructure has changed dramatically as the Web and Internet technologies have driven the global expansion. In Europe, the European Programme for Critical Infrastructure Protection (EPCIP) is concerned with the protection of critical infrastructure in the EU. The EPCIP developed a procedure for identifying and designating European Critical Infrastructure (ECI), which is implemented by the European Commission’s directive EU COM (2006) 786. This directive indicates that European critical infrastructure represents a situation that in case of a security incident or violation, which may affect a hosted country and at least one other European Member State.

Critical infrastructure systems are increasingly being targeted by attackers. This is due to the fact that most of such systems rely on weak security mechanisms. Cyber security threats include such issues as energy and power generation failures, online banking systems malfunction, transportation accidents, and hazardous material accidents. Figure 1 shows different infrastructure that were commonly referred to as “critical”. In December 2011, the FBI’s cyber division released the news that the infrastructure systems of three US cities have been attacked. FBI reported that hackers hit key services and had accessed crucial water and power services (BBC News, 2011)

"We just had a circumstance where we had three cities, one of them a major city within the US, where you had several hackers that had made their way into SCADA systems within the city.” and “Essentially it was an ego trip for the hacker because he had control of that city's system and he could dump raw sewage into the lake, he could shut down the power plant at the mall - a wide array of things"

In 2010, another major security violation incident took place that was the spread of Stuxnet malware. Stuxnet is a complex piece of malware believed to be the first to target a real critical infrastructure such as nuclear power station. It is considered as one of the most sophisticated worms ever detected that uses six different methods that allowed it to spread (Fildes, 2010). Unlike most malware, Stuxnet aims to target specific industrial control systems that are traditionally not connected to the internet for security reasons using USB keys. It is designed to spy on and reprogram industrial control systems and to seek out a specific configuration of Siemens made SCADA (Supervisory Control and Data Acquisition) systems. Once SCADA system is hijacked by Stuxnet, the worm can reprogram PLC (Programmable Logic Controller) to give new instructions to linked machine. This is to cause damage to motors used in uranium-enrichment centrifuges. The PLC is an electronic device that generates control signals, for example, it monitors temperature and turn on coolers if a gauge exceeds a certain temperature as part of an industrial process. Stuxnet is able to inject code into the ladder logic of PLCs, monitor Profibus protocol and then manipulates the operations of the PLC to interrupt processes and modify output (Knapp, 2011). Stuxnet is a kind of malware that cannot be detected until it has been deployed and it infects parts of the control system that is uneasy to monitor. Therefore, security professionals need to change their perception and attitude toward critical infrastructure security to be able to deal with such malicious incidents (Symantec, 2010).

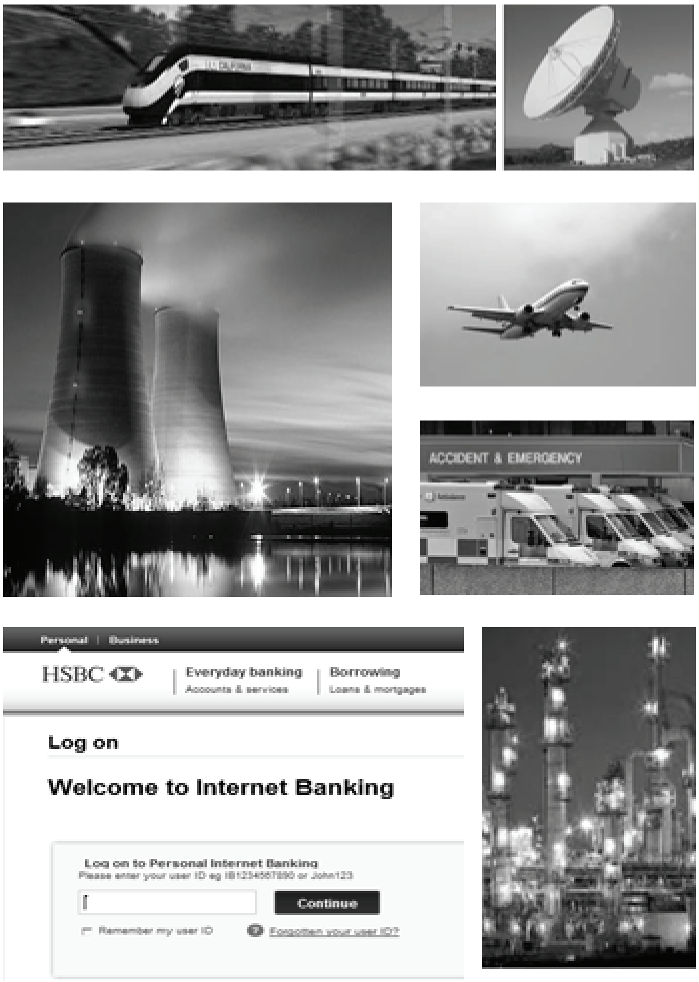


Figure 1: Examples of critical infrastructure

In 2010, Symantec carried out a critical infrastructure protection study. This study included 1,580 private businesses that are involved in industries that are considered providers of critical infrastructure services. The respondents are companies from 15 countries worldwide, with median company had between 1,000 and 2,499 employees. Figure 2 shows the results of one of the companies' responses to a question about the company's experience with four different types of attacks (Symantec, 2010). The results show that average of only 29% were completely sure these attacks never happened in their companies. The rest i.e. about 71% were either not completely sure, suspect or pretty sure that those attacks have happened to their companies. Such statistics indicate that there is a lot of work needs to be done by all parties involved including management, security professionals, governments … etc. to improve the situation.

Security experts predict that there will be an increase in such attacks and malicious activities due to lack of knowledge about cyber security threats and lack of proper security measures. Therefore, proper cyber security training and intelligent security measures need to be considered in order to be able to address recent sophisticated kind of threats. This includes monitoring application sessions and defining up to the level security policies to control all internal processes and communications (Knapp, 2011). In Europe, Member States are pursuing Critical Infrastructure Protection (CIP) initiatives aimed at working with different organisations and industries to address cyber security threats.

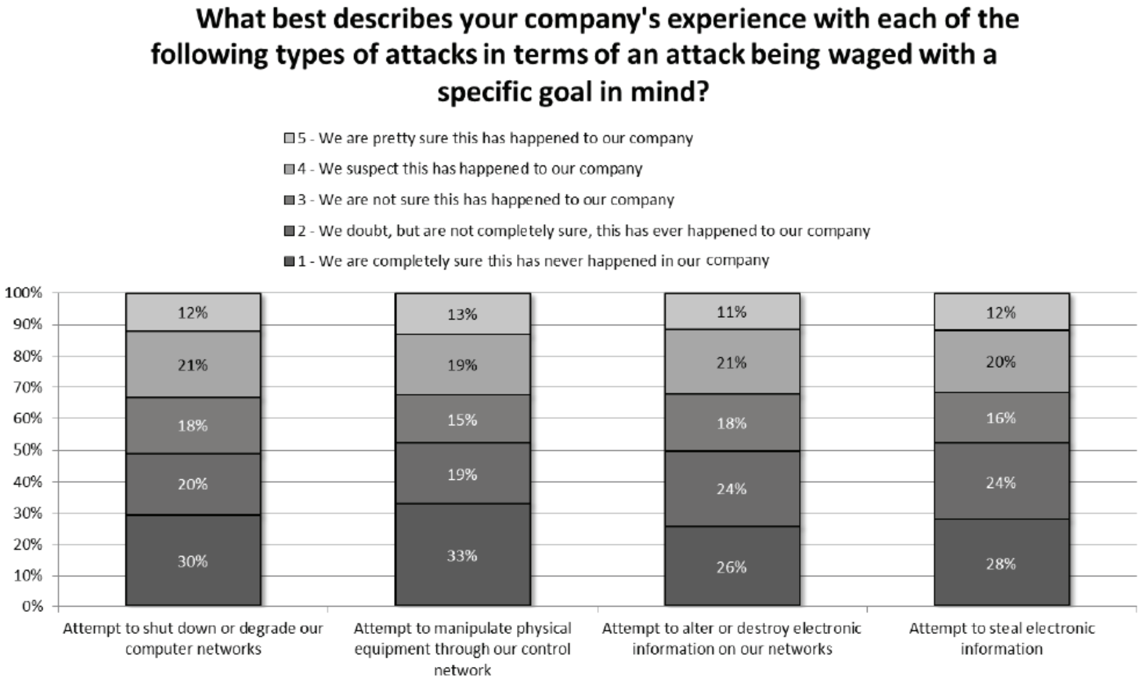


Figure 2: Symantec survey of critical infrastructure based companies in 2010

2. Critical infrastructure security threats

Critical infrastructure represents a system or a number of systems that perform critical functions and operations. Such systems are considered critical if they could impact any other critical processes and/or devices, or provide a pathway/channel to other critical system(s), or are used to protect critical systems (Knapp, 2011) (US (NRC), 2010). Figure 3 shows a general logical diagram provided by the US. NRC (Nuclear Regulatory Commission) for identifying critical systems (US (NRC), 2010).

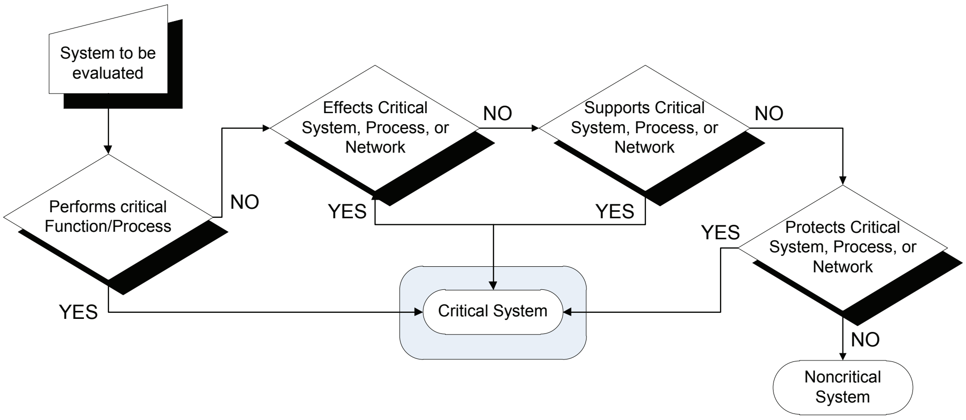

System to be evaluated
Performs critical Function/Process
NO
YES
Effects Critical System, Process, or Network
NO
YES
Supports Critical System, Process, or Network
NO
YES
Protects Critical System, Process, or Network
YES
NO
Critical System
Noncritical System

Figure 3: NRC flow diagram for identifying critical systems

Manipulating a particular process in a critical system could cause certain threshold levels to build beyond safe operating parameters which then could result in loss of life and/or loss of critical services. Such a manipulation event could be performed using a Man-in-the-Middle attack to change control process parameters and its feedback loop using a targeted malware. For example, a successful cyber-attack can block, delay or manipulate the intended operation, thereby preventing a service provider from generating necessary energy output or from obtaining production metrics. This section presents various security threats and violations over the past recent years on different critical infrastructure domains including industrial networks, healthcare services, telecommunication networks, and banking systems.

2.1 Industrial networks

An industrial network performs an operational process of a control or manufacturing system to carry out a particular operation. It consists of a supervisory network, business network of enterprise operations and control process networks (Knapp, 2011). The increasing persistence and sophistication of attacks on industrial networks in general and energy systems in particular requires effective solutions that are capable of mitigating such attacks. In 2009, the International Chief Security Officer (CSO) of the American Society for Industrial Security (ASIS) reported that (Ghansah, 2009)

> *"The electric grid is highly dependent on computer-based control systems. These systems are increasingly connected to open networks such as the internet, exposing them to cyber risks."*

Various entities such as hacking individuals, organizations and even states are involved in probing U.S. power grid systems on a daily basis. In 2009, the Department of Homeland Security (DHS) has reported that (Ghansah, 2009)

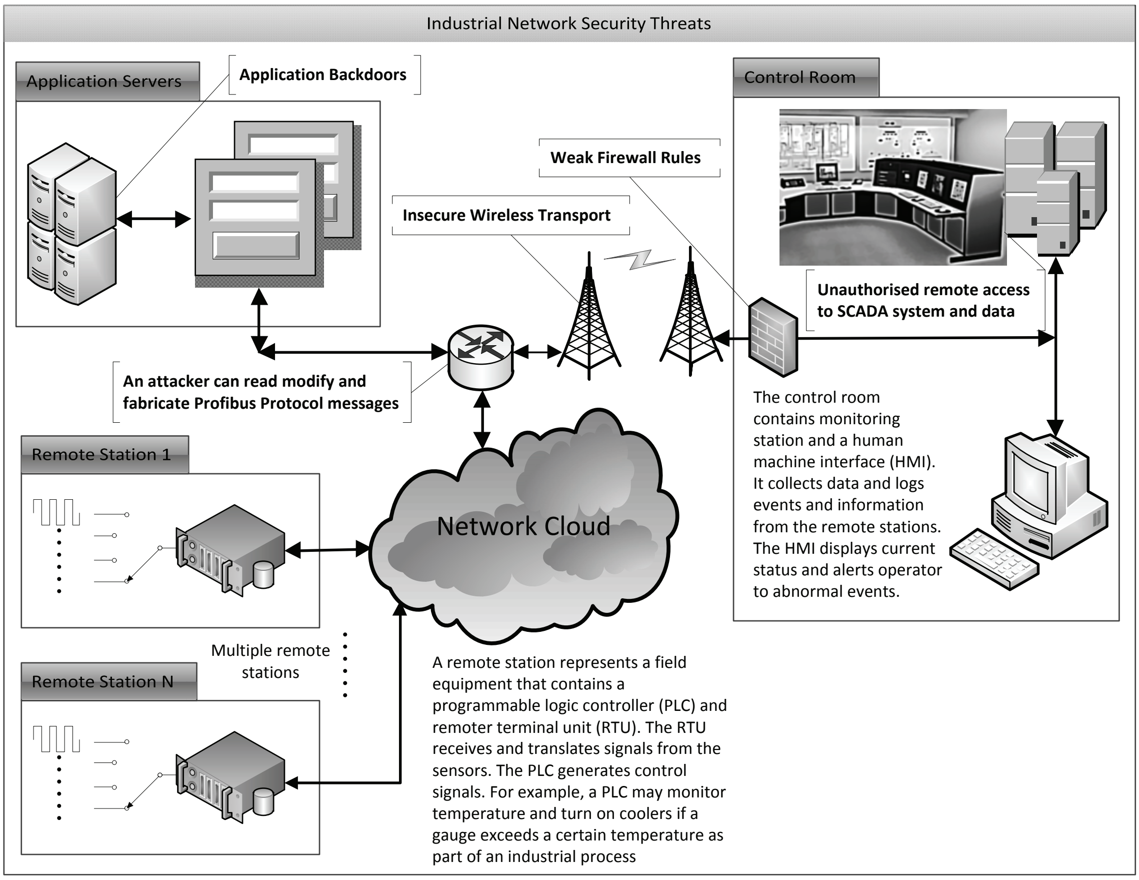
> *"Cyber spies, likely from China and Russia, have managed to inject malicious software into the electric grid, water, sewage, and other infrastructure control software. This software could enable malicious users to take control of key facilities or networks via the Internet, causing power outages and tremendous damage to all sectors of the economy."*

Satellite imagery of nuclear power stations and power grids can easily be located online using Google map. Online vulnerable systems and components such as unsecured servers, SCADA systems and network resources are remotely accessible to anyone with an Internet connection and with a basic knowledge of using attacking tools. The SQL Slammer Worm is one of those tools that are able to disrupt electric system control systems. Cyber-attacks incidents could result in shutting down portions of power plants, breaking into electrical utilities, disturbing cities lights and electricity, grid failures or catastrophic problems (Vaas, 2012) (Andres and Loudermilk, 2012) (Ghansah, 2009). IOActive discovered security vulnerability in many Smart Meters, where a malware managed to spread quickly throughout a neighbourhood, affecting the electric system controls, causing power disruptions and calibration modifications rendering the power meters inoperable (Davis, 2009).

Industrial networks have moved towards more effective mechanisms of managing industrial systems such as power generation and distribution. Such systems have become to rely on networked SCADA systems that use network protocols and about 85% of all analogue relay systems such as meters, demand response systems, control systems … etc. are now digital (Andres and Loudermilk, 2012). Most of the industrial network protocols are sensitive to DoS attacks that using a significant amount of overwhelming traffic could lead to protocol failure. Improper digital network configurations often lead to information leaks between SCADA systems, business networks and the Internet and pose a significant threat to network reliability. Network information leaks can allow worms and/or hackers to disabling safeguards and have a direct access to vulnerable SCADA systems (Ghansah, 2009). The end result could be taking a service offline, production failures, financial losses, life-threatening incident due to misinformation.

The SCADA systems are built using public or proprietary communication protocols such as Profibus. Those protocols are used for communicating between an MTU (Master Terminal Unit) and one or more RTUs (Remote Terminal Units). The SCADA protocols provide transmission specifications to interconnect master station and substation computers, RTUs, IEDs (Intelligent Electronic Devices) (Ghansah, 2009). Profibus is one of the common industrial protocols which were developed to achieve interoperability among systems in the energy utility. An attacker with the appropriate network reconnaissance techniques can access captures and analyses Profibus messages. This attack provides the attacker with information about network topology, device functionality, memory addresses and other data. A hacker can launch a replay attack with knowledge of normal Profibus traffic patterns simulates responses to the master while sending fabricated messages to outstation devices (Ghansah, 2009). Figure 4 shows some common industrial network vulnerabilities and security threats such as poor firewall configurations, insecure wireless links, weak control access mechanisms, remote access vulnerabilities … etc.

As an intelligent malware Stuxnet manages to inject code into the ladder logic of PLCs, monitors Profibus protocol and then manipulates the operations of the PLC to interrupt control processes and modify operation results. Profibus is an industrial protocol developed by the Central Association for the Electrical Industry in Germany. It is a Master/Slave protocol that uses a token for communications between a master and one or more slaves. A master Profibus node represents a PLC or RTU and a slave is sensor or other control system device. One of the major limitations of Profibus is lack of authentication to many of its functions allowing for unauthorised control over all slaves. This could result in disrupting the protocol functions or injecting code into a slave node. Stuxnet is able to exploit Profibus and compromise a PLC as a master node allowing Stuxnet to issue commands to the relevant slave nodes to sabotage the process (Knapp, 2011) (Jonathan, 2010).


Industrial Network Security Threats
Application Servers
Application Backdoors
Control Room
Weak Firewall Rules
Insecure Wireless Transport
Unauthorised remote access to SCADA system and data
An attacker can read modify and fabricate Profibus Protocol messages
Remote Station 1
Network Cloud
Remote Station N
Multiple remote stations

The control room contains monitoring station and a human machine interface (HMI). It collects data and logs events and information from the remote stations. The HMI displays current status and alerts operator to abnormal events.

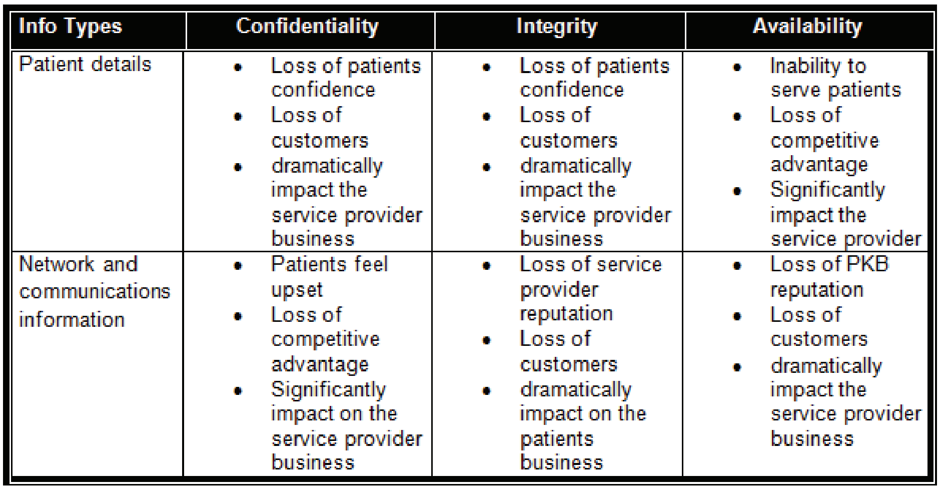
A remote station represents a field equipment that contains a programmable logic controller (PLC) and remoter terminal unit (RTU). The RTU receives and translates signals from the sensors. The PLC generates control signals. For example, a PLC may monitor temperature and turn on coolers if a gauge exceeds a certain temperature as part of an industrial process

Figure 4: Industrial network security threats

2.2 Healthcare

The emergence of Web-based Healthcare applications has generated various risks to patients information security. Malicious software and operations pose a major threat to the security of EPHI (Electronic Patient Healthcare Information), especially those supporting medical identity theft and healthcare fraud. Moreover, the proliferation of handheld devices, such as smart phones, has created an environment in which patients' wireless communications and healthcare staff emails can be intercepted. Lack of effective policies and security controls by healthcare service providers poses a security risk in terms of accessibility to patients' files, such as valid diagnosis and treatment information. Recent developed malware is able to exploit various healthcare system vulnerabilities and continue to grow. Such problems could have a negative impact on patients and affect the proper use of their medication and drugs. This makes healthcare service providers in general, and hospitals and patients in particular, at risk. The critical information attributes which have an impact on a healthcare service provider operations are patients' details/information and network communication information. Table 1 includes definition for the assigned impact levels and their possible effects on a healthcare service provider. Definition of impact levels in table 1 are meant to help healthcare service provider's management team to understand that the loss of security attributes to those pieces of information (patients' details and network and communications information) can impact the service provider in different ways and degrees.

Table 1: Definition of impact levels



Info Types	Confidentiality	Integrity	Availability
Patient details	• Loss of patients confidence • Loss of customers • dramatically impact the service provider business	• Loss of patients confidence • Loss of customers • dramatically impact the service provider business	• Inability to serve patients • Loss of competitive advantage • Significantly impact the service provider
Network and communications information	• Patients feel upset • Loss of competitive advantage • Significantly impact on the service provider business	• Loss of service provider reputation • Loss of customers • dramatically impact on the patients business	• Loss of PKB reputation • Loss of customers • dramatically impact the service provider business

Healthcare service providers are leveraging the networked nature of the Internet and want to take the full advantage of the Internet and distributed computing to serve their customers/patients. Nowadays healthcare services providers are connected to the Internet, have various systems to worry about and are facing an increased number of vulnerabilities and security threats. Such threats represent conditions with a potential to cause damage to an organisation's business and/or system resources including databases and communication links. Threats may come from a system's vulnerabilities, unauthorised access, an insider performing illegitimate activities, natural disasters such as earthquakes, flooding, storms, lightning … etc. Threats are divided into two types, external security threats and internal security threats. Examples of external security threats include denial of service (DoS) attacks, remote brute-force, man-in-the middle attack. Password sniffing, Trojan horses, Data tampering are examples of internal security threats. Such attacks are a direct threat to the confidentiality, integrity, and availability of a healthcare service provider's information assets. The HIPAA (Health Insurance Probability and Accountability Act) security rules and procedures have introduced various solutions to minimize such threats and risks.

2.3 Telecommunication networks

Computer networks, satellite communication systems and links allow unauthorized users to gain access to private information and critical resources. The satellite networks represent one of the major communication systems that face significant security challenges. Attacks such as DoS (Denial of Service) on satellites could cause business and military communications to become unavailable at critical moments and prevent legitimate clients from accessing necessary service(s). Space assets satellite systems are increasingly vulnerable to various attacks, such as RF jamming and network traffic spoofing. Jamming involves intentionally masking a target signal with another RF signal using a little jamming power, which can result in a signal degradation or total signal loss. Spoofing involves transmitting false information to the satellite in order to overpower the intended signal. This is to send the receiver a malicious signal and fooling it into using a false signal for further processing (Northcutt, 2007). Such attacks could lead to disruption of communications and prevention of service access.

To attack a satellite does not require a state space capability. The Tamil Tigers Liberation Front (LTTE), a Sri Lankan separatist group classified as terrorist group, has recently been blamed for illegally using INTELSAT satellites. The LTTE used INTELSAT to broadcast radio and TV transmissions via the use of an empty transponder (Ma et al. 2010). Satellite transponders represent the access points that are configured to retransmit any signal being sent to them and are susceptible to various vulnerabilities. If a transponder has unused bandwidth, a hijacker could easily identify a vacant place on the transponder, using a spectrum analyzer, to broadcast their own transmissions. An attacker can create a DoS condition by turning on an uplink carrier with a great enough signal-to-noise ratio (SNR) into the victim's satellite on the same frequency as the intended signal (Daly, 2007). The victim transponder processes the incoming carrier frequency along with the intended signal and re-transmits both of them down to the receiver. The attacker's signal may impair the intended signal at the receiver making it unable to distinguish the intended signal from the attacker's signal. Moreover, the attacker's signal raises the background noise of the transponder and causes a reduction in the

SNR of all the intended signals, which makes them uneasy to recover. With today's DSP (Digital Signal Processing) systems it is becoming trivial to launch such attacks. The uplink signal from the hijacker is transmitted to the satellite in a highly directed beam, which makes finding the attacker extremely difficult (Daly, 2007). This ability to hack into commercial satellites could lead to a disastrous situation in global communications. The development of effective security models and solutions is a viable response to the rapidly increasing number of malicious activities on satellite systems and Internet services.

3. Security measures

It is important for organisations managing critical systems to deploy all necessary security solutions and carry out a regular security assessment and auditing. Security assessment must be carried out by experienced professionals to gather information and to perform necessary tests. The security assessment is a process that includes interview with key personnel managing the critical system, review of available systems and documentation and carry out comparisons with relevant standards. A security audit is a systematic technical evaluation of an organisation system(s) and service(s) by measuring how well they confirm to standards and guidelines provided by different organisations such as British Standards (BSs), OSI, NIST … etc. Through these efforts, the assessment team should be able to plan a strategy to identify security vulnerabilities and propose solutions to meet the critical infrastructure's security needs. The assessment process includes reviewing the security requirements for critical network architecture and addressing the issues of end-to-end communication infrastructure uniquely pertaining to critical system communications and access control mechanisms. The assessment process should consider technical standards in studying and specifying the requirement details and evaluation of system architecture and services in terms of meeting the necessary security needs. Effective security assessment includes vulnerability assessment and penetration testing services which must be performed regularly to suit critical infrastructure systems as follows:

- Vulnerability assessment is concerned with the evaluation of network configurations, firewalls, vulnerable critical services and/or systems ... etc. using vulnerability scanners. Vulnerability scanners are useful in terms of ensuring the security of services and systems. For example, vulnerability scanners could be used to determine if there are any unauthorised activities are occurring or information leakage is taking place in a power grid network or in a SCADA system. Figure 5 shows a proposed flow chart for a vulnerability assessment to critical infrastructure.

- Penetration test is carried out to perform an external penetration test on all the network systems including servers, databases, communication links ... etc. In general, the testing team should not be given any prior information about the network architecture.

Such assessment should involve using necessary tools to evaluate, test and analyse the security operations and infrastructure. It also should define the countermeasures to security threats and violations. Stallings (2011) had identified the necessary countermeasure to major Outsiders and insiders' threats/attacks. Some of the attacks are common for both outsiders and insiders' threats. In Europe, Member States are required to conduct a security assessment of the threats and violations relating to the designating ECI. Member States have to report to the EC every two years on the security threats, risks and vulnerabilities the various European Critical Infrastructure (ECI) services are facing (The EC directive, 2008). The final outcomes and results of the assessment and auditing are intended to provide the organisation with recommendations to improve security. As many organisations are migrating to cloud services, much of their infrastructure will now be controlled by a third-party Cloud Service Provider (CSP). The extensive use of virtual machines (VMs) in developing cloud infrastructure presents various security concerns for organisations as customers of a public cloud service (Winkler, 2012). Migration to cloud environment brings unique security challenges to critical systems such as virtual threats. Virtualisation uses more complex processes than traditional systems and DoS attacks to VMs have become equally more complex. Therefore, relying on protection techniques traditionally implemented against DoS is insufficient. The operating system (OS) vulnerabilities on the host system can flow upward into the VM OS. Therefore, a compromise of the host OS would allow an attacker to access all VM processes and services. Table 2 lists the major security Cloud-siders threats and necessary countermeasures (Winkler, 2012) (Krutz and Vines 2010) (Winkler, 2011).

Effective critical systems security is based on various factors such as proper policies, procedures, management support, and appropriate implementation. Therefore, critical infrastructure requires a comprehensive security policy that details not only physical security requirements but also includes information protection and systems security considerations. This includes preparation of an acceptable-use policy addressing the appropriate use of corporate technology resources and the actions management will take resulting from the violation of this policy. Password policies are important to specify when passwords must be used, how strong they must be, and how they must be stored and processed. The lack of a password policy and appropriate password controls could lead to unauthorized access to systems and information. Polices are required to conform to specific professional organization such as the OSI's regulations applicable to the communication or use of data. Security policy is the foundation of critical systems security plan and implementation. The lack of security policy could have a negative impact on a critical service provider's performance and ability to meet the industry standards and regulations. In Europe, the EPCIP developed a procedure for identifying critical assets of the European Critical Infrastructure (ECI), which is implemented by the European Commission's directive 2008/114/EC. This directive insists that Member States must ensure that an operator security plan (OSP) is in place for each designated ECI to identify the critical assets of the ECI and the available security policies and measures for protecting them (The EC directive, 2008).

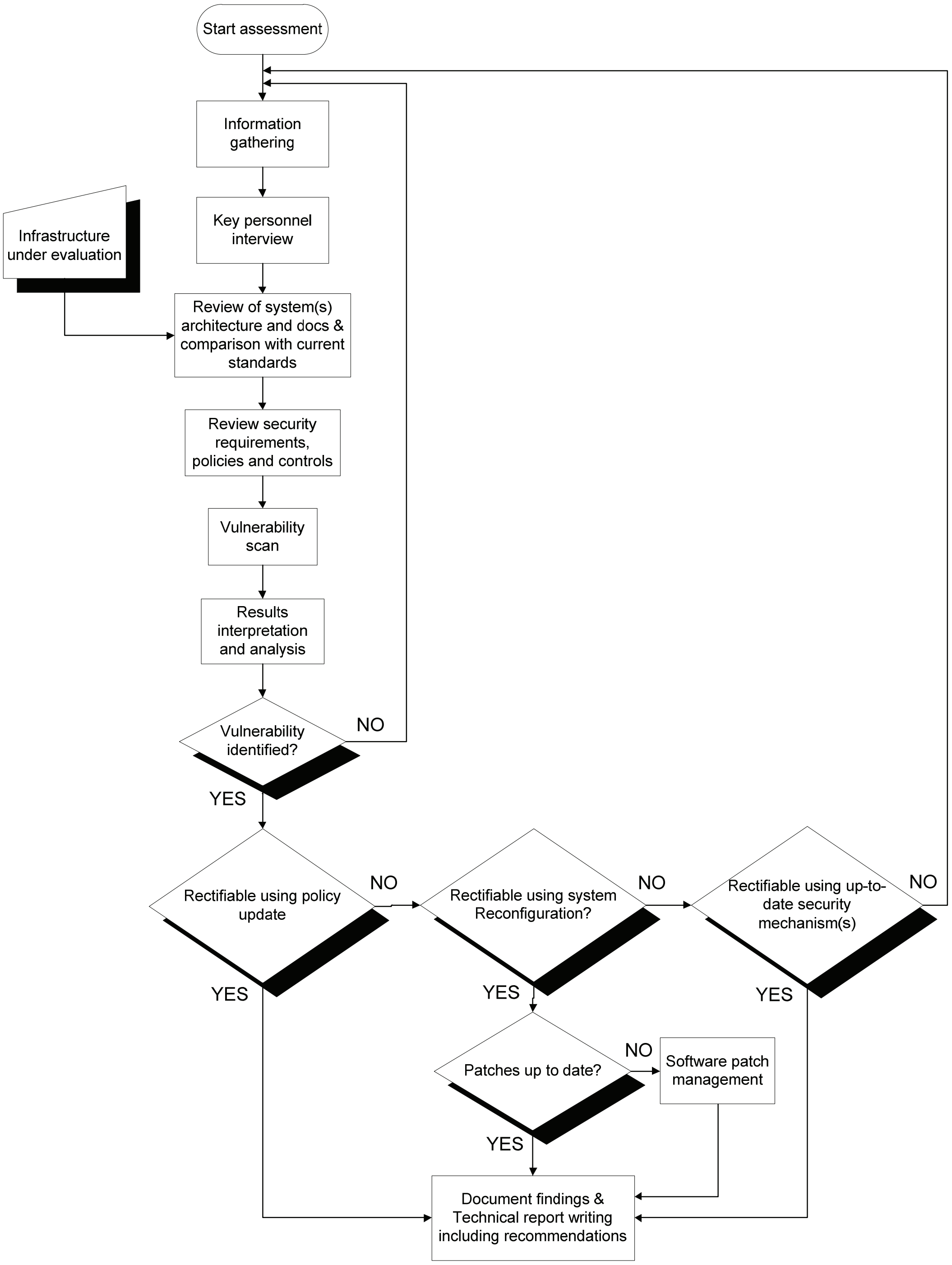
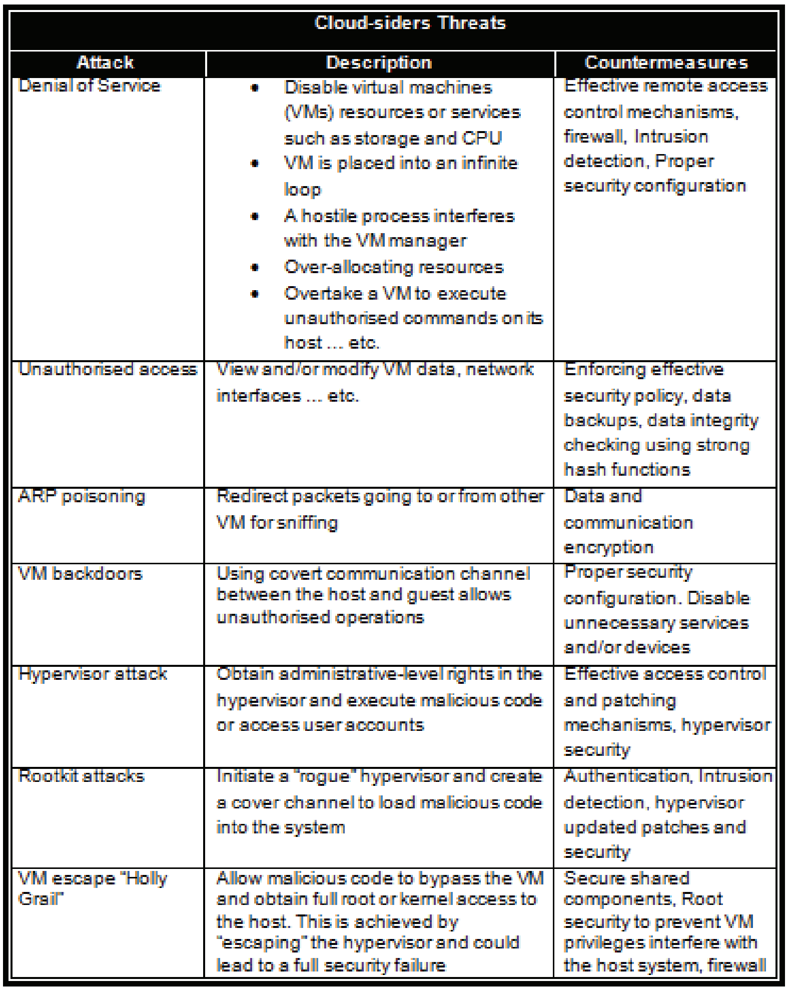

Start assessment
Information gathering
Key personnel interview
Infrastructure under evaluation
Review of system(s) architecture and docs & comparison with current standards
Review security requirements, policies and controls
Vulnerability scan
Results interpretation and analysis
Vulnerability identified?
NO
YES
Rectifiable using policy update
NO
Rectifiable using system Reconfiguration?
NO
Rectifiable using up-to-date security mechanism(s)
NO
YES
YES
YES
Patches up to date?
NO
Software patch management
YES
Document findings & Technical report writing including recommendations

Figure 5: Vulnerability assessment to critical infrastructure

4. Conclusion

Due to the fact that different critical infrastructure systems reply on weak security mechanisms, such systems are increasingly targeted by attackers. Various complex malwares have been developed to target improperly protected critical infrastructure. Critical infrastructure service providers must seek implementing cost-effective and comprehensive secured solutions for their system operations. Various critical industries have demonstrated a desire to ensure the security of their systems architecture and infrastructure. This paper provided a brief survey to recent security threats and vulnerabilities to different critical infrastructure systems including industrial networks, healthcare systems, telecommunication services and online banking. For example, Stuxnet is able to inject code into the ladder logic of PLCs, manipulates the operations of the PLC to interrupt processes and modify output. In this paper, various security measures have been presented including assessment process to infrastructure architecture and systems in terms of vulnerability assessment, security policies, procedures and solutions. In Europe, the European Commission's directive insists that Member States must carry out a security assessment of the threats and violations relating to the designating ECI.

Table 2: Cloud-siders threats and countermeasures



Cloud-siders Threats		
Attack	**Description**	**Countermeasures**
Denial of Service	• Disable virtual machines (VMs) resources or services such as storage and CPU • VM is placed into an infinite loop • A hostile process interferes with the VM manager • Over-allocating resources • Overtake a VM to execute unauthorised commands on its host … etc.	Effective remote access control mechanisms, firewall, Intrusion detection, Proper security configuration
Unauthorised access	View and/or modify VM data, network interfaces … etc.	Enforcing effective security policy, data backups, data integrity checking using strong hash functions
ARP poisoning	Redirect packets going to or from other VM for sniffing	Data and communication encryption
VM backdoors	Using covert communication channel between the host and guest allows unauthorised operations	Proper security configuration. Disable unnecessary services and/or devices
Hypervisor attack	Obtain administrative-level rights in the hypervisor and execute malicious code or access user accounts	Effective access control and patching mechanisms, hypervisor security
Rootkit attacks	Initiate a "rogue" hypervisor and create a cover channel to load malicious code into the system	Authentication, Intrusion detection, hypervisor updated patches and security
VM escape "Holly Grail"	Allow malicious code to bypass the VM and obtain full root or kernel access to the host. This is achieved by "escaping" the hypervisor and could lead to a full security failure	Secure shared components, Root security to prevent VM privileges interfere with the host system, firewall

Migration to cloud computing based services and environment brings unique security challenges to critical systems such as VM threats. A summary of cloud-siders threats and countermeasures are listed in section 3. The lack of proper security policy could have a negative impact on the performance of critical systems and ability of critical service providers to meet the industry standards and regulations. European Member States must ensure that an OSP is in place for each designated ECI. This is to identify the critical systems and assets as well as the available security measures for protecting them. Critical infrastructure service providers can improve their security posture by taking into consideration the proposed security measures and assessment strategy. The final outcomes and results of security assessment and auditing are intended to provide organisations with recommendations to improve security.

References

Andres, Richard B. and Loudermilk, Micah J. (2012), National Security & Distributed Power Generation. livebetter Magazine Issue Number 24, Sep 2012

BBC News, (2011), FBI says hackers hit key services in three US cities, http://www.bbc.co.uk/news/technology-16157883 [accessed on 22nd December, 2012]

Daly, John C. K. (2007), LTTE: Technologically innovative rebels, http://www.isn.ethz.ch/isn [accessed on 3rd Oct 2012]

Davis, Mike (2009), IOActive Unveils Smart Grid Security Research, http://www.ioactive.com/services_grid_research.html [accessed on 26th January, 2013]

Erwin, I. Sandra, Stew Magnuson, Dan Parsons and Yasmin Tadjdeh, November, (2012), Top Five Threats to National Security in the Coming Decade, http://www.nationaldefensemagazine.org/archive/2012/November/Pages/TopFiveThreatstoNationalSecurityintheComingDecade.aspx [accessed on 1st Oct 2012]

Fildes, Jonathan (2010), Stuxnet worm 'targeted high-value Iranian assets, BBC News, http://www.bbc.co.uk/news/technology-11388018 [accessed on 22nd December, 2012]

Ghansah, Isaac, (2009), Smart Grid Cyber Security Potential Threats, Vulnerabilities and Risks, California Energy Commission, PIER Energy-Related Environmental Research Program. CEC-500-2012-047

Knapp, Eric D. (2011), Industrial Network Security, Syngress.

Krutz, Ronald and Vines, Russell (2010), Cloud Security: A Comprehensive Guide to Secure Cloud Computing, Wiley.

Ma, Ting Hui, Yee L. and Ma, Maode (2010), Protecting Satellite Networks from Disassociation DoS Attacks, Communication Systems (ICCS), IEEE International Conference

Northcutt, Stephen (2007), Are Satellites Vulnerable to Hackers? http://www.sans.edu/research/security-laboratory/article/satellite-dos [accessed on 1st Oct 2012]

Northcutt, Stephen (2007), Denial of Service. http://www.sans.edu/research/security-laboratory/article/denial-of-service [accessed on 1st Oct 2012]

Stallings, William (2011), Cryptography and Network Security: Principles and Practices, Fifth Edition, Pearson.

Symantec (2010), Symantec Critical Infrastructure Protection Study – Global Results

The EC directive 2008/114/EC on the identification and designation of European critical infrastructures and the assessment of the need to improve their protection, http://europa.eu/legislation_summaries/justice_freedom_security/fight_against_terrorism/jl0013_en.htm [accessed on 22nd January, 2013]

US Nuclear Regulatory Commission (NRC), Regulatory Guide 5.71 (2010), Cyber Security Programs for Nuclear Facilities, Washington, DC

Vaas, Lisa (2012), Nuclear power plant cybersecurity warnings silenced by legal threats, http://nakedsecurity.sophos.com/2012/10/31/nuclear-security-silence/ [accessed on 3rd Oct 2012]

Winkler, Vic (2011), Securing the Cloud: Cloud Computer Security Techniques and Tactics. Waltham, MA USA: Elsevier. ISBN 978-1-59749-592-9

Winkler, Vic (2012), Cloud Computing: Virtual Cloud Security Concerns, Technet Magazine, Microsoft

Strategic Communication Before “Strategic Communication”: Germany’s Security Police in Estonia 1941–1944

Kari Alenius
Department of History, University of Oulu, Finland
kari.alenius@oulu.fi

Abstract: This study analyses the manifestation of strategic communication in situations and at a time when the concept did not yet exist, nor were the ideas it is based on formulated yet in the way they are today. There is no single universally accepted definition of strategic communication, but variations on the theme are fairly similar. The theoretical starting point chosen for this study is the “Principles of Strategic Communication” compiled by the U.S. Department of Defense in 2008 in which the concept and the principles of applying it are condensed into nine main categories:

(1) Leadership-Driven. Leaders must lead communication process;
(2) Credible. Perception of truthfulness and respect;
(3) Understanding. Deep comprehension of others;
(4) Dialogue. Multi-faceted exchange of ideas;
(5) Pervasive. Every action sends a message;
(6) Unity of Effort. Integrated and coordinated;
(7) Results-Based. Tied to desired endstate;
(8) Responsive. Right audience, message, time, and place;
(9) Continuous. Analysis, planning, execution, assessment.

Keywords: communication, history, Germany, Estonia, WWII

1. Introduction

Estonia was occupied by Germany from the summer of 1941 to autumn of 1944, during which time surveillance of individual’s sentiments and propaganda work were coordinated by Germany’s Security Police (Noormets, 2002, 11-15). Analysis of secret documents compiled by the Security Police shows that the organisation’s situation assessments and the operating plans compiled on that basis correspond nearly 100 percent to the principles compiled by the U.S. Department of Defense almost 70 years later. This indicates that strategic communication is based heavily on viewpoints that are not dependent on time, place or culture. The German Security Police – used here as an example – followed these universal principles internally even though they had not yet been formulated into a theoretical programme. In addition to theoretical contemplation, this study also analyses the main points of what taking these principles into consideration meant in practice in Estonia during World War II. For example, the fact that the German ruling power had to cooperate with the Estonian Self-Administration, created as a local ancillary organisation for the occupying administration, posed its own challenge. The viewpoints and goals of the Estonian functionaries did not always coincide seamlessly with those of the Germans. The same problem applied to ordinary Estonians’ opinions and hopes, which were often contradictory to the views of the Germans (Nurmis, 2011, 128–131). Thus, getting the Estonians to support the war objectives of the Germans required carefully deliberated propaganda work and utilisation of the principles of strategic communication.

In the summer of 1942, Dr. Martin Sandberger, the head of the German Security Police in Estonia, compiled an extensive report on matters within his own field. The material was acquired during the first year of occupation, between July 1941 and June 1942, but the report also included a detailed evaluation of matters on which the German occupation administration should concentrate attention in the future in order to retain and strengthen the willingness of Estonians to cooperate (the original document has been published in EJA, 2002, 21-86). This report is a key document for the analysis of Security Police opinions concerning the condition of 'strategic communication' and its developmental needs. In other surviving archival material from the Security Police, for the years 1941-1944, the same concerns often arise, though in a simpler form. The most essential part of the prevailing material (bi-monthly, monthly and annual reports) has been published as a source collection (EJA, 2002, 87-288). A comprehensive report from the summer of 1944 is otherwise similar to that from the summer of 1942 but makes fewer recommendations for further action (the original document has been published in EJA, 2002, 289-363).

The following account systematically compares a U.S. Department of Defense document (Principles, 2008) with documents from the German Security Service. The analysis focuses on clarifying briefly the extent to which the 'strategic communication' priorities of the two organizations mirror each other and what the practical consideration of these priorities essentially involved in German-occupied Estonia during World War II.

2. The comparative analysis

(1) Leadership-Driven. To ensure integration of communication efforts, leaders should place communication at the core of everything they do. Successful Strategic Communication – integrating actions, words, and images – begins with clear leadership intent and guidance. Desired objectives and outcomes are then closely tied to major lines of operation outlined in the organization, command or joint campaign plan. The results are actions and words linked to the plan. Leaders also need to properly resource strategic communication at a priority comparable to other important areas such as logistics and intelligence (Principles, 2008, 4). This principle is not mentioned directly in the documents from the German Security Service, but its importance is indirectly revealed. For example, in the summer of 1942 the extensive report (Sandberger) stressed that propaganda should, in the future, be subject to more detailed control. Communication with the Estonians had not succeeded as well as it might have during the first years of occupation, but weak leadership was not mentioned as a reason for this (EJA, 2002, 21-31).

The fact that the leadership was not criticized may be associated with two factors. First, propaganda was managed, in principle, by the chief of the German Security Police, who would thus have had to criticise himself in his report. Another option would have been to criticize leaders at a still higher level of the occupation regime, the Commissioner-General of Estonia (Karl-Siegmund Litzmann) or the leader of the Reichskommissariat Ostland (Hinrich Lohse), which would also have been difficult. Neither military nor civilian organizations generally make it easy to criticize the activities of superiors. On the other hand, the Security Police chief did refer to the importance of good leadership by thanking the Commissioner-General for some effective propaganda measures. According to Sandberger's interpretation, Litzmann's personal intervention in grievances had particularly helped to maintain Estonian confidence in the willingness and ability of the German occupation regime to look after Estonian interests (EJA, 2002, 22-23; see also 114-115, 119, 128, 135).

(2) Credible. Credibility and consistency are the foundation of effective communication; they build and rely on perceptions of accuracy, truthfulness, and respect. Actions, images, and words must be integrated and coordinated internally and externally with no perceived inconsistencies between words and deeds or between policy and deeds. Strategic Communication also requires a professional force of properly trained, educated, and attentive communicators. Credibility also often entails communicating through others who may be viewed as more credible (Principles, 2008, 5).

The importance of credibility appeared continually in material collected by the German Security Police throughout the occupation period. One of the tasks of the Security Police was comprehensively to clarify Estonian opinions and attitudes on a variety of subjects. One of the most frequently mentioned problems in the collected material was that Estonians were sensitive to contradictions between words and deeds (EJA, 2002, 21, 87, 149, 157, 174, 184-186, 202-203, 209, 214-215, 259-260). To ensure the acceptance of propaganda it was essential to minimize conflicts, which in most cases meant adjusting the content of propaganda to provide a better fit with reality. Evidently the Security Police was more pragmatic in this respect than the German Propaganda Ministry, led by Goebbels. Goebbels seems to have believed that the drastically altered truth, even outright lies, could convince the target audience, provided the propaganda was skillfully planned and its messages repeated sufficiently often (Doob, 1950, 419-442). The Security Police stated that this model did not work, at least in Estonia, and sought to remedy its own actions accordingly.

(3) Understanding. An individual's experience, culture, and knowledge provide the context shaping their perceptions and therefore their judgment of actions. We must understand that concepts of moral values are not absolute, but are relative to the individual's societal and cultural narrative. Audiences determine meaning by interpretation of our communication with them; thus what we say, do, or show, may not be what they hear or see. Acting without understanding our audiences can lead to critical misunderstandings with serious consequences. Understanding subjective impacts of culture, language, history, religion, environment, and other factors is critical when crafting communication strategy for a relevant population. Building relationships

and collaboration with the interagency, coalition, host nation, academic, non-profit, and business communities can facilitate better understanding of audiences (Principles, 2008, 5).

The German Security Police were fully aware that mistakes had been made in this area and that improved attention to cultural differences would be essential to the future success of their communications. Although Estonia and Germany were basically quite similar societies, some noteworthy differences were observed in public attitudes. The Estonians, for example, were found to be clearly more suspicious of propaganda in general; the kind of open propaganda that provoked no resistance in Germany sparked irritation among Estonians (EJA, 2002, 28-30). The Security Police realized that propaganda designed for Russians or Ukrainians did not work in Estonia, even though they were operating within what was technically 'Soviet territory'. In Estonia, the standard of living was much higher before the war, and the population had greater opportunities to influence public policy than their counterparts in the Soviet Union. The general-purpose propaganda prepared by the Reich Ministry for the Occupied Eastern Territories did not, therefore, appeal to Estonians; themes and methods had to be adjusted, sometimes considerably, before the material could be used in Estonia (EJA, 2002, 28, 199, 215-216, 246-249, 268-274).

(4) Dialogue. Effective communication requires a multi-faceted dialogue among parties. It involves active listening, engagement, and the pursuit of mutual understanding, which leads to trust. Success depends upon building and leveraging relationships. Leaders should take advantage of these relationships to place policies and actions in context prior to operations or events. Successful development and implementation of communication strategy will seldom happen overnight; relationships take time to develop and require listening, respect for culture, and trust-building (Principles, 2008, 5).

In Estonian conditions, dialogue specifically meant building cooperation with Estonians. As early as the first year of occupation, the German Security Police found that Estonians working in the administration with Germans, including Estonians involved in preparing propaganda, wanted more responsibility and decision-making power. The leadership of the Security Police therefore recommended greater respect for the views of Estonians, provided that direct benefits to Germany were not affected (EJA, 2002, 29, 51-54, 272). Reports on the public mood clearly showed the effectiveness of Commissioner-General Litzmann's habit of building confidence by frequently mingling with the people and listening to their concerns (EJA, 22-23, 180, 193, 199, 206-207, 212-213). Wider dialogue between Germans and Estonians, however, remained largely unachieved owing to the dogmatic attitudes of the German leadership: they viewed the peoples under German rule in the East, without distinction, as valueless assistants to whom they did not wish to give real decision-making powers and whose national-cultural aspirations they believed required only a minimal response (Isberg, 1992, 148–153; Myllyniemi, 1973, 286-292).

(5) Pervasive. Communication no longer has boundaries, in time or space. All players are communicators, wittingly or not. Everything the Joint Force says, does, or fails to do and say, has intended and unintended consequences. Every action, word and image sends a message, and every team member is a messenger, from the 18-year-old rifleman to the commander. All communication can have strategic impact, and unintended audiences are unavoidable in the global information environment; therefore, leaders must think about possible “Nth” order communication results of their actions (Principles, 2008, 5).

The German Security Police were well aware that communication involved a combination of behaviour and all the tools of the media. Security Service reports criticized the arrogant, insolent, and openly self-serving behaviour of some Germans. This created in Estonians a picture of Germans as greedy and reckless exploiters who did not appreciate the Estonians' willingness to cooperate and who did not care about Estonian well-being. One of the key messages that the Security Police presented in their reports was that Estonians must be made to feel that they were valued and considered as equals (EJA, 2002, 29-30, 289-297). Otherwise, relations between Estonians and Germans would be further worsened, and this would conflict with the overall interests of Germany. The Security Police clearly understood that the creation of a positive image of the German occupation regime was the fundamental key to preserving an Estonian desire to cooperate and to inspiring Estonians with a wish to work for the promotion of German war aims. The problem once again was the unwillingness of the German leadership to take account sufficiently of the viewpoints and measures recommended by the Security Service.

(6) Unity of effort. Strategic Communication is a consistent, collaborative process that must be integrated vertically from strategic through tactical levels, and horizontally across stakeholders. Leaders coordinate and synchronize capabilities and instruments of power within their area of responsibility, areas of influence, and areas of interest to achieve desired outcomes. Recognizing that your agency/organization will not act alone, ideally, all those who may have an impact should be part of communication integration (Principles, 2008, 5-6).

In the case of Estonia, the German Security Police tried to take this viewpoint into account on two levels. Firstly, the Security Police allocated adequate human resources to adjust propaganda to local conditions. This meant that the basic material provided by the Reich Ministry for the Occupied Eastern Territories was screened and converted into a form in which it would better serve German targets specifically in Estonia. Secondly, the Security Police closely monitored the propaganda activities of the Estonian Self-Administration (the Directorate) and sought to unify its messages with those conveyed by the German occupation regime (EJA, 2002, 27-31, 289-296). For example, there were regular guidance meetings for Estonian newspaper journalists, or in some cases direct briefing sessions, in which journalists were instructed in writing articles that would be appropriate from the German viewpoint. For the Germans, a persistent problem was that Estonians constantly strove to write articles in which Estonian national interests were 'too strongly' displayed. The Germans also thought it detrimental that Estonians wished to follow the communications of neutral countries, and even of enemy countries such as Great Britain, and to summarize them too sympathetically in the Estonian media (EJA, 2002, 24-28, 292, 372; see also Nurmis, 2011, 57).

(7) Results-Based. Strategic Communication should be focused on achieving specific desired results in pursuit of a clearly defined end state. Communication processes, themes, targets and engagement modes are derived from policy, strategic vision, campaign planning and operational design. Strategic communication is not simply “another tool in the leader’s toolbox,” but must guide all an organization does and says; encompassing and harmonized with other functions for desired results (Principles, 2008, 6).

German short-term goals in Estonia involved harnessing Estonian natural and human resources as efficiently as possible in support of the German war effort. Secret scenarios for the post-war future were more indefinite, but in any case it was intended that Estonia should be unified with Germany and would be ‘Germanized’ within a few generations (Myllyniemi, 1973, 145-169). War-time propaganda sought to conceal plans for the future and to focus on the demands of war. The German Security Police in Estonia had a clear vision of these priorities. The Security Police sought to guide both Germans and Estonians to co-operate on propaganda which would concentrate all forces on winning the war. Comments on the future were limited to vague references to the view that Estonians, as well as other peoples fighting on the side of Germany, would find that their part in the 'New Europe' would be directly proportional to their national contribution to victory: the more diligent the effort, the better their part would be (EJA, 2002, 21-30, 184-193).

(8) Strategic Communication should focus on long-term end states or desired outcomes. Rapid and timely response to evolving conditions and crises is important as these may have strategic effects. Communication strategy must reach intended audiences through a customized message that is relevant to those audiences. Strategic Communication involves the broader discussion of aligning actions, images, and words to support policy, overarching strategic objectives and the longer term big picture. Acting with adversaries’ decision cycles is also key because tempo and adaptability count. Frequently there will be a limited window of opportunity for specific messages to achieve a desired result.

An organization must remain flexible enough to address specific issues with specific audiences, often at specific moments in time, by communicating to achieve the greatest effect. All communication carries inherent risk and requires a level of risk acceptance within the organization. Leaders must develop and instill a culture that rewards initiative while not overreacting to setbacks and miscues. While risk must be addressed in the form of assumptions in planning, it should not restrain leaders’ freedom of action providing it has been taken into consideration appropriately (Principles, 2008, 6). The Security Police saw correct allocation and scheduling as a very important element in communication. Collected reports paid constant attention to the confusion and anger caused in recipients by outdated, poorly designed messages, and by messages inappropriate to the situation. For example, when the war situation developed quickly in the summer of 1944, German propaganda in Estonia quickly lost credibility when the public continued to see news films prepared in winter and spring rather than up-to-date and realistic information about the decisive summer battles fought in Belarus and in Normandy (EJA, 2002, 289-296). The Security Police were aware that news unfavourable from the German standpoint would in any case reach Estonia, for instance via BBC radio broadcasts, despite bans on listening. The conclusions of the Security Police therefore stressed that German propaganda should be more realistic and that it should more quickly and fully take into account the developing situation as well as the attitudes and level of knowledge of the target audience (EJA, 2002, 25-26, 28, 297, 372).

(9) Continuous. Strategic Communication is a continuous process of research and analysis, planning, execution, and assessment. Success in this process requires diligent and continual analysis and assessment feeding back into planning and action. Strategic Communication supports the organization's objectives by adapting as needed and as plans change. The Strategic Communication process should ideally operate at a faster tempo or rhythm than our adversaries (Principles, 2008, 6).

It may be said that one of the primary tasks of the Security Police was to fulfil ongoing functions of data collection and analysis. One of the starting points emphasized for the organization's activities was that data collected in 'the field' should be submitted unchanged to the higher levels of the Security Police – not selectively, not concealing information or people's opinions, and neither exaggerating nor understating. According to those instructions, all intelligence work would seek to ensure that the Security Police could build up as truthful and up-to-date a picture as possible of Estonian opinions and attitudes (Noormets, 2002, 15-17).

3. Conclusions

Overall, it is clear that, in 1941-1944, the German Security Police in Estonia applied principles of strategic communication to the monitoring, analysis and design of propaganda, even though the concept did not yet exist. When one compares the reports and strategies of the Security Police with the strategic communication program assembled by the U.S. Department of Defense in 2008, a very high level of consistency is evident. Out of nine main points, eight came strongly to the forefront in the work of the Security Police. The only point not directly mentioned, owing to factors described above, was the need for a leadership-driven approach. Conceptually, this point in the program was nonetheless visible in the background of the other recommendations and practical actions of the Security Police. This analysis confirms the conclusions of those researchers who have suggested that *successful* communication is inevitably based, consciously or unconsciously, on attention to strategic principles (Halloran, 2007, 5-6; Kellermann, 1992, 288-300). Strategic communication is therefore not really a matter of new approaches, independent of the old, but rather of holistic and purpose-driven design. In addition, it may be considered that successful communication is based on universal structural principles that are essentially independent of time, place, and culture. Naturally, there may be considerable variance in the practical application of these principles according to operating environment, temporal context, and target audience (Halloran 2007, 4-14; Mahoney, 2011, 143-153; Murphy, 2008; Paul, 2010).

References

Doob, L. (1950) 'Goebbels' Principles of Propaganda', *The Public Opinion Quarterly*, Vol. 14, No. 3.

EJA (2002) *Eesti Julgeolekupolitsei aruanded 1941-1944*, Noormets, T. (ed.), Tallinn: Riigiarhiiv.

Halloran, R. (2007) 'Strategic Communication', *Parameters*. US Army War College Quarterly, Autumn 2007.

Isberg, A. (1992) *Zu den Bedingungen des Befreiers. Kollaboration und Freiheitsstreben in dem von Deutschland besetzten Estland 1941 bis 1944*. Stockholm: Almqvist & Wiksell.

Kellermann, K. (1992) 'Communication: Inherently strategic and primarily automatic', *Communication Monographs*, Vol. 59, Issue 3.

Mahoney, J. (2011), 'Horizons in Strategic Communication: Theorising a Paradigm Shift', *International Journal of Strategic Communication*, Vol. 5, Issue 3.

Murphy, D. (2008) 'The Trouble with Strategic Communication(s), *Center for Strategic Leadership, U.S. Army War College, Issue Paper*, vol. 2-08, http://www.carlisle.army.mil/DIME/documents/IP2-08TheTroubleWithStrategicCommunication(s).pdf

Myllyniemi, S. (1973) *Die Neuordnung der baltischen Länder 1941–1944. Zum nationalsozialistischen Inhalt der deutschen Besatzungspolitik*. Helsinki: Suomen Historiallinen Seura.

Noormets, T. (2002) 'Saateks', in Noormets, T. (ed.) *Eesti Julgeolekupolitsei aruanded 1941-1944*, Tallinn: Riigiarhiiv.

Nurmis, K. (2011) *Das fein geschliffene Glas. Saksa okupatsiooni aegne propaganda organisatsioon Eestis 1941–1944*. Magistritöö (unpublished MA thesis). Tartu: Tartu Ülikool.

Paul, C. (2010) '"Strategic Communication" Is Vague. Say What You Mean', *Joint Force Quarterly*, Issue 56; http://www.au.af.mil/au/awc/awcgate/jfq/paul_sc_is_vague.pdf

Principles (2008) *Principles of Strategic Communication*. United States of America, Department of Defense, http://www.au.af.mil/info-ops/documents/principles_of_sc.pdf

Cyber Macht - Laying the Groundwork for a new Comprehensive Academic Theory

Leigh Armistead[1] and Scott Starsman[2]
[1]Peregrine Technical Solutions LLC, Yorktown, USA
[2]Avineon, Inc., Alexandria, USA
larmistead@gbpts.com
sstarsman@avineon.com

Abstract: This paper begins the process of laying the groundwork for a new comprehensive academic theory on Cyber Macht (Cyber Power). This new proposed new theoretical construct will not just be an update on Soft Power or Noopolitik, but instead will also include elements of Information Operations (IO) and the practical aspects of diplomacy and warfare. This new theory is all about the communication paths and changes in connectivity focused around a central theme that power has now been globally distributed. This is caused by the huge increase in access to information, for all people around the world, with increased connectivity and the ability to influence events far beyond the previous normal range. This new theory will therefore reference power and influence operations including the ability to shape the information in this new cyber era while still protecting the assets and maintaining the ability to get accurate and relevant information from multiple sources and channels. For what has changed in this new era is the range in which one person can extend the influence of their actions or ideas. This is evidenced in the increased rise in cyber attacks and the need to always be ready, with robust Computer Network Defense (CND) or Information Assurance (IA) efforts. Most succinctly, success in this era can be characterized by the ability to counteract or reject adversaries' global influence while still projecting your own. To succeed, one must maintain the required connectivity and logistics, to include networks, computers, and fiber, to ensure access, while isolating the adversary by countering his ideology propagation and cutting off communication access as well. Military theories in the past often focused on protecting interior lines of communications. This is still applicable in the information age, although the author's believe the focus is not about controlling the networks or information, but rather about ensuring open access and the free flow of data. It is critical to maintain continuous and ubiquitous connectivity, where horizontal integration across different organizations is a key ingredient of this new theory, with an emphasis on sharing vice closed, on open architecture, on open minds, all of which the authors believe equals total transparency.

Keywords: cyber, power, theory, strategy, academic, comprehensive

1. Introduction

The United States government lacks a defined and coherent National Strategy on Cyber Operations, particularly in the IA realm to address real and persistent challenges. This deficiency has been noted at the strategic level in President Obama's latest State of the Union Address (12 February 2013) and in recently released Government Accountability Office (GAO) report on CyberSecurity (GAO-13-462T dated 7 March 2013). This paper attempts to start a discussion on the development of a comprehensive academic theory on power in the information age that we have entitled Cyber Macht. The lack of a strategic theory or academic model to serve as a basis to explain and understand the rise and application of information power around the globe ultimately endangers the overall stability of how cyber in particular is employed today. We believe that the use of an overall comprehensive academic theory often serves as a foundation, a basis on which to build a model of a complex subject such as power so that it can be better understood and analyzed. Unfortunately, an overarching academic theoretical construct, on the order of realism or international liberalism, which explains power in the information age with sufficient rigor does not exist. That is not to say that there have not been influential academics who have set forth theories for discussion and review such as Soft Power and Noopolitik, however, there has not yet been an overwhelming acceptance of either of these constructs (Nye, 2005; Arquilla and Ronfeldt, 1999). This paper is therefore an attempt to outline the necessary theoretical foundation to understand how power is being focused and transformed as the world enters the next phase of the information age.

2. A grand theory on cyber

We are in a new era …a chaotic period where disconnects between theory and reality are at their most pronounced. The authors' research has revealed a marked inability to develop a comprehensive strategic theory that accounts for the changes in the power structure of the federal government relative to the citizenry and opponents; most noticeable and very evident in the United States military and federal agencies. So while earlier theories such as Soft Power and Noopolitik may have struck a chord within the Department of Defense

(DoD) and a number of federal agencies at some point … to date, none of these attempts to develop an overall encompassing academic theory for what is happening with regard to information and power have been formally adopted across the United States as a whole. Even the authors of Noopolitik themselves, note as much in a recap to their book *The Promise of Noopolitik*, published eight years after the original publication of Noopolitik (Arquilla and Ronfeldt, 2007). Their initial enthusiasm for their initial theoretical construct has been dampened considerably not only by the events of 9/11, Operation Enduring Freedom (OEF) and Operation Iraqi Freedom (OIF), but also by the manner in which the Internet and the intellectual community have evolved over the last decade. The hopeful optimism of the early 1990s with regard to the World Wide Web and the Internet has instead turned in the last few years to the awful realization that many hostile individuals and groups have instead used these powerful technologies to their advantage, whether for their political, financial or social gain (Ibid). Likewise Arquilla and Ronfeldt also admit in their postscript that the early promises of a global community are instead being overwhelmed by the day-to-day events, which tend to mitigate the promise of revolutionary change in the use of Information as a force of good. Although they still believe that Noopolitik is an idea for the future and, while they remain optimistic, they are also dismayed by a number of trends shown below that have effectively derailed much of the promised potential of their early theoretical construct:

- Notions like Noopolitik are gaining credibility, but all too slowly
- Soft Power lies behind them all, but the concept needs further clarification
- Activist Non-Governmental Organizations representing global civil society are major practitioners of Noopolitik, but the most effective may be the global network of jihadists
- American public diplomacy would benefit from a course correction (Ibid)

Therefore, none of these early concepts or theory can be properly considered an overall comprehensive rigorous academic theory on power in the Information Age, but instead, they are forming a foundational series of ideas around similar topics that one can use to start to define this radical change in power. The authors believe that a revolution in warfare is and has been occurring with regard to power and that, without an overarching theoretical construct, the United States will not be able to fully maximize this new capability and will fall behind our adversaries.

The authors' concept will be motive-neutral … ie we do not assume that Cyber Macht will be used benevolently; instead it is simply a force for good or for evil, that can be harnessed to influence national policies, actions, and perceptions. Shown below are some early theoretical concepts that the authors are researching in their attempts to develop this unified grand strategy:

- **Global Influence** - The ability to influence events no longer resides primarily at the national or even governmental level. Small groups of people and even individuals with a potent message and a well-chosen audience are able to broadcast their message, excite a population, and even initiate an attack.
- **Knowledge Accessibility** - Information that was previously difficult to find or access is now fully indexed, searchable, and downloadable over the internet. Even information that is intended to be restricted can often be obtained through direct, targeted action.
- **Perception Shaping** - Truth and the internet have long been uncomfortable partners. Public opinion can often be easily shaped by tailoring information for specific consumer groups and by not presenting or suppressing information not supportive of the desired public impression.
- **Pervasive Information** - It has long been known that information has value and that it can influence the mind. In prior ages, information was ephemeral, but with the advent of electronic communication, and the widespread adoption of networking, the ability to conduct long term perception campaigns rises.
- **Distributed War** – The growth of networks has led to the distribution of control and information that is extremely difficult and perhaps impossible to ever lock down again. In fact, there may exist an inverse relationship between the amount of control exerted over a network and its informational value. In the future, the ability to conduct warfare will be spread out far beyond the military forces.

3. Earlier efforts to develop an overarching Cyber Macht theoretical construct

Cyber Macht is not a part of the liberalism or realism theoretical academic theories, but instead it is something else … something that is in between, where it is much more oriented around power. Cyber Macht has its own language such as virus' or worms', which is somewhat metaphoric in nature, but it also can be very technical, especially when concerned with IA issues. This dichotomy of needs and requirements has hampered the ability to date to develop an overarching theoretical construct, because for some, the use of Cyber Macht is so easy to visualize and yet so hard to accomplish? It is often the 'softer areas' of power, mainly the concepts involved that affect the mind, in the form of perception management and strategic communications that the United States government appeared to have the most difficulty in conducting operations. These skill sets are considered an art, with many believing that one needs to take the long view for success in this area, and yet these same research participants also noted that, in the United States, federal organizations normally relied on technology to answer the questions and to overwhelm the adversary quickly. Cyber Macht is not that radical, and, in fact, some academics have suggested that instead it should just really be entitled as "Operations in the Information Age". This is an interesting concept as one attempts to normalize these huge shifts in power, but a new title doesn't solve the need for an overarching construct. Instead the authors believe that developing new academic theory often hinges on radical concepts such as those espoused in the *Third Wave* or Noopolitik (Toffler, 1984; Arquilla and Rand, 1999). It is these concepts along with Soft Power that are perhaps the best examples successful attempts to develop a nascent academic theoretical construct for the Department of Defense (DoD) policy (Nye, 1990).

From the official federal government perspective, a number of documents have been promulgated that lay out in different ways, the vision of the key organizations that will be affected by or using this new form of power incorporated in the concept of Information Operations (IO). The IO Road Map and to a lesser extent the new Joint Publication 3-13 *Information Operations* (2006) are not the only way ahead for the federal bureaucracy with respect to the future of IO within the United States. In September 2004, a Defense Science Board (DSB) Task Force of the Report on Strategic Communications was released as a follow-on effort to an earlier study by the DSB in October 2001. Many critics felt the first study was overshadowed by the tragic events of 11 September and the opening campaign of OEF.

Taken together, the current efforts by the federal government show a huge dichotomy in the goals of these two main policy attempts at developing strategic IO academic theory, with the more pragmatic DoD (*The IO Road Map*) and State Department (*DSB Report on Strategic Communications*) documents released in 2003 and 2004 respectfully, that represent the way in which IO is conducted today throughout the federal bureaucracy. Because the *IO Road Map* has a much narrower focus than the mandate from the DSB, it tends to highlight the huge mismatch between the strategic transformational promise of IO doctrine, with the operational reality of how the DoD tactically conducts information activities and campaigns. So in reality, the *IO Road Map* may very well be just a pragmatic solution to the difficulties in trying to conduct these types of Cyber Macht or information campaigns on a day-to-day basis, as opposed to the lofty and somewhat more ambitious goals of the DSB report, which while utterly correct from a perception management perspective, may in fact never occur due to political and fiscal reality.

4. Case study – 9/11

A number of historical and modern events demonstrate the dramatic impact information and how Cyber Macht has made on international relationships and the conduct of military operations. Because of their temporal compactness and the amount of detailed open source information available, the events of 11 September 2001 provide a well-documented example of the transformative capability of Cyber Macht in this new age and will be used by the authors to contextualize their comprehensive theory. This case study revolves around the five Cyber Macht tenets introduced earlier and how they can be applied to 9/11 to help understand power in the information age

As indicated in Figure 1, through 1967 annual commercial aircraft hijackings were relatively rare, averaging less than 4 per year (Aviation Safety Network 2013). In 1968, hijackings began to be used by groups seeking to publicize their cause or otherwise terrorize commercial travelers. From 1968 to 1972, the average number of commercial airline hijackings rocketed to more than 60 per year. The introduction of metal detectors in airports, the increased criminalization and prosecution of hijackers, and the renunciation of hijacking as a tool by some organizations (such as the PLO) led to a decrease in hijackings and from 1973 to 2011 they average roughly 20 per year. During this time, law enforcement and public safety officials determined that the best reaction for victims of hijackings was to comply with the hijackers and not resist. This quiet understanding that developed between law enforcement, hijackers, and the flying public persisted until 2001 and provided fertile ground for the 9/11 plotters.

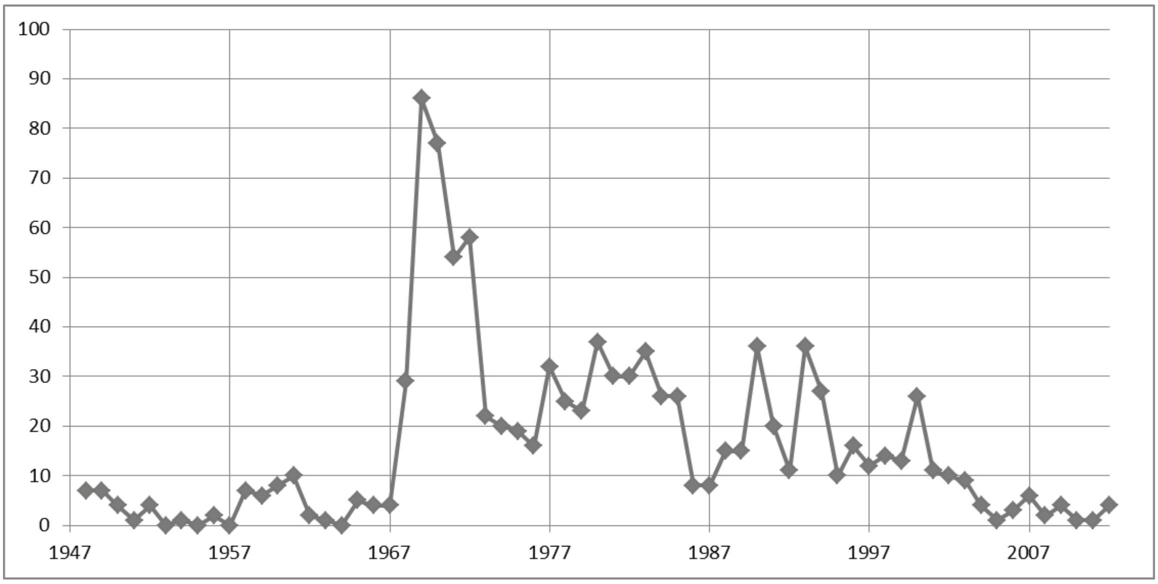


Figure 1: Number of aircraft hijackings per year

The plot to hijack four commercial aircraft and use them to attack large ground targets relied on the passengers and even crew members putting up little resistance. With only four to five hijackers on each flight, the operation planners knew they would be outnumbered by ten or twenty to one; hence passivity of the victims was an essential part of the plan (Kean and Hamilton 2004). The pattern of each hijacking was generally similar. Each of the hijackers had a seat in First or Business class and was close to the cockpit. After the flight had reached cruising altitude, they would gain access to the cockpit and murder, disable, or otherwise remove the pilots and anyone interfering with their operation. Two of the hijackers would take control of the plan in the cockpit and the remaining team members would force the passengers to the rear of the plane with mace and/or violence. They would inform the passengers that they had a bomb on board and that they should remain in their seats and they would not be harmed. One of the messages intended to be passed to the passengers but instead accidentally broadcast on an air traffic control channel was *"Nobody move. Everything will be okay. If you try any moves, you'll endanger yourself and the airplane. Just stay quiet."* (ibid) This message was calculated to reassure the passengers and crew that they should continue to follow their now-outdated perceptions of aircraft hijackings and, if they, complied, they would be unharmed. Of course, this was not to be the case. Unbeknownst to the victims of the first three flights, they were about to become part of a very damaging and deadly terrorist attack and they were riding, not in aircraft, but in guided missiles.

Like most major Al Qaida operations, this attack was characterized by multiple, near-simultaneous and spectacular attacks intended to inflict numerous casualties and massive damage. Figure 2 provides a graphical reference of the events described in the following paragraphs. The four targeted aircraft were scheduled to depart within 30 minutes of each other and, while all were delayed by 10 minutes or more, they all took off within 40 minutes of each other. Of note, United 93 had the longest delay of more than 25 minutes and was the last of the four planes to take off (ibid).

The hijacking of United 93 began around 9:28 AM, nearly 45 minutes after the first plane had crashed into the North Tower of the World Trade Center (WTC) and 25 minutes after the second plane had crashed into the South Tower of the WTC. While passengers and/or crew established communications with the ground via non-standard means such as mobile telephones or GTE Airphones on each of the aircraft, only aboard United 93 did the passengers and crew have the time and information to come to the realization that this event would not be following the anticipated hijacking event sequence and that, by following the convention guidance of passivity, they were supporting the hijackers' plan. By 9:57 AM the passengers and crew had broken with the previous hijacking response paradigm and began an attempt to retake the plane. The battle for control of the aircraft lasted around five minutes and ended when the hijackers realized that their primary mission was defeated and they dove the plane into the ground in Shanksville, PA at 10:01 AM.

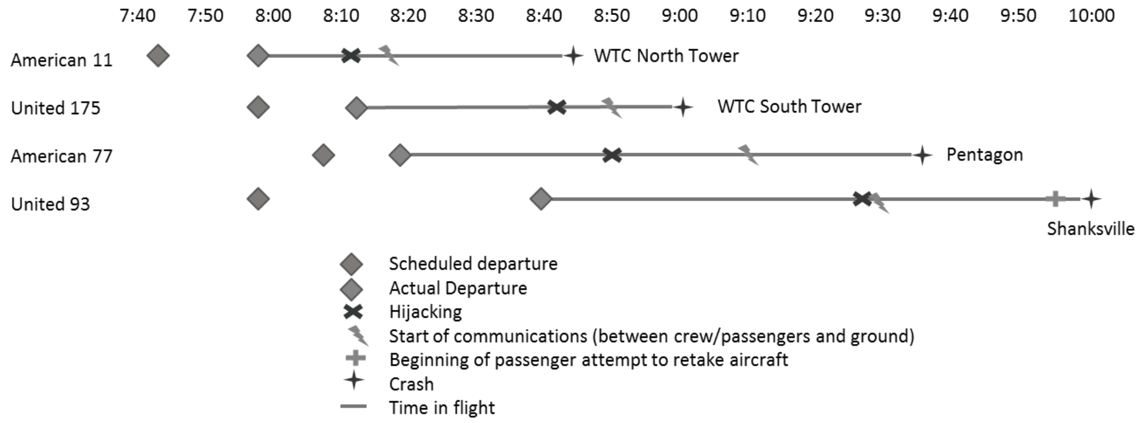


Figure 2: 9/11 hijacking timeline

Often missed in the analysis of this event is the magnitude of the passenger paradigm shift that occurred during the brief flight of United 93. They began the day as did all other airline passengers with a deep-rooted belief on the proper response to a hijacking developed from three decades of experience. With meager access to communications, they were able to:

- assess their situation
- determine that their previous beliefs about proper hijacking response were deeply in error
- formulate a reaction knowing that their lives and those of their fellow passengers were bound in their decision
- execute a plan that prevented further deaths or injuries on the ground

That the passengers and crew were able adapt to this highly unusual situation in such a rapid manner is illustrative of the transformative power of information and Cyber Macht. In military terms, a group of unarmed and untrained civilians improvised an ad hoc communications system and worked inside the decision loop of the adversary, interfered with their plans, and effectively reduced the power of their attack by 25%. Those that doubt the potential of Cyber Macht in the conduct of our operations, disrupting our adversaries', and protecting ourselves from disruption need to understand the power of information through this and other examples. For those of us that understand the potential that we are alluding to in this new theory, this example serves as both a guidepost and a warning of the impact should we misunderstand the application of this power. Specifically, it is capable of rapidly transforming a situation and disrupting detailed plans, and when analyzed in the context of a new Cyber Macht theory, the planning, execution, and response to the 9/11 attacks can be decomposed in the following way.

- **Global Influence** – 9/11 was an attack equaling Pearl Harbor in magnitude and impact on the national psyche. However, this attack was not perpetrated by a nation-state but rather a small band of ideologues with limited resources. Conversely, the hijackers assumed that, once the hijacking of an aircraft was complete, the balance of power would remain in their hands based on decades of social training. However, they failed to recognize the transformative power of Cyber Macht and allowed the crew and passengers to remain connected and receive information. This oversight ultimately foiled the fourth attack.
- **Knowledge Accesibility** - The collection and assimilation of information was a critical element in this attack. The terrorists were able to train to fly airliners, gather information about airline schedules, research targets, and gather many other pieces of information important to their success. As discussed above, the ability of the crew and passengers and crew to acquire and analyze information about the very event in which they were enmeshed proved vital in interrupting an attack for which no one was prepared.
- **Perception Shaping** – 9/11 was an exercise in shaping information on many levels. Strategically, the entire operation was grandly staged to emphasize Al Qaida's capability, embolden radicalized Muslims to act against the US, and damage the image of the US as invulnerable to attack. Operationally, Al Qaida recruits are fed a steady diet of hatred of the US, Israel, and infidels in general and believe that, if they die participating in an attack on any of these targets, they will receive an eternal reward. Tactically, their goal was to shape the information that the passengers had (that they were being hijacked) into a scenario that

the victims would believe and would encourage passivity. While successful in three of the four hijackings, ultimately enough information leaked onto United 93 to overcome tactical information shaping and permit the passengers to see the true situation and take countermeasures.

- **Pervasive Information** – Because information has such value and transformative power, the ability to influence and spread that information can be vital. Because the US values freedom of expression, there is little protection or policing of the vast majority of open networks containing personal and commercial information. Al Qaida took advantage of this and was able to acquire a vast amount of planning and operational information from open sources without arousing suspicion. Al Qaida was able to spread information across their networks with little impunity as they limited their use of various communication networks and practiced other operational security techniques to disguise their use.
- **Distributed War** – This is a great example of the distribution of warfare, well outside the boundaries of military forces. Here you have unarmed citizens protecting the President of the United States from attack, something that no unit in the DoD could. All communications were conducted out of band, on non-approved networks and in the end, the passengers on Flight 93 succeeded in defeating the attack.

5. Suggested approaches to developing a comprehensive Cyber Macht theory

The transformational ideas inherent in Cyber Macht are crucial and must become a reliable capability of the American arsenal because, as the events of 11 September 2001 indicate, military, political or economic power are often simply ineffective in dealing with these new kinds of threats to the national security of the United States. In this new era, all factors of power must be utilized for, as some academics argue, in the future networks will be fighting networks (Arquilla and Ronfeldt, 1999). Good examples of this abound in OEF and OIF, where networks in the form of information campaigns fought networks made up of perceptions, and the side that will ultimately emerge from this epic conflict as the victor is the one that can best shape and influence the minds of not only their adversary, but their allies as well (*Advisory Group on Public Diplomacy for the Arab and Muslim World*, 2003).

Unfortunately, the shift from the industrial age to the information environment may mean that the United States will not forever remain the dominant player in the political arena. Arquilla and Ronfeldt also write that nation-states are losing power to hybrid structures within this interconnected architecture, where access and connectivity, including bandwidth, will be the two key pillars of any new organization. They posit that truth and guarded openness are the recommended approaches to be used in both the private and government sectors to conduct business and, in their opinion, time zones will be more important than borders. It will be an age of small groups, using networks to conduct 'swarming' attacks that will force changes in policy (Arquilla and Ronfeldt, 1997a). Key features include:

- Wide open communication links where speed is everything
- Little to no censorship, the individual controls his own information flow
- Truth and quality will surface, but not initially
- Weakening nation-states and strengthening networks (Ibid, 1997b)

So while the changes that are mentioned in their book Noopolitik are truly revolutionary and describe a profound shift in the nature of power, unfortunately this transformation has not been translated from a strategic concept to tactical actions (Kuusisto, 2004).

In this vein, a thread has emerged from the researcher’s data that the reason that no overall theory on power has emerged in the information era, is because Cyber Macht is a concept that supports so many different and disparate academic areas, it makes it difficult to unify a community around a single concept. The sheer diverseness of this transforming idea is easily seen at academic conferences where the hard and soft topics are instantly separated into separate streams and only rarely touching each other at the plenary sessions. Computer security, psychological operations, electronic warfare, public affairs and the other portions of IO by themselves are all incredibly complex areas. To find a single comprehensive academic theory that can encompass the use of these warfare areas and the others that comprise Cyber Macht is incredibly difficult as one can imagine.

However, it is the goal of the authors to attempt this very feat. It is our belief that taken together, all of these ideas can be formed the basic concepts of a new Cyber Macht theory. We believe that as information decentralizes it becomes less of a monolithic model and more like an onion, with layers of complexity and large communities of interest distributed around the world coming together, much like diasporas where geography previously interfered. The new power of information has spread power to the masses, instead of being held by a select few. Cyber Macht integrates the impact of the rise of connectivity with the knowledge that people, and specifically populations can no longer be grouped into monolithic categories and this provides the basis for the new model to emerge. We believe that this developing theory can best be understood, where people will not act as a singular unit, but only in their best self-interest which constantly changes with shifting goals. The rise in technology, with all of the tremendous advances in computers, cell phones, TV, and the internet, where improvements are changing so fast, and the technology is constantly shifting, has led us to realize that the platform is not the key, but instead it is the information that is important. This is the key to the new Cyber Macht theory that we plan to research and develop, namely that because information is always on, it is ubiquitous, it is persistent, it is living, and it is constantly changing. Power has now been distributed to the masses, with the genie out of the bottle, and it cannot be put back in. We believe that a new Cyber Macht theory, must take these new ideas and concepts into consideration, to understand that any model must work in a chaotic system, one that is characterized by persistent change where constant adaptation is the key to success. Thus the intent of this new effort is to develop a comprehensive overarching Cyber Macht theory to fill this void. The Internet and other emerging communication networks (wireless, peer-to-peer, etc) have forever destroyed the power formerly resident only in the government and that asymmetry now gives the power of information to all.

6. Summary

In conclusion, the authors of this paper believe the only way to have a comprehensive approach to the development of theory on Cyber Macht, is to involve not only the academic community but also the various strategic centers of excellence across the DoD and State Department. A tremendous amount of talented and innovative research on these topics is being conducted outside of the United States and so a collaborative approach is suggested, where the three main IO and information warfare academic conferences are utilized as the backbone for this effort. These three gatherings are held yearly, and typically have many of the same participates attend from around the world. This makes for a nice setting in which over time, a vigorous debate can be held, in which a number of aspects and options to developing a strategic Cyber Macht theoretical construct are analyzed with sufficient academic rigor.

References

Advisory Group on Public Diplomacy for the Arab and Muslim World. Changing Minds, Winning Peace: A New Strategic Direction for U.S. Public Diplomacy in the Arab and Muslim World. 1 October 2003.

Armistead, Edwin L. “A Tale of Two Cities: Approaches to Counter Terrorism and Critical Infrastructure Protection in Washington, DC and Canberra, Journal of Information Warfare, Vol 3, Issue 1, Spring 2004.

________. “Back to the Future: Strategic Communication efforts in the Bush Administration” Journal of Information Warfare, Vol 2, Issue 3, Fall 2003.

________. “Fall from Glory: The Demise of the USIA during the Clinton Administration.” Journal of Information Warfare, Vol 1, Issue 3, May 2002.

________. ed. Information Operations: Warfare and the Hard Reality of Soft Power. Dulles, VA: Brassey’s, Inc., 2004.

________. ed. Information Warfare: Separating Hype from Reality. Dulles, VA: Potomac Books, 2007.

Arquilla, John and David Ronfeldt. The Advent of Netwar. Santa Monica, CA: RAND, 1996.

________. The Emergence of Noopolitik: Toward an American Information Strategy. Santa Monica, CA: RAND, 1999.

________. The Promise of Noopolitik. First Monday. 20 July 2007. Accessed at http://www.firstmonday.org/issues/issue12_8/ronfeldt/index.html on 23 January 2008.

Aviation Safety Network (2013) “Aircraft Hijackings”, [online], Flight Safety Foundation, http://aviation-safety.net/statistics/period/stats.php?cat=H2.

Kean, Thomas H. Hamilton, Lee H., et. al (2004) The 9/11 Commission Report, Government Printing Office, Washington, DC.

Keohane, Robert O. and Joseph S. Nye. Power and Interdependence. 2d ed. Boston: Longman, 1989.

Kuusisto, Tuijo, Rauno Kuusisto and Edwin Armistead. ”System Approach to Information Operations”, 3rd European Conference on Information Warfare and Security, Royal Holloway University of London, 2004, pp. 231 – 239.

Nye, Joseph S. Bound to Lead: The Changing Nature of American Power. New York: Basic Books, Inc., 1990.

________. “Redefining the National Interest.” Foreign Affairs 78 (July/Aug 1999): 22-35.

________. Soft Power: The Means To Success In World Politics. New York: PublicAffairs, 2005.

Nye, Joseph S. and William A. Owens. “America’s Information Edge.” Foreign Affairs 75 (March/April 1996): 20-36.

Toffler, Alvin. The Third Wave. New York: Bantam Books, 1984.

Toffler, Alvin and Heidi Toffler. War and Anti-War. New York: Warner Books, 1993

U.S. Advisory Commission on Public Diplomacy, A New Diplomacy for the Information Age (Washington, DC: State Department, 2000)

U.S. Department of Defense, Defense Science Board Task Force of the Report on Strategic Communications, 2004.

U.S. Department of Defense Directive (DoDD) S3600.1, Information Operations (2006).

U.S. Department of Defense. Information Operations Road Map (2003).

U.S. Department of Defense. Joint Chiefs of Staff. Joint Doctrine for Information Operations (February 2006), Joint Publication 3-13.

U.S. Government Accountability Office, Cyber Macht Security: A Better Defined and Implemented National Strategy is Needed to Address Persistent Challenges (GAO-13-462T), 7 March 2013.

Anomaly Detection via Manifold Learning

Amir Averbuch[1] and Pekka Neittaanmäki[2]
[1]School of Computer Science, Tel Aviv University, Tel Aviv 69978, Israel
[2]Department of Mathematical Information Technology, Agora, University of Jyväskylä, Finland
amir@math.tau.ac.il
pekka.neittaanmaki@jyu.fi

Abstract: The basic approach to protect and secure critical infrastructure and networking data against cyber attacks of the last 45 years called "walls and gates" (barriers between trusted and untrusted components, with policy-mediated pass-through) have failed. There is no reason to think that they will be more successful in the future. Rule based methodologies that govern firewalls and IDS/IPS are irrelevant today to detect sophisticated malwares (viruses, SQL injections, Trojans, spyware and backdoors) that pretend to be regular streaming and penetrate every commercial barrier on the market that are based on signatures of intrusions that detect yesterday attacks but fail to detect zero day attacks. The focus is on detecting zero day malware**.** We describe a prototype security system that automatically identifies and classifies malware. The core technology is based upon manifold learning that uses diffusion processes, diffusion geometries and other methodologies that find geometric patterns that deviate from normality. The main technology core is based upon training the system to extract heterogeneous features, to cluster the normal behavior and then detect patterns that deviate from it which are malware anomalies. The proposed technology offers behavioral analysis of heterogeneous complex dynamic networking data to that maintains and preserves networks' health. The system uses efficient computation that is based on multiscale dictionary learning and kernel approximation, patch processing, adaptive subsampling and clustering and profile updating. These are universal generic core technologies for anomaly detections algorithms that are based on well founded deep unification between different mathematical theories from different disciplines that emerged recently. Promising preliminary results increase the potential of the proposed system to fill the gap that current state-of-the-art IDS/IPS and firewalls are unable to fill.

Keywords: anomaly detection, manifold learning, clustering, malware detection, training, behavioral analysis

1. Introduction

"Cyber-weapons well suit terrorists. Fortunately, perhaps, the likes of al-Qaeda have mostly used so far the internet for propaganda and communication..." (Economist (2010)). It appears that the mouse and the keyboard are the new weapons of conflict and are posing new forms of threats from terrorists and crime organizations everywhere. They can easily use the internet virtually from everywhere in the globe including their "bedrooms" to launch fatal attacks on every institution and every infrastructure.

Cyber war affects all of us in stealing and destroying from financial intuitions: banks, insurance, pension, monitoring system ...; crippling **critical infrastructures (CIs)**: gas, electricity, transportation (air traffic control, train, boats), water supply...; In penetrating governmental and intelligence agencies security agencies; In crippling large and medium enterprizes; In penetrating telecom companies, mobile stations, mobile base stations, smart phones...; Military installations; Internet providers; Social networks, and the list goes on and on

Cyber Espionage is becoming a major threat on nations, governmental institutions, CIs and business organizations. Intruders implant a silent agent inside the "brain/core'' of the organization that steals and sabotages critical data. That agent is an organism that nests inside the core and enables to glance to the most sensitive "knowledge and thoughts'' of the attacked "core''. The agent enables to "paralyze the body'' whenever required. The agent is **sophisticated** (engineered by the best hackers brains of malicious adversaries - individual and states) and **stealthiest**. Intelligence agencies, as well as criminal organizations and knowledgeable hackers are investing fortune to develop new types of Cyber Trojans, which send metastasis within short period of time prior to disappearing into cyberspace. Currently, it is virtually impossible to detect these Trojans since they keep mutating continually. Some Trojans leave a back door.

We have developed methods to achieve malware detection. The core technology is based upon processing high dimensional data via diffusion processes, diffusion geometries and other methodologies that we have developed for finding meaningful geometric descriptions that represent normal behavior of the data and then identifying deviations from normality. The main core of the methodology is based upon training the system that extracts heterogeneous features, identifies their characteristic (normal) and then finds patterns, which did not participate in the training, that deviate from it. These deviations are the malware anomalies we are after. This methodology offers behavioral analysis of heterogeneous complex networks to produce a unified threat manager to maintain and preserve networks' health. It is based on well founded deep unification between different mathematical theories from different disciplines that emerged in the last couple of years with classical mathematics such as applied and computational harmonic analysis, differential geometry, stochastic processing and classical analysis.

1.1 The current state-of-the-art

Intrusion detection systems (IDS) and Network intrusion detection systems (NIDS) also called intrusion prevention system (IPS) have become an integral component in any security systems. The challenge is to perform online IDS and NIDS without miss-detections and false alarms. To achieve it, most systems are based on signatures of intrusions that are developed and assembled manually after a new intrusion is exposed and distributed to the IDS clients. This approach is problematic because already known intrusions (yesterday's attacks) are detected but they fail to detect new attacks (zero day attacks). In addition, they do not cover a wide range of high quality new sophisticated emerging attacks such as Trojans horses that exploit innovatively many vulnerabilities. They fit specific attacks. Trojans are like a spying ring, they can emerge only after a long time while assimilated in the system to be looked as a legitimate process. We are looking for a pattern in ocean of data that that their behavior is deviated from the normal behavior of the system. A comprehensive survey is given in Dua S., Du X. (2011).

1.2 Outline of manifold based solution for malware detection

We have developed new data mining based methodologies to detect attacks that are classified as anomalies. In contrast to other current methodologies, which can be classified also as data mining and use for example statistical classification and clustering, support vectors machines, etc, we present a unified threat manager algorithms and system to be applied to dynamically changed data via non-linear reduction of the inspected data to handle. These non-linear transformations identify geometric patterns (manifolds) in these high-dimensional datasets and the connections among them while projecting them into low dimensional spaces that can be processed more efficiently and visualized. It automatically detects anomalies that deviate from normal behavior where anomalies mean malicious use (such as intrusion) or some abnormal behavior such as Trojan. The main core of the proposed methodology is based on training the system on extracted features that automatically generates at least one cluster that represents the normal behavior (normal profile of the data). The algorithms are capable to classify newly arrived data point to be either normal (belong to a normal cluster (manifold)) or deviate from this cluster (called abnormal, intrusion). This classification, which detects anomalies, is done in a low dimensional space that was embedded from the source multi-dimensional data. The diffusion distances, which dictate affinities among data points in the source multi-dimensional data, become Euclidean distances in the embedded space. The classification of multi-dimensional data points as normal or abnormal is done by the application of an out-of-sample extension algorithm which provides coordinates (parameterization) for each newly arrived data point in the embedded space. "Out-of-sample extension" (Bermanis, A., Averbuch, A. and R. Coifman (2013), Coifman, R.R. and Lafon, S. (2006)) is defined as the action of providing diffusion coordinates to each newly arrived data point in the embedded space. The processing of the multidimensional data points is based on a kernel method (in our case we are based on the Diffusion Map methodology by Coifman, R.R. and Lafon, S. (2006)) that assigns distances (affinities) among the data points. These affinities, which constitute a metric space, have to be converted into coordinates in order to identify their locations. Thus, the application of out-of-sample extension enables to determine whether a newly arrived data point belongs to clusters of normal activities or lies outside these clusters (called deviations, abnormal, anomalous). The organization of the empirical observations into simpler low-dimensional structures is enabled by replacing the diffusion distance (in the source multi-dimensional space) with Euclidian distance of the non-linear embedding and by the application of the out-of-sample extension. We get that the sought after malicious patterns lie in the embedded space outside the normal behaved cluster.

We are able to identify abnormal behavior in the monitored heterogeneous data that is constantly sensed from different networking sources to provide cyber threat management that finds anomalies in this data. Our preliminary results (David,G. (2008)) show that the proposed algorithm (methodology) outperforms in our lab the existing commercial and non-commercial (public domain) known systems. In summary, we propose a comprehensive approach for feature-based organization and analysis of heterogeneous complex networks using newly developed methodologies.

1.3 Related work and the insufficiency of the current cyber protection

A broad class of dimensionality reduction methods is kernel-based methods. The kernel encapsulates a measure of mutual affinities (or similarities) between data points. Spectral analysis (Chung, F. R. K. (1997)) of the associated integral operator of a kernel matrix via singular values decomposition (SVD) obtains an embedding of the data into an Euclidean space, where hidden structures are discovered. The dimensionality of the embedding space is affected by the spectrum's decay rate. Two classical kernel-based methods for dimensionality reduction are principal component analysis (PCA) (Hotelling, H. (1933), Jolliffe, I.T. (1986)) and multidimensional scaling (MDS) (J.B. Kruskal, J.B, (1964). PCA projects the data points onto a space spanned by the significant singular vectors of the data's covariance matrix. The MDS does the same but it uses the Gram matrix of the inner products between data points. Both embeddings preserve significant directions of change in the data, i.e. its directional variances. Other kernel methods use the same mechanism with different kernels (Gram matrices). Examples of kernel methods are diffusion maps (DM) (Coifman, R.R. and Lafon, S. (2006)), local linear embedding (LLE) (Roweis, S.T. and Saul, L.K. (2000)), Laplacian eigenmaps (Belkin, M. and Niyogi, P. (2003)), Hessian eigenmaps (Donoho, D.L. and Grimes, C. (2003)) and local tangent space alignment (Yang, G., Xu, X. and Zhang, J. (2008)). Classical dimensionality reduction techniques fail on datasets that appear to have a complex geometric structure (but a low intrinsic dimensionality). Recently, a great deal of attention has been paid to the so-called "kernel methods" like Local Linear Embedding, Laplacian eigenmaps, Hessian eigenmaps, Local Tangent Space Alignment. These algorithms exhibit two major advantages over classical methods: they are nonlinear, and they are locality-preserving. The first aspect is essential as most of the time, in their original form, the data points do not lie on linear manifolds. The second point is the expression of the fact that in many applications, distances of points that are far apart are meaningless, and therefore need not be preserved. These methods were unable to produce the same quality results. Currently, there is no operational system that does robustly and reliably anomaly detection (Chandola, V., Banerjee, A. and Kumar, V. (2009)) in complex networking as we propose.

Most of the current security systems (IDS, NDS, IPS, firewalls) are based on signatures of intrusions that are developed and assembled manually after a new intrusion is exposed and distributed to security clients who are subscribed on this service. This approach is problematic because these systems detect only already known intrusions but they fail to detect new attacks (zero day attacks). Zero day attacks also called Advanced persistent threat (APT) of malware are our primary interest. Trojans for example cannot be detected by signatures. They have to be detected according to their behavior. Because they come from `legitimate' domains, they jump randomly from one place to another, they can communicate from social networks, there many different types such as bots, downloaders, backdoors, etc. Trojans are like a spying ring, they can emerge only after a long time while assimilated in the system to be looked as a legitimate process. In addition, the current IDS/IPS do not cover a wide range of high quality new sophisticated emerging attacks such as Trojans horses, which do not have signature but a complex unexpected behavior, and SQL injections that exploit innovatively any vulnerabilities. The proposed system will look for a malware pattern that shows that their behavior is deviated from the normal behavior (represented by a normal cluster from the training phase) of the data activities in the system. Therefore, only anomaly characterization, which monitors online and dynamically the behavior of the data, can detect them and is the only way to go. Common methods for anomaly based detection are described in Dua S., Du X. (2011) (Chapter 4).

All the current state-of-the-art IDS/IPS and firewall systems do not supply protections and coverage against a range of different malwares. They work on standard already known attacks and are configured to meet a specific malware. This is a major disadvantage because there is long list of sophisticated malwares that have different behaviors and different targets which penetrate every IDS/IPS and firewall. For example, none of the commercial IDS/IPS can detect Trojans and backdoors and hardly can find viruses and SQL injections and firewalls such as clients (there are more than 100 new each year and they contain many vulnerabilities), chats and IP phones have a way to transfer files that by pass all the available protections and mobile telephony to give some typical examples. Only constant smart monitoring of data behavior can succeed. In addition, all the current commercial IDS/IPS and firewalls are not based on real sound scientific work to enable them to advance. Based on our knowledge and experience as active developers and practitioners in the field, the current commercial available devices are incapable to supply any adequate protection against cyber on CI by the most advanced IPS from CISCO, CheckPoint, Impreve, etc.

2. Solutions to some of the problems

We have used manifold learning via diffusion geometry, which is based on Applied and Computational Harmonic Analysis, that provides a comprehensive synthesis of recent novel tools in signal processing, machine learning, data mining and fast numerical analysis.

We have developed an effective and flexible "empirical network-based" modeling system to enable efficient anomaly detection. In particular, we have built observational models of large computer networks for anomaly detection in network security monitoring. This approach detects new unknown intrusions by learning and modeling the normal "behavior state" of the networks. Detection of Trojans and back doors will be done constantly offline. In other words, we have developed smart clustering (David, G. and Averbuch, A. (2012) – 2 papers) that represents normal behavior, which describes the behavior of the modeled system, and an identification of deviations from this normal behavior which indicate malwares existence.

2.1 Patch processing

Patch processing also called vector processing is the way to choose when we want to manipulate high dimensional data. We assume that the processed data have been generated by some physical phenomenon. Therefore, the affinity kernel will reveal clustered area. In other words, these high dimensional data points reside on several patches located in the high dimensional ambient space. On the other hand, if the data is spread sparsely over the high dimensional manifold, then the application of an affinity kernel to the data will not reveal any patches/clusters. In this case, the only available processing tool is the use of some type of nearest neighbor algorithms. Therefore, data points on high dimensional manifold can either reside in patches and then the methods in Salhov, M., Wolf, G. and A. Averbuch (2013), Wolf, G. and Averbuch, A. (2013), David, G. and Averbuch, A. (2012) are applicable to process it or scattered all over the manifold and thus they were not generated by some coherent physical phenomenon and they are of no interest to us. In general, all the tools that extract intelligence from high dimensional data assume that under some affinity kernel there are data points that reside on patches otherwise no intelligence will be extracted from the data and it can be classified as noise of uncorrelated data points. The clustering and out-of-sample extension Bermanis, A., Averbuch, A. and R. Coifman (2013) computations will be expedited by processing patches instead of single points. Patches enable also to localize spectral processing that is typically global.

2.2 Dictionary building

We combined patch technology in manifolds, dictionary learning and multiscale sampling Salhov, M., Averbuch, A., Bermanis, A., Neittaanmaki, P. and G. Wolf (2012), Salhov, M., Bermanis, A. and Averbuch, A (2012), Wolf, G., Rotbart, A., David, G., and Averbuch, A. (2012) to enable efficient processing of large kernel matrices that appear in data mining processing of massive high-dimensional applications including automatic processing (classification and fusion). These methodologies are combined to form synergy between them to achieve efficient processing of high dimensional data using diffusion geometries of patches. Usually, the kernel that represents the affinities among data points is too big to fit into a computer RAM and a dictionary provides compact description of the processed data points.

2.3 Efficient profile updating

The normal profile, which was generated in the training phase, has to be updated constantly. It can be computational expensive if the whole process of profile generation has to be repeated every time some parameters are modified and thus the profile has to be updated frequently. We have developed procedure how the modified parameters affect the whole process by perturbation analysis of the eigenvectors, eigenvalues and singular values in order to reduce the update time of the modified parameters - see Shmueli, Y., Wolf, G. and Averbuch, A. (2012).

3. Theoretical background

We measure, receive, sense many parameters at every pre-determined time interval - it forms high dimensional data. The data can be heterogenous - not on the same scale: huge database, e-mails, logs, communication packets, computers' data, electronic data, intelligence data, etc. The challenges: How to cluster and segment high-dimensional data that eventually represent the `normal' profile of the data? How to find deviations (malwares) from normal behavior? How can we determine whether a point belongs to a cluster/segment or not? How to process it to find abnormal patterns in ocean of data? How to find distances in high-dimensional data using diffusion processing? We identify data points that deviate from normal behavior which do not reside in the normal cluster/segment/manifold. We have developed methods to process big high dimensional data that is dynamically and constantly changes. We have developed constructive solutions to the above challenges for processing and analyzing multi-dimensional heterogeneous data to find Malware. Our new and disruptive technology is capable of detecting deviation from normal behavior of networks for network health to fuse data that come from heterogeneous sources. The system's breakthrough involves the ability to very quickly learn the normal behavior of the system by clustering and to report any anomalous behavior. This enables the protection of complex networks against a huge variety of new threats, not just classical cases while predicting the emergence of performance degradation. The basic algorithm has two sequential steps: 1. Training step: Study and analysis of the behavior of the datasets while projecting them onto a lower dimensional space. This is done once and updated as the behavior of the training set changes. 2. Detection step: The output from the training step enables online/offline detection of anomalies to which we apply automatic tools for Malware detection. It classifies each new arrival of data to be normal or abnormal via the application of out-of-sample extension (Bermanis, A., Averbuch, A. and R. Coifman (2013)) that is based on online prediction.

Here are the main guidelines of proposed methodology: It builds observational models of large heterogenous data for detection of anomalies and other risks. The proposed approach detects new anomalies by learning from the known normal data by modeling the normal "behavior state" of the captured (sensed) data. The fundamental ingredient enabling such tasks is the ability to organize and model into simple (reduced dimension) geometries a large number of observable quantities in the source multi-dimensional data that are called vectors of observations. We consider the collection of vectors of observations and organize them as a graph in which various vectors of observations are linked by their similarity which were determined by a cluster. And a second graph in which the actual entries in the observation vector are linked through their mutual dependence. It is the spectral and harmonic analysis of the similarity matrix (or dependence matrix) that enables the organization of the empirical observations into simpler low dimensional structures. Each similarity matrix leads to a related diffusion or inference geometry on which statistical analysis and detection of anomalies is much simpler. Various types of observations each leading to their own model can be fused into a hyper model relating all the critical model parameters. These methods provide a far reaching nonlinear extension of conventional linear statistical tools such as principal components analysis (PCA), and independent components analysis (ICA). Our methods reduce the observed data to allow a small number of parameters (features) to model all the variability in observations. Moreover, our method computes automatically the parameters associated to new data as it sensed by the system, and updates dynamically the observations. There are several key enabling ingredients:

- A robust similarity relationship between two observation vectors is computed as a combination of all chains of pairs linking them. These are the diffusion inference metrics;
- Clustering in this metric leads to robust segmentation of observations and characterization of network regimes;
- Various local criteria of linkage between observations lead to distinct geometries. In these geometries, the user can redefine relevance and filter away unrelated information;
- Self organization of network observations is achieved through local similarity modeling. Several discriminating eigenfunctions of the matrix defining the pair linkages provide global organization of the given set of observations;
- The diffusion maps Coifman, R.R. and Lafon, S. (2006) embed the data into low dimensional Euclidean space and convert isometrically the (diffusion) relational inference metric to the corresponding Euclidean distance;
- Diffusion coordinates can easily be assigned to new data without having to recompute the map as new data streams in;
- Diffusion metrics can be computed efficiently as an ordinary Euclidean distance in a low dimensional embedding by the diffusion maps/bases (total computation time scales linearly with the data size, and can be updated on line);

- Data exploration and perceptualization is enabled by the diffusion maps/bases since it converts complex inference chains to ordinary physical distance in the perceptual displays to provide situational awareness of the estate of the observed system/data;
- The diffusion geometry, which is induced by the various chains of inference, enables a multiscale hierarchical organization of regional folders of observations corresponding to various states of the system/data;
- Data fusion is achieved through the tensor product of embedded diffusion embeddings of heterogeneous data;
- Data/system dynamics are evaluated by building the network of dynamic temporal observations. For sensor networks and communication networks the dynamics of the links impact the network embeddings and enables analysis.

This is a comprehensive approach to feature-based organization and analysis of heterogeneous complex networks. The key enabling methodology is derived from diffusion geometries, which provide a comprehensive synthesis of recent novel tools in signal processing, machine learning, data mining and fast numerical analysis. In the context of analysis on network graphs (which provides the unifying structure for geometrization and processing of networks), we have developed methodologies for geometric analysis and automated feature optimization, including adapted dynamic analysis of changing networks. We will develop a range of tools to assist the network analyst by providing systems of network observables and a model-building environment enabling massive data information extraction and assessment. In particular, we see inferential methods of diffusion harmonic analysis as a toolkit to enhance and reinforce most network-processing tasks, enabling efficient automation of integration and information fusion, from heterogeneous sources.

3.1 Diffusion geometries

Diffusion geometries are various geometries that can be defined on an abstract graph or network, or on a cloud of (digital) data points in the weights on the edges connecting two points (vertices) enabling the definition of diffusion processes (of "affinities", "inferences", "uncertainties", "relevance", etc.). This diffusion process leads to the introduction of multiscale diffusion geometries as well as diffusion distances. Moreover, it enables the organization of the nodes of a network into a tree hierarchy of "affinity folders" (David, G. and Averbuch, A. (2012),) or sub networks, at different scales, (generalized network quad trees). In particular, the eigenfunctions of the diffusion operators, or sometimes a Laplacean on a graph, provide useful empirical coordinates, which enable an embedding of the network to low-dimensional spaces so that the diffusion distance at time t on the original data becomes Euclidean distance in the embedding, providing nonlinear coordinates (generalizing the SVD) as well as a powerful dimensional reduction methodology. The diffusion at different times leads to a multiscale analysis generalizing wavelets and similar scaling mechanisms.

References

Belkin, M. and Niyogi, P. (2003), “Laplacian Eigenmaps for Dimensionality Reduction and Data Representation”, *Neural Computation*, Vol. 15, pp. 1373-1396.

Bermanis, A., Averbuch, A. and R. Coifman (2013), “Multiscale Data Sampling and Function Extension”, *Applied and Computational Harmonic Analysis*, Vol. 34, 15-29.

Chandola, V., Banerjee, A. and Kumar, V. (2009), “Anomaly Detection: A Survey”, *ACM Computing, Surveys*, Vol. 41, No. 3, pp. 1-58.

Chung, F. R. K. (1997), *Spectral Graph Theory*, AMS Regional Conference Series in Mathematics, 92.

Coifman, R.R. and Lafon, S. (2006), “Diffusion maps”, *Applied and Computational Harmonic Analysis*, Vol. 21, pp. 5-30.

Coifman, R.R. and Lafon, S. (2006), “Geometric harmonics: a novel tool for multiscale out-of-sample extension of empirical functions”, *Applied and Computational Harmonic Analysis*, Vol. 21, pp. 31-52.

David, G. (2008), “Anomaly Detection and Classification via Diffusion Processes in Hyper-Networks”, *Ph.D Thesis*, Tel Aviv University, March 2008.

David, G. and Averbuch, A. (2012), “Hierarchical Data Organization, Denoising and clustering via Localized Diffusion Folders”, *Applied and Computational Harmonic Analysis*, Vol. 33, pp. 1-23.

David, G. and Averbuch (2012), “SpectralCAT: Categorical Spectral Clustering of Numerical and Nominal Data”, *Pattern Recognition*, Vo. 45, No. 1, pp. 416-433.

Donoho, D.L. and Grimes, C. (2003), “Hessian Eigenmaps: New Locally Linear Embedding Techniques for High Dimensional Data”, *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 100, pp. 5591-5596.

Dua S., Du X. (2011), *Data Mining and Machine Learning in Cybersecurity*, CRC Press.

Economist (2010), Cyberwar: The threat from the internet, July 3, pp.22-24.

Jolliffe, I.T. (1986), *Principal Component Analysis*, Springer, New York, NY.

J.B. Kruskal, J.B, (1964), "Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis", *Psychometrika*, Vol. 29, pp. 1-27.

Hotelling, H. (1933), "Analysis of a Complex of Statistical Variables into Principal Components", *J. of Educational Psychology*, Vo. 24.

Roweis, S.T. and Saul, L.K. (2000), "Nonlinear Dimensionality Reduction by Locally Linear Embedding", *SCIENCE*, Vol. 290, pp. 2323—2326.

Salhov, M., Wolf, G. and A. Averbuch (2013), "Patch-to-Tensor Embedding", *Applied and Computational Harmonic Analysis*, Vol. 33, pp. 182-203, 2012.

Salhov, M., Bermanis, A., Averbuch, A., Neittaanmaki, P. and Wolf, G. (2012), "Dictionary construction for patch-to-tensor embedding", submitted.

Salhov, M., Bermanis, A. and Averbuch, A (2012), "Dictionary based multiscale expansion for spectral methods", in preparation.

Shmueli, Y., Wolf, G. and Averbuch, A. (2012), Updating Kernel Methods in Spectral Decomposition by Affinity Perturbations, Linear Algebra and its Applications, 437(16), 1356-1365.

Wolf, G., Rotbart, A., David, G., and Averbuch, A. (2012), "Coarse-grained localized diffusion", *Applied Computational Harmonic Analysis*, Vol. 33, pp. 388–400.

Wolf, G. and Averbuch, A. (2013), "Linear-projection diffusion on smooth Euclidean submanifolds", *Applied and Computational Harmonic Analysis*, Vol. 34, pp. 1-14.

Yang, G., Xu, X. and Zhang, J. (2008), "Manifold Alignment via Local Tangent Space Alignment", International Conference on Computer Science and Software Engineering.

Strategic Communication of the EU: The Counter-Terrorist Aspect

Darya Yu. Bazarkina
Sholokhov Moscow State Humanitarian University, Moscow, Russia
Bazarkina-icspsc@yandex.ru

Abstract: The urgency of the strategic communication problems («the synchronization of words and deeds and how they will be perceived by selected audiences, as well as (b) programs and activities deliberately aimed at communicating and engaging with intended audiences, including those implemented by public affairs, public diplomacy, and information operations professionals» (White House, 2010, p. 2) keeps under the conditions of contemporary conflicts as high as never before. "War on Terror" which is conducted today by the West countries, became a part of their strategic communication which is oriented both on internal, and on external for the EU target audiences. At the same time discussion on a problem of terrorism has affected many people who felt themselves as victims of stereotypes, for example, certain Moslems groups in Europe. It has a negative impact on the image and reputation of law enforcement agencies of the European Union. The key roles in the European authorities' "War on Terror" interpretation play the supranational structures which are responsible for the regional security. In the present paper we make an attempt to analyze the basic problems of strategic communications of those European structures, caused by "War on Terror" contradictions. The author of the present paper tries to examine the following questions: To analyze the basic messages of the antiterrorist strategy in the Organization for Security and Co-operation in Europe (OSCE) documents. To characterize in brief the main problems and contradictions of the communication struggle against terrorism in the European Union. It is used in the paper both methodology of historical research, and methods of communication management which in system allow to track the development of the governmental strategic communication while the crisis spread all over the world.

Keywords: terrorism, strategic communication, communication management, European Union, OSCE, SIPRI

1. Premise

Amid the global economic crisis, the spreading of propaganda of terrorist organizations acquires new relevance, which demands the re-evaluation of communication strategies aimed at combating terrorism. Manipulative use of the idea of social change is inherent in terrorists, representing different ideologies, as indicated by the texts by al-Awlaki, A. Breivik, "Informal Anarchist Federation" from Italy, etc. At the same time, the ideas of transformation proclaimed by terrorist organizations, are in fact the Marxist ideology or the idea of national liberation, arbitrary distorted in the interests of terrorists. It's pretty well allows terrorists to mislead large masses of those people, dissatisfied with the current political and economic system, thus provoking them to violence. Perhaps only the fascist ideology is an organic basis for the activities of several right-wing organizations.

At the same time, various strategies to respond to terrorist activities offered by experts in the fields of public relations, journalism, political science, etc., after having been tested in Europe, have shown both the advantages and the objective weaknesses. In particular, the discussion of the problem of terrorism within the "religious" discourse has not resulted in active, uniform involvement of different religions and faiths to the counteracting the spreading of terrorist propaganda.

In recent years, the need in revising the basic concepts of the struggling against terrorism, that have been implemented since 2001, is recognized by both the scientific community (for example, see Crelinsten R., 2009) and the representatives of the law enforcement agencies, as well as by the journalists who report safety problems. "Experts from the U.S. and European intelligence agencies say that for the five years of the war against terrorism, we have learned a good deal about the mechanism of functioning and the weaknesses of terrorist cells ... Paradoxically, the so-called "war on terror" has caused more moral damage to the member-states of the belligerent coalition than that caused by terrorist groups against whom it is directed", (Unigovskaja S., Edynak Ja., 2012, pp. 25 – 26).

Insufficient degree of development of the EU strategic communication concept, in particular in the security context, is noted in the West too: "If the EU can deliver messages to local publics in the countries where its mission are deployed, domestic audiences seem to be, paradoxically, more difficult to reach. Indeed, member states will not allow supranational institutions to influence their own population. But member states do not relay much of EU's communication either. About countries that host an EU CSDP (Common Security and Defence Policy) mission, member states give little information about the ongoing mission. On the contrary, member states release more information about their operations in Afghanistan. Such a difference is explained by the command structure. Contributing nations to NATO operations can present caveats and keep a certain control over their troops. Seconded personnel to CSDP missions are under direct control of the Head of mission, which creates a certain distance between domestic public and their fellow citizens serving overseas under EU flag» (Jacquemet O., 2011).

Antulio J. Echevarría II, leading military analyst from the US Army War College, identifies four main types of ideological warfare in the modern world – the intellectual debates, wars over religious dogma, ideological wars and advertising campaigns. The various elements of the "war on terror" may be in one or another degree attributed to a particular type of ideological warfare. For example, in discussing the intellectual debates, which are "the disputes in which opposing sides advance their arguments, support them with evidence, and endeavor to refute the reasoning and conclusions of the other" (Echevarria A. J. II, 2012, p. 41), one should remember, that the focus of an intellectual debate today is in some cases determined by the economic and political priorities of the parties, by the factor of influence of domestic and foreign policy agenda and other factors, that is even a scientific discussion, while striving to remain objective, cannot be free from external effects.

Media is in even greater degree susceptible to these effects, as today they operate under the influence of the "CNN effect" and thus they are often unable to adequately analyze the current events in a short period of time. Political communications, whose subjects are vested with the special responsibility for the discourse, existing in the communication space with regard to a particular problem, are the tools of these external influences. At the same time the importance of the communication aspect of activities of the authorities is growing: the world community recognizes the inefficiency of military methods to combat terrorism (SIPRI Yearbook, 2005, p. 282), and thereby the analysis of the mechanisms of ideological wars seems particularly relevant.

Crisis communication, the example of which in the "war on terror" is the reaction to the terrorist attacks, is carried out in conditions of communication management, under the influence of a number of complicating factors, including, but not limited to, the element of surprise, the additional threats posed by the crisis, the shortage of time to respond (Ulmer R., Sellnow T., Seeger M., 2011, p. 18 – 20.), the lack of information (Aleshina I. V., 2006, p. 375), the need for decision making under the stress caused by the instantaneous revealing of shocking information (Fearn-Banks K., 2011, p. 35), etc. According to some experts, the 9/11 attacks exhibit some fundamentally new properties, including the scale, the character of attacks (sudden, occurred in the American territory, no weapons or sophisticated technologies used), a new type of enemy (non-state criminal organization), the choice of the target (the mightiest world power) (SIPRI Yearbook, 2002, p. 4). In our view, the problems that have arisen in the EU in implementing the crisis communication in response to terrorism, are in a large extent conditioned by these and by other factors, which have prevented the European governments to promptly identify the target audience and to adjust their key messages in line with the expectations of this audience. This led to the fact that the discussion of the problem of terrorism has affected the feeling of lots of people, who had felt themselves victims of stereotypes, for example, those that have formed in respect of the European Muslims.

2. The communication aspect of the counter-terrorist struggle in OSCE documents

The reaction to the events of 9/11 can be considered a starting point in the developing of a new European strategy to combat terrorism, as well as the beginning of a new period of understanding of the phenomenon of terrorism. At the same time a key role in the interpretation of the "war on terror" by the European authorities is played by the supranational agencies responsible for security in the region.

For example, the Bucharest Declaration of the OSCE Ministerial Council, adopted in response to terrorist attacks in the United States, condemns all the acts of terrorism, "whatever its motivation or origin". Also, "Reflecting the OSCE's solidarity, the Ministerial Council has adopted today a decision and Action Plan on Terrorism" (OSCE, The 9th OSCE Ministerial Council Meeting, 2001). It should be noted that although the governments of the EU have supported the "war on terror", there are fundamental differences in the methods of conducting of this "war" between the U.S. and the EU, that can be traced from the very beginning. For example, the experts from the Stockholm International Peace Research Institute (SIPRI) have differentiated the understanding of the objectives of counter-terrorism:

- the mission of the US is to defeat and to destroy the Al-Qaeda;
- the mission of the European countries is to implement the wide spectrum of measures aimed at the "elimination of sources of terrorism with focus on non-military methods" (SIPRI Yearbook, 2002, p. 2).

The process of the rethinking the situation in Europe proceeded in parallel with the establishment of the new security institutions, that has resulted in the delayed development of counter-terrorism strategy, primarily as the response to the already existing and rapidly evolving terrorist threat. Gudrun Steinacker notes, that "Before 11 September 2001, combating international terrorism was just one among many aspects of promoting security in the OSCE. The attacks on the World Trade Center and the Pentagon, however, led to the immediate adoption in Bucharest at the annual OSCE Ministerial Council on 3 and 4 December 2001 of a Plan of Action for Combating Terrorism. The Action against Terrorism Unit (ATU) was established within the OSCE Secretariat in Vienna, and the post of Co-ordinator on Anti-Terrorism Issues was created within ODIHR (Office for Democratic Institutions and Human Rights) in Warsaw with the task of co-ordinating all projects and joint activities related to terrorism with the Secretariat" (Steinacker G., 2003 (Russian edition, 2005), p. 69). Thus, crisis response plans had to be developed in the new environment, in the framework of the new organizations in a fairly rapid pace. An additional problem were the words of George W. Bush, who announced a "crusade" on the Arab World.

Originally the communicative component was not part of the anti-terrorist priorities, unlike for example, the counteracting the funding of terrorism. The resolution on counteracting terrorism says, that "there must be no safe haven for those perpetrating, financing, harbouring or otherwise supporting those responsible for such criminal acts» (OSCE, The 9th OSCE Ministerial Council Meeting, 2001, p. 7).

The governing factor in determining the communication strategy as part of the terrorism combating was the provision on compliance by the OSCE in the course of its counter-terrorist activities with the standards of international law and the human rights: "The OSCE participating States will not yield to terrorist threats, but will combat them by all means in accordance with their international commitments... They will defend freedom and protect their citizens against acts of terrorism, fully respecting international law and human rights. They firmly reject identification of terrorism with any nationality or religion and reconfirm the norms, principles and values of the OSCE" (OSCE, The 9th OSCE Ministerial Council Meeting, 2001, p. 7).

Initially the development of any radically new directions of activities was not presumed: "The aim of the Action Plan is to establish a framework for comprehensive OSCE action to be taken by participating States and the Organization as a whole to combat terrorism... The Action Plan seeks to expand existing activities that contribute to combating terrorism, facilitate interaction between States and, where appropriate, identify new instruments for action". The Action Plan provided for the three categories of the anti-terrorist activities:

- implemented immediately;
- medium term activities;
- long term activities.

The main areas of cooperation, as set forth by the Ministerial Council of the OSCE, first of all provided for the practical activities within the frameworks of the counter-terrorism operations, as well as the measures to control the movements of people and financial flows. Communication component was not yet in the list of the main measures, that included:

- police and judicial co-operation;
- prevention and suppression of the financing of terrorism;
- denial of other means of support; border controls including visa and document security;
- access by law enforcement authorities to information (OSCE, The 9th OSCE Ministerial Council Meeting, 2001, p. 9).

The actions, in one or another way connected with the communication component, within the Bucharest Actions Plan act as the applied ones to the major actions. In our opinion these actions are initially motivated by the striving to ensure the respect of the human rights and the international law, as declared above. The preventive counter-terrorism measures include those, aimed at the cooperating with the media, promotion of appropriate values, and so on: "Participating States/Permanent Council/ODIHR/High Commissioner on

National Minorities (HCNM)/Representative on Freedom of the Media: Will promote and enhance tolerance, co-existence and harmonious relations between ethnic, religious, linguistic and other groups as well as constructive co-operation among participating States in this regard. Will provide early warning of and appropriate responses to violence, intolerance, extremism and discrimination against these groups and, at the same time, promote their respect for the rule of law, democratic values and individual freedoms..." (OSCE, The 9th OSCE Ministerial Council Meeting, 2001, p. 10).

Exclusively communication component, that is the problems and information-sharing mechanisms, have been isolated, mainly within the expert community in order to facilitate the analysis. Thus, the OSCE Permanent Council on the Bucharest Actions Plan had to consider the issue on the arranging of "regular meetings of law enforcement officials of participating States and, where applicable, of OSCE experts with relevant experience in the field to exchange best practices and ways of improving co-operation" (OSCE, The 9th OSCE Ministerial Council Meeting, 2001, p. 11). This activity became the second priority (along with propaganda) in the implementing of communication not only by the OSCE, but by other major international organizations. On March 6, 2003 the special meeting of the Counter-Terrorism Committee was held in the the UN Security Council in New York.

3. Problem of the “religious” terrorism discourse

Thus, the preventive measures to combat terrorism in part of the communications, provided for in the OSCE documents, are first of all dictated by the wish to demarcate the religion, the nation and the threat of terrorism. In our opinion, despite the obvious need in achieving this goal, the declaring of this goal in the context of an open discussion of the communication strategy of the state or the supra-national law enforcement structures is not quite rightful. In discussing, for example, the "Jihadi" terrorism as a phenomenon, which has been formed under the influence of the differences in the religious debates between the supporters of traditional and radical Islam, the target audience gets the message that the anti-terrorist structures "between the lines" recognize the problem of ethno-religious terrorism as a problem, the roots of which must be searched *exclusively in the religious domain*. Against this background, the assurances that European governments are not against religion, but only against crime and political violence, are less convincing.

To some extent, such a choice of the strategy had been conditioned by the influence of "political" interpretation of terrorism, formed after the 9/11, when the discussions of the criminal essence of terrorism have been overpowered by the discussions of its political motivation. Widespread discussion of the political component of terrorism presented the masterminds and the perpetrators of the terrorist attacks, primarily as the communicators of a particular political point of view. And any political power inevitably finds supporters among ordinary people, intellectuals and even the elites.

A good example of perception of the combating terrorism as the problem of the religious discussions is the study of Abd al-Hakim Carney, the scholar of the School of Oriental and African Studies in London, who interprets the activities of Osama bin Laden and his associates as the actions within the Salafi movement - the fundamentalist sect in Islam, which declares its goal the return to traditions that existed at the times of the formation of the Islamic religion, that is as the actions, conditioned exclusively by the religious factors. Abd al-Hakim Carney defines two broad groups in the Salafi movement:

- The official Salafism of Saudi Arabia, epitomized by religious leaders such as Sheikh Abdul Aziz bin Abdullah bin Baz or Sheikh Muhammad Nasiruddin al-Albani. These groups are avidly apolitical. They have compromised firmly with the Saudi regime and have no revolutionary ambitions. They are far more concerned with correcting what they feel are heretical practices among the Muslims (so-called Islamic clothing as well as an extremely close scrutiny of minor ritual practices and declarations of takfir).
- "The so-called jihadi wing of the Salafis is epitomized by Osama bin Laden: the Saudi regime is considered to be one of the worst cases of bid’a and should be replaced by a pure Islamic state modeled on the rashidun period. The apolitical Salafis are basically a quiet religious sect; the Bin Laden movement believes in the full and free use of violence in pursuing their goals" (Carney A. al-H., 2005, p. 179).

This standpoint correlates with the assessment, the SIPRI (Stockholm International Peace Research Institute) experts made in 2002: "From the strategic standpoint the September 11 terrorist attacks may be treated as part of a kind of civil war in the Islamic world. Terrorists were striving to force the US to eliminate their presence in the Islamic world: to this end they wanted to excite the neo-isolationism mood and thus to put strain on the US government from the part of the American people" (SIPRI Yearbook, 2002, p. 5).

This understanding in the West of the terrorist activity of Al-Qaeda as a practical expression of the "political Islam" became an additional tool of terrorist propaganda, that manipulates the Islamic values in the interests of terrorists. When some of the scientific community' members being affected by rapidly evolving crisis, began to express a point of view that the new type of terrorism is in fact the phenomenon of a religious nature, this has obviously created certain stereotypes of the negative perception of Muslims even among well educated residents of the EU countries.

It should be noted, however, that the manipulative nature of terrorist propaganda begins to be discussed in the scientific community at a fairly early stage. For instance, the SIPRI experts have in 2002 recognized the distortion of the Islamic standards: "There is no religion, that would be based on evil, but in the course of multiple conflicts the teachings of the world religions have been twisted and employed into evil" (SIPRI Yearbook, 2002, p. 10). However, the estimates inherent in the rhetoric of the "war on terror" are becoming the cause for some absolutization of the problem: "Under the influence of September 11, 2001 there are concerns, that wherever there is any conflict, it may lay the stage for the creation of the terrorist Diaspora". Thereafter in 2006 the decisive role of education in the modern “religious” terrorism was recognized. The fact of destruction of the monolithic terrorist organizations and the transformation of jihadi terrorism in the extensive network with no solid core (SIPRI Yearbook, 2006, p. 141, 150) has been also recognized.

Thus, the standpoint, which in many respects determines the discourse of a new type of terrorist threat, it is a "religious" concept of terrorism, which in turn determines the "military" perception of combating against terrorism. Statement by George W. Bush ("this crusade, this war on terrorism is going to take a while" (citation taken from: Shakleina T. A., 2002), as well as public recognition of the fact that the assessment of the problem of terrorism today is a prerogative above all of experts in religion, have had certain managerial effect and caused growth of the stereotypical identifications of terrorism with religious practice.

One can assume, that the "religious" understanding of terrorism is currently the efficient tool aimed at the diverting attention from its economic component, which tool is being widely used by the subjects of terrorist propaganda. In particular, this helps to expand the target audience of terrorists through the engaging of those who are about to interpret the "religious" discourse in the anti-Islamic manner. In defending the discriminated these people take an anti-Western stand, which is the goal of the terrorist propaganda. In so doing the primary target audience of such propaganda is replenished by potential extremists from the number of people moving to Islam only by reason of its "political" interpretation. "The appeal of the Global Islamic Media Front was posted at one of the extremist web sites ... It states that "the new Al-Qaeda soldiers were born in Europe from European and Christian parents. They drink alcohol and eat pork, but Al-Qaeda takes them as they practice Islam in secret, share the philosophy of Al-Qaeda and are prepared to use weapon. They walk through the streets of Europe and America while they are preparing the new attacks" (Grinevskiy O. A., Gromyko A. A., 2009, p. 15).

A certain lag in the development of a communication strategy could be observed after the terrorist attacks in New York and in analyzing the EU documents. In developing the specific practical measures the emphasis was made on peacekeeping operations - the military missions in countries of interest to the U.S. and its European allies. The communication component had been developing at the level of the "improving the system of information and the early warning systems" (Zhurkin V. V., 2005, p. 51), but not at the level of the "ideological war" against terrorist organizations.

Crisis communications of governmental organizations of the EU immediately after the attacks on the WTC conformed with classical recommendations, according to which, in addition to the focusing the attention on specific anti-crisis measures, it is required to express condolences to the countries whose citizens died as a result of the crisis. Thus, "On September 12, 2001 the meeting of the European Council was urgently convened and held in Brussels. The Council has adopted the Declaration of the European Union, that expressed full solidarity with the United States Government and the American people ... ". The Council has declared the September 14, 2001 the day of mourning throughout the EU (for the first time in history)" (Zhurkin V. V., 2005, p. 49).

On September 20, 2001 an emergency meeting of the EU summit held in Brussels, adopted the first comprehensive anti-terrorist program - the Action Plan, followed by the development of the respective Road Map, consisted of 68 counter-terrorism activities and measures of the European Union. The priority sector was the "earliest possible achievement of operational readiness of military forces of the EU".

V. V. Zhurkin notes, that the European Security and Defence Policy (ESDP) has not undergone any radical changes in connection with the attacks of Sept. 11, 2001. "The EU has actively and unanimously supported military campaign against al-Qaida and the Taliban, launched by the U.S. and Great Britain on October 7, 2001. And in early 2002 ... another three EU member states - France, Germany and Denmark - sent their special purpose units to Afghanistan" (Zhurkin V. V., 2005, p. 50).

Ch. Grant, Director of the London-based Center for European Reform, agrees with increase of the military expenditures, that steadily declined before the September 11th attacks. He thinks, that Post-September 11th, it is easier to argue that countries need effective armed forces, and that the EU needs to be able to deploy troops. Evidently, a war against terrorism requires many capabilities, including more and better trained spies, but it needs military hardware too. One reason why the US has chosen to run a largely unilateral military campaign in Afghanistan is that the Europeans do not have many of the most useful kinds of military asset…" (Grant Ch., 2002, p. 62). However, to reduce the weaknesses existing in combating terrorism, exclusively to the lack of military force is extremely dangerous for modern Europe, since the idea of the "war on terror" is undergoing the crisis, since it has passed into protracted military conflict, primarily aimed to pursue their own economic and political interests of different political groups in the Middle East.

Intrinsic weaknesses of the EU communication strategy as part of the counteracting terrorism result the deepening of the crisis of public and political life, which manifests itself in a variety of aspects.

According to experts the Institute of Europe of the RAS, the multi-ethic society (the term introduced by American politologist Arend Lijphart), which is forming in the EU, "can be highly unstable, prone to ethnic nationalism as opposed to civic nationalism" (Grinevskiy O. A., Gromyko A. A., 2009, p. 54). In these circumstances, the discussion of terrorism is closely connected to the discussion of the relationship of the indigenous European population and migrants.

In a situation where "... on the one hand, there is a strengthening of the right and the center-right forces in Europe, which trend has been once again demonstrated by the European Parliament elections in June 2009, and where on the other hand, the yesterday' s migrants (Grinevskiy O. A., Gromyko A. A., 2009, p. 54) increasingly use violent methods in asserting their rights and the worldview", the political arena is increasingly infiltrated by such people as the "Toulouse shooter" or Anders Breivik, which fact has been for already for a long time considered as the symptom of the crisis of the policy of multiculturalism, and in some cases even the representatives of the major political parties of Europe declared its apparent collapse. Against this background, the extremist and terrorist propaganda in Europe is likely to be even stronger. "Religious" interpretation of terrorism (search for its roots in the religious sphere only) on a background of crisis of multiculturalism in its present form (i.e. multiculturalism with no practical outcome, unable to offer common ideology for EU citizens belonging to different nations and religions) results in weak ability to bring together representatives of Eastern and Western cultures to actively combat the communication effect of terrorism.

4. Conclusion

In view of the foregoing, one of the possible measures, aimed at the minimizing of ethnic and religious radicalization as part of the communication policy of the authorities, is to promote the perception of terrorism as a force hostile to the working population of multicultural Europe. Of course, the implementation of such a concept is only possible within a coherent communication strategy aimed at positioning of the secular state (within the frameworks of this state the gradual unification of the efforts of European and foreign employees in achieving common prosperity becomes possible). In so doing, in implementing of this strategy the statements of the authorities and their real deeds must comply with each other: availability of equal labor conditions under which the general criterion of recruitment is the qualification of people, as well as the opportunities to enhance the qualification in conditions of enhancing employment. In other words, the basic objective of the implementation of counter-terrorism is the overcoming the socio-economic crisis.

We believe, that another relevant component of the communication strategy would be the promotion of the idea of the destructiveness of terrorism regardless of its ideology and the stressing of regressiveness, barbarism of the new round of "religious wars" in the XXI century, as they are seen by terrorists, who manipulate Islamic, Christian or other religious values. Moreover, in our view, there is a need in developing of messages, that would motivate the potential target audiences of terrorist propaganda in promoting the development of critical perception of reality. This task is adequate for both the European and the current Russian situation.

References

Carney A. al-H. (2005). Analyzing Political Islam: The Need for a New Taxonomy. *Annual Report on OSCE Activities*, 2003 (Russian edition, 2005). Moscow.

Crelinsten R. (2009). Counterterrorism. Cambridge: Polity Press.

Echevarria A. J. II (2012). Wars of Ideas and the War of Ideas. *Politika.* No. 96.

Fearn-Banks K. (2011). Crisis Communication. A Casebook Approach. 4th Edition. New York – London: Routledge.

Grant Ch. (2002). Strengthening the Common European Security and Defense Policy. *Europe after September 11th.* Moscow.

Jacquemet O. (2011). Should the European Union develop its Strategic Communications capabilities? *Echo Sierra – Thoughts on Conflicts, Peace and Defence policies.* Available from: echo-sierra.net/2011/08/20/should-the-european-union-develop-its-strategic-communications-capabilities/ [Accessed 20 January 2013].

OSCE, The 9th OSCE Ministerial Council Meeting (2001). The 9th OSCE Ministerial Council Meeting, Bucharest, December 3-4, 2001. Available from: www.osce.org [Accessed 20 January 2013].

SIPRI Yearbook (2002). Armaments, Disarmament and International Security. Moscow: IMEMO RAS.

SIPRI Yearbook (2003). Armaments, Disarmament and International Security. Moscow: IMEMO RAS.

SIPRI Yearbook (2004). Armaments, Disarmament and International Security. Moscow: IMEMO RAS.

SIPRI Yearbook (2005). Armaments, Disarmament and International Security. Moscow: IMEMO RAS.

Steinacker G. (2005). The Role of the OSCE as a Regional Security Organization in Combating International Terrorism. *Annual Report on OSCE Activities*, 2003 (Russian edition, 2005).

Ulmer R., Sellnow T., Seeger M. (2011). Effective Crisis Communication. Kharkov, Humanitarian Center.

White House (2010). *Strategic Communications report to Congress "National Framework for Strategic Communication"*, 16 Mar 2010, released 17 March 2010. Government Information Earl Gregg Swem Library.

Aleshina I. V. (2006). Public Relations for Managers. Moscow, Eksmos.

Grinevskiy O. A., Gromyko A. A. (2009). The Problems of the Extremism and Terrorism in Europe: Roots and Consequences. *RAS Reports.* №239. Moscow, 2009.

Zhurkin V. V. (2005). The EU in 21st Century: European Security and Defence Policy. *RAS Reports.* №170. Moscow, 2005.

Unigovskaja S., Edynak Ja. (2012). Terrorism: Roots, Development, Forecasts. *Politika.* No. 96.

Shakleina T. A. (2002). Republicans' New "Crusade": How Bush Doctrine Appeared. *International Trends.* Available from: www.intertrends.ru/three/017.htm [Accessed 28 October 2012].

Unrestricted Warfare Versus Western Traditional Warfare: A Comparative Study

Grégory Commin and Éric Filiol
ESIEA-OUEST, (C + V)^{O} Laboratory, Laval, France
gcommin@et.esiea-ouest.fr
Eric.FILIOL@esiea-ouest.fr

Abstract: The rise of the cyber dimension as well as the emergence of new strategic/economic leadership in the world, like China, is currently changing not only the face of the world but also are about to upset the strategic balance in the world. At the same time, it is the definition of the concept of warfare itself that must be redefined. In this paper, we intend to analyse the concept of new warfare precisely, then we take up our development on the new art of warfare and finally, we establish the redefinition of the stakes.

Keywords: warfare, doctrine, cyberwarfare, China, Qiao Liang, Wang Xiangsui, modern warfare

1. Introduction

The word "warfare" draws its etymology from the inheritance of the Middle-East and the Latin name "*bellum*". At the root, it has caused over some years the most abhorrent violence and the gloomiest stupidity. It is hideous and it is the mother of all crimes *[1]*. It features in the preoccupations of statesman and it is considered sometimes as the pursuit of politics by the other means *[2]*. Justifications for warfare were made until the XIX^{e} sc. In the course of its evolution, it has turned in the XX^{e} sc., into demographic warfare whose goal is the massive destruction of people and commodities, for example cold warfare. It takes all possible forms of confrontation, from espionage to technological competition in the domain of the conquest of space (The cold warfare had taken place in the entire world with indirect fights between two power states through their respective partners. It is characterized by a bipolarization of world). That is the principle of traditional warfare. Consequently, it has usurped a new identity during the XXI^{e} sc. Technology has become one of the crucial means of the conquest of the new space, which itself brings some sense of globalization and in parallel, the sense of time is also essential for this evolution, so the idea of acceleration is coming. Warfare has died out in its old form but it has not been totally abolished.

By this mind-blowing multiplication of these methods and techniques to improve warfare, what about the warfare to peace? This jutting concept of absence of perturbation appears insufficient to Spinoza (In the XVIII^{e} s.c, the philosopher Spinoza criticizes the traditional theology and the idea of a God. He designs the materialists improperly). This ideal is always considered but the goal is almost never reached, put aside with the intervention of organizations such as the United Nation (The United Nation is an organization which we found all the states of the world. Its aim is the international peace. Their targets are making the cooperation easier in the field of international rights, the international security, and the economic development, the social in progress, the human rights and the realization of the peace world eventually) or League of Nations (The League of Nations is an international organization in order to protect the peace in Europe at the end of the first warfare world. Their targets include the disarming, the warfare prevention through principle of collective security and the resolution of conflicts by the negotiation. The quality of life is on the upswing) which made an essential point of honor. In this way, it could be examined distinctly this success toward the cool-down that we could define like a win. This feeling is achieve either by the physical extinction of the enemy, or by a giving up of his claims and of his surrender. This international harmony come through besides the borders but will be fraught with difficulty in the coming years.

The sense of limit during this hardship of strength organized in the struggle form between the States is a more controversial subject for the coming years. The state is the organized society; it is endowed with a government and considered as a moral authority in the regard the others identically organized. Two high ranking officers of the Chinese military, Qiao Liang -- Qiao Liang is a Major General in the Chinese Army. He has been member of writer Union of China. He was assistant director for the political department to Army air corps. He made several researches on the military theory and he is author of book "Unrestricted Warfare" (Qiao & Wang, 1999) with another Chinese senior officer -- and Wang Xiangsui -- Wang Xiangsui was a political instructor, a political commissar. He has been the position of Major for Army air. He collaborated on writing of book

"Unrestricted Warfare" -- allude principally to the cold warfare in their book to expose the elaboration of the future Chinese strategy with regard to new conflicts and tension in the world.

The question which we will risk facing in following the years is: Under which face will "*the warfare*" be presented? And what are the consequences for different great powers?

This article gets organized as follows: Firstly, we will broach the concept of new warfare, then we take up our development on the new art of warfare and finally, we establish the redefinition of the stakes.

2. On new warfare

Chinese military officers, Qiao Liang and Wang Xiangsui redefine warfare in their book "Unrestricted Warfare", vis-à-vis many hostile acts carried out in the upstream and downstream section of cold warfare. They show us that warfare is no more "*the use of armed force forcing the enemy to bend to our wishes*" but rather use "*of all the ways, which the armed force or not armed, military or not military […] to force the enemy to submit to its own interests*" *[3].* It is at this moment that humanity envisages replacing the creator. Along the reading of this exceptional book, the authors value the sense of initiative to the Chinese military to manifest a hard wish to this point of view.

The Chinese military is trying to proceed differently following many failures have gone to the front and faced with U.S fire power on the battlefield. Indeed, everything started off with the revolution of weapons. The American military has greatly demonstrated its superiority on the ground with military arms and advanced technologies more and more exact. We stand out they are the only ones to put "*high cost, exceed sometimes the value of the gold bar*", so as to stay at the height of technology". The authors show an increase in spending on development in favor of armies that raises to a paradox whether we want that weapons should be always in the forefront of a battlefield. In this way, for American the military arms are seen as being a claim in connection with their military hobby, whereas the Chinese adopt the military arms as a mean. The consequence of this determination to being armed is illustrated by the United States becoming "*slaves of technology*", production of military arms can be a ruinous means because of the research of technological prowess. The more it invents, the more the link with the role of weapon reduction particularly in warfare, evidenced in the inherent paradox between military arms and warfare. The prompt evolution of technology on the ground is thus described as the first objective and the states in warfare are in constant quest of new techniques. On the other hand, humans get to worship blindness the innovations and move away to the old. This problem has caused the change with which their development does not depend on performance ameliorations for every individual weapon but rather to know if the weapon presents the characteristics allowing it to be associated and to match the other weapons. Nevertheless, if we respect the principle written by Qiao Liang: *the "new" of today will become the "old" of tomorrow (*The word high technology can lead to confusion because the high technology of a day is the low technology of tomorrow. An interesting quotation has been reflected for weapons of high technologies: "During the prehistoric time, the bow and the arrow will be a determinant weapon, like the sword during the time of barbarism and the weapons during the time civilized". Engels (*Oeuvres choisies* de Marx et Engels, vol 4)), we will be leading up to say it does not never weapons that should stay to sit on the throne of high technology. Our authors put into perspective the importance of high technology in all possible fields.

Secondly, the globalization which is to say the notion of space. The space occupied on the battlefield becomes limited. It will seem the battlefield achieve inevitably its extreme limit. We arrive at a paroxysm to the art of warfare on the battlefield. As our authors, the Chinese envisage an evolution of the battlefield, an original dimension beyond all borders despite the real ones (sea-ground-space). In this perspective, we hit on a more or less rhetorical question: Where will you find the battlefield? And we would all be brought to answer "everywhere". It should be noted that the battlefield take an on-dimensional. The addition of this, producing the control that has become strategic. In preparation with this cul-de-sac on the ground, it reflects a new concept of arming binding the life of populations, from the moment, we hear talk of warfare without limits. As the point of view of the philosopher Clausewitz (Clausewitz, 1832), warfare without borders includes the greater part in the problems "fog and friction" binding the difference between the ground and the map, between the strategy and its concrete application to the real world. As they describe him so well "*all the difficulty of the new warfare is to know how to plan classic military arms and new military arms*". The principle is to go over the borders, that is to say to pass not only the physical borders, but also domains such as the economy, finance, religion and culture to attain the enemy. Such a combination will go over all the limits of conflict leading until here for the militaries, for example the movement of *Indignés* -- Indignés is a movement of group and non-violent, grouping millions people in a hundred cities, extending for different types of actions up to now. The name Indignés has inspired by the manifesto "Ignorez-vous". One of principal thing of this movement is the wish to transcribe the political speech in the reality, to put into practice in the political theory -- and the "*Occupy Wall Street*" -- Occupy Wall Street is a movement of pacific contesting opposing the breach of financial capitalism. The movement is particularly inspired by the Egyptian revolutions in the same time the movement of Indignés in Europe. Their principle demand is enhanced by this phrase: we are the 99% that not stand for avidity and the corruption the remaining 1%, this is the thing in common. The heterogeneous character of movements exceed the setting of Spain's Indignés and some people riche have join the movement. It shares the bleakness of social inequalities, the crisis and the refusal of the ultraliberal system. There is a phenomenal acceleration in the number of actions and growing politicization of these actions. We are base on the old member of group Anonymous, Julien Assange, a perfect representative of this movement (Bardeau & Danet, 2012). He has gone from the logic of the technical to the logic of the journalist. The technique is not an end, it is a way *[4]*.

As a result, the warfare will not be only military but "*the warfare wholly will become civil*". That raises on two critical points:

- The future warfare will do less physical "death" but a lot of "deaths" of other types.
- Warfare will stay permanently

The nuclear age and the information age of the XX^{e} sc., gave place to the computer science warfare and precision warfare which is giving up its place to a new type of warfare: the cyberware. Back in 1990, John Arquilla -- John Arquilla have a PhD in International Internship graduate at Stanford in 1991. During several years, he has worked for RAND Corporation, then he has joined the USA University Naval Postgraduate School in 1993. He is an author of a lot of articles and books on the future warfare such as "Networks and Netwars: the future of terror, crime and militancy" CA: RAND Corporation -- and John Ronfeldt designed the pleasant concept of cyberwar *[5]*. In this way, it is the major component of security strategy and power. We are in an era of permanent instability. This is way, the Chinese doctrine deals in depth with the research on the way to impact directly on the neuralgic center of the enemy without damaging the rest. The best way is to control and not to kill. They mention to us at this moment the objective "*zero deaths*" but it is paradoxical. In this way, we can take as an illustration the use of the U.S neutron bomb to take possession of Baghdad airport. This thermonuclear bomb which detonated with minimum explosion, gives out radiation which penetrates both in buildings and in tanks and is instantly fatal to humans. The Chinese want to discreetly a new art of warfare which has become the combination of all the ways to achieve their ends. They do not necessarily seek to inflict maximum civilian casualties on an enemy, but to obtain sufficient losses other than to human casualties within the limits of what is tolerable to the public. They anticipate a changing of the future vision warfare on all continents. Moreover, the United States, even with a perfect power on the ground during fighting, this time will be the most vulnerable.

> *"2007-2009, Michael McConnell proclaimed in front of the American Senate: if a cyberwar broke out today, we would lose it" [6].*
>
> *"3rd September 2007, the Pentagon recognized that an informatics attack targeted on its servers had made specialists have to disconnect for some days a part of the informatics network used by the Defense Department" [7].*

Because of the vulnerability, they also created a commission to study economics and security in China, which has sounded the alarm in a report presented to the Chief of Congress and the House of Representatives explaining that Beijing had put in place specialized units in cybernetic combat whose objective was to develop computer viruses able to pass the defenses of Uncle Sam -- Uncle Sam is symbolic and an allegorical person in United States. Uncle Sam comes from the warfare in 1812. Troy militaries received the meat in the box tagged "US" and this initial letter is interpreted like "Uncle Sam" in honor of their provider Samuel Wilson. These threats related to cyber-security were initiated by Bill Clinton, the former President of the United States (1993-2001), who wanted to eliminate the vulnerabilities of computer systems.

In the light of this principle, the new protagonists on "terra incognita" are hackers, whose principal objective is to seriously threaten the safety of an army or a country. They receive no vocational military training nor do to exercise a military profession. Hackers stand-alone in understanding and controlling the technology and have therefore extended the battlefield. Hide, know, delude, persuade; four necessities inherent in any conflict, recognized by antique strategies, but these can add another new dimensions to the capabilities of the technique. If the computer was out of control and incomprehensible, humanity would be at the mercy of Ace computer specialist and Humanity would be caught in a huge net. All notions of width, depth and height of the operational space from now on seem to have gone. Facing threats, armies will be much impoverished. It turns out that we cannot let national security stand only on military strength Qiao Liang and Wang Xiangsui. There is also a cultural and sociological aspect. The U.S. and the Western world prefers the individuality, which divides and weakens... In China, the group takes precedence over the individual whereas the State in private. We are forced to admit that the modern concept undeniably favours the Chinese side.

Moreover, the Chinese doctrine recommends the annihilation of humanity which leads to the use of pleasant dissimulated weapons. In connection with recent facts, they would seem Chinese have becoming:

> *"Nowadays, the attacks which are considered as coming from China are evaluated as of necessary seriousness to constitute the other days of warfare." [9].*

This high ranking officer of the Chinese military, after an analysis of this phenomenon of warfare, we break the new question that is not clear: Who is? Where is the enemy in the cyber-world? And if other countries join China, we would have created a new form of threat such as the cyber-terrorist.

3. A discussion of new methods of operation

We are witnessing a metamorphosis with this new form of warfare. This new image is seen as an art. The conduct of warfare is an art similar to that which auscultates its patient --J.F.C Fuller was an officer of British Army, military historian. He has been the creator of "artificial clair de lune", a high light of battlefield allow to localize the enemy during the nocturnal attacks thank projectors. According to Clausewitz (Clausewitz, 19832), a nation must now engage fully in any conflict, and the goal should be the complete destruction of the enemy. This idea involves the combination of all means military or non-military, to achieve its purpose. Our authors perceive the evolution of the art of warfare by the coming of new players in the current conflicts. The extension of the battlefield and the fact of carrying fighting without limits are some of the main actors that take in changing. According to them, increasing the field of view means *"going to the other side of the hill to greet the rising sun" (« Go the other side of the mountain to take the rising sun ».* This metaphor put into practice the idea of soldier that attacks the enemy in face after combining the strategist of surprise*).*

The most beautiful aspect of this change is that a soldier can successfully attack five targets. The emergence of new tactics and reference manœuvrers greatly increases, in fact, the possibility of non-military action threating the security of states. *"The international community, face with non-military threats of destruction no less serious than those caused by warfare, lack the minimum resources and effective measures to limit it".* Objectively, this has accelerated the appearance of situation *"civil"* warfare. Thus, each domain can tomorrow, at any time, be the trigger for warfare between different groups. Obviously, no military in any country is adequately mentally prepared for this new type of warfare that totally exceeds military space. Therefore, we could proclaim the old warfare die out under the trampling of a future fresher but it still remains in warfare. You get the same damage as in "hand-to-hand warfare" and even worse.

Therefore, humanity has discovered afresh that peace efforts can be negated with a single hit; so, we realize that we have never *"controlled warfare"*. Contempt for rules and states results in a loss of legitimacy of territories. Indeed, the visible borders of States, the invisible space of the Internet, international law, the rules of conduct and ethical principles have no discouraging effect on a certain types of hacktivist. Their very discreet movement is the cause of major scale damages. Internet has been helping to detract from the social order still more. And as we saw in the previous section, the facts are justified to prove that warfare will be constant and it will entail less *"death" [10]*. Warfare will no longer be traditional warfare but rather something that we have never considered as warfare, such as an exchange of blows on the Internet, a battle between mass media, a conflict on the futures market and a currency risk which will leave us stunned. These are the principal type of games. The difficulty of locating the opponent or of understanding the rules of the game also shows the difficulty of localizing warfare.

Therefore, the sense of mistrust denounced by Qiao Liang and Wang Xiangsui, invites us to understand that there is a presumption some events but also the thought of the Master of warfare: "*Nothing is more secure.*

The only certainty, uncertainty ..." -- Neither the enemy nor the weapons, nor the battlefields won't be the same of the beginning. In this situation with many uncertainties, it must define the new rules of play.

After this famous denunciation, we arrive at the point where the *"addition"* of ways would dominate this game. Indeed, it is interpreted according to these high ranking officers of the Chinese military, as the art of combination. It is perceived as a lack of understanding from these warlords. It resulted in the failure to understand that crossing all boundaries and all borders is just the preliminary to a revolution in thinking. Faced with this entirely new conception of warfare, there is no doubt that the vision of the warfare to which we were accustomed will be shaken. He points out *"these men who only know to dispose of impressive numbers of troops [...], who also cry that warfare is to kill, the art of warfare is the technique to kill".* It is in this vision that they specify a concept already well known but often forgotten: "*If an army prepares too specially against an enemy, it neglects what is outside of its field of vision*". They must get out of rut made over thousands of years by the tank. To achieve this, they can only use additions. But before doing this, they must overcome all obstacles: political, historical, cultural and moral and engage in thorough reflection.

Clearly, the development of techniques by the U.S. military is still insufficient, on the ways, especially in terms of military theory. We can see that the art of traditional warfare and combinatorial art are totally different but approach each other very respectively. Indeed, the aim of one is the retreat of its enemy's army and the other the total falling down of a state, while triggering social panic and a governmental crisis. China's strategy is based only on combinatorial art, especially, on the changed definition of the battlefield in which the notion of retirement is removed. There is nowhere where you can take refuge. This is part of globalization. In addition, Yue Fei -- Yue Fei is a famous patriot. He has fight for the dynasty Song of South against the army of dynasty Jin des Jurchen, gives a detailed explanation of the exact use of the combinatorial method: "*The excellence of its use come from the existence of exceptional will*".

At this stage, they expose us to a quest for the rule of victory where the idea is to bring the sword to the side of the opponent. Li Shimin, described it very succinctly:

"When I make the surprise a rule, the enemy expects a surprise, and then I attack according to the rule. When I make the rule a surprise, the enemy expects an attack by the rule, then I attack by surprise" like the tragedy of September 11, 2001... At this point, the authors get into long unconvincing discussions: indeed, they hope to persuade for regulate the means of the warfare, it should refer to the amount of gold. All statements in accordance follow the rules in this secret. During searching the terrorist Bin Laden, the U.S. military has used enormous resources, but the most effective was a cyberguerre[7]. We conclude that warfare should affect all aspects of life in countries without us having to call for military action. They use other means besides military means to complement and enrich, replacing military means to achieve goals which are unattainable by only force of arms.

The purpose is reduced to one way which is the combination of anything off limits. The idea, the end justifies the means is the most important spiritual legacy that warfare can give. However, all available means to achieve its goals, it is not the earliest source of *"thinking out of limits"* but it is the clearest. At this moment, we understand the limits have a sense of the relative. From outside the limits is the exceeding which is designated or to be understood as limits. Whether it is physical, spiritual or technical limits, scientific, psychological, ethical, or moral. The notion of overtaking of the limits is exposed by a passing ideology. Warfare of the past, as we have seen previously, is the combination of strict meaning. In a few words, we could confirm that the recipe for victory is not real but if we want to win the warfare, we must follow some rules of survival: *"Combining all of the resources warriors - Requirements rules of victory - A hostile hand to pick the victory."*

The combination out of limits results in various types. Firstly, the supra-state combination that appeals to a new paradox where exceeding the limits consists in tolerating any restriction and going over and above it. For China, the state is equal to the general concept equivalent to the entire civilized world. Then, the combination out of area is a vital link in the reflection on the innovative idea of exceeding limits. It situates between the concept of supranational combination and means out of combination. Then, the means out of combination which is to say keep away from the method and tools necessary to achieve a goal. The supranational action describes a country as a means and the national action reads into the army as a means and the country as its goal. Finally, the combination out of degrees shows the war out of bounds, it is thinks to take precedence on the method. This leaves us with a vague idea of the necessary basic principles to this type of warfare, as it noted by G.KENNAN -- G.KENNAN is a diplomat, political scientist and American historian which the ideas should be a big influence on the politic of United State. He became a model of loyalty in the Chinese culture -- and SUNZI -- SUNZI – art of warfare He describe the Chinese strategy or how inquire, estimate, divide, beat "*without encountering any opposition*"): *"Principles constitute a code of ethics, but it is not an absolute value", "Hitting the enemy where it does not expect it and taking him by surprise, avoiding the full and attacking the empty".*

The adoption of these principles does not guarantee victory, but to not comply would lead with defeat. The authors cite the most likely and the most efficient. Firstly, we formulate the concept of omnidirectional which is the main starting point for the ideology. The general principle or the basic requirement is to consider all the factors linked to at warfare such as social wilderness. The implementation of such synchrony controlled by computer science or can lead to actions in the same space-time and in different places simultaneously. Thereafter, include limited objectives and unlimited resources. All objectives are limited by the means and they should not be expanded. One thing is certain as soon as the objectives beyond the means, we are forced to inevitable defeat. As for the unlimited means which consist of filling the limited objectives, expanding the type of resources used and combined to achieve the goals but does not use immoderate excessive resources, or does not use absolute means. Last principle, we use the idea of loss of balance. This is looking for nodal points of the action in a direction opposite to the balanced symmetry, i.e. avoiding the brutal face-to-face with an opponent but exploiting their weaknesses.

4. Redefined stakes

We have noticed that the art of warfare is the pooling of power between Nations. This is the main rule to successfully defeat a state. Indeed, the only difference is that the combinations and alliances are made at several levels simultaneously: multi-, supra-state and off-state. The example of the Asian crisis in 1997, illustrates an ultra-modern battle and the financial crisis of 2008. The United States interfered in imposing the IMF and laying down the conditions for their own interests. Among the examples cited by our authors, we understand in a few words it is a weakness to strength. Certainly in 1940, Germany attacked the Allied armies from unfavorable terrain: the Ardennes.

It was form movements such as the one formulated previously that the Monroe Doctrine was born. This is summarized in a popular saying: *"America for the Americans".* However, the Chinese vision does not share the same opinion and says that *"the world belongs to the Chinese".* Monroe was established on the occasion of conflicts between the American continent, the United States and European powers. He published his doctrine after seeing the danger threating on the membership of some states, it translates as a potential danger. According to him, *"It is only when one encroaches on our rights or we are seriously threatened, we feel insults that we make preparations for our defense",* then the states remain united and strengthen. We detect through its doctrine, three fundamental principles: any American intervention in European affairs would be excluded - any intervention by a European power in the American continent would be considered like an unfriendly demonstration against the United States - the American continent must now be considered closed to any subsequent attempt to colonization by European powers.

In fact, we are witnessing the exclusion of Europe. Thus, the Barrroso -- Jose Manuel Barroso is a Portuguese politician and Chief of European Commission - speech on the state of the union, shared a few rooms for maneuver that Europe must consider. On one hand, it reveals Asia as an emerging continent and Europe as a country overwhelmed. This proves that despite all efforts, their responses have not yet convinced the citizens, markets and their international partners. On the other hand, they also need a trade pro-active policy in opening new markets. Free trading is the DNA of the European Commission. And the last point on which it relies is the creation of a federation of nation states, not a super-state, just a federation of nation states. It will be post-national. Moreover, in this commission, Berlin, recommended the convening of a convention to be composed of several hundreds of representatives to discuss the objectives and the institutional functioning of the European Union.

All of the principles are in tow stakes:

- Contract of mutual non-intervention
- Isolationism and interventionism

- American Ambiguity

Stakes of cyberspace demonstrate the invisibility and the pervasiveness of warfare. Who says battle in space, speaks in relation of warfare, defense and attack, between states or between non-state actors, in terms of sovereignty and rivalries. This virtual space is the subject of a strategy of control or influence as seen with the laws on Internet and Information flow through appearing in different countries. Multiple terms in cyberspace transpose the traditional terms of warfare: cyberattack, cyberdefence, cybercops, cyberpolice. The human dimension must not be obscured by a technical vision which is dominant today. The computer arm is only a tool and has no positive or negative effect on the human will that animates it.

The Monroe Doctrine is part of two European dangers:

- Russian ambitions in North America
- The threat of intervention of the Saint Alliance on the old Spanish colonial empire

It was not put in place soon as he arrived, but only at the mid-warfare of the XIX^{e}s.c.

An image of a still very superfluous Occident perceived negatively by the non-Occidental world is born. Indeed, tensions between the United States and China following the Chinese cyberattacks in 2009 are perfectly expressed by a comment from the Chinese press reported in *Le Monde* (January 2010): *"The campaign of the United States by the free flow of uncensored information on an unrestricted Internet is a disguised attempt to impose their values on other cultures in the name of democracy".*

An ultimate vision is described by Eric Przyswa (Przyswa, 2010) -- Eric Pryswa said the cyber criminality became a phenomenal world. Millions people are tricked by the cyber fraud or the counterfeits. He explains us through his book "*Cybercriminalité et Contrefaçon*" -- through a story about counterfeiting and cybercrime. Indeed, it considers the fact that weaknesses in terms of expertise are a problem of education. CNAC (National Anti-Counterfeiting Committee) does not have enough experts to monitor this traffic. It is feared that the analysis devices and control implemented by the majority of experts give priority to especially a *"strategizing about the daily dangers, an unreassured world where the risk is always read as a danger and not as an opportunity".* The fight against counterfeiting can be seen as a new or a different type of warfare.

5. Conclusion

Finally, "humanity is making progress, and it does not imagine warfare could be a possible court of appeals". In fact, between the last trails of fog of the XIX^{e} sc, the dusk of the XX^{e} sc, and the dawn by XXI^{e} sc, we make our point about the opening of a new era. Humanity has no reason to be relieved, because we have done absolutely nothing except replace as far as possible bloody warfare by no bloody warfare -- the traditional warfare has often been assimilated to several corpses and the blood flooding the different battlefields. The new image of warfare uses the civil technologies and Internet. Consequently, the notion of battlefield becomes out of date. It would explain the absence of blood trail through this virtual fight. The world has become a huge battlefield. Military arms are more modern and according to the means, they have become more sophisticated. However, warfare has come out of the domain of the military and become the story of politicians, scientists or ecologists. This is demonstrated by a relative reduction in military violence. The future will be done with the technological union and apart from the unrestricted warfare, we cannot find the key the most appropriate for this news warfare.

References

Clausewitz, C. v (1832). *On War*. Translation by Howard M. and Paret P. in 1989. Princeton University Press; Reprint edition. ISBN-10: 0691018545

Danet N. and Bardeau F. (2012). Anonymous: Pirates informatiques ou altermondialistes numériques. FYP Editions, ISBN-10: 2916571604

Col. Qiao*,* L. and Wang X. (1999) "*Unrestricted Warfare".* People's Liberation Army. Litterature and Arts Publishing House, Beijing. [online] http://www.terrorism.com/documents/TRC-Analysis/unrestricted.pdf

Przyswa E. (2010) "*Cybercriminalité et contrefaçon*". FYP Editions. ISBN-10: 2916571477

Ventre D. ed, (2010) « *Cyberwar and Information Warfare* » ISTE Ltd and John Wiley & Sons Inc., ISBN-10: 1848213042

Secret Sharing for High Security Codes on Smart-Cards

Paul Crocker[1, 2] and Adolfo Peixinho[1]

[1]Department of Computer Science, University of Beira Interior - 6201-001 - Covilhã, Portugal

[2]Institute of Telecommunications, Covilhã, Portugal

crocker@di.ubi.pt

m4067@ubi.pt

Abstract: This paper discusses the use of secret sharing cryptographic protocols for top secret and high security codes which may be used in the context of military and information warfare and Cyber-Security for Cloud Computing and e-Health. In particular a framework for secret sharing using threshold cryptographic methods in order to distribute and share high security and top secret access and authorization tokens on Java Cards is presented. An implementation of this framework, a proof of concept, is described that uses as a base an existing European Electronic Identity Smart-Card, more commonly known as e-ID card or Citizen Card, that makes use of the associated national Public Key Infrastructure (PKI) and time stamping infrastructure. Details of the cryptographic sharing algorithm based on an existing well known secret sharing scheme that uses polynomial interpolation over a finite field are given and also the secret distribution and secret recovery protocols are specified. Since Smart-Cards often have only a relatively small writable memory we shall discuss key sizes and the impact this has on the amount and size of secret shares that can be stored. A secret sharing scheme is a collaborative system and this paper shows how this may be implemented using the concept of a Chat Room whose members are able to collaborate in the process of secret sharing and recovery.

Keywords: secret sharing, threshold cryptography, smart-card

1. Introduction

The distribution and management of Secrets and Cryptographic Keys has historically always been a difficult problem. The process is even more complicated when a secret such as a top secret code needs to be shared amongst a group of entities or when separate codes must be authenticated at a given time in a group. Many times such as in battlefield and other critical systems redundancy must also be built into the system and thus an access right must be shared between a group of people, however to increase security its often necessary that a subset of key holders simultaneously authenticate themselves in order to gain the access right or secret.

The role of anonymity and identity manipulation in the field of Information warfare is well documented (Jacobs et al 2010) and of particular importance in a military context with the particular difficulties of Battlefield Code distribution, authentication and access rights. Criticism of current schemes based entirely on Public Key Infrastructures (PKIs) and Digital Certificates are widespread and new robust authentication mechanisms across key network security services should be explored. Also in cyber security in areas such as e-Health and Smart-Grid settings group security is an important aspect during planning and implementing solutions in these areas. (Byeong et al. 2011). In e-Health services ensuring security, privacy and confidentiality amongst a group of stack holders is of critical importance when critical life changing decisions often need to be undertaken by diverse groups of medical practitioners. In Smart-Grid scenarios a critical concern of the end user is how their data is stored and who has access rights.

Secret Sharing Schemes were announced independently by Shamir (Shamir 79) and Blakley (Blakely 79) in 1979. Informally we can define a threshold secret sharing scheme as a method to divide a secret into a set of shares by an entity known as the dealer and then distribute these shares amongst a set of participants, where only qualified subsets of participants, at least equal to a given threshold, can recover the secret. They were called threshold (*k,n*) schemes by Shamir where *k* is the threshold value and *n* the size of the participant set. A scheme in which any subset of participants bellow a given threshold does not obtain any "information" about the secret, other then their own share, is called perfect. Under the scope of theory of information we can say, if the length of every share is the same as the length of the secret, which is the best possible case, these schemes are ideal (Karnin et al. 1983). Shamir's basic scheme is based on polynomial interpolation, this scheme is used to implement an application in order to distribute and share high security and top secret access and authorization tokens on Java Cards and create a protocol for secret recovery in the sense of authorization rights or key recovery. We also aim to show how the use of a national e-ID card and associated time stamping infrastructure may strengthen the overall system. Also since a secret sharing scheme is inherently a collaborative system this paper shows how this may be implemented using the concept of a Chat Room whose members are able to collaborate in the process of secret sharing and recovery.

The rest of this brief paper is organized as follows. Firstly a description of Shamir's threshold secret sharing scheme is given. Secondly a description of the overall system developed is given, in particular the secret distribution and recovery protocols. We then present the practical implementation of the scheme using Web Based Technology and ActiveX components to read and write from the Smart-Card. Finally we end with some conclusions.

2. Secret sharing schemes

Shamir's secret sharing scheme is based on polynomial interpolation and the fact that given k distinct points in a 2-dimensional plane $(x_i, y_i) : i = 1,2,..k$ there is one and only one polynomial $q(x)$ of degree *(k - 1)* such that $\forall i, q(x_i) = y_i$. To divide the secret data S and create n shares, assuming that the data S is (or can be made) a number, we pick a set of random coefficients $\{a_i\} : i = 1,..k-1$ for a *(k - 1)* degree polynomial $q(x) = a_o + a_1 x + ... + a_{k-1} x^{k-1}$ and set $a_o = S$ (the secret).

The *n* shares $\{S_i, x_i\}$ are calculated by evaluating the polynomial at the *n* points

$$S_1 = q(x_1); S_2 = q(x_2); ..; S_n = q(x_n)$$

To recover the secret note that given any subset of shares $\{S_i, x_i\}$of size *k* one can find the coefficients of $q(x)$ by Lagrange interpolation,

$$q(x) = \sum_{i=1}^{k} f(x_k) L_i(x) \text{ where } L_i(x) = \prod_{\substack{j=1 \\ j \neq i}}^{k} \frac{(x - x_j)}{(x_i - x_j)}$$

and hence evaluate

$$q(0) = a_o = S = \sum_{i=1}^{k} S_i \prod_{\substack{j=1 \\ j \neq i}}^{k} \frac{-x_j}{(x_i - x_j)}$$

In practice the points $x_i = i$ are chosen or fixed depending on the recipients and arithmetic is over some finite field, we shall use modular arithmetic over Z_p is used. Therefore given an integer value S the prime, p, is chosen such that it's bigger than both n and S, in our case *n* is usually a small value. The coefficients of the polynomial $\{a_i\} : i = 1,..k-1$are randomly chosen from the uniform distribution $[0, p)$ and the values $S_1, S_2,..S_n$are calculated modulo p and $S_i \in [0, p)$

3. Collaborative secret sharing algorithm using smart-cards

While there are several existing software packages available for secret sharing these normally consist of a desktop or web application where a single user enters a secret and then the dealer distributes all the shares to the user who is responsible for there storage and distribution. An example of such software can be found for example at http://point-at-infinity.org/ssss. Another related software is Nightingale from RSA Security which uses secret splitting to secure sensitive data [Brainard 2003] by dividing a user's password (or other key) into shares for two independent servers, the scheme is used in password authentication. Another example is Finnegan Lab System, a graphical desktop application designed for a single user to learn and explore how secret sharing works [Olsson 2004]. However there seems to be no practical implementation of a collaborative mechanism to share secrets between a group of mutually authenticated users. We will now describe such a mechanism.

3.1 Secret sharing generation

When designing and implementing the secret sharing algorithm we must take into account the proposed storage medium on which the secret shares will be stored. The basic problem is the size of any proposed secret in relation to the space available on the Smart-Card and the maximum number of secret shares that we wish to store. For instance using the Portuguese e-ID card (known as the Citizens Card or Cartão de Cidadão) the maximum available writable space is only 1024 bytes. Therefore we must define a maximum size for the secret that can be shared. If the initial secret data is larger than this critical value then the preferred solution is to encrypt the secret data using a symmetric cipher (only one key), store the encrypted data on some mutually accessible location and distribute the key as the secret.

For the scheme to be practical we must first choose a prime number and also convert any secret into a sequence of data blocks of size D such that $P > D$. We must therefore fix the maximum size of individual data blocks D and also the maximum size of P. Also the result of the secret sharing algorithm is simply a binary digit which and since we must read and write text to the Smart-Card it makes sense to use an encoding, such as Base64, to convert the binary number produce by the secret sharing algorithm to text that may be read and written to the card. Base64 represents binary data in an ASCII string format by translating it into a radix-64 representation and hence as a 64 character alphabet (the standard alphanumeric characters and the '+' and '/' characters). For instance consider data blocks as strings of 12 characters in the Base64 alphabet this would therefore be encoded as a 9 byte binary number. Hence a suitable prime number would be any 96 bit (12byte) prime number. The resulting secret, at most a 12 byte number, would therefore be decoded as a 16 byte Base64 string. This means for example that a 24 character secret will be divided into shares of size 32 bytes. These are the value chosen for the implementation presented later. In the application implemented for the secret sharing we have used a prime of length 12 bytes. However, considering RSA keys are typically 1024 bits and the product of the primes of similar bit length, 512 bits (64bytes), and also the fact that many Smart-Cards contain coprocessors for modular arithmetic then it fact a prime of 64 bytes would be a more appropriate value to use in a production environment.

The algorithm to generate and distribute the shares by the dealer is then:

```
Algorithm 1: Generation of the shares by the Dealer
    (1) Input : Secret  (S) Number of shares, (n) and threshold value (k)
    (2) Calculate size of Secret  |S| and compare to SMAX
            a.  If     |S| <= SMAX then continue
            b.  Else S is Encrypted to the cipher text S' using Algorithm E and key KS.
                    S' = E(S, KS), S' is stored, and S<- KS
    (3) Generate the secret shares : S –> $S_i$  i=1...n
    (4) Generate $T_i$,  i=1..n the timestamp of the share $S_i$
    (5) Generate a plain text description (P) of the secret for management purposes
    (6) The Secret can now be destroyed
```

3.2 Protocol for the distribution of the shares

In order to distribute the shares to the various users, we can distinguish between several different cases, one where all the participants are physically in the same place together and secondly where the shares are distributed over some channel, which could be an insecure channel such as email or a previously configured secure channel (such as IPSec) or a more complex process but more flexible approach involving online authentication of the participants in an online collaborative application, which may involve all the participants being actively online for the distribution process

- The Receiver Authenticates himself to some trusted third party using their e-ID credentials. This guarantees that the user is authenticated to this trusted third party which for convenience can also be the dealer. Authentication is done over the secure socket level (SSL) protocol using server certificates and significantly in this case client certificates (SSL does not require client certificates)
- Send the message <P || i ||S_i || T_i >, to the recipient using the secured channel.

3.3 Reception of the Secret Share

Once the data <P || i ||S_i || T_i > has been received by the recipients client program then the timestamp can be verified and then the secret share Si can then be written onto the card. The timestamp can be kept or disposed of as the user desires. Notice that if each user is attributed a unique integer identifier it wont be necessary to store the value <<i>>.

In order to be able to distinguish the secret shares written on to the Smart-Card we need to define a plaintext metadata, a tag, containing a plain text description of the secret and some means for the before writing the data onto the card and in order to identify it a later stage the share S_i is also encapsulated with a leading Tag (#S#||P#) describing the application and concatenated with the plain text description (p) of the secret as and also a terminating "#" symbol as shown below.

DATA = TAG + SHARE → #S#123# jmMS9NFh0GVXAejj#

The size of data that will be stored on the card to store is therefore 5 bytes for the tag plus the size of the share and the plain text description. (5+|P|+|Sk|). For instance a 12 byte secret with a 3 byte plain text description will be stored as 5+16+3=24bytes. In order to create a practical scheme we fix the maximum size of the plain text description (P) at 15 bytes and maximum size of the secret (S) at 120. In this way the smallest number of shares that the Smart-Card can store is at least 5 (As each share is at most 180 characters).

The use of the timestamp Ti means that the receiver can verify that the share Si has being generated at a certain time and also that it was generated from the trusted body and is a simple measure to introduce in order to protect against replay attacks and also to enhance user credibility of the process.

3.4 Distribution of symmetric cipher keys

Often the secret to be shared is the key of a symmetric cipher. We choose to discuss two widely used ciphers namely the Advanced Encryption Standard (AES) and Blowfish.

AES is a widely adopted symmetric block cipher, standardized by the U.S. National Institute of Standards and Technology in 2001 [NIST01]. It has a fixed block size of 128 bits and a variable key size of 128, 192, or 256 bits. The US government recommends the use of any of the key sizes for information classified as secret and 192 or 256 bit key sizes for information classified as Top Secret. In light of this this means that in the case of an AES cipher being used (Key size 192 bits (24bytes) + ID TAG) would permit up to 20 secret keys being stored on the e-ID Card.

Blowfish is also a symmetric block cipher thas a fixed 64-bit block size and a variable key length from 32 bits up to 448 bits, is free and unpatented [Sch93]. The OpenSSL documentation [http://www.openssl.org/docs/crypto/blowfish.html] states that 128 bit (16 byte) keys are considered good for strong encryption. Both ciphers are suitable, the only constraint being the maximum number of keys we may wish to share and store.

3.5 Secret recovery protocol

Considering the general case where all the participants may not be physically together. Then again, either some trusted third party application or simply one of the participants may initiate the process of secret recovery sending the shares to a trusted third party (dealer). The protocol is thus:

1. System or a Participant specifies Tag of the secret to recover and threshold *k*
2. Dealer require a subset of *{Si}* whose size is >= *k*
3. Each participant in the group places their e-ID cards in the card reader in order to read the share and reads the appropriate data.
4. Each participant authenticates themselves at the trusted third party, authentication using their e-ID card.
5. Notice that the trust and security of the client certificates is provided by the national PKI, for instance it is possible to check for revocation of the client certificates.
6. Data, the shares, are then sent over the secure connection to the server.
7. The Secret is recovered by the dealer and its value displayed to one or more participants involved.

4. Implementation and example of the secret sharing scheme

In this section we will give a brief resume of the prototype application implemented which implements the schemes specified in the preceding section.

4.1 Architecture

The Secret Share scheme is implemented using a Web Application similar to a web based chat room. Users log onto the application (Authentication using certificate stored on their Smart-Card) , enter into a chat room and then the group of users in this chat room can send and receive typical text based chat messages and are also able to distribute and recover secrets, written and stored on their Smart-Cards. The users use the supplied buttons and texts fields which read/write to the Smart-Card and generate messages with a specific syntax which are sent to the chat room, interpreted at the web sever and appropriate messages displayed on the users browsers. The chat room server plays the role of dealer in the secret sharing.

The Web application is built using ASP.Net 3.5, LINQ and AJAX written in C#, it is modelled on an application described in (Vicencio 08) where further details may be found on the basic chat application. The application makes use of a small database which is needed to keep track of the users connected to the various chart rooms and the messages sent. The application was deployed using a server with Windows XP Operating System. The web server is Microsoft IIS and the SQL server MS SQL Server 2008.

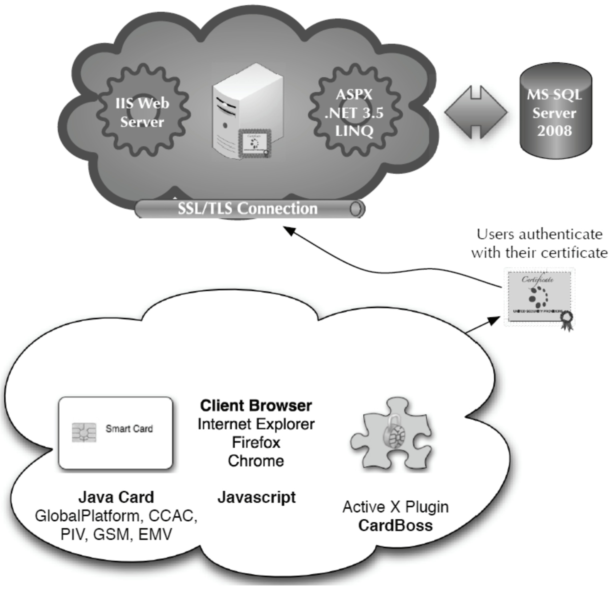


Figure 1: Architecture of the application

4.2 Smart-card communication

The Smart-Card used is the Portuguese e-ID card (CC). The on-card chip contains the citizens personal identification information, a private writable information space (1Kbyte) called *Notebook* or *Personal folder* and three *Personal Identification Numbers* (PINs), for authentication, digital signatures and accessing the address information. As a Java Smart-Card it executes Java Applets via the JCRE (Java Card Run Time Environment, or more simply, Java Card Manager) that runs on top of the native card operating system. Communication with card readers is made via standard ISO 7816 APDUS (Application Protocol Data Units).

The standard middleware for accessing the Portuguese e-ID card is available from (http://www.cartaodecidadao.pt/) however the middleware is available only for stand alone applications, not web applications. Low level communication with the card is possible by sending the appropriate APDU's, correctly formed, directly to the card reader and hence to the card and therefore proprietary middleware can be constructed if necessary (Crocker et al 2011). In order for web applications to interact with Smart-Cards several alternatives are available the majority of which make use of a Public-Key Cryptography Standards #11 (PKCS #11) library, for instance a developer can use signed Java Applets or a product such as SConnect by Gemalto, etc. For this project an ActiveX browser plugin was used from CardBoss (http://cardboss.cometway.com/) that is installed in the clients internet browser and permits interaction with JScript on the clients browsers and also provides secure access to the card from the web application acting as the bridge between Smart-Cards and any page on the web server (Figure 2). It supports any standard card reader with Firefox, Chrome or Internet Explorer as client browsers. Card Boss scripts can control the card by sending the APDU commands that the card already supports and receiving the responses.

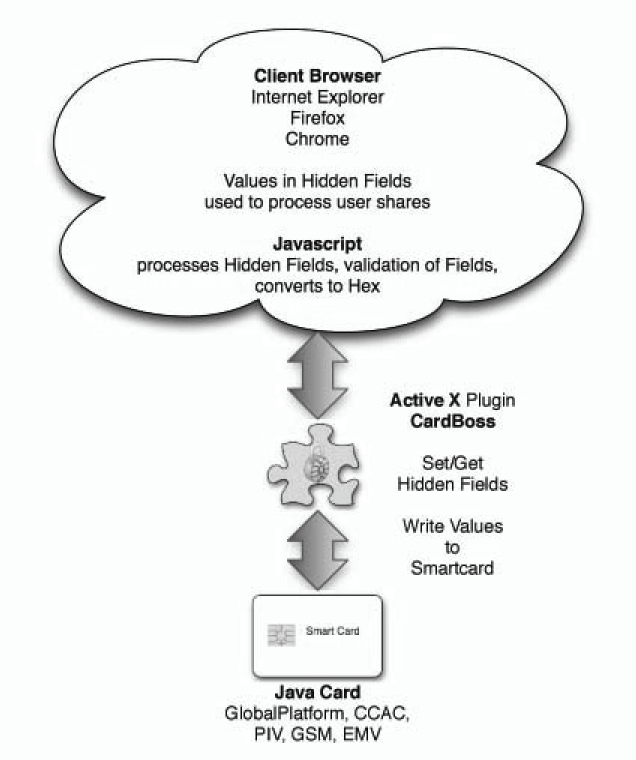


Figure 2: Web page interaction with the card

The APDU's we need to use for this application are for the following operations

- Select Applet,
- Send PIN (Authentication PIN)
- Select File,
- Read File (The Notebook Area)
- Update File (which can only be done after sending a valid PIN)

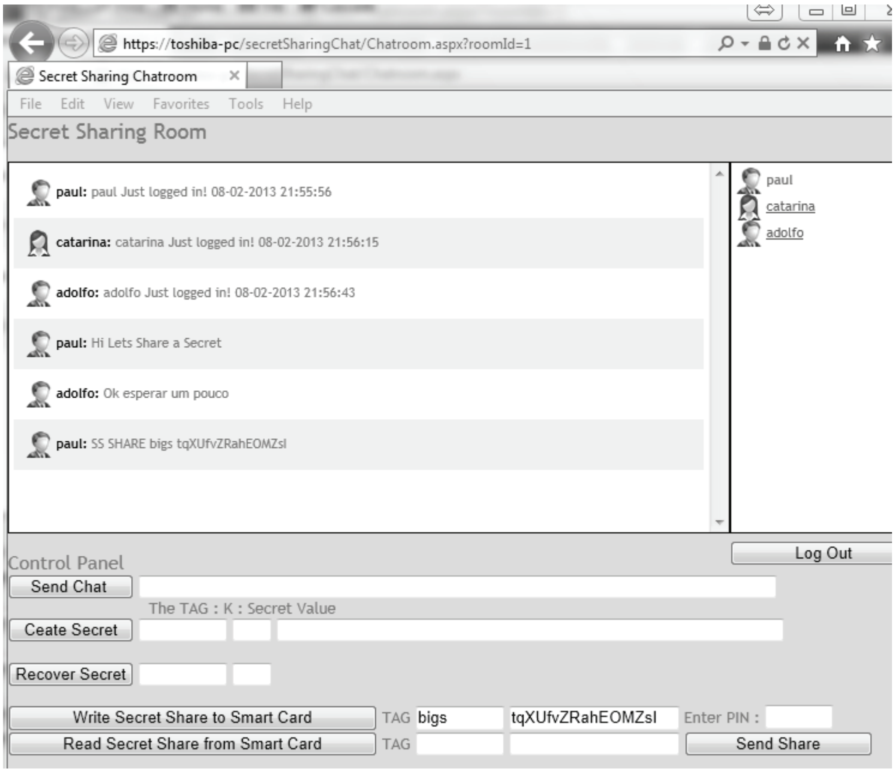
The following Figure 3 shows a snippet of the CardBoss Script used to send the APDU's to select the Identification Applet and then send the pin. The script uses variables constructed from the raw APDU's , which are hexadecimal strings and shows how to obtain data, in this case the PIN number, contained on a field in the current page, the command GET FIELD field, variable. The user enters the pin as a 4 digit string, client side JScript then converts this string into an appropriately formed hexadecimal string (the pin 1234 -> 31 32 33 34). Data returned from the Smart-Card is stored in the script variables [SW1] and [SW2] for the return code statuses and [RESPONSE] for the application data.

```
SET VARIABLE selectappletid, 00 A4 04 00 07 60 46 32 FF 00 01 02
SEND [selectappletid]
GET FIELD pinField, pin
SET VARIABLE apducommand, 00 20 00 01 08 [pin]  2f 2f 2f 2f
SEND [apducommand]
```

Figure 3: CardBoss Script

4.3 The web application

One of the users logged into a chat room with a group of *(n)* other users introduces the secret and parameters (the tag and threshold k) and sends a message to the server using the “create secret” button , there is a separate button for normal chat text messages.



https://toshiba-pc/secretSharingChat/Chatroom.aspx?roomId=1

Secret Sharing Room

paul: paul Just logged in! 08-02-2013 21:55:56

catarina: catarina Just logged in! 08-02-2013 21:56:15

adolfo: adolfo Just logged in! 08-02-2013 21:56:43

paul: Hi Lets Share a Secret

adolfo: Ok esperar um pouco

paul: SS SHARE bigs tqXUfvZRahEOMZsl

TAG bigs tqXUfvZRahEOMZsl Enter PIN :

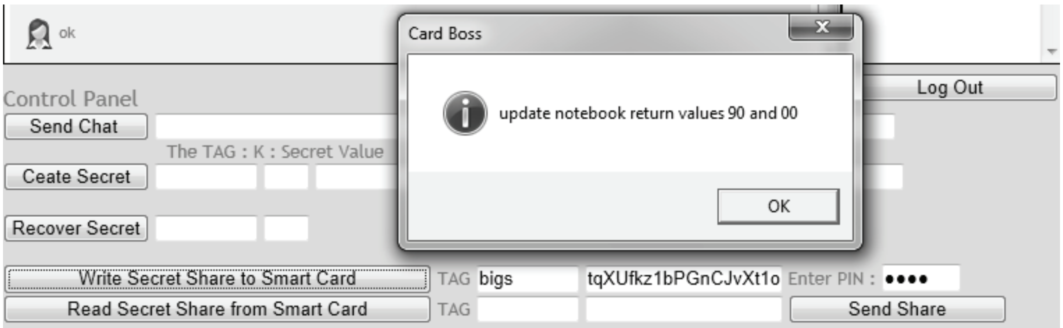
Secret Sharing Chatroom
File Edit View Favorites Tools Help
paul
catarina
adolfo
Control Panel
Log Out
Send Chat
The TAG : K : Secret Value
Ceate Secret
Recover Secret
Write Secret Share to Smart Card
Read Secret Share from Smart Card
TAG
Send Share

Figure 4: The web application

This message is interpreted by the server, the application then generates the shares and distributes the shares by inserting the messages in the message table, these messages will subsequently be picked up by the users browsers, see Figure 4. These messages appear in the form

SS SHARE tag share

The message is displayed in the message display area and also the tag and share values are written to separate fields on the page. Then all the user then has to do to write the values onto his Smart-Card is to introduce his (authentication) PIN code into the appropriate field. The button activates the client side JScript and CardBoss scripts necessary to complete the operation. The result is shown in figure 5.



ok

Card Boss

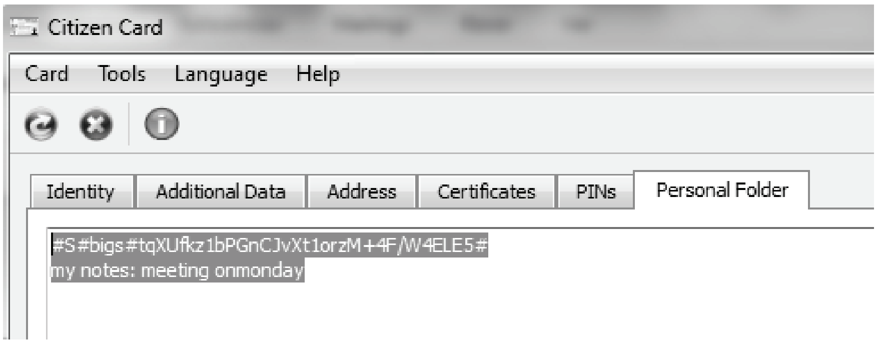
update notebook return values 90 and 00

TAG bigs tqXUfkz1bPGnCJvXt1o Enter PIN : ●●●●

OK
Control Panel
Log Out
Send Chat
The TAG : K : Secret Value
Ceate Secret
Recover Secret
Write Secret Share to Smart Card
Read Secret Share from Smart Card
TAG
Send Share

Figure 5: Application message after successfully writing share to card

In the next figure we show the secret share and tag sucessfully written to the card using the desktop application available that comes with the standard middleware package.

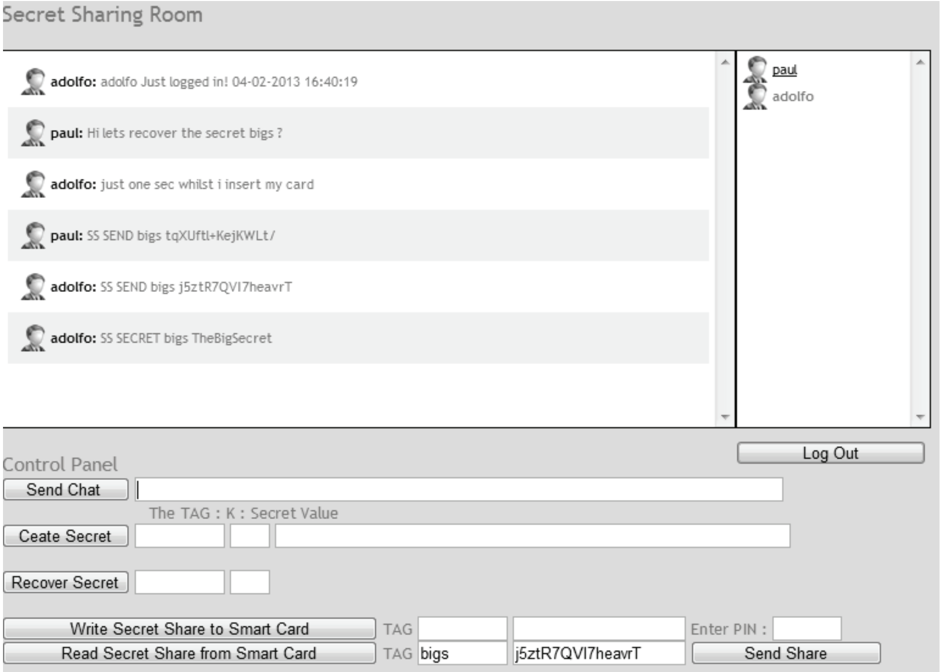


#S#bigs#tqXUfkz1bPGnCJvXt1orzM+4F/W4ELE5#
my notes: meeting onmonday

Figure 6 Personal folder accessed by a desktop application

For the secret recovery one of the users in the chat room sends the message to request a secret recovery with a specific tag using the message syntax generated by a button for the purpose. The button will generate a message for all members of the chat room with the syntax SS GET tag k. The appropriate values are written into fields on the users web page. Users can then read their Smart-Card for their share and then send this value back to the chat room using the appropriate page buttons. The message generated has the syntax: SS SEND tag share

Each time a users sends his share to the server, the server can check if the appropriate number of shares has been received, in which case the secret can be recovered and be shown to the group using the message syntax: SS SECRET tag secretValue



adolfo: adolfo Just logged in! 04-02-2013 16:40:19

paul: Hi lets recover the secret bigs ?

adolfo: just one sec whilst i insert my card

paul: SS SEND bigs tqXUftl+KejKWLt/

adolfo: SS SEND bigs j5ztR7QVI7heavrT

adolfo: SS SECRET bigs TheBigSecret

TAG bigs j5ztR7QVI7heavrT

Figure 7. Secret sucessfully recovered by two of the users in the chat room

If the required number of shares is not received inside a specified time frame (10 minutes) the request message and shares already sent are deleted from the database.

The web application uses SSL, Client based Certificates as authentication although, this can be configured to be required or be made optional in which case the user should be given a username and login. Notice that, the IIS server can be configured to map a client certificate to a specific user account (a one-to-one mapping, where the specific certificate is mapped to a single user account) see AMA (2008), MSDN (2006) and Microsoft Corporation (2003).

5. Conclusions

Secret Sharing is inherently a group collaborative action as it requires individuals working together in a coordinated fashion, towards a common goal, namely sharing and recovering of secrets. Chat applications are well suited to such types of scenarios where only a few number of users are involved and we have shown how such a scheme can be successfully implemented solving the problem of storing these secret shares and also of mutual trust and authentication using e-ID cards, in this case the Portuguese e-ID card and presented a working prototype of such a scheme.

The prototype is built using web based technologies that can be deployed in a scalable and fault tolerant architecture and the user access is based on familiar browser based and Smart-Card environments. This prototype still needs to be assessed in terms of performance and usability under real conditions and for use in production environment arithmetic over a Galois field GF (2^n) should be used. Also since Smart-Cards have extremely reduced writable memories it seems appropriate to consider other encoding schemes, such as base85.

References

AMA (2008) Manual Autenticacao com Cartão de Cidadão, Agência para a Modernização Administrativa, 2008, [online] http://www.cartaodecidadao.pt

Blakley G,R (1979). Safeguarding cryptographic keys. Proceedings of the 1979 AFIPS National Computer Conference, page 313--317.Monval, NJ, USA, AFIPS Press.

Brainard, J, Juels,A, Kaliski,B and Michael Szydlo. (2003). A new two-server approach for authentication with short secrets. In Proceedings of the 12th conference on USENIX Security Symposium - Volume 12 (SSYM'03), Vol. 12. USENIX Association, Berkeley, CA, USA, 14-14.

Byeong Gi Lee; Galli, S.; Brunner, M.; Tsong-Ho Wu; Hsiao-Hwa Chen (2011) New technical areas: exploring our future, Communications Magazine, IEEE Volume: 49 , Issue: 10

Crocker, P., de Sousa, S. M. & Nicolau, V. (2010) Sniffing with Portuguese Identity Card for fun and profit, Proceedings of the Ninth European Conf. on Information Warfare and Security.

Jacobs J,Chitkushev L and Zlateva T, (2010) Identities, Anonymity and Information Warfare, , Proceedings of the Ninth European Conf. on Information Warfare and Security.

Karnin, E. Greene, J. and Hellman, M (1983). On secret sharing systems. IEEE Transactions on Information Theory, 29(1):35–41.

Microsoft Corporation (2003) Building Secure Microsoft® ASP.NET Applications By Publisher: Microsoft Press

MSDN (2006) Building Secure ASP.NET Applications Authentication, Authorization, and Secure Communication - [online], http://msdn2.microsoft.com/en-us/library/aa302412.aspx

NIST (01), National Institute of Standards and Technology (2009), Announcing the Advanced Encryption Standard (AES), [online], http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf. [Accessed 7 January 2013]

Olsson, F, (2004) A Lab System for Secret Sharing, [online] http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.58.411

Schneier, B (1994) Description of a New Variable-Length Key, 64-Bit Block Cipher (Blowfish), Fast Software Encryption, Cambridge Security Workshop Proceedings, Springer-Verlag, pp. 191-204

Shamir, A. (1979) How to share a secret. Communications of the ACM, 22(11):612–613.

Vicencio, J (2008) [VIC08] Build a Web Chat Application using ASP.Net 3.5, LINQ and AJAX (in C# 3.5) [online]http://www.junnark.com/blog/detail/2

ECENTRE – Education as a Defensive Weapon in the war Against Cybercrime

Denis Edgar-Nevill
Canterbury Christ Church University, Canterbury, UK
denis.edgar-nevill@canterbury.ac.uk

Abstract: Since the EU Cybercrime Convention in 2000 (EU 2001), there has been a clear recognition of the accelerating threats to society posed by those who would exploit computers for crime and the logical progression to cyber-terrorism and cyberwarfare. In 2002 a committee of UK Members of Parliament highlighted how unprepared the country was to deal with this problem:

> *"We have around 140,000 police officers in the UK. Barely 1000 have been trained to handle digital evidence at the basic level and fewer than 250 of them are currently with Computer Crime Units or have higher-level forensic skills."*

Since then the capacity to deal with digital crime has improved but the problem has grown alarmingly. In the last decade the author has been working with The College of Policing (formerly named the NPIA) and developed a jointly validated Masters programme for UK law enforcement, intelligence and Government agencies. A variety of courses have now been taken by thousands of police officers in this area acting as first–responders to crime scenes and more specialist High Tech Crime Units. The European Commission has committed significant research and development funding in seeking to protect the community from Cybercrime and Cyberwar. This paper discusses a new European Commission funded project ECENTRE – England's Cybercrime Centre of Excellence Network for Training, Research and Education. On 20th December 2012 the European Commission signed the €0.935million (£760,000) contract for the project. The contract is awarded under the Programme Prevention of and Fight against Internet Crime Targeted Call – ISEC 2011 Action Grants– Project Number HOME/2011/ISEC/AG/INT/4000002226. The author is the Project Manager and Principal Investigator for the project. ECENTRE forms part of a wider European network of centres of excellence to share expertise, promote best practice and provide training opportunities for law enforcement across the EU. The challenges in establishing effecting cooperation and sharing are discussed. The considerable problem of keeping pace with the fast-developing, complex, problem posed by threats to national infrastructure, organisations and individuals is examined; highlighting the role of education as a fundamental weapon in the fight. The more we know about a threat (real or potential) – the better protected against it we become.

Keywords: cybercrime, cyberwar, education, defence

1. Cybercrime, cyberwar and cyber terrorism

Cybercrime and Cyberwar are inextricably linked. Give Cyberwar is covert in nature it's not really clear who is being attacked by who, or if the aggressor is a national government, organisation or individual. To that extent all are manifestations of cybercrime with different levels of severity (figure 1).

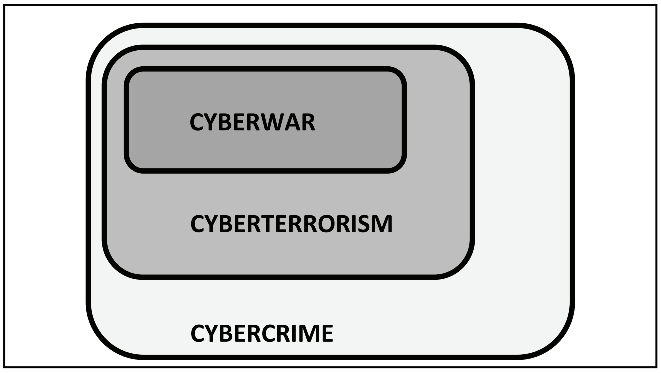


Figure 1: Levels of cybercrime

The recorded rates of computer crime and in particular, Internet-based Cybercrime, is growing at a staggering rate in the European Union and across the rest of the World (SDI 2012) (Paganini 2012). The ability of governments, intelligence agencies and police forces to cope with this tidal wave is very limited. In the last few years we have seen a lot of money being committed to establishing national, CERT teams (Cyber Emergency Response Teams) (EU 2011) (AMEinfo 2012) (TTOI 2012) (CERT Australia 2013), Cyber Commands (US 2013)

(Gilad 2012) (Segal 2011) (Leyden 2011) or Cyber Militias (Arsene 2012) (Segal 2012). Particularly at the lower-levels of Cybercrime, the situation is less favourable with police bodies being subject to major budget cuts as part of wider government austerity measures such as those in the UK (Burn-Murdoch 2012). The lack of spending on crime education, prevention, detection and prosecution make countries more vulnerable to major cyber-attacks. It creates weaknesses and opportunities allowing the way in for larger-scale attacks. The obvious example here would be facilitating phishing to establish large botnets in the target country. The July 2009 and March 2011 cyber-attacks on major government, news media and financial targets in South Korea (estimated to involve up to 166,000 computers) were launched using compromised computers mainly located in South Korea itself (BBC 2011).

If we picture Cybercrime as a pyramid with lower level crime (small thefts involving small sums of money, 'card not present' fraud) at the base and national/international crime and Cyberwar at the peak (organised crime, major fraud, counterfeiting, drugs, people trafficking), there is a rising threshold of the non-investigation and prosecution of crime (figure 2). Many Cybercrimes are so commonplace and involve small amounts of money that the costs of investigation far out-weigh the money recovered (Bracey et al 2007). The investigation of many crimes have now been reduced to the level where reporting them is only done to obtain a crime number for insurance claims purposes. No policeman will call. No fingerprints will be taken. No arrest is likely to be made. The general public, however, still have a false expectation that their personal crime report will involve major forensic investigations; an expectation fostered by television CSI (Crime Scene Investigation) dramas telling stories of hi tech detection costing tens of thousands of euros.

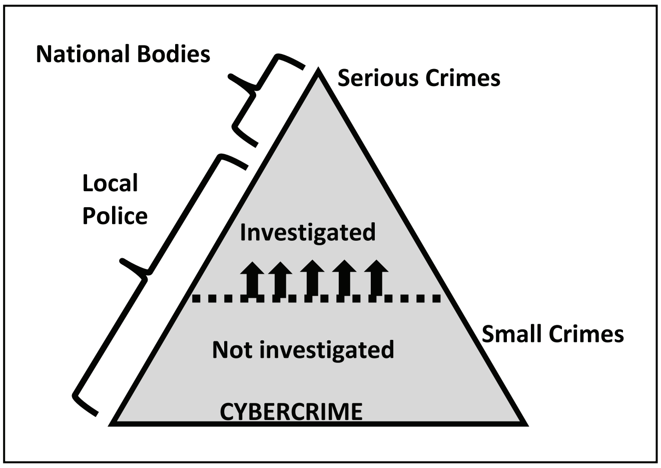


Figure 2: Increasing non-prosecution of low-level cybercrime

Because we cannot expect, or come to rely on, law enforcement agencies protecting us from crime cybercrime, society must become better educated to avoid becoming victims in the first place. Police themselves must learn to recognise the many facets of cybercrime even to begin to protect society from its consequences.

2. Education and cybercrime

There is general agreement on the need to establish EU-wide agreement to fight cybercrime, and that high-quality educational provisions are required to train law-enforcement. The availability of opportunities to receive training and education in cybercrime forensics is growing across the EU. In the UK the small number of universities offering awards 10 years ago, has now expanded to become more than 50 UK universities and over 100 other companies offering training courses and awards. There has also been a scramble to create a programme which can become 'the standard' course which everyone 'must' take to become a recognised professional working in this area. Most efforts here come unstuck because of the fast-moving rate of technological change. This is also the reason that many governments have been reluctant to try to impose mandatory standards (Coelho 2012).

The major problem to address is the wide variability and quality of cybercrime forensics courses offered. Many of these are missing essential input from law enforcement to make them credible and useful. In many cases, because of the rapid development of technology, courses become out of date very quickly and of declining relevance to practitioners. For a course provision to be successful it must be designed to meet the real needs of practitioners and informed by the practical experiences in an on-going dialogue. Many courses have no strong development, evolution and quality control underpinned by pro-active research. As always we fail to grasp the simple truth that we don't share what we know and we don't trust what we have not created. The 'not invented here' syndrome is alive and well and flourishing in the 21st Century. We also still cling onto a 'silo mentality' where different sectors (commercial/law enforcement/academia) fail to trust each other, communicate or coordinate their actions to improve the situation for everyone.

These goals will not be achieved unless an infrastructure supporting the on-going collaboration of law-enforcement, commercial organizations and universities is established to exploit the subject and process expertise which exists within these sectors. In particular this infrastructure must encourage cooperation and sharing teaching materials and resources; avoiding the frequent 'reinventing the wheel' which happens too often. Such a national infrastructure is proposed by this development based on the EU-Wide 2Centre (2Centre 2013) project standard under the ISEC Programme (ISEC 2013).

3. The ECENTRE EU project

ECENTRE (England's Cybercrime Centre of Excellence Network for Training, Research and Education) objectives are:

- To establish a network of 5 regional groups each based on the 2Centre membership from law-enforcement, academia and commerce;
- Develop a range of new teaching materials (case studies, presentation DVDs, software tools) which support training and new research which can contribute to the 2Centre EU-wide network;
- Exploit materials produced by ECENTRE and 2Centre in course provisions to enhance the quality and applicability of training courses;
- Deliver workshops and training to law enforcement practitioners from the UK and wider EU;
- Establish ECENTRE as an important body of reference for expertise in developing high-quality cybercrime forensics practitioner training and research (ECENTRE 2012)

The idea for this project began with one of the call for expressions of interest meetings in Brussels in 2009 run as part of the 2Centre development. The notion of only one university being a centre of excellence in the UK is problematic; given the population of the UK and 50 universities with an active interest in this field. The author circulated a document in the UK to all universities suggesting a framework for a network centre of excellence representing the interests and providing an infrastructure to support the wider goals of the 2Centre project to improve the overall standard of Cybercrime Forensic education across the UK and contribute to its development over the member states of the EU. The British Computer Society Cybercrime Forensics Specialist Group (BCS CFSG 2013) will act as a main contact point for the centre and provide an additional professional body dimension to its operation. This document was shared with the 2Centre team and well-received and encouraged as a possible way forward for the UK. A number of national a local regional meetings have taken place between universities, police bodies and commercial companies and some groupings have secured national funding to move forward. Regional groupings in Scotland and Wales have launched centres of excellence in the last two years with funding from their respective government bodies.

The ECENTRE consortium is made up of 18 member organisations drawn from UK law-enforcement, commercial and university sectors (figure 3).

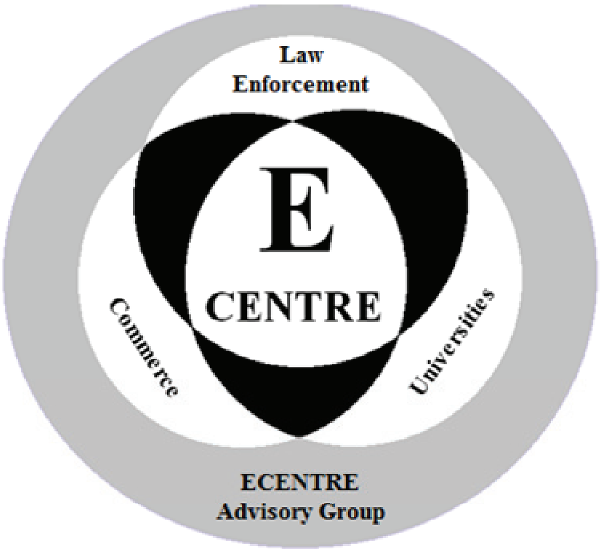
Each of the five local regional groups is a cluster led by a university (figure 4).

ECENTRE is a network of these regional clusters sharing resources using a common repository (figure 5).

The major ECENTRE project tasks include:

- Developing common standards for use by the consortium for artefacts;
- Evaluation of current and future developmental needs of partners;
- Developing new training courses for law enforcement nationally and from the EU member states;

- Developing new Cybercrime Forensic tools;
- Capture and sharing teaching materials, case studies;
- Research into developing areas of Cybercrime Forensics;
- Building information sets on current learning opportunities & people and organisations actively working in Cybercrime Forensics;
- Liaison with other centres of excellence in the EU sharing materials and toolsets;
- Project internal and external project peer review;
- Project research/teaching dissemination at a number of levels and hosting an international conference..



England's Cybercrime Centre of Excellence Network
For Training, Research and Education

Law Enforcement Agencies
The College of Policing
(formerly the NPIA
National Policing Improvement Agency)
ACPO eCrime Training
PCeU (Police Central e-Crime Unit)
(Advisors to Project)
Cheshire Constabulary (Consultants to Project)

Universities
Canterbury Christ Church University
Anglia Ruskin University
Kings College London
University of Bedfordshire
University of Greenwich
Liverpool John Moores University
University of Plymouth
De MontFort University
University of Coventry
University of Staffordshire

Companies
Evidence Talks Ltd
Technology Risk Ltd (Consultants to Project)
First Cyber Security Ltd
n-Gate Ltd
ManageMyProject (Consultants to Project)

Figure 3: ECENTRE project consortium

The aims of ECENTRE and 2Centre are complementary in that both seek to improve the standardization of Cybercrime Forensics education, training and research across the EU's member states. In the case of ECENTRE it is to create and infrastructure for the wealth of expertise, experience and good practice within England to be shared with the wider 2Centre network and to benefit from the wealth of expertise, experience and good practice within captured and disseminated by the members of the 2Centre network.

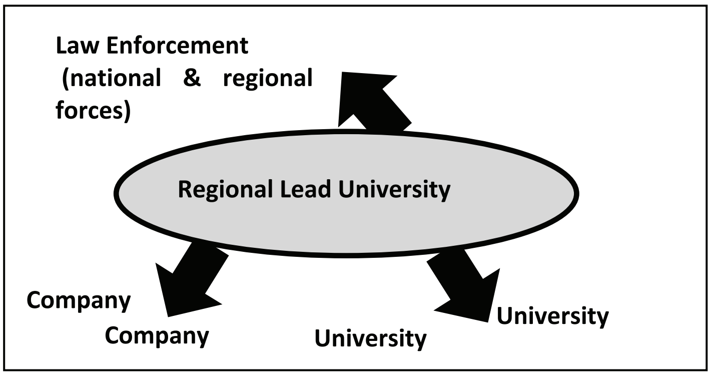


Figure 4: ECENTRE regional clusters

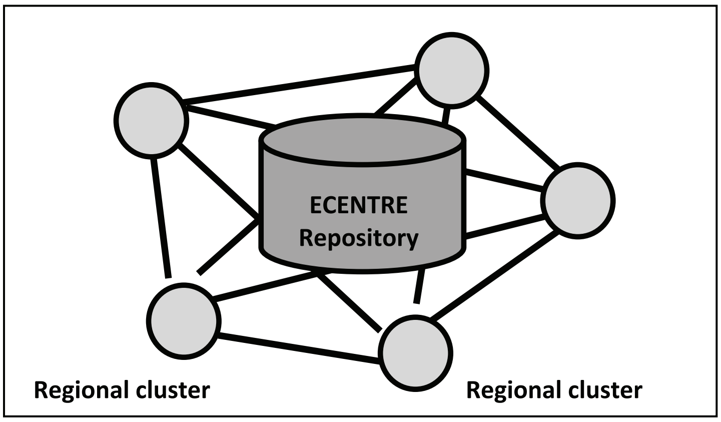


Figure 5: Common ECENTRE repository

The European network, at present, consists of centres of excellence in Cybercrime in France, Ireland, Belgium, Bulgaria, England, Estonia, Greece, Romania and Spain. The growing network covers a large geographic area in the EU (figure 6). This ECENTRE project will result in a large number of training opportunities for police officers from the UK and across the EU member states. The main focus of this is with The College of Policing in the UK who have the primary responsibility for providing such specialist training for the individual police forces. This is centralised within one body geared up to deliver such a provision; hence achieving a large economy of scale.

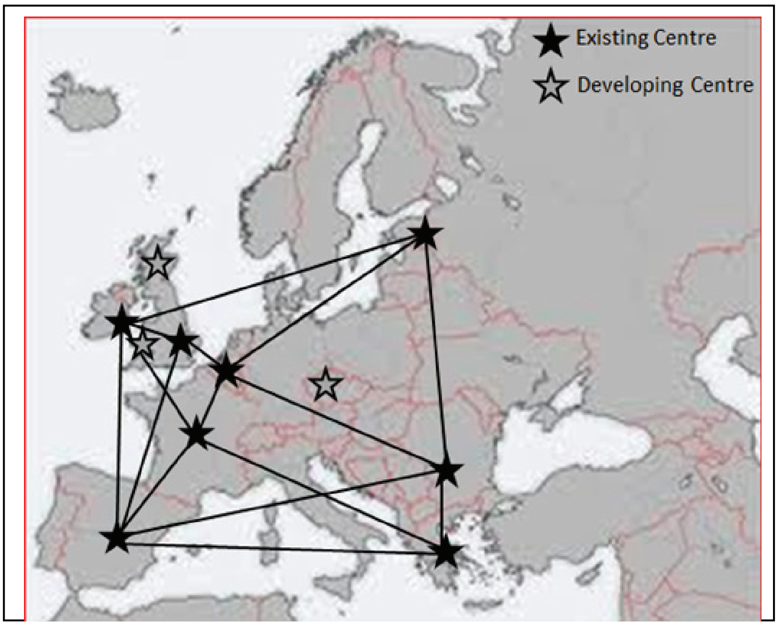


Figure 6: EU centre of excellence in cybercrime forensics

The diversity of the ECENTRE network across five regions of England identified extends the reach of dissemination activities by providing local regional opportunities for engagement with the public. the universities, in particular, have an outward public focus for informing the public about developments in the fight against cybercrime.

4. European cybercrime centre (EC3)

With the creation of the new European Cybercrime Centre in Europol in January 2013 (EC3 2013) the European Commission have created a focus of expertise and reference within the EU. Already discussion and meetings are taking place on how the national centres of excellence such as ECENTRE can interact and mutually assist EC3 in its role of fighting organised cybercrime. The new EC3 centre has been created at a time when budgets are very restricted and shows the importance and urgency of the problem. The funding for EC3 is not large; only a small fraction of Europol's total budget of €84 million. Compare this to the very large sums being spent by individual national governments on cyber security and it begins to look like very good value for money. For example, the UK Government alone has allocated an additional €800 million (£650 million) for its own national; cyber security strategy (PostNote 2011).

5. Not sharing information, ideas and solutions

There is clearly a lack of joined-up thinking in international cyber security. As discussed earlier, you cannot consider Cybercrime as being distinct from Cyberwar or Cyber Terrorism. They are all different views of connected and overlapping problems. There is considerable common ground but the opportunity attack the problem needs to reflect this to be effective. This is not happening. As soon as organisations of different types are involved, law enforcement agencies find it difficult to interact effectively without organisational styles, conventions, control and ownership issues getting in the way. But this is nothing compared with the problems of confidentiality, responsibility and trust issues which arise when they interact with other organisations such as universities. In a very high percentage of cases no details or information is released being classed as confidential. Intelligence agencies use a blanket banner of 'secret' to restrict practically everything. Within the Ministry of Defence in the UK the well-known saying is ensuring information is only shared on a 'need to know' basis. How do you determined if you 'need to know' or if there is indeed anything to ask that question about?

No solution has emerged which solves this fundamental problem. It would be naive to believe that information can be shared freely. It is, however, sensible to reflect carefully on what restrictions exist and if we are missing opportunities to learn from each other at different levels/sectore.

The ECENTRE project has also to avoid many more mundane and very practical questions to address.

6. Practical problems faced by the ECENTRE project

Obtaining funding does not guarantee success. As EU funded projects have found before this one, great care has to be taken not to get lost in discussions on intellectual property rights and copyright to the point where sharing does not take place. The whole point of the project is to share information, software, teaching materials, case studies and ideas. There is no easy solution to this. If you have spent a lot of time and money creating training materials, making them freely available can result in a third party copying, rebadging and using the materials themselves in commercial courses for profit. Examples of this have already happened in the 2Centre project to the point where a fairly restricted view is taken on some teaching materials. However, there are some good examples of sharing educational materials openly such as can be found in iTunes U where significant free resources from many subjects are available to anyone:

> *"Share your courses with the world.*
>
> *If you teach at an institution with a public iTunes U site, your courses can join the world's largest online catalog of free education content on iTunes U. Stanford, Yale, Oxford, UC Berkeley, MoMA, and the New York Public Library are among the hundreds of institutions that are sharing courses, lectures, videos, books, and other resources with students and lifelong learners all over the world."*
>
> *(Apple 2013)*

This, of course, to a great extent works because of the multimedia form of what is made available. A mistake made is to place too much emphasis on the teaching materials and forget that as important is the interaction with the teacher; the person telling the story and answering your questions.

Just as difficult is the balance between perceived quality of materials and artefacts produced and the volume and spread of what is available. Is it right to impose a very strict editorial control where the process of making something available in the repository is akin to journal publication or is it better to allow anyone to publish anything? For example, Wikipedia is a very useful resource as an overview of a subject and index into other sources but most academics would think twice before quoting it as a definitive source. Everyone strives for high quality but it can quickly limit what you are able to achieve. Many of the people actively engaged in Cybercrime Forensics (particularly in the law enforcement and commercial sector) do not have experience of academic publication or publication at any level. However, these are the people who have most to talk about and contribute with direct experience and ideas working at the cutting edge of the application of the subject.

As difficult is the simple question of what is, and is not, valuable to include in our materials we gather together. Try to include too much and you miss opportunities to develop the appropriate depth and detail; to give learners the techniques and tools to solve real problems. This brings us back to the difficulties identified with many existing training courses in the UK. Many courses lack the input from law enforcement or commerce to make them credible. It is also possible to over-value some aspects of this subject. In the experience of the author, many courses just focus on the technological skills to perform computer forensics, without giving appropriate time to the wider context of why things are done this way or the implications of actions. We also spend too little time on the history of Cybercrime and the different forms and effects seen in attacks.

7. Summary

What we would like to achieve is a wider range of understanding across all sectors of society of the nature of the problem of Cybercrime and Cyberwar and the possible solutions which are available. Without education in a particular subject people make poor decisions and make themselves more vulnerable to exploitation and the threats posed by computer misuse.

The term 'Zero-Day' flaw is used to describe the situation when we are confronted by a problem we have no experience of. Today is the day zero of trying to deal with the problem and find a solution. Without appropriate education and training (especially in this case for police officers) we all end up dealing with every problem as a new 'Zero-Day' problem (even if it is a well understood occurrence with known solutions). What is worse, we don't take the necessary simple steps to avoid becoming a victim in the first place.

Acknowledgement

With the financial support of the Prevention of and Fight against Crime Programme
European Commission - Directorate-General Home Affairs

This Project has been funded with support from the European Commission. This publication reflects the views of the author, and the European Commission cannot be held responsible for any use which may be made of the information contained therein.

References

2Centre (2013) *Cybercrime Centres of Excellence Network for Training Research and Education*, EU, [online], http://www.2centre.eu/

AMEinfo (2012) *Oman's CERT designated as Regional Cyber Security Centre in Arab World*, 15th December 2012 [online], http://www.ameinfo.com/omans-cert-designated-regional-cyber-security-322828

Apple (2013) *iTunes U,* Apple in Education, [online], http://www.apple.com/education/itunes-u/

Arsene, L (2012) *China's Cyber Militia Threatens US Cyberspace, Hot for Security*, 7th November 2012, [online], http://www.hotforsecurity.com/blog/chinas-cyber-militia-threatens-us-cyberspace-4313.html

BBC (2011) *South Korea Hit by Cyber Attacks*, BBC News, 4th March 2011, [online], http://www.bbc.co.uk/news/technology-12646052

BCS CFSG (2013) British Computer Society Cybercrime Forensic Specialist Group, [online], http://www.bcs.org/category/10468

Bracey, C & Edgar-Nevill, D (2007) *Prosecuting Low-Level Computer Crime in the UK*, CFET 2007, 1st International Conference on Cybercrime Forensics Education and Training, Canterbury Christ Church University UK, 6th & 7th September 2007, ISBN 1899253-041

Burn-Murdoch, J (2012) *The Latest Police Cuts Data: Where Have They Hit Hardest?,* The Guardian, 11th September 2012, [online], http://www.guardian.co.uk/news/datablog/2012/sep/11/police-cuts-reduce-force-sizes-data

CERT Australia (2013) *CERT Australia*, [online], https://www.cert.gov.au/

Coelho, H (2012) *Mandatory Cyber Security Standards Pose Risk to Competitiveness, Experts Say,* Business Technology report, distributed with The Daily Telegraph & The Sunday Telegraph, 17th December 2012, [available online], http://biztechreport.co.uk/2012/12/experts-warn-mandatory-standards-on-cyber-security-harm-competitiveness/

EC3 (2013) European Cybercrime Centre (EC3), Europol, [online], https://www.europol.europa.eu/ec3

ECENTRE (2012) *ECENTRE (England's Centre of Excellence for Cybercrime Training Research and Education),* EU, Prevention of and Fight against Internet Crime Targeted Call – ISEC 2011 Action Grants– Project Number HOME/2011/ISEC/AG/INT/4000002226

EU (2011) *Cyber Security: EU Prepares to Set Up Computer Emergency Response Team for EU Institutions*, Europa, {online], http://europa.eu/rapid/press-release_IP-11-694_en.htm

Gilad, A (2012) *Israel Prepares Itself for Cyber Attacks*, ALMONITOR, [online], http://www.al-monitor.com/pulse/security/2013/01/cyber-attacks-are-a-new-form-of.html

ISEC (2013) *Prevention of and Fight Against Crime*, EU, [online], http://ec.europa.eu/dgs/home-affairs/financing/fundings/security-and-safeguarding-liberties/prevention-of-and-fight-against-crime/index_en.htm

Leyden, J (2011) *Spooks Take the Wheel in UK's £650m Cyber-War Operations,* The Register, 28th November 2011, [online], http://www.theregister.co.uk/2011/11/28/cyber_security_strategy_analysis/

Paganini, P (2012) *Ponemon Statistics on Cost of Cybercrime for 2012*, Infosec Island, [online], http://www.infosecisland.com/blogview/22541-Ponemon-Statistics-on-Cost-of-Cybercrime-for-2012.html

PostNote (2011) Cyber Security in the UK, Houses of Parliament – Parliamentary Office of Science and Technology, Number 389, September 2011

Segal, A (2011) *Idea's About China's Cyber Command*, Asia Unbound, 27th December 2011, [online], http://blogs.cfr.org/asia/2011/12/27/ideas-about-chinas-cyber-command/

Segal, A (2012) *The Rise of Asia's Cyber Militias*, The Atlantic, 23rd February 2012, [online], http://www.theatlantic.com/international/archive/2012/02/the-rise-of-asias-cyber-militias/253487/

SDI (2012) *Strategic Defence Intelligence Yearbook 2012*, Berenice Baker (Ed), [online], http://viewer.zmags.com/publication/cc63a17c#/cc63a17c/1

TTOI (2012) *CERT-In, Electronics Companies Team Up to Combat Cyber Threat*, The Times of India, 6th December 2012, [online], http://articles.timesofindia.indiatimes.com/2012-12-06/security/35646489_1_cert-cyber-attacks-electronics

US (2013) *U.S. Army Cyber Command*, [online], http://www.arcyber.army.mil/

The Control of Technology by Nation State: Past, Present and Future: The Case of Cryptology and Information Security

Eric Filiol
ESIEA - Operational virology and cryptology laboratory, France
filiol@esiea.fr

Abstract: Since the end of WWII, strong controls have been enforced to prevent the spread of military-grade technology or dual use technologies and, since the end of the seventies, especially of Information Security science. The rises of the Internet phenomenon as well as the rise of terrorism make this control even stronger yet more subtle. Contrary to the common belief, the freedom of technology and science is just an illusion. Recently the emerging hacker phenomenon has upset and thwarted this balance between the need of freedom and the need of State security requirements. The main issue lies in the fact that these controls originally focus on homeland and international security purposes (e.g. protection against terrorism or mafia activities). But the fall of the iron curtain and of the Soviet block has dramatically changed the rules of the game. The enforced controls aim at first organizing an economic dominance of a very few Nation States (e.g. G-8 countries) whose real intent is to organize the strategic dominance over the ever-growing technological societies. As an example (among many others), we could mention the case of Cisco vs Huawei/ZTE companies. Based on his own military experience and on his academic work, we are going to explain how this control has been and is organized and will explain the role of the four major actors: Nation State, Industry, Academics and Hackers. We will take the domain of cryptography and of network equipment as illustrating cases.

Keywords: cryptography, dual technologies, export control , economic dominance, CoCom, Wassenaar arrangement, strategic dominance

1. Introduction

Since the end of World War II (in fact, in the mid-40s), the technology has been under strict controls especially regarding the export towards foreign countries. Although it concerned the military worlds at the very beginning - especially in the emerging context of the cold war era - this concerned many other technologies as well. The public opinion is totally unaware of that. The goal – which is understandable and compelling in itself - was the need to preserve, through the control of exports, the sphere and the regal power of the States which have an obligation to protect their interests (in particular military as well as economic ones), their citizens and their values. Terms like CoCom, Wassenaar, counter-measures, cryptology laws and regulations, backdoors ... illustrate this trend.

But things have gone much further and in a more pernicious way, thus preparing almost the absolute control by a very small number of entities (states and multinationals) on everything related to computer security and Internet. The purpose of this paper is to present the dark side of the computer industry, computer security and through it, of the Internet, and how could end a game of chess started in the 40s and to show the consequences - dramatic - both at the strategic level for the Nation states and regarding the issue of freedom for citizens.

To illustrate this and without loss of generality, we consider, as an illustration, the field of information technology, security, in other words - to be fashionable - the cyber defense and all areas that depend on that particular domain. But what is about concerns extends in fact to any possibly sensitive technology known or said "dual". The word "dual" should put question to any citizen, especially in a context of fierce international economic competition and of the rise of terrorism – the classical one but also the economic one. The information and the systems that process it are THE most critical dimension, which, nowadays, determines and is at the heart of everything: who is the master of information is the ruler of the world. The proposed restriction is therefore highly relevant.

The paper is organized as follows. Section 2 will present the four different players who are involved in a way or another in the different controls in place. Then Section 3 will explain the four steps that of those controls from WWII to the forthcoming years. Section 4 will summarize the consequences that we can draw from all that and will present what we can expect in the future before concluding in Section 5.

2. The four key players

Three players have led the different existing controls since the mid-40s. But the rules of this game have been fixed by the United States from the early beginning. Since the late 90s, the landscape has changed – as a noticeable singularity of history -- with the emergence of the hacker phenomenon. This fourth player, which can be considered as a real protest troublemaker that nobody really expected, entered the game with the will - and especially the technical ability - to upset the delicate balance in place, according to the principle that power is, according to Frank Herbert "energy that learns" (Herbert, 1965)

2.1 The Western Nation States

More than any other conflict, World War II was a war of technology (electronics, chemistry, nuclear science, radio communications ...) especially in the field of information. The failure in controlling the export of equipment sales and the dramatic consequences in the conduct of the war until 1942 (when the first successful cryptanalysis of the Enigma were possible) (Kahn, 1996) have sensitized the U.S. and Western countries to the need for export control and dissemination of sensitive technologies or dual use technologies, especially against the new threat posed by the communist bloc. A number of treaties and controls, similar to real selective embargoes were put in place. These controls, since the late 40s, have taken different forms and have been diversified with the development of society, of the economic and geostrategic balances and business practices. But they are still present today and the future will only finalize a long-planned program (see Section 3).

These control mechanisms are based on an increased capacity in the field of theoretical and applied research (R & D) in order to have a substantial advance regarding science and technology, whose purpose was to anticipate and orientate the direction of future controls. This research is heavily subsidized for years. These research entities publish very little, sometimes declassify in time, revealing a systematic and substantial advance compared to the academic research. Finally, their role is also to control the industrial and academic research, to fund them, to guide and to influence them as well as most of standardization bodies.

2.2 Industry (manufacturers, software publishers…)

Industry (manufacturers, publishers, service companies) are in fact subject to the law of their country or international laws when they are ratified nationally. This applies in particular to sensitive technologies in the case of technology export. In fact, the industry merely implements and sells technical choices that follow (real or *de facto*) national or international standards. According to Bernard Carayon, a former French MP in charge of economic intelligence issues, "*The power of a country lies in the ability to impose standards*" (Carayon, 2003).

In the case of cryptography, the block[1] cipher encryption technology or the statistical processes for the evaluation of the quality of randomness (FIPS-140, 2001) are among the best examples. Government entities in charge of the upstream control in key countries (e.g. G-20 countries) work in complete synergy with the industry, which somehow acts as the armored arm of the controls (especially for export). Except that the globalization of the IT industry and its concentration in a few non-national actors only mean that control is no longer possible by most of the nation states which must deal with the offer imposed by the very few dominant nation states (government and its industry). And all clients of those actors - we and our own nation state -- are de facto foreign customers of those dominant powers, thus are subject to certain considerations regarding the export of technology. This offer is also largely funded at the level of R & D by the governments of these dominant powers.

Beyond this relationship between governments and industry, which may be other factors and means of control? The weaknesses of implementation (backdoors disguised as programming or implementation flaws, the incompetence of developers can always be invoked to conceal malicious intent) promoted by the

[1] The technology of block cipher consists in mixing the plaintext and the secret key so that the links between the two become inextricably enough in order to prevent the cryptanalysts to access the plaintext. The problem is that the intractability is also for those who want to analyze the security of the process. No evidence of reliable or proven security to date has been ever published and, worse, their combinatorial richness and complexity can easily conceal backdoors mathematics. These systems have supplanted under the influence of the USA, the other encryption techniques. They are found everywhere nowadays (computers, networks, banking, smart cards ...).

organization of unbridled commercial competition, the absence of an obligation, from the governments, to enforce secure development processes, the existence of undocumented features allowing backdoors, the extreme variability of versions ... and a global cycle of evolution of systems requiring the user to run behind a commercial movement that places him de facto in a context of uncontrollable security. We have also to consider issues regarding intellectual property that protect against reverse engineering, most of the black boxes we buy - in other words, it is prohibited to analyze the products we buy, and even watch inside it.

An efficient and powerful backdoor does not summarize to a single piece of code: it is the combination of different factors (technical, organizational, human ...) which are known to be realized with a very high probability.

2.3 The academic community

The academic world is the third level of control. It is often used as scientific backing and therefore as a smokescreen. Why? Because in the field of security of information systems and, most underlying problems are so complex (in the sense of computational term) that any operational advance and any evidence of security is impossible to produce. The number of internal states of a cryptographic system for example is greater than the number of particles in the universe (according to the closed model). It is therefore impossible at least to store and explore any real system. The proof of security that the academic world is trying to provide is in fact out of reach. We have reached a point where the failure to provide proof of insecurity has become a security proof in itself. Under these conditions, proving the existence of a backdoor (particularly a mathematical backdoor, that is to say, at the conceptual level) is like looking for a needle in a million haystacks. Only the one who has put this backdoor knows where to find it and then how to exploit it.

Part of this control goes through

- the promotion and organization of scientific orthodoxy (coring and controlling program committees of international conferences, thematic orientation towards fashion research topics who are more likely to be published, subtle but perverse exploitation of the "*publish or perish*" syndrom...);
- The control by money (state funding but also by manufacturers who are able to define and influence academics to what fashion research topics and thematic are, see above) and by research funds and grants (National Science Foundation [NSF], National Security Agengy [NSA], Seventh EU Framework Programme [FP7] ...);
- The control by law: patents, intellectual property, scientific research themes that are potentially contrary to the law and the national security (e.g. France's Article 323 of the Penal Code).

The best example is the mathematical problem of factoring[2]. An easy method of factorization (still unknown at this time, or unpublished if it exists) would be prohibited from publication because of the huge implications on the security of all systems in the world. All the security of those systems would collapse causing global chaos. Such a discovery would be quickly identified at the stage of its premises thanks to the various control in place (the first one being the many internal evaluations of research), and banned from publication. The scientific community is anything but independent. If it can be force for scientific proposals - and often brilliantly, admittedly - it is in fact the standards (enforced at the international or national level, mainly by the United States, the latter monopolizing the standardization bodies and entities), the nation states and the industry which lead and control the game This is the reason why, despite an academic wealth in the field of cryptographic algorithms, we undergo an actual hegemony of block cipher systems and in particular of the U.S. algorithm AES-256 (FIPS PUB 197, 2001): a single algorithm to tie us all. This result has been obtained by combining political and industrial lobbying and industrial, strategic influence, threat of economic retaliations, exploitation of the fear of non-interoperability with the dominant technological power. And the academic community has just served as scientific caution or has sold his soul for some publications and honors.

2.4 The rise of the hacker phenomenon

The State / Industrial / Academic triptych worked well until very recently: the nation states choose and control which technology can be proposed to citizens, the industrial manufacturers or software vendors build and

[2] The integer factoring problem – to split integers into a product of prime integers; a prime integer p is divisible by 1 and p only – is at the heart of most security systems. Solving this problem – for large integers – would put the security of all those existing system into question.

market approved products while the academic community bring a semblance of academic scientific backing. But since the late 90s, the hacker community has risen and has put everything upside down, such as a real singularity of technological history. Hackers - unlike academics - favor the results over the methods. Academic and hackers are at both ends of the activity of scientific and technical research as explained René Thom, an eminent mathematician and philosopher of science: "Nature is such that understanding and action are not synonymous." (Thom, 1991) For the academic community, the fact that it works in practice is not valid until it does not work in theory! For the hacker, it is the law of the efficiency and operational realism operational which must rule science and technology.

The problem is that hackers find backdoors very quickly (unless they are of a mathematical nature) even when they are hidden at the silicon level (Nohl, 2009). Because they are innovative, creative, free from conceptual straightjacket. For them no subject is dangerous or threatens their careers and image. Everything is happening now in hacking conferences (Black Hat, CCC, Defcon, Brucon, Hack.lu, Syscan, HIP...). And in a society where ultra-computerized technology grinds and crushes citizens more and more instead of freeing them, hackers stand as resistants and whistleblowers in an economic and strategic warfare which is increasingly evident.

2.5 A simple but illustrative case

In order to illustrate the way things are currently managed, let us take the example of encryption systems. If we have look at the Wassenaar Arrangement dual-use list, category 5, part 2 (Information security), on page 3, paragraph 5.A.2.a.1.a, we can verify that “symmetric algorithm employing a key length in excess of 56 bits;” are encryption technology under control. As far as the AES (whose secret key has an entropy ranging from 128 to 256 bits) is concerned, the publication of the AES is a clear violation of the Wassenaar Arrangement as well as of the different national regulations of G-8 coutries that have been derived from this international arrangement.

3. The four phases of controls history

3.1 The “prehistory”: From 1942 to 1975.

The control of sensitive or dual-use technology seemed obvious from the beginning of the Second World War. The technological advance was THE most critical dimension of that era. Any technical advance was a strategic advantage indeed on both sides. At the end of the war, Western countries under the influence (or pressure) of the United States signed in Paris in 1949, the *Coordinating Committee for Multilateral Export Controls* (CoCom) (CoCom, 1949) whose role was initially to prevent countries under influence of the communist bloc (USSR and China) to purchase goods, materials and technologies that would really or potentially represent a military, strategic or economic interest. This included, for example, computers, software, sophisticated equipment for telephone, GPS technology, technology of chemistry, physics (electronics)... If the context of the Cold War could explain such a control at that time, it is more difficult to accept when it was enforced by the United States against European countries such as France. For example, among others, we can mention the market of supercomputers that has long been closed to Europe and has started to open up at the dawn of the 90s.

It is interesting in this context to consider an unprecedented event that occurred in 1977: the publication for the first time ever, of a ciphering algorithm, the DES (Data Encryption Standard) (Fips Pub 46, 1977), by the U.S. government with technical support of the NSA. This algorithm was presented as a highly secure one. To make it clear to readers the significance of this event, you can draw a parallel, which is entirely appropriate, as follows: to publish such a know-how in the field of highly secure communications would have been equivalent to publish the plans of the nuclear weapon (H-Bomb).

Who can believe seriously that such a publication in the CoCom context was possible without any form of control upstream? This publication would otherwise have constituted a serious violation of the CoCom. In fact, later in 1992 the publication of certain works of researchers (Biham, 1990) and the corresponding embarrassed response from the NSA have provided hindsight regarding this issue. NSA acknowledged more or less explicitly that it began to work on DES-like technology since 1966 at least, although officially its official birth was in 1975. The control probably lies at mathematical backdoors level. Note that so far, it is still impossible to investigate and explore computationally DES exhaustively and therefore find these backdoors.

Finally, the publication of this algorithm gave birth to an actual academic community in cryptology. This community was indeed *de facto* strongly influenced, therefore, "directed and oriented" by mathematical concepts of US and government origin and that nobody has been able so far to prove the actual security, due to the huge combinatorial complexity of these concepts. The "proof" of security for these algorithms, which can be seen as a “default” proof, lies in fact in the inability of the academic community to produce a single operational attack.

This publication has given rise naturally to all other ciphers marred with the same conceptual backdoors thus making them automatically controllable: for example the IDEA algorithm - directly inspired by the same mathematical concepts in DES – which is included in the common user-oriented encryption software PGP. The PGP case was at the heart of what must considered a "true-false” or fake legal case designed to encourage people to use an alternative software -- however still under control -- and hence to manage a growing mistrust *vis-à-vis* the DES. Philip Zimmermann was probably sincere and convinced of the security of PGP, and he was himself the victim of a subtle control technology policy, prepared years ahead.

It is important to remember that, unlike almost all countries and transnational organizations (approximately 120 in 1995), General de Gaulle made mandatory, for any nation state (military, diplomatic, economic, political) needs regarding communication security and encryption means[3], that France uses national equipment designed (from the mathematical concepts to the very final implementation) and built by national entities only. The Hans Buehler case (Strehle, 1994; Filiol, 2006), in 1995, showed that all the countries that had not done so, had seen their encrypted communications heard by the U.S. for over 50 years, thanks to mathematical and implementation backdoors put in all enciphering devices that had been sold to those countries. It proved, once again, the extremely clever vision of General.

3.2 The transition phase: From 1975 to 2001

The birth of a scientific community in the field of information and system security as well as the evolution of society itself (especially with the spread of computer and networks for all citizens), the fall of Eastern bloc in 1989 and the various geopolitical upheavals ... made from the late 70s, the CoCom type controls difficult to explain, maintain and manage. The period 1977 to 2001 was marked by the end of the Cold War, the rise of terrorism and the beginning of the global economic war that no longer hides its name. It became therefore necessary to diversify the controls according to the principle of eggs and basket.

CoCom has been dissolved in March 1994, and quickly replaced by the Wassenaar Arrangement (Wassenaar, 1996), which defines 12 lists of materials and technologies subject to controls (42 countries have signed this agreement up to now). Other controls, less visible but however very efficient, are also taking place: the Socrates project, GATT and later WTO with the aim of building a global world economy (see next step), to organize and lead the technology standardization under U.S. influence (particularly regarding technology related to the Internet), the mutation of national laws and regulations (supranational laws [European treaties or laws] become automatically national laws), the development of monitoring networks like Echelon... So the world is changing, the controls are organized and adapted to follow those mutations. The target is no longer a few communist states, but every citizen of the world, equipped with a computer that is both a potential consumer in a world under globalization and/or a potential terrorist (Islamic, anti-globalization, occupy Wall street, Anonymous...). The "threat" becomes diffuse, polymorphic, so fine in granularity that is necessary to globalize controls and their management.

3.3 The globalization phase: from 2001 to 2012

With the rise of terrorism, the 9/11 attacks on the one hand and the alter-globalization on the other hand, another dimension has emerged: that of emerging countries (China, India, Gulf countries ...), which most often have not signed the Wassenaar Arrangement. The control of technology which is spreading everywhere must become a control set up well in advance. It is necessary to globalize in order to concentrate technology and its marketing in the hands of a very few multinationals easier to control. To reach that goal, from 2001 to 2012 many news control levels have been implemented:

[3] The reintegration of France into the integrated NATO command raises "questions" from a few of our decision makers on the need to maintain such a French specificity. To maintain full interoperability with NATO, why not adopt the tools of our "American friend" and therefore reduce our costs in this area. It illustrates that culture and sociology may be another control tool.

- The dominance of the private sphere over the public sphere. Nation states can no longer close their market: they are open to competition (WTO effect under U.S. influence while the latter are closing their own market...). It is the political level of controls. The aim is to weaken the current political landscape, consisting of nation states exercising their economic and political power and hence still own pieces of sovereignty.
- take-over bids, M&A (Mergers and Acquisitions), elimination of competitors, concentration of production and services (eg the market of routers with Cisco vs Huawei). Here lies the economic control: the free competition imposed by the WTO and the subsequent various deregulations - with a U.S. protectionism which becomes even stronger in parallel) have shattered the last control capabilities at national level by the gradual disappearance of national technology champions.

Without loss of generality one example will suffice to illustrate this phase: the liberalization of cryptography. It has been initiated in 1997 with the Hourtin speech of Prime Minister Lionel Jospin under the pressure from the U.S. A, has in fact led to the hegemony of a single cryptographic algorithm: the “Advanced Encryption Standard (AES)” (FIPS PUB 197, 2001), which once again, has been offered to the world for free and openly, by the U.S. Department of Commerce, with the technical assistance of the NSA. Let us recall that the official publication of the algorithm occurred in a rather critical context: the rise of terrorism, of rogue state like Iran, North Korea. This would mean a direct and clear violation of the Wassenaar Arrangement and of any embargo in place at that time! Anyone who uses this algorithm *de facto* protects its communications and data against eavesdropping (naïve view) … or fall under the control of the one who knows the backdoors in place (realistic view).

It is possible to give many other examples: RIM/Blackberry, the microprocessor industry (disappearance of AMD, Sun, powerpc, ... RISC processor in favor of Intel products only), operating systems, Facebook/Google/Twitter which results in the mass surveillance of at least one billion of people ...In 2012, we finally obtain a massively computerized, networked society, which enable to spread thousands of backdoors of very different types: 0-day vulnerabilities relentlessly renewed day after days, sophisticated state malware (as spying malware like Magic Lantern (USA), Bundestrojan (Germany), Stuxnet ...). In short, a world that makes Orwellian vision a simple fairy tale or a children story.

3.4 The “legal” phase: from 2013 to…

What will follow next? This whole edifice cannot stand and be viable without a final dimension: the dimension of law. From 2012 now on, we are witnessing to the rise of supremacy of laws over the international business primarily in two main dimensions:

- Intellectual property regulations. Multinationals in place, their power, interest and influence must be protected. PIPA/SOPA/ACTA, software patents, patent wars (Samsung vs Apple case but previously NOVELL vs France Telecom), standards war ... are only the beginnings of this trend;
- The rise of law related to cyber defense and the demonization of actors who stay beyond the existing controls (such as hackers, for example with Vupen or CoseInc cases): the extension of the geographical territoriality to the digital territoriality (of a computer server within a U.S. domain name, even located outside the U.S., is considered as a part of the U.S. territory), the Patriot Act and its various variations from Bush to Obama, cybercrime laws…

4. Consequences and discussion

In the end, all of the above has led to the concentration of technology and services in the hands of a very few states and multinationals that have absorbed and/or eliminated competition. They are now in control of everything, putting the rest of the world in an obvious dependency state that must be the ultimate control. In the growing context of cyber warfare (defensive but also offensive) this economic domination in fact hides the strategic dominance (both at political and military level). The confrontation Cisco vs Huawei - except for the naive - should not be considered as a mere technological and economic confrontation. This is the strategic dominance of the information passing through their routers - accessible via the backdoor they can hide inside. This is the real dimension. It is the same for many other cases: Samsung/Apple, Microsoft /Android/Linux ... In this context, the Intel/Microsoft[4] agreement around UEFI technology (with the consequence of binding

[4] It should be noted that these companies of U.S. law, are subject to the national laws of their country and many critics, often unfair in fact against them forget this important aspect.

equipment [computer] to a unique operating system thus preventing the installation of any other alternative system) is a clearly worrying and terrifying trend.

In this gigantic worldwide game, it seems that China and other emerging countries (South Korea, India, Brazil, the Gulf States ...) are poised to put 60 years of U.S. hegemony, and, by extension, Western countries influence, into question. From the Monroe Doctrine -- however revised by the U.S. intelligence at the end of the Second World War -- to the Chinese vision of globalized confrontation as defined in the book "Unrestricted Warfare" (Qiao & Wang, 1999), it appears that the world could change of game and switch from the chess game to the game of Go.

We should therefore consider the rapid economic development of emerging countries which have not signed the Wassenaar arrangement under a new light. Those countries have been the victims of this arrangement for over 50 years. It is very likely that they will choose to take their revenge and consequently not to ratify the Wassenaar arrangement. Like China, the growing economic dominance could become a strategic dominance. It is not sure that the declining influence of the U.S. is strong enough to impose controls as it was the case for example regarding the laser printers in 2001 (the *Electronic Frontier Foundation* in 2001 revealed[5] that most laser printers marked printed material with micro-points. From any printed page, it is possible to trace back to the printer and then to its owner).

5. Conclusion

In addition to the results mentioned in this paper, the evolution of these controls has a direct impact on innovation in Western countries and the medium to short term, those countries may become nations of second order. In fact, for the innovator, it is wiser to create his company in a non-Wassenaar country.

It is therefore essential to quickly restore the capacity and willingness of national sovereignty or at least in Europe in the area of sensitive technologies, driven by SMEs/SMIs. If Wassenaar-like controls are undoubtedly necessary to protect the safety and interests of the state and its citizens, the latter must receive assurances from their governments that control is strictly enforced by their governments only not by outside actors (another State [USA, China..], multinational companies…). In other words, as citizens, we should still have the choice of the backdoor, if there should be. Backdoors are not inevitable. It is possible to provide truly secure tools - as are national versions of the technology when they still exist - without backdoor. The needs of a sovereign country can be assured at national level, more by law and through the respect of citizens for their nation state. The problem is not technology but the power and the consideration that the citizen gives the state.

It is necessary, even vital, to have a truly independent and active academic community which is still the master of its own destiny and that is not subject to the dictates of the Shanghai ranking (another form of control where the size outweighs the quality).

References

Biham E., Shamir A. (1990). "Differential Cryptanalysis of DES-like Cryptosystems". Advances in cryptology CRYPTO 1990, Lecture Notes in Computer Science, pp. 2-21.

Carayon, B. (2003). "Intelligence économique, compétitivité et cohésion sociale". France National Assembly Report, July 2003. Available on http://www.ladocumentationfrancaise.fr/rapports-publics/034000484/index.shtml

CoCom Archives (1949 – 1994) http://www.diplomatie.gouv.fr/fr/ministere_817/archives-patrimoine_3512/fonds-collections_5143/organismes-internationaux_11594/cocom-1949-1994_25985.html

Federal Information Processing Standards Publication (1977). „Data Encryption Standard", FIPS PUB 46, National Bureau of Standards.

Federal Information Processing Standards Publication (2001). „Security Requirements For Cryptographic Modules" FIPS PUB 140-2. US Dept. Of Commerce. Available on http://csrc.nist.gov/publications/fips/fips140-2/fips1402.pdf

Federal Information Processing Standards Publication (2001). „Advanced Encryption Standard" FIPS PUB 197, US Dept. Of Commerce. Available on http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf

Filiol E., Richard P. (2006). "Cybercriminalité – Les mafias envahissent le web". Dunod ISBN 978-2100502783. Available under Creative Common License on

[5] https://www.eff.org/pages/list-printers-which-do-or-do-not-display-tracking-dots

https://docs.google.com/viewer?a=v&pid=explorer&chrome=true&srcid=0B6BlkqAoxXq1ZDIzZDVjY2QtMjZjNi00NTNmLTkzNmItMDQ4NDc5YWYwYjdk&hl=en_US

Herbert F. (1965) “Dune”. Chilton Book Co Publishing. ASIN B003GRQJ86.

Kahn, D. (1996).“ The Codebreakers - The Comprehensive History of Secret Communication from Ancient Times to the Internet”. Macmillan Publishing. ISBN 0-684-83130-9

Menezes A J, van Oorschot P C and Vanstone S A. (2001) « Handbook of Applied Cryptography », CRC Press, ISBN 0-8493-8523-7.

Nohl K., Starbug (2009) Silicon Chips : No More Secrets. PacSec 2009. Tokyo.

Col. Qiao, L. and Wang X. (1999) “Unrestricted Warfare”. People’s Liberation Army. Litterature and Arts Publishing House, Beijing. [online] http://www.terrorism.com/documents/TRC-Analysis/unrestricted.pdf

Strehle R. (1994). “Verschluesselt – Der Fall Hans Buehler”. Werd Verlag. ISBN 978-3859321410

Thom, R. (1991). “Comprendre n’est pas expliquer – Entretiens avec Emile Noel”. Flammarion.

Wassenaar Arrangement (1996) Official website http://www.wassenaar.org

The Issues of Software Being Classified as Malicious by Antivirus False Positive Alerts

Grigorios Fragkos[1], Olga Angelopoulou[2] and Konstantinos Xynos[3]
[1]Senior Consultant - Penetration Tester, Sysnet Global Solutions, UK
[2]School of Computing and Mathematics, Faculty of Business Computing and Law, University of Derby, Derby, UK
[3]Information Security Research Group, Faculty of Advanced Technology, University of Glamorgan, Wales, UK
greg.fragkos@sysnetglobalsolutions.com
o.angelopoulou@derby.ac.uk
kxynos@glam.ac.uk

Abstract: The continuous development of evolving malware types creates a need to study and understand how antivirus products detect and alert the user. This paper investigates today's antivirus solutions and how their false positive alerts affect the software development and distribution process, which in the long term could even lead to loss of business. It is discussed and demonstrated how antivirus detection deals with bespoke applications and how this can be reversed and manipulated to evade detection, allowing to be used by malicious software developers. The paper also presents ideas that would enable antivirus products to overcome these detection issues without altering their detection engines but by focusing on the developer's source code submission. The potential lack of essential and in most cases obvious steps in malicious software detection is also examined. The paper concludes that the inconsistencies between different antivirus detection engines along with the introduction of reputation based detection, allows more sophisticated and undetectable malicious software to be created and spread.

Keywords: antivirus, false positive alerts, software, malware, reputation systems, APT, EMEA

1. Introduction

People who are keen on programming languages tend to write custom code in order to perform specific tasks. A small stand-alone application is sometimes all that is needed in order for a developer, a researcher, or even a computer enthusiast to effectively overcome a time-consuming task. Thus, it is not uncommon for people who are familiar with one or more programming languages to come up with an algorithm which in turn is compiled into a piece of executable code. In some cases, minor code tweaking and debugging is required for the application to be distributed as freeware (Graham, 1999) to the public.

Antivirus (AV) software is *a program that aims to detect and remove computer viruses and malware.* According to Skoudis and Zeltzer (2003), malicious software or malware is

> *...a set of instructions that run on your computer and make your system do something that an attacker wants it to do.*

The distribution of freeware could result in a number of unpredictable detection scenarios when installed on a computer system that is using an AV software solution (Hawes, 2013). As it is presented in the paper, there is high possibility that original and legitimate applications will most likely produce a number of alerts, false positives, similar to those seen in malicious software, when scanned by AV products. Thus, any legitimate application can be unreasonably classified as malicious by AV products.

As already mentioned, this becomes an even bigger problem if the application has already been distributed to multiple users and then it is being flagged up by a number of AV solutions as malicious. Jacobson and Idziorek (2013), correctly identify that:

> *a false positive occurs when an antivirus program incorrectly determines that a legitimate piece of software is actually a malware program.*

The classification of software as malicious affects not only the application's credibility, but also the developer's reputation. In reality, the AV solution classifies as malicious any application that matches a specific signature in its rule-based system (Wang, 2006). Dealing with tiny applications -in regards to their file size-, which contain only a few simple and straightforward functions should be a trivial task. In reality this is not the case, especially when the application behaves in a manner that is not expected by the AV. The results vary as the software being deemed as suspicious is classified differently by the various AV vendors.

The paper discusses the aforementioned problems and investigates how "black box" scanning affects not only legitimate software. It also examines how this can be manipulated without any advanced methods for the purpose of hiding malware. These methods utilized by malware authors are clearly outlined by Symantec (Symantec 2013).

2. Dealing with false positives in the real world

AV products tend to produce false positive and false negative alerts (Christodorescu and Jha, 2004). Jensen et al. (2009) discuss the weaknesses of antivirus software in terms of the signatures they adhere to. We aim to support that false positive alerts are problematic for the wider distribution of legitimate freeware applications. Therefore, we focus our work on the demonstration of scanning small applications, developed on purpose to perform the experiments for this work and with the intention to validate our argument. Jensen et al. (2009) also discuss about the weaknesses of antivirus software and modern endpoint products (EPP) (Abrams R., et al, 2013). As much as the numbers of false positives need to be minimized by the AV vendors, it is not possible to eliminate them completely as the current state of the art research demonstrates (Bayer et al., 2010, Schneider and Griffin, 2011).

The intent of the paper is not to describe or assess in a technical level the limitations of the current state of the art of AV detection techniques, but to provide some thoughts concerning the way these limitations could alternatively be dealt with. When a legitimate piece of software is being classified as malicious by a number of antivirus products, the developer has limited options. She can modify the source code hoping that the end products will not be flagged up again. Another solution would be to go to each antivirus vendor's website and report a false positive exception. In any case the procedure is time consuming both for the developer and for the antivirus code reviewers.

Users may use the Avira (2013) service where they can either submit suspicious files for analysis or report false positives to be re-examined. However, it appears that not all antivirus vendors have a straightforward way of reporting a false positive (Liu and Ormaner, 2009). Thus, going through approximately over forty AV vendors can be very time consuming, if not impossible. Getting into the process of reporting a false positive alert to each antivirus vendor, does not mean that it will necessarily end up being classified as legitimate. For example, the authors reported for analysis a false positive the xIP.exe application, which will be further discussed in section 3.1. The response which came back from the AV vendor was as follows:

> *Thank you for your email to Avira's virus lab.(26/Jun/12)*
> *Tracking number: INC01188731.*
>
> *xIP.exe 6 KB – MALWARE*
>
> *The file 'xIP.exe' has been determined to be 'MALWARE'.*
> *Our analysts named the threat TR/VB.Downloader.Gen. The term "TR/" denotes a trojan horse that is able to spy out data, to violate your privacy or carry out unwanted modifications to the system. Detection will be added to our virus definition file (VDF) with one of the next updates. Please note that Avira's proactive heuristic detection module AHeAD detected this threat up front without the latest VDF update as: TR/VB.Downloader.Gen.*

Figure 1: Avira's response after filing a false positive alert

Consequently, going through this process could take days or even weeks and eventually end up without the desired results.

There have been cases where the antivirus vendor recognized the false positive alert and issued an update [Figure 2] to their virus definition file.

In a couple of months' time the same piece of software was suddenly once again flagged as malicious, by the same AV product which had only recently excluded it from its threat list. In this case, one of the executable files contained within the main setup file were deemed suspicious. Again, going through the process of reporting it as a false positive [Figure 3] resolved the problem.

Thank you for your email to Avira's virus lab. (11/Nov/11)

Tracking number: INC00884907.

A listing of files alongside their results can be found below:

File ID	Filename	Size (Byte)	Result
26391328	SetupFile.exe	5.67 MB	**FALSE POSITIVE**
26391927	lame_mod.exe	216.5 KB	CLEAN
26391928	MSNPSM3.exe	294.86 KB	CLEAN
26391929	MainApplication.exe	264 KB	CLEAN
26391931	SScreen.exe	120 KB	CLEAN
47057	…	144.27 KB	KNOWN CLEAN

Figure 2: Avira’s response recognizing False Positive alert

Thank you for your email to Avira's virus lab. (26/Jan/12)

Tracking number: INC00961360.

A listing of files alongside their results can be found below:

File ID	Filename	Size (Byte)	Result
26527344	SScreen.exe	120 KByte	**FALSE POSITIVE**

Figure 3: Avira’s response after filing a second false positive alert

This indicates that having an AV vendor withdraw an alert due to a false positive doesn’t mean that it will not suddenly reappear in future AV definitions. Christodorescu and Jha (2003) discuss about the effective detection of malicious code. It can be applied in the identification of malicious desktop applications, as well as the currently popular mobile applications. Mobile applications available for download from Apple iOS App Store as well as from Google Android Play Store are in total approximately at 700,000 each (Oliver, 2012). The number of desktop applications is undefined due to the inability to account for the different kind of software available, open or closed source (Laurent, 2004). It is interesting enough to see that in this new era of modern mobile applications the responsibility for detecting suspicious or malicious applications has shifted from the AV vendors towards the OS vendors, such as Apple(R) with iOS and Google(R) with Android.

In both cases, the aforementioned companies are trying to protect the end users from malicious applications. Especially those which contain suspicious code or those that might have suspicious behaviour and/or activity, before being deployed and made available for download (Ketari and Khanum, 2012). Effectively, the AV products in these cases have to work with applications which have been approved in advance (Prince B., 2013) compared to the chaos of desktop applications being scanned on a daily basis.

Thus, it is not clear enough in the literature if the detection engine of antivirus software for mobile devices has been actually improved in detecting concealed threats (Raggo and Hosmer, 2012). There is also no evidence suggesting that by removing a big portion of the responsibility in identifying malicious activity allows the antivirus software to suggest it performs better than usual. On the other hand, due to the vast amount of desktop applications, antivirus products commence detection by relying on regularly updated signatures, which describe potentially malicious software patterns. Effectively, this paper makes an effort to discuss and suggest a similar process for the AV vendors to better filter desktop applications.

The process of just regularly updating the virus definitions file has proven inadequate to meet the purpose of successfully defending the end systems from unknown threats (McMillan R., 2012). It would seem that the AV vendors have turned towards software reputation in order to deal with rising threats (Microsoft, 2012) and to classify unknown applications. For example, Symantec in a recent blog post about the New York Times hack (Muncaster P., 2013) states clearly:

> *The advanced capabilities in our endpoint offerings, including our unique reputation-based technology and behaviour-based blocking, specifically target sophisticated attacks. Turning on only the signature-based anti-virus components of endpoint solutions alone* ***are not enough*** *in a*

> *world that is changing daily from attacks and threats. We encourage customers to be very aggressive in deploying solutions that offer a combined approach to security.* ***Anti-virus software alone is not enough.***

Despite the fact that individual developers or small software houses cannot be compared to the reputation of larger software giants, there is also something else that is not fully taken under consideration. It surfaced in the media throughout last year's reports (i.e. 2012) that Middle-East countries were heavily targeted by sophisticated malware such as Stuxnet, Duqu, Flame, Gauss, Red October and other variations. In these countries any variations of the malicious software would be extremely difficult to be classified based on software reputation simply because the rules of engagement are different. These countries have a relatively small Internet footprint as they are isolated, in a sense, due to local laws and Internet restrictions (such as countries of the Middle-East, China, Pakistan and others).

Effectively, the different variations of threats created by an attacker could be customized to inflict damage to domestic targets, not necessarily allowing the threat to spread out of the "digital borders" of these countries. Consequently, for reputation-based antivirus software solutions it would become very difficult, if not impossible in some cases, to classify arising threats. Most importantly it won't be long before the reputation system could start working in favour of malicious software writers. All they would need to do is produce software which might seem useful to a large number of people, such as a new web browser, which in turn could actually be hiding a far more sinister behaviour once it gained a viable reputation.

Thus, one may think that today's antivirus false positive alerts along with the forthcoming reputation based antivirus solutions are trying to protect end users from unknown authors, instead of pushing the barrier higher by developing even better detection engines.

3. Creating and testing simple executable files against up-to-date AV products

As already stated, false positive alerts with custom made applications initiated the idea to look closer into this issue and the decision was made to investigate it further. In order for this to be achieved, a number of custom made application were created which are not malicious and the source code is made available for verification upon written request to the lead author. The online service Virus Total (2013) was used for testing and comparison of the results. This service incorporates a large number of antivirus products and it was preferred compared to other similar online file scanners from individual AV vendors (avast!, 2013, Bit Defender, 2013, Metascan Online, 2013). Virus Total is a widely used tool among the Information Security community (Bishop et al., 2011, Kelchner, 2010, McMillan R., 2012) and scans a file against many more antivirus engines in comparison to any other online file scanners (Threat Expert, 2013).

The applications created and used in our investigation are:

- xIP.exe
- OnlineFileProperties.exe
- wget.exe
- KeepMeIn.exe

A simple testing method was used. Each one of the applications was scanned twice by the Virus Total (2013) service in a six month period. Each application used for the experiment has two analysis dates along with their respective findings. It was interesting to notice the different results produced by the same antivirus product over that six month period.

3.1 Application I: xIP.exe

In our first test, six out of forty-two results classified the application xIP.exe as threat, (Table 1 - Analysis date 1). This is a command line tool which displays, as output on the console, the external IP address by accessing a hard-coded URL in the source code. The application does not execute anything in the background, it does not look for write or read access to the hard drive and it does not contain any malicious code. It requires Internet access in order to display the external IP address.

Table 1: Analysis results for xIP.exe on two different dates

File name:	xIP.exe
File type:	Win32 EXE
File size:	7.0 KB (7168 bytes)
SHA256:	9757cac64238c14d5f970711f3f67465cc4be04161070fe0e65c7e428cf95546
SHA1:	8af15a14bd4ad06c0a9fbdec14a43a67539e61c3
MD5:	50ee45f4855257f978f4a442b3823bce
Analysis date 1:	2012-07-12 07:51:18 UTC
Detection ratio:	*6 / 42*
AntiVir	TR/VB.Downloader.Gen
Emsisoft	Trojan-PWS.Win32.QQPass!IK
Ikarus	Trojan-PWS.Win32.QQPass
PCTools	HeurEngine.ZeroDayThreat
Symantec	Suspicious.Emit
TheHacker	Posible_Worm32
Analysis date 2:	2013-01-31 22:40:30 UTC
Detection ratio:	*18 / 46*
AntiVir	TR/VB.Downloader.Gen
BitDefender	Gen:Variant.Kazy.117115
Comodo	UnclassifiedMalware
F-Secure	Gen:Variant.Kazy.117115
GData	Gen:Variant.Kazy.117115
Ikarus	Trojan-PWS.Win32.QQPass
K7AntiVirus	Riskware
McAfee	Artemis!50EE45F48552
McAfee-GW-Edition	Artemis!50EE45F48552
MicroWorld-eScan	Gen:Variant.Kazy.117115
Norman	Suspicious_Gen4.AQUEQ
Panda	Suspicious file
PCTools	HeurEngine.ZeroDayThreat
SUPERAntiSpyware	Trojan.Agent/Gen-Koobface[Bonkers]
Symantec	Suspicious.Emit
TheHacker	Posible_Worm32
TrendMicro-HouseCall	TROJ_GEN.R47B1I4
VIPRE	Trojan.Win32.Generic!BT

When the xIP.exe application was scanned on a consecutive date, (Table 1 - Analysis date 2), eighteen alerts were raised. It is quite clear the exponential increase in the alerts generated during the second scan. It is also worth noting that in the first scan a number of the most well-known AV vendors are not found. Knowing beforehand that this application is not malicious and having products detecting it as one can clearly cause problems with small developers. On the other hand, on both dates, AV products such as Avast, AVG, Kaspersky, Microsoft, NOD32, and Sophos have not raised an alert for this specific file. This can be either due to the intelligence of their detection engine or due to the lack of classifications that would pick up the application as a threat. From a developer's perspective this scan result can cause serious problems in the deployment of the application and in this particular case there is almost nothing the developer could effectively do differently. On the other hand, the large number of false positive alerts raised for such a simple application brings more questions to the table. For that reasons, more applications were developed and more tests were carried out in order to have a better idea of the matter at hand.

3.2 Application II: OnlineFileProperties.exe

In this example, three out of forty-two results, as outlined in Table 2, classified the application OnlineFileProperties.exe as threat. More specifically, this is a simple application which comes with a GUI. It accepts as input a URL address, e.g. http://www.iana.org/, and performs a 'get' request to the server, which returns the HTTP HEADER information of the reply. Thus, given a valid URL will return status '200 OK' and the relevant information contained in the response packet is displayed. If a URL points to a file e.g.

http://www.iana.org/favicon.ico, all relevant information of that file contained in the HTTP HEADER are displayed.

Once again the application does not execute anything in the background, it does not look for write or read access to the hard drive and it does not contain any malicious code. It only requires Internet or network access in order to request data from the provided web server.

Table 2: Analysis results for OnlineFileProperties.exe on two different dates

File name:	OnlineFileProperties.exe
File type:	Win32 EXE
File size:	36.0 KB (36864 bytes)
SHA256:	bbef0d99e39a844191d14f6c54d3228ca4c2df9d09c329d24ea42592ce9af7c4
SHA1:	630f0adab6376b4985c35282efd446561ecd44da
MD5:	c9e75ef93de9910b42a15120102f1c29
Analysis date 1:	2012-07-12 08:11:22 UTC
Detection ratio:	*3 / 42*
BitDefender	Gen:Trojan.Heur.VB.cm0@d0TVztdi
F-Secure	Gen:Trojan.Heur.VB.cm0@d0TVztdi
GData	Gen:Trojan.Heur.VB.cm0@d0TVztdi
Analysis date 2:	2013-02-01 00:43:06 UTC
Detection ratio:	*2 / 46*
AntiVir	TR/Spy.36864.1241
Comodo	UnclassifiedMalware

In this case, as Table 2 illustrates, interestingly enough the number of alerts decreased over time. The classification of the software has changed over time from being a generic Trojan to being spyware. In each analysis date, the alerts raised are given by completely different AV products. It should also be noted that none of the applications have been changed in any way. Thus, it is interesting to see this piece of software, which was initially classified as malicious, having its classification changed over a six month period. Despite the fact the application is being classified as malicious, it should be noted that AV products such as Avast, AVG, Kaspersky, McAfee, Microsoft, NOD32, Panda, Sophos, and Symantec, amongst others, did not raise any alerts on both analysis dates. Therefore, from a developer's perspective, these few alerts can cause some problems to the developers and their users who make use of the particular AV software.

3.3 Application III: wget.exe

In this particular example, three out of forty results, Table 3, classified the application wget.exe as a threat. More specifically, this is a Windows implementation of the 'wget' command line tool under Linux created from scratch by the authors. More specifically, given a web address of a file as an argument, the tool downloads that file to your system. The application does not execute anything in the background, it writes to the hard drive the file requested into the working directory and it does not contain any malicious code. It also requires Internet or network access in order to request data from the web server.

Table 3: Analysis results for wget.exe on two different dates.

File name:	wget.exe
File type:	Win32 EXE
File size:	7.0 KB (7168 bytes)
SHA256:	f55ac16a01c9e8ab10c904d12f20f2d2958eb935b264bc2abfc7e9ec19e5c012
SHA1:	30bc522efacd3bf5fa37358ed39973c4fe0e0956
MD5:	7016f0d536593b4f0a560f1a08e523c5
Analysis date 1:	2012-07-12 08:38:42 UTC
Detection ratio:	*3 / 40*
Emsisoft	Trojan-Spy.Win32.VB!IK
Ikarus	Trojan-Spy.Win32.VB
TheHacker	Posible_Worm32
Analysis date 2:	2013-02-01 00:46:54 UTC
Detection ratio:	*3 / 46*

File name:	wget.exe
Ikarus	Trojan-Spy.Win32.VB
SUPERAntiSpyware	Trojan.Agent/Gen-Koobface[Bonkers]
TheHacker	Posible_Worm32

Similarly to the previous examples, alerts were raised for this application as well. In this case, as illustrated in Table 3, the total number of alerts in each analysis date remained the same. The only difference here is that Emsisoft is no longer considering this application as a Trojan, whereas SUPERAntiSpyware does.

3.4 Application IV: KeepMeIn.exe

KeepMeIn.exe does not require Internet or network access to work. It runs locally and it only performs mouse events in regular intervals, in order to stop the screen-saver from running. It does not execute anything else in the background, it does not need write or read access to the hard drive and it does not contain any malicious code. One out of forty-two results classified the application as threat.

Table 4: Analysis results for KeepMeIn.exe on two different dates.

File name:	KeepMeIn.exe
File type:	Win32 EXE
File size:	6.0 KB (6144 bytes)
SHA256:	17fe25e856a706ec6c31368131f04333b08497125c4646d841578e18d309054b
SHA1:	7eb9db1e297b613ed5aa54c8bd7a6fa16d7ebb8d
MD5:	77e6d7a1c8b63efe7c0e22bc20e5fa61
Analysis date 1:	2012-07-12 08:47:38 UTC
Detection ratio:	*1 / 42*
TheHacker	Posible_Worm32
Analysis date 2:	2013-02-01 00:53:39 UTC
Detection ratio:	*2 / 45*
SUPERAntiSpyware	Trojan.Agent/Gen-Koobface[Bonkers]
TheHacker	Posible_Worm32

By observing all the tables so far it is evident that the AV called 'TheHacker' is present in almost every scan performed in all four applications. A real concerning fact is that the tests have shown that TheHacker has classified everything as "Posible_Worm32" and SuperAntiSpyware has classified everything as "Trojan.Agent/Gen-Koobface[Bonkers]". It should also be noted that during all the tests conducted in the first analysis date, in the tables above, SuperAntiSpyware had not been included yet as part of the Virus Total scanning system. Nonetheless, their continuous false positive alerts start to stand out from the rest of the AV products, raising serious concerns about the application's and developer's reputation.

4. Creating and testing the same executable file with a different payload each time

In order to take this false positive issue a step further, it was considered essential to add one more test and discuss the results produced. In this case, an application was created and scanned having a different payload attached to it each time. The nature of the application was a dropper, a piece of software capable of dropping its payload as a file into the target system. In the first instance, it drops the well-known SQL worm known as Slammer. In the second instance, instead of the Slammer worm, it carries and drops a simple batch (.bat) file. The .bat file had the following four simple and harmless commands, as shown in Table 5.

Table 5: Payload which contains four simple commands in a batch file

```
1. @echo off
2. rem 1
3. dir
4. pause
```

Also, in the last instance it had no payload at all, thus it dropped nothing into the system hosting the file. The following table outlines the findings of each instance of the application while carrying a different payload each time.

Table 6: The findings of each instance of the payload carrying software

SlammerCarrier.exe	BatCarrier.exe	NullCarrier.exe
MD5: b132c1b80bfeb1114940a63774d1a902	MD5: 33f8204b22843eae4fc455305cc93cfc	MD5: 10b6937eb674d2694eea8f137ed7568e
Detection ratio: 7 / 46	*Detection ratio: 2 / 46*	*Detection ratio: 1 / 46*
Analysis date: 2013-02-01 01:01:33 UTC	Analysis date: 2013-02-01 01:05:57 UTC	Analysis date: 2013-02-01 01:17:08 UTC
BitDefender Dropped:Worm.Sql.Slammer.Dump.A	-	-
Emsisoft Dropped:Worm.Sql.Slammer.Dump.A (B)	-	-
F-Secure Dropped:Worm.Sql.Slammer.Dump.A	-	-
GData Dropped:Worm.Sql.Slammer.Dump.A	-	-
?	McAfee-GW-Edition Heuristic.BehavesLike.Win32.Suspicious	-
MicroWorld-eScan Dropped:Worm.Sql.Slammer.Dump.A	-	-
nProtect Dropped:Worm.Sql.Slammer.Dump.A	-	-
TotalDefense Malicious	TotalDefense Malicious	TotalDefense malicious

There are some very interesting points to be made based on the findings illustrated in Table 6. Even though the application drops the actual Slammer worm, it is impressive to see that antivirus software from the most well-known and major vendors did not identify it.

Seven out of forty-six different antivirus products flagged it as malicious, which is a very low detection rate to be ignored. When the payload is changed to a simple .bat file the results are completely altered. Instead of being picked up as a dropper, it is being ignored by the antivirus products. McAfee-GW-Edition identified is as suspicious even though it is actually carrying a harmless batch file. Effectively, enhancing the source code a little bit to hide the slammer payload would have avoided detection. This is b

eing verified by the third case where no payloads were included. In this case, only one antivirus product flags the software as malicious and it is not one originating from the leading vendors, as it would be expected.

To further support our findings, the Flash Exploit CVE-2013-0633 (NIST, 2013) is an excellent example that drops a number of malicious files and it was scanned using Virus Total. The detection results are outlined in Table 7.

Table 7: The findings of each instance of the payload carrying software

IEEE2013.doc-54322.exe.virus	Update	IEEE2013.doc-301630.exe.virus
MD5: bd19e2f953096a251a3b0f6744cbe7de	MD5: 432dce23d00694b103dd838144253d1b	MD5: 0bb90855eba25441ab3bd2c6b4cf0dec
Analysis date: 2013-02-08 19:30:16 UTC	Analysis date: 2013-02-10 13:35:56 UTC	Analysis date: 2013-02-08 19:24:31 UTC
Detection ratio: 3 / 46	*Detection ratio: 27 / 45*	*Detection ratio: 2 / 46*
AVG Suspicion: unknown virus	Agnitum: Trojan.Delf!6G4SenmRNIM	ESET-NOD32: Win64/TrojanDropper.Agent.U
ESET-NOD32: Win32/TrojanDropper.Agent.QAU	AhnLab-V3: Trojan/Win32.Delf	VIPRE: Corrupted File (v)

IEEE2013.doc-54322.exe.virus	Update	IEEE2013.doc-301630.exe.virus
VIPRE: Corrupted File (v)	AntiVir: ADWARE/Adware.Gen	
	Avast: Win32:Malware-gen	
	AVG: Dropper.Agent.BAKZ	
	BitDefender: Trojan.Agent.AYAF	
	Comodo: UnclassifiedMalware	
	ESET-NOD32: Win32/Plugax.B	
	F-Secure: Trojan.Agent.AYAF	
	Fortinet: W32/Delf.B!tr	
	GData: Trojan.Agent.AYAF	
	Ikarus: Trojan.Win32.Bredolab	
	Kaspersky: Trojan.Win32.Delf.deey	
	Malwarebytes: Backdoor.Poison	
	McAfee: BackDoor-AKU	
	McAfee-GW-Edition: Artemis!432DCE23D006	
	MicroWorld-eScan: Trojan.Agent.AYAF	
	NANO-Antivirus: Trojan.Win32.Poison.bfqxth	
	Norman: Killav.LB	
	nProtect: Trojan.Agent.AYAF	
	Panda: Trj/CI.A	
	PCTools: Backdoor.Boda	
	Symantec: Backdoor.Boda	
	TrendMicro: BKDR_PLUGAX.A	
	TrendMicro-HouseCall: BKDR_PLUGAX.A	
	VIPRE: Trojan.Win32.Generic!BT	
	ViRobot: Trojan.Win32.A.Delf.209852	

In the first column of the table it is interesting to notice that only three out of forty-six AV products eventually detected the threat; while in the third column only two out of forty-six AV products. In the second column of Table 7 there is an impressive increase in the detection rate, which is twenty-seven out of forty-five. Even though the number of alerts in this occasion is much higher, it is worth noticing that it is still only the 60% of the total AV scanners available.

5. Hiding in plain sight

We ran another test by creating an application larger in file size, composed of a large number of dummy functions in many lines of code. This application was programmed to access a few web-pages and return some info after processing the response packets. It also has the ability to execute any commands requested remotely on the client's side. This behaviour should be classified as malicious.

However, the application does nothing suspicious and therefore all antivirus vendors classified it as such (Table 8). There is nothing obfuscated or encrypted in the source code, other than it is composed of many lines of code.

Consider the following example of an application being able to access a number of weather websites and return temperature values from a number of predefined locations. In addition, it makes use of tweets posted from different online weather providers. As long as this application is available online, and used by a number of users posing no threat, it builds up a good reputation (Symantec 2013). In effect, an attacker could reply to a tweet on one of the twitter accounts of the weather information providers. She could either tweet some scrambled text or provide a link to an external webpage. As a result, the malicious application reads that tweet and from that point onwards it is activated to perform any commands given, especially after it has gained a good reputation. A similar example is malware such as Stuxnet and Flame that were spreading for years until formally identified in May 2012.

Table 8: The findings of each instance of the payload carrying software

File name:	WeatherDisplayApp.exe
File type:	Win32 EXE
File size:	264.0 KB (270336 bytes)
SHA256:	55ff027901d6373de29be1f5fb86af3c17f9fb84874f8082115389a486719756054b
SHA1:	b45fa3b239e990c3bf0ba02f407e830a92e4175e
MD5:	af2ca3b0de83e52a7ab97d6ac44bc516
Analysis date 1:	2013-02-01 14:04:13 UTC
Detection ratio:	*0 / 46*

The point that needs to be raised here is that it generally seems that AV vendors have a generic rule that says "*if it is small and weird, even though you don't have proof, classify it as threat*". On the other hand, there are files which are larger than the usual tiny malicious applications. The antivirus looks at these files and identifies a large number of operations that need to be executed by a large number of function calls. In that case, the AV seems to overlook the behaviour, as there is no need to look into it in detail due to the fact that it doesn't have a known signature. The general rule seems to suggest "*if it is large file and I don't really know what it does, just classify it as legitimate and time will show*". However, this hypothetical AV engine behaviour has a serious effect on people who just want to make legitimate applications for everyday use.

6. Conclusion

Looking closer to the AV false positive alerts, it is thoroughly discussed that the more functions an application has as noise, the more difficult it is to detect its malicious intent. Even if an application appears to perform a legitimate task and gains the appropriate reputation, it is probably impossible to detect its malicious intent.

Recent studies show that 140 new mobile malware threats are introduced every day and in general, threats increase by 125,000 per day which translates to 67 million unique threats (Kaspersky Lab, 2013). As reported by Securelist (2012) for the first quarter of 2012, around 81% of all vulnerabilities target Java and Adobe Acrobat Reader. According to the Advanced Threat Report (Fireeye, 2013), malware tries to invade technology companies every 60 seconds. There is a vast amount of threats to deal with, but in some cases the situation could have been improved if AV vendors decided to collaborate. This collaboration could lead towards minimizing the number of false positive alerts. The noise in AV detection could be eliminated and an in-depth real-time analysis of potential threats could become a primary focus. Effectively, having software to go through a screening process will eventually filter out and classify malicious or potentially malicious applications.

The research towards this unified solution of assessing threats using different means, such as a source code submission system, could lead into identifying unknown threats significantly. Not only because of the large number of signatures that will be collected and shared, but also due to the knowledge and intelligence that will be gained from automating the process of source code validation. Having such a system in place could provide a valuable point of reference in software development. This automated process would certify a particular piece of software with a malicious intent or not. Nevertheless, such a system should also respect the developer's copyrights by ensuring that the source code is kept safe while being assessed and not passed on to third parties.

Acknowledgements

The authors would like to thank Mr. Dimitrios Sarris for his kind contribution to the paper. Sharing his thoughts and the industry's concerns were a valuable input for the overall paper especially with his expertise in the Middle East area. Dimitrios Sarris is currently a senior security specialist at the National Bank of Kuwait.

References

Abrams R., Pathak J., Barrera O., 2013, Consumer AV/EPP Comparative Analysis - Phishing Protection, NSS Labs, Available at: http://bit.ly/12c6DSl [Last accessed: 5 Feb 2013]

Avast! Online Scanner, 2013, AVAST Software, Available at: http://onlinescan.avast.com [Last accessed: 14 Jan 2013]

Avira, Online Scanner, 2013, Avira, Availabe at: https://analysis.avira.com/ [Last accessed: 14 Jan 2013]

Bayer, U., Engin K. and Christopher K., 2010, Improving the efficiency of dynamic malware analysis. Proceedings of the 2010 ACM Symposium on Applied Computing 22 Mar. 2010: 1871-1878, Boston, Massachusetts, USA.

Bishop, P., Bloomfield, R.; Gashi, I.; Stankovic, V., Diversity for Security: A Study with Off-the-Shelf AntiVirus Engines. Software Reliability Engineering (ISSRE), 2011 IEEE 22nd International Symposium on 29 Nov. 2011: 11-19.

Bit Defender, Online Scanner, 2013, Bitdefender Security, Available at: http://www.bitdefender.co.uk/scanner/online/free.html [Last accessed: 14 Jan 2013]

Christodorescu, M., & Jha, S., 2003, In Proceedings of the 12th USENIX Security Symposium, Washington, D.C., USA

Christodorescu, M. and Jha S., 2004, Testing malware detectors. ACM SIGSOFT Software Engineering Notes 29.4 : 34-44.

Graham, L. D., 1999. Legal battles that shaped the computer industry. p. 175. Greenwood Publishing Group. p. 175. ISBN 978-1-56720-178-9.

Gu, G., Porras, P., Yegneswaran, V., Fong, M., & Lee, W., 2007. Bothunter: Detecting malware infection through ids-driven dialog correlation. Proceedings of 16th USENIX Security Symposium on USENIX Security Symposium. USENIX Association.

Devine, C., & Richaud, N., 2009, A study of anti-virus' response to unknown threats. In Proceedings of EICAR (Vol. 9)

FireEye, 2013, Advanced Threat Report – 2H 2012, FireEye, Available at: http://bit.ly/10zjaKi

Jaatun, M.G., Jensen, J., Vegge, H., Halvorsen, F.M., Nergard, R.W., 2009, Fools Download Where Angels Fear to Tread, Security & Privacy, IEEE , vol.7, no.2, pp.83-86

Kelchner, Tom, 2010, The (in) consistent naming of malcode. Computer Fraud & Security 2010.2: 5-7.

Kaspersky Lab, 2013, Why Complexity is IT Security's Worst Enemy, A whitepaper analysing how complexity is causing new security challenges, and how best to address this, SC Magazine 8 February 2013, [online] http://bit.ly/12lvEWX

Ketari, L., & Khanum, M. A., 2012, A Review of Malicious Code Detection Techniques for Mobile Devices, International Journal of Computer Theory and Engineering Vol. 4, No. 2http://bit.ly/12lvEWX

Laurent, A. S., 2004, Understanding open source and free software licensing. O'Reilly Media, Incorporated.

Lillard V.T., 2010, Digital Forensics for Network, Internet, and Cloud Computing: A Forensic Evidence Guide for Moving Targets and Data, Elsevier Inc., ISBN 978-1-59749-537-0

Liu, S., & Ormaner, J., 2009, From Ancient Fortress to Modern Cyberdefense.IT professional, 11(3), 22-29.

McMillan R., 2012, Is Antivirus Software a Waste of Money?, Wired Magazine, Available at: http://bit.ly/X91K8b

Metascan Online, 2013, OPSWAT, Available at: https://www.metascan-online.com/en [Last accessed: 14 Jan 2013]

Microsoft, 2012, Microsoft Security Intelligence Report, SIRv14, Volume 14, July through December, 2012

Muncaster P., 2013, Symantec: Don't blame us for New York Times hack, The Register, Available at: http://bit.ly/14untL1 [Last Accessed: 10 Feb 2013]

NIST, 2013, Vulnerability Summary for CVE-2013-0633, US-CERT/NIST, Available at: http://web.nvd.nist.gov/view/vuln/detail?vulnId=CVE-2013-0633, [Last Accessed: 10 Feb 2013]

Hawes J., 2013, How do you know if an anti-virus test is any good?, Naked Security Sophos 2013, Available at: http://bit.ly/15r9jxD [Last accessed: 21 Apr 2013]

Oliver S., Google Android store reaches 25 billion downloads, 675,000 apps, Available at : http://bit.ly/Sz95Oz [Last accessed: 31 Jan 2013]

Prince B., 2013, Massive Android Botnet Built on Backscript Trojan, Security Week, Available at: http://bit.ly/WR0WTL [Last accessed: 31 Jan 2013]

Raggo, M. T., & Hosmer, C., 2012, Data Hiding: Exposing Concealed Data in Multimedia, Operating Systems, Mobile Devices and Network Protocols. Syngress, ISBN: 9781597497411.

Schneider, S., & Griffin, K. (2011). U.S. Patent No. 8,028,338. Washington, DC: U.S. Patent and Trademark Office.

Securelist, 2013, Kaspersky Security Bulletin. Statistics 2011, Available at: http://bit.ly/zZCTM v [Last accessed: 10 Feb 2013]

Symantec, 2013, Turning the table on Malware; A comprehensive approach to unique and targeted attacks, Symantec, Available at: http://bit.ly/yLaeG8

Threat Expert, 2013, Threat Expert Free Online File Scanner , Available at: http://www.threatexpert.com/filescan.aspx [Last Accessed: 1 Feb 2013]

Virus Total, 2013 , Available at : https://www.virustotal.com/ [Last accessed: 11 Jan 2013]

Wang, W., 2006. Steal this Computer Book 4.0: What They Won't Tell You about the Internet. No Starch Press.

Leaks by a Thousand Clicks: Examining the Growing Risk of Mobile Working

Wendy Goucher and Karen Renaud
School of Computing Science, University of Glasgow, UK
wendy@idrach.com
Karen.renaud@glasgow.ac.uk

Abstract: Was there ever a 'Golden Age' when it was possible to construct a fortress for protecting organisational information? Probably not, but if there was the era is gone. Recent developments in mobile computing, and the availability of high speed WiFi and cellular connections, means it is now possible to take the entire office functionality outside protected spaces and to work on material in plain sight in public places. This creates a possibility for data leakage that is not yet fully appreciated. Data now moves across organisational boundaries stored on mobile devices, USB keys, in emails or on paper. Security efforts often focus on protecting against data loss, but an area which is often neglected is that of data being obtained surreptitiously without any trace of the acquisition remaining. This can happen during transfer or via an endpoint while the data is being worked on: often called 'shoulder surfing'. There is plenty of anecdotal evidence of information being overheard or observed in public. So far, hard proof of significant business loss has remained elusive. The recent rapid increase in the uptake of photographically-enabled smart phones has exacerbated this risk. The casual sounding term 'Shoulder Surfing' is slowly changing to 'Visual Data Capture' (VDC) to reflect the fact that the data is no longer merely being observed or overheard but is now being captured in photographic format resulting in information capture. However, it is still challenging to quantify the degree of risk related to VDC as data is not removed or altered by the observer during acquisition. In this paper we report on the results of two experiments which set out to investigate the vulnerability of data on laptops and tablet devices to VDC. We address both capability and likelihood (probability) of such acquisition. The results deliver insight into the size of the VDC risk and suggests possible approaches towards mitigating it.

Keywords: shoulder surfing, mobile computing, mobile devices, information security, visual data capture, sensitive data risk, threat

1. Introduction

Many 21st century employees work outside of the formal office environment for a significant part of their working day. A report commissioned on behalf of the Chartered Society of Physiotherapists [Honan 2012] reported that 65% continued to work on Smart Phones or other mobile devices once outside the office; working for an average of 2 hours 34 minutes in this way. This is made possible by the increasing capability of these devices and means that previously "dead" time spent travelling or waiting can now be utilised, enhancing employee productivity and maintaining the employee's focus on business activities. On the other hand, there is a concern that workers will be observed while working on confidential documents, and leakage thereof could harm the organisation. Another interesting trend is that employees are increasingly purchasing and maintaining their own devices, a phenomenon recognised by the increasing use of the term 'Bring Your Own Device' or 'BYOD'. While ownership is unlikely to have an impact on the employee's likelihood of working on sensitive information in public places, it *is* possible that it will influence exactly what actions organisations can take to reduce the risk. For example, it is harder for an organisation to enforce the use of discretion screens on mobile devices that are privately owned. Both BYOD and mobile working make businesses uneasy, because both are emergent trends with as yet to be discovered propensities, both individually and in tandem. Moreover, increasing numbers of mobile devices incorporate powerful cameras. This means confidential information is no longer merely observed, but very likely to be captured too: *visual data capture* (VDC). At this stage no one knows the magnitude of the risk.

A comparison of Figure 1 (taken in 2008) to Figure 2 (taken in 2012) is enlightening. In 2008 those few who took pictures used cameras (indicated with a square). By 2012 the majority were using smart phones (marked with a ring). While the odd camera appears, these are clearly outnumbered by smart phones

What this and other comparisons suggest is that the emergent response to events of interest is to record them by photo capture. If this is the case, it increases the possibility that someone might photograph an interesting document or screen when they see it. With the quality of the camera on these devices now a key selling point (e.g. rumours of a new iPhone in 2013 having a 12 megapixel sensor incorporated in the main camera), any risk is likely to increase with the increasing likelihood that such a captured image will be legible.

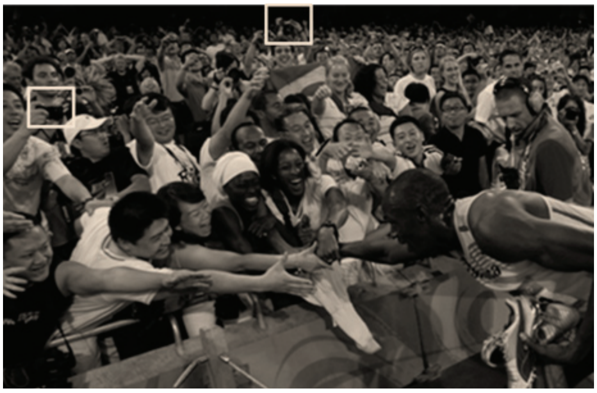


Figure 1: Usain Bolt being congratulated by the crowd at the 2008 Beijing Olympics. Copyright information: REUTERS/Hans Deryk courtesy of alertnet.org

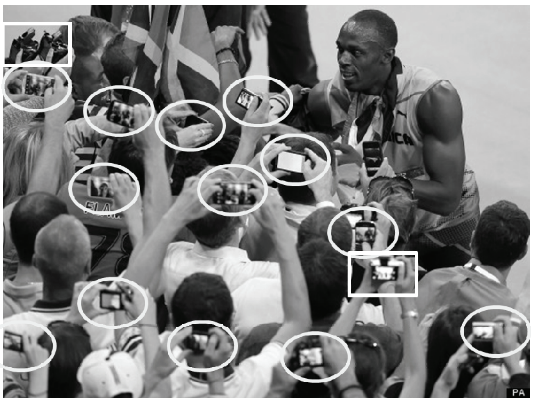


Figure 2: Usain Bolt being congratulated by the crowd at the 2012 Beijing Olympics. Photo from http://www.huffingtonpost.co.uk/2012/08/09/usain-bolt-200m-gold-medal-london-2012_n_1760979.html

It is also important to consider the increasing trend to "work on the move", which seems to be gathering pace, leaving organisations in an anomalous position. They reap the benefits of this working approach, but have no way of assessing the threat that casual observation constitutes. The size of the risk, and effective mechanisms for addressing it, are yet to be discovered.

A series of experiments were designed to give some indication of the size of the vulnerability these interacting factors constitute. This paper reports on two studies to assess:

- *Vulnerability*: examine the visibility of text on photos taken by current smart phone cameras.
- *Probability*: conduct a survey of mobile device owners, to determine how and when they used the camera facility on their mobile devices. We wanted to determine the extent to which people took photos to store information for later use, as opposed to photos of people or places. This tendency would conceivably indicate how likely people would be to record casually observed, yet potentially interesting, information.

2. Background

The vulnerability of documents to VDC is driven by a rise in mobile working. A number of surveys and questionnaires, including one by Good Technology, quantify that risk. They found 93% of respondents reported that they continued to work outside of the office, with 38% believing that their job would be impossible without at least mobile access to Email.

A white paper commissioned by Secure (Honan 2012) reported that 85% admitted overlooking sensitive information that they were not authorised to see. It could be regarded as significant that such a high proportion of people were prepared to admit not just to looking, but to be paying enough attention to be able to categorise what they saw as 'sensitive'.

Okenyi and Owen (2007) looked at various impacts to business of a variety of potential attacks including shoulder surfing, which is of interest here. The major consequences of these events are:

- Loss of public confidence;
- Loss of share value;
- Bad publicity;
- Possible legal proceedings;
- Fines from regulatory authorities; and
- Increased interest and monitoring from regulatory bodies.

This list should also include the cost of corrective action both as a result of internal investigations and investigatory action by the regulatory body. It should be noted that a VDC-enabled leak will make it harder to ascertain how and when the information leaked. This will lead to a more wide-ranging, time-consuming and expensive. Unless the source leak event can be definitively identified, corrective action may be more widespread and expensive than it needs to be. For example, the loss of an unencrypted laptop containing sensitive information has a clear antecedent and obvious consequences. A VDC-enabled leak is harder to demarcate. It is clear that businesses should indeed be concerned about this possibility if their staff work on the move.

2.1 Examples of VDC

In one example, cited by Secure, The European Association for Visual Data Security reports the Vice President of an S & P 500 company took the time during her flight from London to New York to work on her company's profit forecast for the following 6 months. Soon after she landed a newspaper that was going to run a "Splash" on her forecast in the next day's edition phoned her. The leak had come from the person in the neighbouring seat who happened to be a journalist. She had had plenty of time to appraise the data herself and contact her paper with an analysis immediately upon landing.

In November 2008 a civil servant in the Department of Business fell asleep on the train while working on his laptop on documents marked with the security level 'Restricted'. This event was captured photographically by a fellow passenger and led to a story in the Daily Mail newspaper (Owen 2008)

2.2 PIN and password VDC

Much of the existing academic work is focused on the risk of the observation and capturing of PINs and passwords wherever they can be entered. Although there are essential differences, some core relevant findings apply equally to the VDC threat. However, two factors are often lacking in VDC:

- An identifiable loss that can be traced to a specific time, and possibly an actual event, that might have led to the information being leaked.
- The ability to identify and track changes in loss rates after mitigation techniques are deployed.
- Mitigation efforts often being myopic and having unanticipated consequences which potentially negate any positive impact. With investigation into quantifiable financial loss these consequences are more likely to become apparent.

The perceived likelihood of observation is critical in all research of shoulder surfing, including VDC. In the case of passwords entry, the user will implicitly assess the probability of observation. They can then decide whether to amend their behaviour to, for example, shade their keyboard input from observation. However, there is also a social aspect to this decision. Anderson (2006) explained that attempts to input authentication data covertly can be seen as a lack of trust and therefore suggest anti-social tendencies. This means that the user would need to choose between exposing their PIN or shielding it and thereby clearly express a lack of trust in those around them. Renaud and Maguire (2009) examined the use of passwords to protect game or film purchase in a social setting and found users did not want to be seen to distrust their friends and family by covertly inputting their PIN or password in a family situation.

2.3 Bring your own device

When discussing the topic of the security of staff using mobile devices to work on-the-move the evolving 'Bring Your Own Device' (BYOD) tendency becomes pertinent. BYOD is a concern to business as its proliferation means that 'alien' devices, outwith the control of the organisation, are connecting to corporate systems. While there is no evidence to suggest that people use devices differently, depending on ownership, it could be argued that the increase in device numbers due to the BYOD approach is germane to the scale of the VDC problem. Where the ownership is likely to be of greatest relevance is in the consequent lack of control of data as discussed by Miller (2012) who expressed concern that once data is on personal devices it can be accessed and stored anywhere, including, potentially on 'free' clouds such as Dropbox or Google which are outwith the organisation's control.

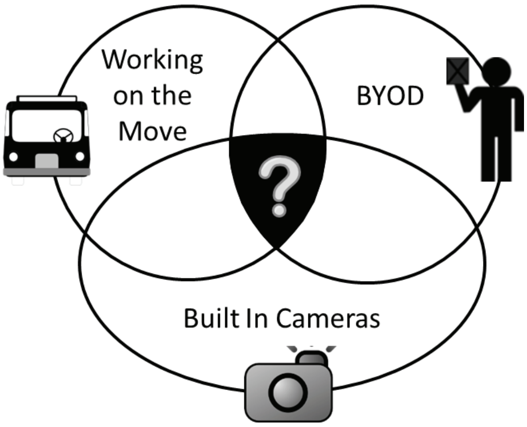


Figure 3: Vulnerability on the move

As we can see in [Fig 3], when the contributory factors of: organisational data being worked on in public, an increase in devices with ever improving screen quality and the increasing quality and ubiquitous use of smart phone cameras are combined, the existence of risk becomes clear. The extent of this risk is, as yet, unknown. However, investigations that bring greater insight into the nature of the risk are required.

2.4 Visual data capture

Shoulder-surfing has quickly morphed from casual annoyance to potential threat. This means of data leakage has been expedited by the recent proliferation of photo enabled smart phones and tablet devices with increasing levels of clarity. When people are working on the move they will probably use the most convenient device. Tablet devices are easier to use than a laptop in a confined space, such as might be available to a commuter, and so are perhaps more likely to be used. Software is also emerging to facilitate capture of documents using mobile devices (see Fig 9).

Even ordinary citizens have engaged in utilising the photo capture techniques in vital exercises such as the Nigerian elections of 2011. Volunteers used their cameras to monitor the process of the election in the hope of helping to make the democratic process fairer and safer

Figure 4: Recording the electoral process in Nigeria. http://www.fordfoundation.org/about-us/2010-annual-report/engaging-citizens-to-strengthen-democracy#volunteers-forreclaimnaija-post-election-information

Academic publications in this area are sparse, which confirms the newly emergent nature of the threat. Investigations are often promoted or sponsored by those organisations having a vested business interest in addressing the problem and benefitting from its remediation. These investigations can still be helpful, but it is sensible to consider the reports in their context and not to accept their findings as gospel.

Honan (2012) sets out the core challenge to this area of study when he says: "*Organisations will often never know if they have suffered a visual data security breach*". Unless the information owner is aware of the leakage there is no reason why they would be aware of its capture by visual means. 3M, a leading provider of privacy screens for mobile devices, emphasize the growing nature of this threat: "*The ability to inconspicuously capture information viewed on screens with camera phones has increased substantially*". Oculis are developers of a software privacy screen and have found that current pressure from HIPAA requirements has led to a growth in interest in this area. They commissioned a survey in the US in which 89% of respondents admitted to reading over someone else's shoulder. There is no suggestion that all of such events applied to sensitive corporate or personal data. However, in a study by Thomson [Thomson 2010] 67% of "Working Professionals" admitted to working on sensitive material outside of a secure office environment.

Evidence supporting the idea that people are increasingly working on the move is important, but is only one element in the 'black hole approach' which seeks to establish the likelihood of significant of data loss. It does not establish whether this information could, or might, be captured and (mis)used. Further research into the possibility of capture of passwords by video in public was carried out by Maggi (2011). They carried out experiments to develop software that processed the video capture images from a user's iPhone. This device was targeted because when the keys pressed on the touchscreen keyboard, they highlight. This has been done so that the user, unable to get the accustomed tactile feedback, is assured that the device had recorded the keystroke. In the experiment the authentication was 'shoulder surfed' and software developed to interpret the captured image and highlighted feedback. In both laboratory and 'real world' conditions the accuracy of the capture was in excess of 92%. The only key that was not magnified, and so was slightly harder to detect, was the space bar. However that key does change colour so in most cases was also detected. The conclusion of the study was that keyboards where the activated key was magnified were unsuitable for authentication and other sensitive input.

With the iPad the activated key is not magnified, but rather is coloured so this will also be detectable. A further study by the Haroon and the Thinkst Applied Research Institute attempted the same capture of passwords and also found it a straightforward procedure when the capturing was from an iPhone. As authentication actions are transitory in comparison with the exposure of a document being worked on it should be even easier to capture an image of a working document.

Earlier in this research project, reports of incidents of data leakage, both business and personal, from mobile working were collected. Some of these examples were notable for the potential for significant repercussions if such data came into the possession of the 'wrong' person, be they thief or business rival (Goucher 2011).

A new development makes the preservation of security for the user even harder: the ability of the smartphone to capture clear images off the screens via the inbuilt camera. This is a potential catalyst for leakage of sensitive information.

Fig 5 depicts the contributory factors for a significant data leak to occur. These include the fact that the observer has to be both able and motivated to make the effort to record the key details and then have a means to use the information themselves or to forward it to someone who can. Many modern cities find that industries such as finance and law are geographically grouped and so rival businesses' employees may share the same public areas such as trains and cafés and be in a position to oversee sensitive information.

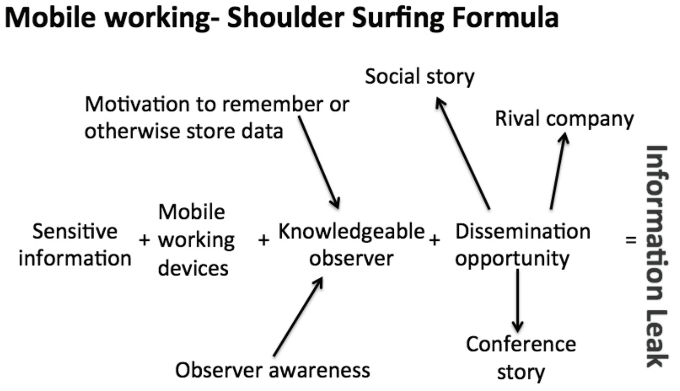


Figure 5: Causatives leading to Information leak

Recent security concerns by US and UK governments led to a drive, reinforced by sanction, to improve the protection of sensitive data. The use of privacy screen covers has been less widespread than might have been expected. In a recent survey conducted for Secure; The European Association for Visual Data Security, only 56% of respondents said their organisations had procedural measures in place to prevent visual data security breaches. This was particularly surprising since 98% believed it was important to educate individuals on the visual data security threat and 32.4% said they had no confidence that users would make the effort to prevent data from being observed when working in public places. When aligned with the findings of research by Herbert H Thompson (2010), which found that 67% of respondents admitting to working on sensitive documents in public in the previous 12 months, there is even more reason for businesses to be concerned.

As Figure 6 demonstrates, the route to potential damage is shorter and carries more weight when a digital image of the data is captured. In order to determine whether photo capture of devices should be considered significant vector for leakage, two aspects should be explored.

- Firstly, the visibility of such screen to smart phone cameras (vulnerability).
- Secondly, the tendency to use smart phone cameras to record information rather than just visually pleasing images (probability).

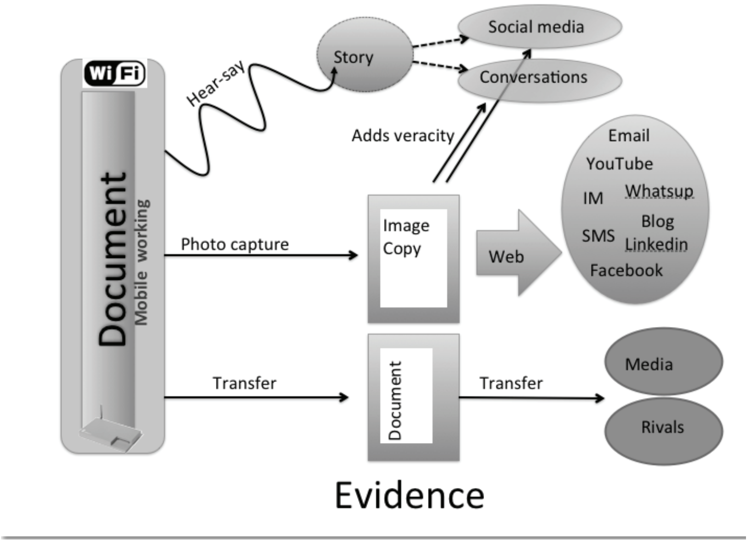


Figure 6: The VDC process

3. Feasibility experiment

The inherent, and inescapable problem with the investigation of the risk of Visual Data Capture is that the loss itself may never be detected. Even if evidence of the leak materialises the method by which the data was leaked may be elusive. If the data refers to a significantly high profile subject, it is likely that it would be made public in some way, but otherwise the path is hard to uncover. The original document remains unaltered by the capture and information can be used without the source organisation even becoming aware of the leak. This lack of evidence is the core challenge for this research overall and has resulted in the adoption of a methodological principle developed from the basic astronomical approach regarding the identification of 'Black Holes'. Black holes cannot be directly observed with current technologies; therefore astronomers observe celestial bodies moving in ways that suggest the presence of an invisible obstruction. By recording these movements scientists can calculate the position, dimensions and character of the Black Hole. This approach is deployed here to identify those attitudes, behaviours, policies and procedures that indicate the presence of a significant risk.

3.1 Methodology

The first step was to establish whether the cameras in current smart phones (capturing device) have the capability to capture textual images of sufficient quality from mobile devices (source device). If they can, they it will be possible to make the information viewable and even reproducible, and therefore potentially useful to an observer.

This was a small scale experiment that tested 3 source devices:

- A laptop – Macbook Pro,
- An HFC flyer – android tablet device
- An iPad 2

The photographic capability of five capturing devices:

- HTC7 Pro
- iPhone 3GS
- iPhone 4
- Blackberry Pearl
- HTC Wildfire

These devices were selected due to their availability and ubiquity in the marketplace.

A layout was devised to resemble a public environment where a user would be sitting at a table using a source device. Capture positions were identified both in sitting and standing positions behind the source device.

3.2 Stage one

A pilot experiment was carried out using an iPhone 4 and a digital camera capturing a screen image from a laptop and an iPad. All possible positions were mapped and the screen captured.

This resulted in ninety two potential capturing positions, but some of these were clearly not going to constitute a risk. Positions which provided either the most common VDC positions were identified but not tested as they provided reliable images so did not require further investigation e.g. from the seat next to the user and these were used. Positions where visibility was so poor as to render any capture unfruitful were eliminated. The remaining positions were those which were on the margin so where either positions not expected to provide a clear image or positions where the clarity was marginal depending on the devices. Consequently the number of required captures was reduced to twenty.

3.3 Stage two

The experiment was conducted in a room with a stable artificial light source, ensuring that identical conditions applied to all captures.

Chairs were placed as in the design [Figure 7] and those chairs not directly required were removed in order that the existing chairs were less likely to be nudged or subtly moved in the course of the experiment.

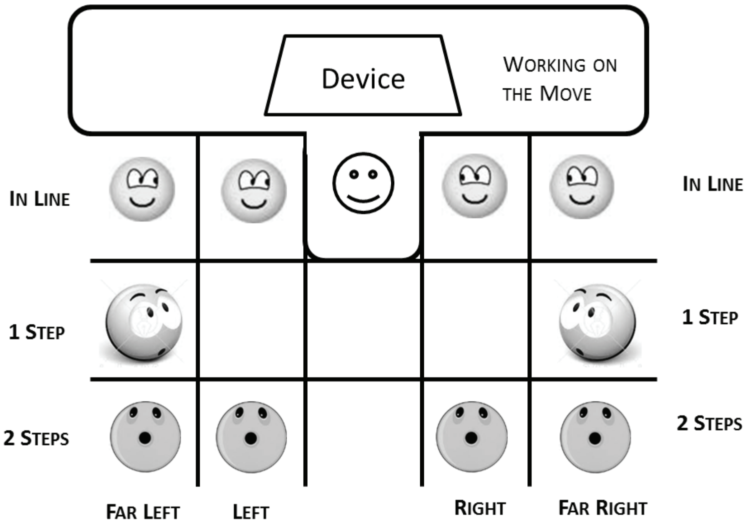
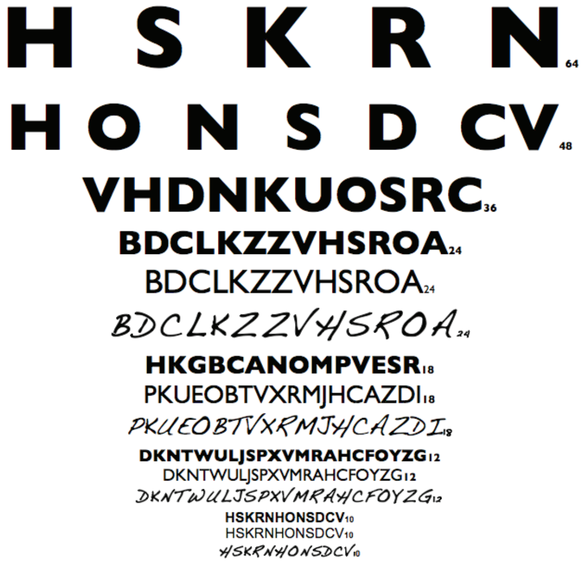


Figure 7: Positions for capture



H S K R N 64

H O N S D CV 48

VHDNKUOSRC 36

BDCLKZZVHSROA 24

BDCLKZZVHSROA 24

BDCLKZZVHSROA 24

HKGBCANOMPVESR 18

PKUEOBTVXRMJHCAZDI 18

PKUEOBTVXRMJHCAZDI 18

DKNTWULJSPXVMRAHCFOYZG 12

DKNTWULJSPXVMRAHCFOYZG 12

DKNTWULJSPXVMRAHCFOYZG 12

HSKRNHONSDCV 10

HSKRNHONSDCV 10

HSKRNHONSDCV 10

Figure 8: Document to be captured

The document to be captured was derived from the standard opticians' eye chart [Figure 8] using Arial font with size of print decreasing down the chart. All photo captures were coded by device and position to aid analytical comparison. Participants were allowed to adjust focus and capture parameters but put under a degree of time pressure to best mimic a realistic context.

The visibility of the captured images was evaluated in terms of the visibility of letters on the test card as evidenced by the captured image. Enhancement of the image was considered reasonable providing the software used was a standard installation as any 'data-thief' would be likely to do the same if sufficiently motivated to take the photo in the first place. For the purpose of this experiment 'iPhoto 11' was used.

3.4 Results

This experiment set out to establish how visible laptop and tablet devices are to VDC. All upright devices were visible to a font size of 12 when captured by someone stood in line with the user.
Table 1 gives the results from the marginal positions used.

Key findings:

- The elevation of the smart phone camera had a greater effect on the quality of the image than the distance. All of these shots were captured with the smart phone user in a standing position.
- Where the screen was displayed at a similar angle to that of the laptop the quality of capture was better in all positions.
- Where the screen was flat the size of font visible to capture was only large or non-existent.

Table 1: Marginal visibility

	Smallest visible font	**Smallest universally visible font**
Laptop- Upright	12	18
iPad- flat	36	-
iPad- upright	12	24
Android - flat	64	-
Android – upright	24	36

Although the risks of this capability have yet to be fully explored, there are products, such as the software that utilises it. Even though this tool is clearly marketed as something that will enhance operations, its use will also facilitate illicit visual data capture.

Kofax
Mobile Capture
Capture Information and Initiate
Business Processes from Mobile Devices
Capture Documents, Photos and Data
Using a Smart Phone or Tablet
KOFAX

Figure 9: Software to facilitate document capture. http://www.tarvos.nl/site/producten/kofax/kofax-mobile-capture

4. Photo capture survey experiment

Having gathered information to demonstrate that visual data capture was technically possible, the next step was to investigate the likelihood of someone exploiting this functionality to record sensitive data. Discussions with organisations that specialise in gathering corporate information indicate that VDC is relatively inefficient as a means of deliberately gathering data as control of the displayed data remains with the user. However, the proliferation of camera-enabled smart phones means that the opportunity for spontaneous capture is growing daily. Evidence that this becoming more common comes from research on the externalisation of memory. In her paper on the ‘Google Effects’ on Memory, Sparrow (2011) argues that having a ubiquitous connection to the Internet means less information has to be individually retained. We only need to know how to access the information, either on computer or mobile device. This is echoed in the work of Clarke (2004) who examined the growing symbiosis with external computing capability. This indicates that the observed behaviour of users capturing information, such as train timetables, using the camera facility on their smart phone as a method of storing such information, is consistent with this concept of external memory storage.

A survey was devised to test the thesis that the ubiquitous availability of effective camera technology has led to a growth in the practice of storing such information externally.

4.1 Methodology

The survey utilised Survey Monkey and was made available to first and second level contacts of the researcher for a period of three weeks.

4.2 Results

The survey attracted 118 respondents.

- Respondents reported storing an average of 458.25 photos on their smart phone. Later questions revealed that this high number was partly due to a conscious use of the phone as storage. Furthermore, it transpired that numbers of photos were often reduced only when storage capacity became an issue.
- 90.5% of respondents reported that they had taken photographs with their smart phone with the express purpose of noting information. Hereafter this sort of shot is called an ‘Informational’ photo capture.
- Of those images currently on their smart phones, an average of 37.66 were informational photo capture shots.
- 67.9% said that they either ‘mostly’ or ‘sometimes’ deleted informational photo capture shots shortly after they had been taken.
- When asked how many informational capture pictures they had taken over the last 6 months the responses showed a significant divide between 33% estimating between 1 and 5, and 23.6% estimating more than 20.

5. Discussion and conclusion

The two experiments reported here explored the vulnerability and probability aspects of the VDC risk. We were able to establish that

- We were able to demonstrate that current mobile phones can easily capture images of sufficient quality to support later perusal (significant vulnerability).
- Photo capture is becoming an increasingly routine method of storing information, demonstrating that probability of such activity is not insignificant.

Our studies have convinced us that there is indeed a growing risk, as facilitated by the three coinciding emerging factors: (1) working on the move, (2) proliferation of own devices, and (3) increasingly ubiquitous cameras. Now that we have satisfied ourselves that there is indeed a problem the next step in the research is to investigate ways of mitigating and reducing this risk.

References

Anderson, R. Bond, M. and Murdoch, S. J. (2006) “Chip and Spin” [online] accessible from http://www.chipandspin.co.uk/ [Accessed 8th February 2013]

Clark, A. (2004). *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*. *Canadian Journal of Sociology / Cahiers canadiens de sociologie* (Vol. 29, p. 471). doi:10.2307/3654679

Goucher, W. (2011) In a World of their Own: Working on the move. HCI 2011. Poster. Newcastle, UK, July 2011.

Honan, B. (2012) “Secure, Visual Security White Paper”, The European Association of Visual Security [online] accessible from http://www.visualdatasecurity.eu/visual-data-security/ 8th February 2013

Maggi, F., Volpatto, A., Gasparini, S., Boracchi, G. and Zanero, S. (2011) “Fast, automatic iPhone shoulder surfing”. In *Proceedings of the 18th ACM conference on Computer and communications security (CCS '11)*. ACM, New York, NY, USA, pp 805-808.

Miller, K. W. Voas, J. Hurlburt, G. F. (2012) “BYOD: Security and Privacy Considerations” *IEEE Security and Privacy Magazine.* September/October.

Okenyi, P. & Owens, T. (2007) “The anatomy of human hacking” *Information Security Journal: A Global Perspective*, Vol 16, No. 6, pp 302-314.

Owen, G. (2008) “The zzzzivil servant who fell asleep on the train with laptop secrets in full view” [online] Mail Online. 1 November, 2008. http://www.dailymail.co.uk/news/article-1082375/The-zzzzivil-servant-fell-asleep-train-laptop-secrets-view.html [Accessed 8th February 2013]

Renaud, K. and Maguire, J. (2009) “Armchair Authentication” In: *Proceedings HCI 2009. People and Computers XXIII Celebrating People and Technology.* Cambridge, UK. 1 - 5 September. pp 388-397.

Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google effects on memory: cognitive consequences of having information at our fingertips. *Science (New York, N.Y.)*, *333*(6043), 776–8. doi:10.1126/science.1207745

Thomson, H. H. (2010) “Visual Data Breach risk assessment study 2010 for 3M and with People Security” [online, assessable from] http://solutions.3m.com/wps/portal/3M/en_US/3MScreens_NA/Protectors/For_Organizations/Industry_Whitepapers/Visual_Data_Breach_Risk_Assessment/ [Accessed 8th February 2013]

Modelling Attribution

Clement Guitton
War Studies, King's College London, UK
clement.guitton@kcl.ac.uk

Abstract: This article posits that attribution is better approached as a process than as a problem. Departing from models for attribution emphasising its technical constraints, the article distinguishes between two distinct attribution processes operating in two different contexts and answering two different questions. On one hand, in a criminal context, the attribution process seeks to identify individuals who launched cyber attacks. On the other hand, in a national security context, the process seeks to identify adversaries which may have sponsored cyber attacks. Five elements can serve as determinants to distinguish between the two processes: the type of target, the severity and scale of the damage, the apparent origin of the instigator of the attack, the means of the attack, and lastly the claims by political groups. After reviewing these elements, the article analyses how the constraints and characteristics of the two different attribution processes vary in terms of responsible authority, standard of evidence, stake in play and relevance of timing for attribution. In a criminal context, the collection of digital evidence is of primary importance to be able to reach a high level of judicial proof. However, in a national security context, investigators will likely not have access to evidence but only to intelligence, and the decision of attribution will not be taken within the judiciary but within the executive (e.g. a government). As such, attribution of national security incidents requires a form of political judgement to heed the political context and circumstances of the incidents that courts do not need to consider when examining the guilt of criminals. Political judgement is also required to balance the consequences of attribution on a political level, potentially resulting in a straining of relations with other international actors. As the constraints between the two different attribution processes differ, their requirements for easing the constraints also vary.

Keywords: attribution, sponsorship, cyber attacks, national security threats

1. Introduction

In August 2012, a malware called Shamoon infected the computers of Saudi Arabia's national oil company Aramco, and wiped out the hard drives of 30,000 computers. The attack prompted US Secretary of Defense Leon Panetta to mention the attack in a widely quoted speech in front of business executives (Panetta, 2012) and to answer further questions in a briefing at the Pentagon (Panetta and Dempsey, 2012). While Panetta fell short in his speech of accusing Iran for the attacks, a journalist, quoting an unnamed US government official mentioned that Iranian hackers carried out the attack with the support from the Iranian government (The Associated Press, 2012). Officially, the attack remains unattributed to this day. In contrast, many other malware have in the past infected many computers without prompting high political officials to react about the identity of the instigators. For instance, in 2004, the Sasser worm infected millions of computers including those of rail systems in New South Wales, the Italian Interior Ministry, the European Commission and British coastguard stations (Jesdanun, 2004) without stirring such a response. The police later found its author, Sven Jaschan, thanks to an information reward program set up by Microsoft (Lemos, 2005).

How and why does the attribution of an incident become politicised? Does attribution require the involvement of political figures? How do the processes of attributing a criminal incident and attributing an incident of national security relevance differ?

As attribution informs states' responses to cyber attacks, finding answers to these questions is essential. But there is currently no overall model for attribution that can answer these questions and explain how attribution occurs by taking into account the disparities between the responsible authorities, the standards of evidence, the stakes in play and the relevance of timing for attribution of cases either similar to the Aramco incident or to the Sasser worm. This article proposes such a model for attribution based on the nature of the attribution process either as apolitical or political. Apolitical cases are cases of purely criminal nature that do not rise to the level of 'national threat'. When an incident crosses the 'national threat' threshold, the question of knowing the identity of the attackers becomes less relevant than knowing who the enemy is, for instance a state actor or a terrorist organisation. The process of reaching conclusions on the identity of the instigator or on the sponsorship of an attack differs in great regards. Assessing enmity is a matter of political judgement, and not for a domestic court to decide. Evidence to support the judgement for enmity hence does not have to reach the level of judicial proof that the conviction of a criminal requires. Distinguishing between the two processes allow the discernment of solutions to the attribution problem that is inherently political in some instances, and not technical as often characterised.

This article presents the model, namely the two attribution processes, in two steps. Firstly, it shows the criteria that lead to the politicisation of incidents. Secondly, it delves into the different characteristics of these two distinct processes.

2. Politicised attribution: A set of conditions for the model

Three main models have previously emerged to characterise attribution, one having its root in computer science, a second one in political science, and a third one in both fields. Firstly, the cyber security expert Tom Parker suggested assigning values to three variables to characterise the attacker: the attacker's intrinsic characteristics (e.g. resources, motivation), the target, and the general environment of the attack (Parker et al., 2004). Within this model, analysts compute values using Bayesian probabilities obliging them to assess the certainty of assumptions and known information. But this model suffers two pitfalls: many variables do not lend themselves to being ascribed numerical values; and Parker's attempt to draw the line between criminal and terrorist cyber attacks based on the unknown motivation of the attacker is unrealistic. The motivation of an attacker usually follows at least partial attribution, and does not precede it.

Secondly, the lawyer Susanne W. Brenner proposed a model for attribution as a political decision based on the trichotomy war – terrorism – crime, a trichotomy whose relevance she acknowledged is eroding in cyberspace (Brenner, 2009: 96). Brenner suggested starting with attempting to classify the incident within this trichotomy also based on three variables: the origin, the target, and the motive for the attack (Brenner, 2009: 87). Yet, determining if an incident is an act of war, hence originating from a state actor, requires solving attribution first, a problem Brenner also acknowledged. Brenner's solution to the eroding relevance of the trichotomy focused on the dilemma the erosion poses on an institutional level between military and civilian entities involved in responding to the threat rather than on the core problem it poses for attribution (Brenner, 2013). Such a solution did not address the main flaw in her model for attribution, namely that knowing the sponsorship of an attack cannot be at the same time the sought solution of the problem, and the variable on which the problem depends.

Lastly, the computer scientists David Clark, Susan Landau and Earl Boebert have separated the attribution problem into two sub-problems: one technical and consisting in identifying the attacking machine, and one human consisting in identifying the machine's owners and potentially their sponsor (Clark and Landau, 2010: 37; Boebert, 2010: 43). The computer scientists delved into technical procedures to attribute a machine and left aside the details of the identification of individuals and sponsors. While much research has been carried out to improve tracing back attacks (Wheeler et al., 2003; Clayton, 2005), many researchers have also acknowledged that technical considerations alone cannot be sufficient to solve the attribution problem (Hunker et al., 2008: 10; Clark and Landau, 2010: 39). It is hence important to continue expending the model for attribution of cyber attacks without focusing on its technological aspects. The focus away from technology is further justified as political considerations dwarf technical ones when the incident meets certain criteria constraining the access to evidence.

To remedy the shortcomings of each of the three existing models, attribution can be approached as two distinct processes: on one hand, as an apolitical process following well-established criminal procedures; and on the other hand, as an act requiring a political decision, such as classifying an incident as an act of war or terrorism. Considering attribution as a process and not as a problem allows the identification of steps where different questions arise with different saliency within the process. For instance, only when attribution is politicised does the question of the different degrees of involvement of a state arise (Healey, 2012), as well as strategic questions about instrumental use of cyber attacks as covert means to achieve foreign policy objectives (Libicki, 2012).

The determinants indicating which process an incident fall within need to be assessable from the outset of the investigation and cannot rely on the formulation of early assumptions, as Parker's and Brenner's model required for determining the motive of the attack. Five determinants influence the attribution of an incident as becoming the resort of politics rather than of the criminal justice system.

2.1 The target

The profile of the attack rises in function of the target according to two different elements: firstly, the political nature of the target, and secondly, the threat to national security that the target can constitute. Attackers can focus on these two types of targets in an instrumental way to try to weaken a state and achieve a political objective.

Targets creating a threat to national security are attacks either on critical national infrastructures (e.g. Stuxnet), or on companies losing their intellectual property on a great scale. The loss of intellectual property impacts companies' productivity and, in turn, harms the national economies of countries hosting them. In the UK, tackling the threat to the economy is the very first objective of its national cyber security strategy, aiming at placing the UK as 'one of the most secure places in the world to do business in cyberspace' (Cabinet Office, 2011: 8). The theft of intellectual property via cyber attacks is therefore a topic taken seriously as a threat to national security, and is as such, approached by high-ranking officials.

Not all attacks on companies that correspond to corporate espionage will fall within the national security threat category. Rogue corporations, even within the same country, can attack each other to gain economic advantages. Solely considering a company victim of espionage may therefore not be sufficient to raise the political profile of the attack. But the systematic loss of intellectual property from a range of companies is sufficient to constitute an economic threat for the victim country.

2.2 The severity and scale of the damage

One of the reasons the targeting of critical infrastructure moves attribution to the higher end of the political spectrum is because of the high severity of damage an attack can potentially cause. The assessment of the severity of an attack is not always evident. In espionage cases, the severity is dependent upon the information accessed by the hackers. But in most cases it is difficult to know which information the attacker accessed, and what the attacker will do with this information. Based on the information that a victim company provides to law enforcement agencies, state officials can attempt to gauge the severity of an attack and assess if it crosses the threshold of threat to national security. In the case of an attack on a ministry or on a private defence contractor for instance, the sensibility of the information potentially accessed automatically scales up the potential threat it poses. State officials do not particularly need to feel confident that a state actor is behind the attack to become involved in the attribution process. The mere potentiality that the damage can be important is sufficient to reflect a deviation from the criminal attribution process.

2.3 The apparent origin of the instigator

State officials are often cognisant that the IP address showing the origin of an attack from a particular state can be deceptive. Hackers can easily falsify the address, and use it for instance to pit one state against another, or to use underlying existing strained relations to spark an escalation of tensions. The apparent origin of the instigator can justify the implication of state officials for at least two reasons. Firstly, officials from the victim state may need to intervene to coerce the seemingly state of origin into providing more information about the attackers. Secondly, the apparent origin of the attack may be from a country that the victim state regards as an adversary. Each state has at least a vague idea of who their enemies are, or who they regard as a threat. The instability of the political situation with the enemy state can lead even a small incident with low damage to an escalation of tensions. Again, the involvement of political actors will be necessary to make a trade-off in order to decide if pursuing the investigation is possible and worth taking the risk of a political setback in case of refusal for cooperation.

2.4 The means of the attack

Instigators of cyber attacks have a wide variety of tools at their disposal to commit cyber attacks. Attackers can discover a new vulnerability in a system and exploit it. While finding new vulnerabilities is difficult and time-consuming, it also indicates a level of technical sophistication to which analysts of cyber attacks have not remained insensitive. Stuxnet, the worm that destroyed Iranian centrifuges in Natanz in 2010, utilised in its latest version four of these vulnerabilities. When the cyber security company Kaspersky started examining Stuxnet, it very quickly declared that 'Stuxnet's sophistication, purpose and the intelligence behind it suggest the involvement of a state' (Halliday, 2010). Currently, only certain states and few criminal groups if any possess the means to carry out very sophisticated attacks alike to Stuxnet. Considering the means of attacks, according to analysts' reasoning, allows the group of potential suspects to be narrowed down, and may even indicate state sponsorship. Although it is disputable if sophistication is really of any help to achieve attribution as many different actors can have the skill and resources to engineer 'sophisticated' attacks, it is incontestable that associated with the type of victim organisation it raises the profile of the attack.

2.5 Claims by a political group

Any cyber attacks claimed to have been carried out for a political reason, and substantiated by enough evidence, will have to be assessed within its political context. The claim may be unfunded, unrealistic, and an entity other than the claimant may truly be behind the attack. But they will necessitate a certain assessment involving political judgement, and not solely the analysis of technical and forensic evidence. Yet, within a purely criminal context, the police and cyber security firms will lack this political judgement and the authority to wield it. The mere claim by a political group may not be completely sufficient for it to be credible. The level of confidence in the credibility of the claim plays a role in considering the attack still within the realm of criminal justice or as a threat to national security.

3. Descriptions of the attribution processes

What is so different between the attribution process of criminal act and the attribution process of threat to national security? Firstly, it is necessary to recognise that the two processes have different goals. While the attribution process of criminal apolitical cases seek to unravel the identity of the individuals launching the cyber attack, the attribution process of cases threatening national security focuses on discovering the sponsorship of an attack. The underlying processes to reach these two different goals are fundamentally different.

3.1 The criminal context

Both attribution processes start with the discovery of a breach of an information system and its subsequent classification as an attack (i.e. as malicious and intentional). Following the discovery of the attack, and if none of the five aforementioned criteria apply, the case will follow the path of a criminal investigation, a well-understood process dependent on the country's good functioning of its judicial system. The next steps are then as follows: investigators from cyber security companies or from law enforcement agencies attempt to gain information to find the attacker; investigators find evidence by seizing incriminating material on the person's computer; and lastly, a court validates the inferences made from the collected evidence proving that the person was indeed the criminal using the computer from which the cyber attack was launched. In this regard, the way digital forensics are collected need to respect high standards to show that the police did not tamper with them, and that the defendant's actions, and not only his machine, caused the alleged attack 'beyond reasonable doubt'.

The access to evidence is predicated on first finding information about the identity of the individual. Finding information can broadly occur from three different ways: from discovering forensics left behind the attack, from information tracing back the attack to its origin (e.g. an IP address, the place where a malware first appeared), or from informants. This information does not usually prove 'beyond reasonable doubt' the individual's involvement in the attack, but evidence seized on his computer can. Two complications in the acquisition of information and evidence often arise, as many cases are spread across several jurisdictions. Firstly, the victim country's law enforcement agencies will require cooperation from the state hosting the Internet service provider which assigned the attacker's IP address. The state where the attack originated will have to request the service provider to give the match between the IP address and a name, before relaying this information to the victim country. Both steps are strongly contingent upon the state of affairs between the two countries and the tools at the disposal of the victim state to incentivise or coerce the other state into cooperating. Secondly, gathering evidence, which requires physical access to the accused's computers, can also be sensitive. Law enforcement in the originating country may be willing to seize evidence based on the information provided by their foreign counterpart, but in function of their decision to prosecute the attacker in their country or not, they may be unwilling to share the collected evidence. In case they refuse to extradite the attacker, the country in which the attacker was located will still rely upon the victim country to exchange information about the attack it suffered in order to successfully convict the accused. These legally time-consuming and twisted procedures are mostly formalised within mutual legal assistance treaties. However, the lack of an existing treaty covering cyber attacks does not have to thwart the attribution process, as much as the existence of a treaty between two countries does not ensure its applicability. Politics often trump the law, even for the attribution process in a criminal context where cooperation to exchange information or to extradite attackers is mainly dependent on political will.

Lastly, the conviction of the attacker by a court represents the highest level of attribution possible in a criminal context. But even if the court does not convict the attacker, it can remain clear that he is the one to whom the cyber attacks are attributable. The lack of legal framework criminalising cyber attacks or the high cost of the trial involving flying witnesses over from another country have in the past prevented conviction despite the accused admitting to the charges. A conviction by a court, while preferable, is hence not an absolute necessity.

3.2 The national security context

Within the national security context, the access to evidence is much more difficult. The four-step process for attribution of criminal cases (discovery of the attack, finding information, finding evidence, and conviction) becomes the following four-step process: discovery of the attack, finding information, formulating an assessment, and publicly acting upon the assessment. The first two steps are similar to the attribution process in a criminal context.

The entities formulating the assessment of attribution in cases of national security are more various than in criminal cases where only the court has the monopoly of doing so, but formal attribution still only occurs via a single formal channel, a representative of the government. The assessment by state institutions and the decision to attribute are the only forms of 'binding' attribution that will bear consequences, as the state is the only institution with the authority to take a decision of attribution accepted and recognised by whom they affect. Placing the onus on governmental institutions rather than on judicial entities for attribution has several implications. The methodology changes from the mere assessment of digital forensics to using political judgement; the standard of evidence becomes lower and proving 'beyond reasonable doubt' is not required anymore; the stakes in play are by nature political; and the timing for attribution, while important when the judiciary would pronounce a verdict, is far more crucial when it is a form of political act carried out by a political entity.

Accepting political judgement as a methodology for attribution (Lin, 2012: 350) also implies accepting that attribution in some cases fall between beliefs and reasons (Arendt, 1992: 4), is prone to expectation biases, and operates in an environment mainly defined by an acute lack of certain information (6, 2011: 21). The difficulty of accessing evidence and the dependence on intelligence rather than evidence for attribution in this context confirms the appropriateness of resorting to this methodology. Good judgement will however rest upon the analysts being able to show the thought process as being as removed from feelings and preconceived ideas as possible, and following strict methodologies to ensure it (Heuer, 1999).

In this light, the most appropriate organisation to formulate judgements will be intelligence services. Intelligence services bring together both the political and technical expertise required for attribution. Analysts from private companies specialising in cyber security, such as Kaspersky and Symantec, can also mention potential authors for the attacks when they report new attack vectors. But while these companies are at the forefront of technical developments in cyber security, they certainly lack the experience of heeding political elements to ascertain the origin of attacks. On the other hand, investigative journalists have the experience of taking into account in their analysis political elements, but not technical ones. Furthermore, as they use non-technical investigation techniques, they also bring other pieces of information relevant to the assessment of sponsorship of attacks. For instance, the sponsors of Stuxnet, namely the US and Israel, only emerged when unnamed US officials with knowledge of the operation leaked information to David Sanger, a journalist at *The New York Times* (Sanger, 2012).

Journalists very often protect their sources and it is not possible to independently verify their sources' credibility. The lack of verifiability of the information, either collected by investigative journalists or its analysis by state officials, is a defining constraint for incidents falling within the attribution process of national security incidents. Relevant information feeding into political judgement, apart from being non-testable in contrast to evidence presented in front of a court, is also mainly non-technical. Technical evidence is indeed not best suited to inform on sponsorship, which is what attribution is set to achieve in this context. Instead of using technical forensics, human intelligence gained from informants and whistle-blowers is more appropriate. This means that investigators of incidents of national threats are more interested in intelligence than in digital forensic evidence, when the opposite is true for the attribution of criminal cases.

Intelligence services are also well acquainted with the task of estimating their adversary's strength and capabilities to launch attacks, which *a priori* could help narrow down the number of potential attackers if one assumes that sophistication can inform on state-sponsorship – a contestable thesis regarding that criminals can also engineer sophisticated attacks. The 'strength' variable can however be difficult to estimate for cyber attacks, especially in comparison with estimating the military strength of a country. It is much easier to subcontract cyber attacks and buy zero day exploits in a hacker underground market than it is to buy military grade weapons. Furthermore, hiring a single very talented hacker can be sufficient to engineer malware with far ranging effects. It is hence difficult to assess with high certainty any groups' capability to launch cyber attacks. Such a difficulty hinders the work of intelligence services to neatly differentiate between potential enemies with capabilities and those with only low resources to launch cyber attacks. Further blurring the picture, it is strategically advantageous for actors to hide their offensive capabilities as much as possible. Such a conclusion can appear at first counter intuitive considering that on the other hand displaying one's capabilities can act as a deterrent measure against an adversary. However, if the adversary knows that the likeliness for a threat of retaliation is low due to the uncertainty of attribution, deterrence is unlikely to function. Keeping possibilities of using covert cyber attacks as a tool for foreign policy hence becomes a reasonable strategic decision.

But once intelligence services attribute an attack, and as the intelligence services are by design an opaque structure, the decision of making the attribution public is somewhat a completely different decision. This decision differs from the judgement pronounced by a court in two regards. Firstly, the decision, by the nature of the parties potentially involved as state actors or as terrorist organisations, becomes a political decision for a state official to take. The decision to make attribution public also requires political judgement to balance the risks of misattribution, certainty, timing, and potential benefits of attributing the attack. The decision is likely to be taken by senior and high elected officials only. As Francis Maude, British Minister for Cabinet Office, pointed out in a hearing in the House of Commons (2012), 'if something looked like it could be a sovereign attack', the judgement of deciding sponsorship 'would clearly be for the Prime Minister' to make. Secondly, a legal judgement aims at instating justice, while a political judgement has wider possible ramifications. Following an assessment of the potential instigators of an attack, state officials can refuse to publicly name an attacker as a state with which they have strategic partnerships in order to preserve economic ties with the country. Such a decision will have followed the different steps of the attribution process, and stopped at the third one. Refusing to publicly name the attacker would be morally wrong, as the victim state would fail in its duty of owing transparency and the truth to its citizens. But it could also be a politically shrewd decision. As the level of damage caused by cyber attacks increases, the political argument will increasingly trump the moral one, leading to more frequent formal attribution.

4. Conclusion

The new proposed model for attribution challenges three previously accepted arguments. Firstly, it challenges that attribution is a problem and posits that it is better described as a process. Secondly, it challenges that attribution is mainly a technical issue. Lastly, it challenges that identifying a criminal and identifying the sponsor of an attack follow similar characteristics. Brought under the common umbrella term of 'attribution problem', many officials and researchers have as such advocated the same solutions for what appears in fact as two distinct attribution processes. The process within which a case falls determines the responsible authority (judiciary vs. executive), the methodology (digital forensics vs. political judgement), the standard of evidence ('beyond reasonable doubt' vs. dependent upon judgement), the stake in play (individual vs. political), and finally, the timing (irrelevant vs. critical) that will apply for attribution. The separation of attribution into two processes leads to the conclusion that they face different constraints with different requirements to relax these constraints.

References

(2012) Defence and Cyber Security. *House of Commons.* London: Parliament.

6 P. (2011) *Explaining Political Judgement,* Cambridge: Cambridge University Press.

Arendt H. (1992) Postscritum to Thinking. *Hannah Arendt Lectures on Kant's Political Philosophy.* Chicao: The University of Chicago Press.

Boebert WE. (2010) A Survey of Challenges in Attribution. *Workshop on Deterring Cyber Attacks: Informing Strategies and Developing Options for U.S. Policy.* Washington DC: The National Academies Press.

Brenner SW. (2009) *Cyber Threats: The Emerging Fault Lines of the Nation State,* Oxford: Oxford Scholarship Online.

Brenner SW. (2013) Cyber-threats and the Limits of Bureaucratic Control. *Minnesota Journal of Law, Science, and Technology* 14.

Cabinet Office. (2011) The UK Cyber Security Strategy - Protecting and promoting the UK in a digital world. London: Cabinet Office.

Clark DD and Landau S. (2010) Untangling Attribution. *Workshop on Deterring CyberAttacks: Informing Strategies and Developing Options for U.S. Policy.* Washington, DC: The National Academy Press, 25-40.

Clayton R. (2005) Anonymity and traceability in cyberspace. *Computer Laboratory.* Cambridge: University of Cambridge, 189.

Halliday J. (2010) Stuxnet worm is the 'work of a national government agency'. *The Guardian,* 24 September.

Healey J. (2012) Beyond Attribution: Seeking National Responsibility for Cyber Attacks. Washington, DC: Atlantic Council, 1-7.

Heuer RJ. (1999) *Psychology of Intelligence Analysis,* Washington, D.C.: Centre for the Study of Intelligence.

Hunker J, Hutchinson B and Margulies J. (2008) Role and Challenges for Sufficient Cyber-Attack Attribution. *Institute for Information Infrastructure Protection.* Hanover, New Hampshire: Dartmouth College.

Jesdanun A. (2004) New virus snarls hundreds of thousands of machines worldwide. *The Associated Press*, 3 May.

Lemos R. (2005) Microsoft to reward informants after Sasser conviction. *SecurityFocus*, 8 July.

Libicki M. (2012) The Specter of Non-Obvious Warfare. *Strategic Studies Quarterly* Fall: 88-101.

Lin H. (2012) Thoughts on Threat Assessment in Cyberspace. *I/S: A Journal of Law and Policy for the Inormation Society* 8: 337-355.

Panetta L. (2012) Text of Speech by Secretary of Defense Leon Panetta. *Business Executives for National Security*, 11 October.

Panetta L and Dempsey M. (2012) News Briefing. *Departement of Defense*, 25 October.

Parker T, Devost MG, Sach MH, et al. (2004) *Cyber Adversary Characterisation,* Rockland: Syngress Publishing.

Sanger DE. (2012) Obama Order Sped Up Wave of Cyberattacks Against Iran. *The New York Times,* 1 June.

The Associated Press. (2012) Official: US blames Iran hackers for cyberattacks. *The Washington Post*, 11 October.

Wheeler DA, Larsen GN and Leader T. (2003) Techniques for Cyber Attack Attribution. Alexandria, Virginia: Institute for Defense Analyses.

Cyberspace from the Hybrid Threat Perspective

Håkan Gunneriusson[1] and Rain Ottis[2, 3]
[1]Swedish National Defence College, Stockholm, Sweden
[2]University of Jyväskylä, Jyväskylä, Finland
[3]Tallinn University of Technology, Tallinn, Estonia
rain.ottis@jyu.fi

Abstract: Hybrid threats use conventional and unconventional means to achieve their goals. In this paper we explore the cyber threats as one possible aspect of hybrid threats. We describe three ways of approaching cyberspace (operations) from the hybrid threats perspective: supporting conventional operations, exploiting non-military systems, and exploring the opportunities provided by this environment. In particular, we highlight the aspects that are or likely will be relevant to the military community.

Keywords: hybrid threat, cyberspace, cyber operation, Internet, military

1. Introduction

One of the problems with the concept of hybrid threats is that it is very difficult to define. Hybrid threats are not defined by the actors, since states, non-state actors and even individuals might be considered (part of) hybrid threats. They are not about some specific technology, since the list here keeps growing as new technologies become available. They are not about specific effects, as a hybrid campaign may result in casualties, changed decisions, altered public perception, etc. Perhaps the best way to put it, hybrid threat is a manifestation of total war. It is about making the other side submit to one's will, with any means available.

Threats from or using cyberspace are similarly difficult to define, and can in fact be viewed as a subset of hybrid threats. Cyber threats come in the form of state actors, criminal groups, terrorist organizations, hacktivists, professional hackers for hire (mercenaries), etc. The list of exploitable technologies also keeps growing – aside from servers, personal computers and laptops we also have to worry about smart phones, smart meters (electricity distribution to our homes), wireless-enabled pacemakers, industrial control systems, etc, and that is just the hardware side. Possible effects can range from tongue-in-cheek publicity campaigns to destruction of critical infrastructure components and potentially – deaths.

Cyber threats operate in a man-made environment. As such, they are constrained by the capabilities built into that environment. However, this is in fact an enabler for the cyber attacker. Non-trivial man-made systems (such as computers, airplanes, etc.) are rarely perfectly implemented and may be based on flawed design assumptions. These assumptions may be about type of input, length of input, number of simultaneous user sessions, etc. Cyber attacks work by exploiting design assumptions or implementation flaws. It is important to realize, however, that while cyberspace is the "home" environment for cyber threats, they can and do affect other environments as well. Consider, for example, the case of StuxNet, where a cyber-attack disrupted the uranium enrichment process and caused a number of physical devices to break in Iran. (See, for example, Falliere, Murchu and Chien 2011, Sanger 2012, The Economist 2010a) Or consider the case of cyber-attacks against Georgian government and news sites during the 2008 Russia-Georgia war, which hampered the Georgian government's ability to communicate with the citizenry. (Markoff 2008)

These risks from cyberspace have a serious effect on society. On one hand, we are concerned with various threats – hacktivists, criminals, spies, etc. In order to protect ourselves (our systems, our data, our way of life) we are constantly endeavoring to improve the security situation in cyberspace. On the other hand, however, we are concerned with the opportunities and liberties associated with cyberspace – freedom of speech, privacy, etc. Unfortunately, in many cases, an increase in security tends to undermine the open and liberal society. Therefore, it is important that (military) security professionals are aware of these concerns and will take steps to minimize the adverse effects of new security solutions.

In this paper we describe three ways of approaching cyberspace (operations) from the hybrid threats perspective. In an effort to better explain this new type of threat to commanders, planners and soldiers, we highlight the aspects that are or likely will be relevant to the military community.

2. Hybrid threats in relation to cyber threats

Hybrid threat as a concept has changed over time. This is not unusual when it comes to the combination of the military culture and theoretical concepts. The now dead acronym EBO (Effects Based Operations) had a similar story. (See, for example, Mattis 2008 or Ho 2005) The stake holders agreed that the term held some truth but they could not come to an agreement of the content or meaning of the concept. The terms used in EBO were so hollow yet so widely discussed that it was considered better to leave that debate open and to concentrate on developing our theoretical thinking on military operations instead.

The initial meaning of hybrid threat was described as a non-state actor wielding a conventional capability as if it were a state-actor. (Matthews 2008) The concept has evolved to a catch-all phrase for unconventional and unexpected threats which strike asymmetrically. Now, as with EBO, NATO has abandoned the development of the concept of hybrid threats. However, this does not mean that the underlying concept is not of any use.

In the early days of the Internet, many wondered about the possibilities that global networking would bring. In historical terms, this was similar to the time when electricity was harnessed for the benefit of society. Such advances in science and technology bring about all-encompassing effects to the entire society, not just specific markets, businesses or governments. Instead of talking about, for example, *the electricity threat* or *the cyber threat,* it might be better to use the hybrid threat concept as a way to describe the interplay between conventional and unconventional threats to our society. In this paper we have taken this route in order to explore the cyber threat from a new perspective, since the military is already somewhat familiar with hybrid threats.

Although this is not a firm rule, hybrid threats tend to target the civil society rather than the military. This is a double asymmetry as it both strikes in unconventional ways and targets parts of society that may not be prepared for the attack. Defending against cyber threats requires a comprehensive approach, involving all relevant stakeholders from responsible government agencies (including the military) to private companies to individuals.

3. Cyberspace and cyber operations

Cyberspace is the extension of some of the greatest technological developments of the 20th Century: the electronic computer, the Internet and the World Wide Web. In 1948 Norbert Wiener coined the word cybernetics, which refers to “communication and control in the animal and the machine” (Wiener 1948). The discipline of cybernetics plays an important role in understanding and developing the underlying infrastructure of cyberspace. In 1984 William Gibson, a science fiction writer, first used the term cyberspace to describe the “consensual hallucination” of a new domain formed by interconnected computers. (Gibson 1984) Over the last decade the term has been widely adopted, but there are numerous ways of defining, interpreting and using the underlying concept.

One of the most prominent concepts of cyberspace is the one that has emerged in the national defense and security sector. It refers to cyberspace as a new domain of (military) operations, on equal footing with land, air, sea, and sometimes – space. (The Economist 2010b) The western military doctrine generally divides the operations in cyberspace into two or three categories. Perhaps the best known is the US approach, which uses the term ‘computer network’ instead of ‘cyberspace’. According to this doctrine, computer network operations (CNO) are a component of Information Operations and break down into computer network defense (CND), computer network exploitation (CNE) and computer network attack (CNA). (Joint Publication 3-13)

While the main purpose of CND and CNA is self-evident, CNE is somewhat more controversial. It primarily refers to covert intelligence gathering, but it is unclear where the ‘exploitation’ ends and the ‘attack’ starts. From the defender’s perspective, it is very difficult to tell if an intrusion into their systems is an attempt to gather military intelligence (CNE), to prepare for a subsequent attack (CNA), to make money (criminals), or to make an ideological statement (hacktivist).

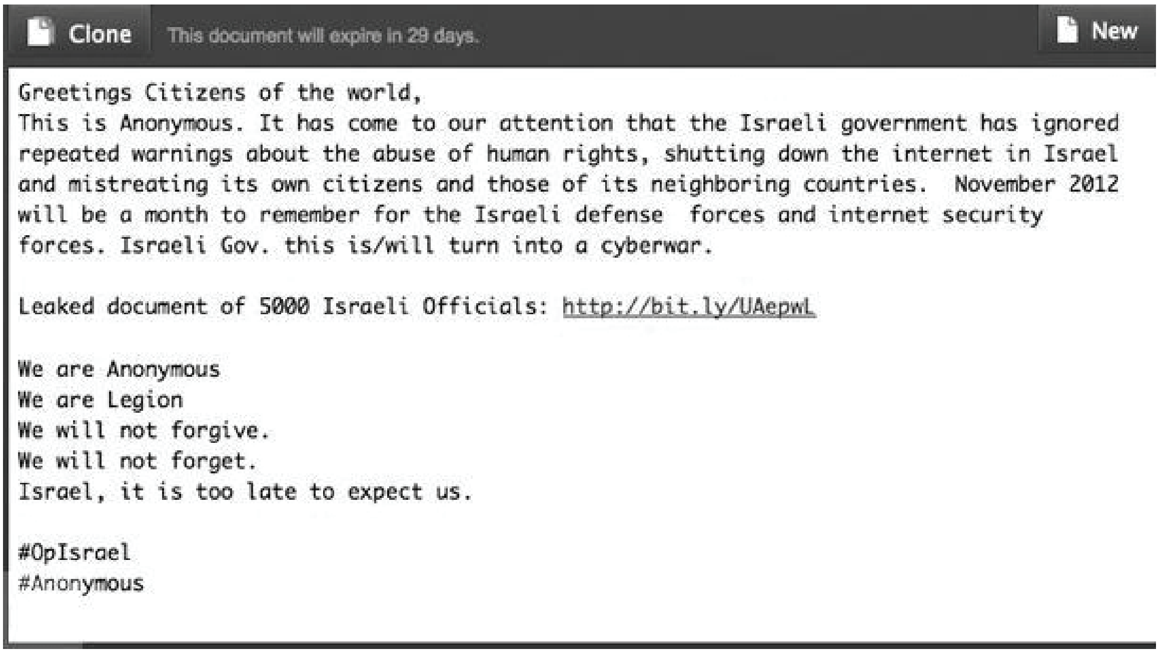
Of the three, CND is the most mature discipline. This does not mean that the art of defense is perfected - just that know-how is available and widespread. The offensive forms of operations (CNE, CNA) are comparatively rarely discussed in public and the actual capabilities of various actors are difficult to assess. It is this emergent quality that raises offensive cyber operations into the hybrid threat discussion.

The western doctrine is by no means finalized, nor is it the only one. For example, the Chinese military has spent nearly two decades of developing ‘informationized warfare’. Inspired (shocked) by the US performance against the Soviet style forces of Iraq in the Gulf War, Chinese scholars and military leaders have blended the techno-centric approach of the US doctrine with the ancient ‘Art of War’ of China. (See, for example, Thomas 2007, 2009) The resulting mix offers a potentially more holistic approach than the often stove piped and limited Western IO doctrine.

The Russian military doctrine (or vision) of the future wars also combines conventional and unconventional approaches. For example, there is strong emphasis on the question of information superiority, both in terms of functionality of systems and of the prevailing content or narrative in the information sphere and the public perception. There is also discussion of using unconventional concepts like nano technology weapons, ‘disorganization’ techniques, affecting people’s thought processes, etc. (See Thomas 2011 for more details) While it is unclear how much and which components of the unconventional approach are mere intellectual musings, it is a strong indication that Russia should be considered as a hybrid actor.

Another view of cyberspace is focused on the opportunities offered by cheap and easily accessible computing devices and global networking. Online shopping, social networks, strong public cryptography and (anonymous) real-time communication are examples of this. These solutions provide asymmetric advantages to actors who have limited resources. For example, it is possible to raise awareness of an issue on a blog, find people who are supportive of the cause through social networks and coordinate group actions on encrypted chat channels. On the other hand, the technology also allows for much greater control by those in power. For example, state (security) services might limit people’s access to the Internet or specific services, eavesdrop unencrypted (or weakly encrypted) communications and even hack into personal computing devices to gather evidence against them.

The possibility to perform these activities with scarce resources enables sub-state actors. This is a big change compared to the Cold War and earlier times. The hacktivist group Anonymous has reported, for example, that its members recently hacked several hundred websites and published information on thousands of Israeli government officials as a response to Israeli efforts to shut down Internet in the Gaza Strip (see Figure 1). (RT.com 2012) This is a non-state actor attacking a state, as in the case with Hezbollah during the Lebanon war (although the scenario was quite different). Israel’s finance minister declared in no uncertain terms that the government was now waging war on a "second front” [in cyberspace]. (RT.com 2012)



```
Greetings Citizens of the world,
This is Anonymous. It has come to our attention that the Israeli government has ignored
repeated warnings about the abuse of human rights, shutting down the internet in Israel
and mistreating its own citizens and those of its neighboring countries.  November 2012
will be a month to remember for the Israeli defense  forces and internet security
forces. Israeli Gov. this is/will turn into a cyberwar.

Leaked document of 5000 Israeli Officials: http://bit.ly/UAepwL

We are Anonymous
We are Legion
We will not forgive.
We will not forget.
Israel, it is too late to expect us.

#OpIsrael
#Anonymous
```

Figure 1: A message from the hacktivist group Anonymous (RT.com 2012)

The war-like reference might seem a bit overdone, but cyber operations against a state are likely to get such reactions, depending on how serious the state thinks the problem is. A US state official has stated that “If you shut down our powergrid, maybe we put a missile down one of your smoke stacks". (Gorman and Barnes 2011)

Cyberspace is a contested environment. In recent years there have been many interesting developments to illustrate this point. The discovery of the StuxNet malware in 2010 created a lot of discussion about government malware and sabotaging critical infrastructure through cyber-attacks. (See, for example, Falliere, Murchu and Chien 2011, Sanger 2012, The Economist 2010a) In Germany, a debate sparked on the use of malware and hacking techniques for law enforcement purposes. (See, for example, Herkner 2007) The so-called Arab Spring demonstrated the dual use of information technology for both the people and the government. (See, for example, Afanasjev 2011)

The international community is trying to find consensus on some of these issues, but so far there is little success. There are efforts to shape or analyze the legal instruments for this area, such as the Council of Europe Convention on Cybercrime (2001) or the Tallinn Manual on the International Law Applicable to Cyber Warfare (2013). In reality, however, state practice is developing the norms for tomorrow.

4. Overlap of cyberspace (operations) and hybrid threats

From the perspective of hybrid threats, cyberspace can be viewed in several ways. First, cyber capabilities in support of conventional forces. Second, as an asymmetric and unconventional attack vector on its own. Third, as an enabler or disabler of events and social movements.

4.1 Supporting conventional forces

Offensive cyber capabilities may one day be considered part of the 'conventional' toolkit. However, for now they are a rarity in military combat operations and can be safely categorized as 'unconventional'. This means that any military operation that includes offensive cyber operations as well as conventional capabilities is by definition a manifestation of hybrid threats. Potential targets of military cyber attacks include sensors, computer controlled systems (drones, guided missiles, etc.), command, control, and logistics systems, etc.

Remotely controlled or autonomous drones operating in air, land, sea or space domains, rely completely on computers and computer networks to function. As such, they also fall into the domain of cyber operations. An interesting example of a possible cyber attack against such systems occurred in 2011, when Iran was able to manipulate a US drone to land in Iran - in effect, performing a remote hijack of the drone. While technical details are not published, there is speculation that the event involved jamming the control signal (forcing the drone into autopilot mode), as well as jamming and spoofing the GPS signal (tricking the drone into landing at the wrong coordinates). (See, for example, Peterson 2011; Rawnsley 2011)

It is important to note that the cyber operation does not have to cause lasting effects. For example, a short disruption in adversary air defense sensors or control systems may be the only goal of a cyber operation that is preceding an air strike. A potential example of such an event is the Israeli air strike against the alleged Syrian nuclear site in 2007, where Israeli (non-stealth) planes flew the mission without being harassed by Syrian air defense. Potential explanations include built-in kill switches in the Syrian systems or advanced EW capabilities. (See, for example, Adee 2008; Fulghum 2007)

However, the most likely target for military cyber operations is not a drone or any other tactical weapon system. The modern military relies heavily on its logistics and communication systems. These systems are more vulnerable to cyber attack, since they are less mobile and less reliant on custom hardware and software (compared to a drone, for example). Consider the strategic and operational effects of a cyber attack that scrambles the data (in a way that is not easy to restore) in the information systems of a major logistics hub: which container is where, what is in that container, who needs what, when they need it, etc. At the very least there will be delays, which could translate into loss of tactical or operational momentum and lives.

The military must be able to protect their own systems from cyber attacks. However, in many instances the military is reliant on or sharing infrastructure (dual use systems) with non-military systems, such as civilian Internet Service Providers. Therefore, the military must also be concerned with the security of these enabling systems. The problem here is that the military is rarely in a position to actively contribute to their defense. The best approach is to map the dependencies and mitigate the associated risks through cooperation and duplication of service providers.

4.2 Non-military targets

A hybrid actor might also target systems that are not directly linked to the military. It is important to note that such attacks could be in violation of international law, depending on the circumstances. However, there are actors who are not (too) concerned with laws, so it makes sense to explore this from cyber hybrid threat perspective. While the military is typically not responsible for the protection of these systems, it is important to realize that attacks against them can significantly change the conditions that the military has to operate in. For example, an attack against civilian infrastructure could cause civil unrest or a mass evacuation in the area of operations.

If an actor wanted to influence the state or the population in general, then an obvious (although probably unlawful) target would be some Critical Information Infrastructure (CII) system. Our societies and economies are very dependent on CII, which are the information systems that enable and maintain our way of life. For example, systems that are used to control the power grid, water treatment plants, air traffic control and banks. In recent decades, critical infrastructure has become more and more automated in order to increase efficiency. Often this has also increased the attack surface of the information systems within.

CII attacks, if successful, could cause serious harm to human life, (critical) infrastructure, economy, ecology, etc. While most of this is hypothetical so far, it is within the realm of the possible. StuxNet is a great example of a cyber attack that causes physical damage against critical infrastructure. It is exceptional, because it exploits four zero-days (vulnerabilities that are only known to the attacker), disrupts industrial control systems and damages physical devices. While nobody has taken responsibility for the attack, there is general consensus that it was developed and deployed by a state or states. It manipulates the parameters of the variable frequency drives (programmable logic controllers - PLCs) that control uranium enrichment centrifuges. This causes the centrifuges to speed up and slow down repeatedly and rapidly. The two primary effects are that the enrichment process is no longer efficient (the isotopes mix again when the centrifuge slows down) and that the centrifuge may physically break due to mechanical stress (the discs inside the centrifuge may shatter). (See, for example, Falliere, Murchu and Chien 2011, Sanger 2012, The Economist 2010)

The same approach can be used against other types of CII, since PLCs are used everywhere, from elevators to nuclear power plants. However, this does not mean that physical damage is always possible. First, there could be alternative systems or safeguards (for example, brakes on the elevator). Second, the system may not involve destructive forces (for example, banking systems deal with information). It is also clear that physical destruction is not the only outcome that has national security implications. An attack against the banking sector that leads to bank runs can cascade into a serious economic problem for a state - more costly, perhaps, than a bomb.

However, CII is not the only concern when facing a cyber-enabled hybrid threat. Everyday life gets more and more entangled with information technology enabled services and devices. Consider, for example, that your smart phone ‘knows’ your location, schedule, contacts, etc. If soldiers take their smart phones to the field, they and their units can be tracked in real time by technologically savvy adversaries.

But information technology is becoming even more intimate. There are medical devices that are surgically implanted into people - pacemakers, insulin pumps, etc. The problem is that those devices are sometimes very poorly secured. Researchers have been able to remotely manipulate such devices (in lab conditions) in ways that could kill or harm a person. (See, for example, Halperin et al. 2008; Kirk 2012) While there are no known examples of lethal attacks against personal cybernetic enhancements, they should be considered in case a key person has one ‘installed’.

With this in mind, it is clear why a hybrid threat actor might consider offensive cyber operations against non-military targets - the list of potential victims keeps growing and very often these systems are not hardened against dedicated attackers. While the military is not the right entity to provide security for these systems, they may be in charge of disabling or eliminating the source of the attacks. In addition, the military should be ready to provide assistance to local crisis management services, upon request and within the existing legal framework.

4.3 Environment

The third reason to consider cyberspace from the hybrid threats perspective is that it offers new ways of accomplishing tasks that were previously prohibitively expensive or complicated. For example, consider the challenges of global communications and self-organization under oppressive regimes before the widespread adoption of Internet.

While cyberspace did not provide the motivation for the so-called Arab Spring, it definitely had a role in the events. On one side, people used the Internet to gather and share information about their governments, and to self-organize using social media and instant communication tools. On the other side, several governments tried to limit access to the Internet or to specific services on the Internet in order to regain control and to quell the unrest. (Afanasjev 2011)

It is well known to the national security and intelligence community that terrorist organizations use cyberspace to facilitate their operations. The Internet provides accessible, cheap (free) and anonymous ways to spread propaganda, identify and shape recruits, share training materials, gather intelligence, plan and coordinate attacks, etc. (Bardin 2010) However, to date there are no publicly known cases of cyber terrorism - cyber attacks that aim to coerce a population or government through terror.

Criminals and criminal groups are also taking advantage of cyberspace. Identity theft (including theft of credit card information), fraud, money laundering, extortion (for example, by using ransomware that encrypts the victim's data, or by performing distributed denial of service attacks), sale of counterfeit goods, and breaking into bank and electronic currency accounts are just a few examples of criminal use of cyberspace. The relative anonymity and problems with international law enforcement cooperation foster a thriving underground community that operates on a global scale.

Cyberspace is also a useful medium for espionage. Since the vast majority of information is stored digitally, the cyber spy can usually operate remotely. This means that there is very little risk of getting caught (although one might be identified), especially considering that there is no international law prohibiting espionage. From the hybrid threats perspective it allows to even the playing field considerably by 'skipping' the research and development phase on new technologies or by getting advance warning of deployments and capabilities of adversaries.

Intelligence agencies are actively monitoring Internet in the interest of national security. This is a rather passive and defensive form of cyber operations compared with cyber attacks. Still, the signs are clear that many states are preparing for cyber conflict. What we can conclude is that cyberspace is an area of conflict where states act in an apparently more direct way than they would when it comes to conventional means. Getting peoples' and organizations' financial information or destroying uranium enrichment centrifuges in conventional ways would stir up a lot more controversy than it does in the cyber world – much because the problems with attribution. The old definition of hybrid threats that non-state actors wield state actor capabilities seems to be in reverse here, as state actors try to masquerade as non-state actors. There are most likely a host of reasons for this, but again – the lack of strong attribution is a key enabler.

Cyberspace in itself can also be attacked with rather conventional means. Consider the problem of supply side vulnerability. In recent years there have been numerous cases of counterfeit microchips and other hardware, which could also contain hidden flaws, back doors or remote kill switches. National security is at risk as modern missiles, airplanes, and even munitions often have microchips in them. In the future one can imagine nanotechnology applied undercover on sensitive hardware, which might result in faulty or even changed functionality.

Once again, the military is not the primary actor in this field. However, it is very important to stay informed of the opportunities that information technology provides and to embrace them where applicable. In terms of social media, the military must practice good OPSEC on one hand, and STRATCOM on the other hand. For example, NATO homepage has links to the Alliance's presence on FaceBook, Twitter and YouTube, in order to reach the demographic that prefers this type of media. (Newsroom 2013)

5. Conclusion

The national security implications of cyberspace are growing. Many states have recognized the importance of being able to operate in cyberspace – if for nothing else, then to boost their economy. Some states have even started developing military capabilities to ensure freedom of action during military conflicts, while suppressing the adversary's capabilities. It is also widely believed that state actors are very active on the cyber espionage front, although this is done in a clandestine manner.

This new focus on cyber capabilities is mirrored by sub-state actors as well. From individual activists to organized crime to terrorist organizations, these actors are seeking ways to benefit from cyberspace. On the one hand this is about using the myriad services available: communication, information gathering, etc. On the other hand, it is about abusing the services – harvesting personal information, stealing money, disrupting other services, etc. Traditionally, this form of activity has not been of much interest for the military. On the modern battlefield, however, cyberspace enables both prospective allies and enemies, engages the global community with local operations, and creates new modes of operating for the military.

For (relatively) like-minded states, such as members of NATO and EU, it is important to develop a common understanding on the opportunities and risks posed by cyberspace. On the defensive side, international cooperation is required to deter or defeat serious cyber threats, whether military or not. Cooperation between military and civilian (government and private) sector sphere is also required, since most of the CII is not owned and operated by the military, but may impact the capability or operations of the military. Therefore, the military must be ready and eager to cooperate and share with various partners that also have 'cyber power' and can affect the mission.

Acknowledgements

The authors would like to thank the Swedish National Defence College for funding this work.

References

Adee, S. (2008) "The Hunt for the Kill Switch", *IEEE Spectrum*, May. Available at: http://spectrum.ieee.org/semiconductors/design/the-hunt-for-the-kill-switch/0. [Last accessed: 13.02.2013]

Afanasjev, M. (2011) *Approaches to Avoiding Government Censorship, Blockade and Surveillance on the Internet*. Master's thesis, Tallinn University of Technology.

Bardin, J. (2010) *Cyber Jihadist Use of the Internet: What Can Be Done?* Whitepaper.

Council of Europe. (2001). Convention on Cybercrime. Available at: http://conventions.coe.int/treaty/en/treaties/html/185.htm. [Last accessed: 13.02.2013]

Falliere, N., Murchu, L.O. and Chien, E. (2011) W32.Stuxnet Dossier. Available at: http://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/w32_stuxnet_dossier.pdf. [Last accessed: 13.02.2013]

Fulghum, D. (2007) "Why Syria's Air Defenses Failed to Detect Israelis", *Aviation Week*, Oct 3. Available at: http://www.aviationweek.com. [Last accessed: 13.02.2013]

Gorman, S. and Barnes, J.E. (2011) "Cyber Combat: Act of War. Pentagon Sets Stage for U.S. to Respond to Computer Sabotage With Military Force", Wall Street Journal, May 30. Available at: http://online.wsj.com/article/SB10001424052702304563104576355623135782718.html. [Last accessed: 13.02.2013]

Halperin, D., Heydt-Benjamain, T.S., Ransford, B., Clark, S.S., Defend, B., Morgan, W., Fu, K., Kohno, T. and Maisel, W.H. (2008) "Pacemakers and Implantable Cardiac Defibrillators: Software Radio Attacks and Zero-Power Defenses", *Proceedings of the 2008 IEEE Symposium on Security and Privacy*. Available at: http://www.secure-medicine.org/icd-study/icd-study.pdf. [Last accessed: 13.02.2013]

Herkner, L. (2007) "Hacken für den Staat", Zeit Online, May 16. Available at: http://www.zeit.de/2007/21/Sicherheitsplaene. [Last accessed: 13.02.2013]

Ho, J. (2005) "The Advent of a New Way of War: Theory and Practice of Effect Based Operations", Johan Elg (Ed.), *Effektbaserade operationer,* Stockholm.

Joint Publication 3-13. Information Operations. (2006) Chairman of the Joint Chiefs of Staff.

Kirk, J. (2012) "Pacemaker hack can deliver deadly 830-volt jolt", ComputerWorld, Oct 17. Available at: http://www.computerworld.com/s/article/9232477/Pacemaker_hack_can_deliver_deadly_830_volt_jolt. [Last accessed: 13.02.2013]

Markoff, J. (2008) "Before the Gunfire, Cyberattacks", NYTimes.com, Aug 12. Available at: http://www.nytimes.com/2008/08/13/technology/13cyber.html?em&_r=0. [Last accessed: 13.02.2013]

Matthews, M.M. (2008) “We Were Caught Unprepared: The 2006 Hezbollah-Israeli War”, *The Long War Series Occasional Paper 26*. U.S. Army Combined Arms Center, Combat Studies Institute Press, Fort Leavenworth.

Mattis, J. (2008) “Commander’s Guidance for Effects-Based Operations”, *Joint Forces Quarterly,* 51:4, Washington.

Newsroom (2013) North Atlantic Treaty Organization. Available at: http://www.nato.int/cps/en/natolive/index.htm. [Last accessed: 13.02.2013]

Peterson, S. (2011) “Exclusive: Iran hijacked US drone, says Iranian engineer (Video)”, Christian Science Monitor, Dec 15. Available at: www.csmonitor.com/World/Middle-East/2011/1215/Exclusive-Iran-hijacked-US-drone-says-Iranian-engineer-Video. [Last accessed: 13.02.2013]

Rawnsley, A. (2011) “Iran's Alleged Drone Hack: Tough, but Possible”, Wired, Dec 16. Available at: http://www.wired.com/dangerroom/2011/12/iran-drone-hack-gps/. [Last accessed: 13.02.2013]

RT.com (2012) “Anonymous leaks personal information of 5,000 Israeli officials”, Nov 18. Available at: http://rt.com/news/anonymous-israel-officials-leaked-002/. [Last accessed: 13.02.2013]

Sanger, D. (2012) “Obama Order Sped Up Wave of Cyberattacks Against Iran”, NYTimes.com, June 1. Available at: http://www.nytimes.com/2012/06/01/world/middleeast/obama-ordered-wave-of-cyberattacks-against-iran.html?_r=2&pagewanted=2&seid=auto&smid=tw-nytimespolitics&pagewanted=all. [Last accessed: 13.02.2013]

Tallinn Manual on the International Law Applicable to Cyber Warfare (to appear 2013). Cambridge: Cambridge University Press. Draft available at: http://www.ccdcoe.org/249.html. [Last accessed: 13.02.2013]

The Economist (2010a) “A worm in the centrifuge”, Sep 30. Available at: http://www.economist.com/node/17147818. [Last accessed: 13.02.2013]

The Economist (2010b) “War in the fifth domain”, Jul 1. Available at: http://www.economist.com/node/16478792. [Last accessed: 13.02.2013]

Thomas, T. (2007) *Decoding the Virtual Dragon: Critical Evaluations in the Science and Philosophy of China’s Information Operations and Military Strategy*. Fort Leavenworth: Foreign Military Studies Office.

Thomas, T. (2009) *The Dragon’s Quantum Leap: Transforming from a Mechanized to an Informatized Force*. Fort Leavenworth: Foreign Military Studies Office.

Thomas, T. (2011) *Recasting the Red Star: Russia Forges Tradition and Technology Through Toughness*. Fort Leavenworth: Foreign Military Studies Office.

Wiener, N. (1948) Cybernetics: Or Control and Communication in the Animal and the Machine. New York: John Wiley.

Combining Research and Education in Tactical Combat Modelling – 5th Sandis Workshop

Juhani Hämäläinen, Bernt Åkesson and Esa Lappi
Finnish Defense Forces Technical Research Centre (PVTT), Riihimäki, Finland
juhani.hamalainen@mil.fi
bernt.akesson@mil.fi
esa.lappi@mil.fi

Abstract: Modelling and simulation is a cost-effective tool to compare and explore possibilities of different military technologies, equipment and tactical using practices. Information warfare can affect initial situations or outcomes of the combat and tactics to be modelled. This leads to demands for studying alternative tactical and technical situations as well as related model and software development and user education. However, the education and development of new software needs realistic test cases and resources for teaching, improving and developing corresponding software. Learning is more motivated if the teaching methods can utilise corresponding test cases as an introduction to software use. By combining research and education on the topic together, experiences and information exchange on modelling and simulation between research specialists and students can be done effectively. The mentioned facts encouraged us to combine previously separately organised research workshops and education events together into a combined 5th Sandis workshop. The workshop was held in 2012 as a part of Advanced Technology and Military Acquisition Course (ATMAC). The workshop provided a one-week package in operational analysis in the study package in operational analysis methods, which is a total of 4 credits. The research part of the workshop was aimed to deal with two different case studies and testing of a new version of the simulation software. The aim of the educational part of the workshop was to introduce tactical combat simulations with the Sandis software for comparative studies including studies with or without electronic warfare (EW). The research parts were focused on development of indirect fire models in realistic terrain and combat casualty evacuation studies. These studies provided deeper understanding and some new scientific results of the topics. The third goal was to test a new version of the software. Experiences and feedback of the participants showed us that by combining research and education into a common workshop, both aspects win. This kind of a working method seemed to provide intelligent and motivating atmosphere and the focus of the participants were on the subjects under consideration.

Keywords: modelling, simulation, education, electronic warfare

1. Introduction

Relevance of modelling and simulation has increased constantly during past decades. Simultaneously, the diversity of methods, techniques and software has increased as well. In a simulation study, it may be necessary to create and implement mathematical models for a specific purpose. The development of such models tends to be an iterative process. The first phase includes specific calculations with general-purpose software like MATLAB or Excel, in order to prototype new ideas. After this prototype phase, it may prove necessary to implement the model in a general-purpose programming language. Usually, this implementation supports well the given question in the simulation study at hand. However, further studies may require new features and computational models to be included in the software, making a software development project necessary. The development of the simulation software Sandis has followed this process.

Modelling and software development need testing and open research case studies in order to validate and verify the models. Education improves the information exchange between users and developers. Open research cases enable scientific refereeing of reported models and results. It also supports testing of the software by providing relevant test cases. We shall consider experiences in combining education, research and software testing in the case of modelling alternative tactical and technical situations by using the Sandis simulation software. This paper reports on research and education at the 5th Sandis workshop on combat and evacuation modelling held 2012 at Päivölän Kansanopisto.

A workshop enables more a diverse and comprehensive study of the topics than lectures alone. This constructive learning method seemed to work well within modelling and simulation education. The co-operative learning among subject specialists is related to the project working method (Poranen 2009) and it enhances the usability of the silent knowledge by experts (Lehtonen et al. 2006). The workshop was a part of The Advanced Technology and Military Acquisition Course (ATMAC) for officers.

Two research cases were considered: indirect fire and field medicine. In addition, a new version of modelling software Sandis 2 was tested. The previous research workshops and ATMAC studies were held separately during 2008-2011, but recognised synergy led us to organize them parallel.

This combination was found to be an effective working method that allowed specialists, customers and co-operation industry partners to exchange ideas, learn and improve modelling and simulation methods, test software in realistic case studies and achieve new scientific results.

2. The advanced technology and military acquisition course

The Advanced Technology and Military Acquisition Course (ATMAC) (in Finnish Tekniikan Lisäopinnot, TLO) is a 60 credit course given annually by the Department of Military Technology of the Finnish National Defense University. It includes 14 different study packages from military technology, operational analysis and military economy. The course is lectured in one year, from August to July. The number of whole-year participants is limited to 20. The course is aimed at giving officers comprehensive technological knowledge and understanding in diverse topics. Furthermore, individual study packages may be taken and civilians are welcome to take the course as well.

The workshop provided a one-week package in operational analysis as part of the study package in operational analysis methods, which is a total of 4 credits. ATMAC 7, lectured in 2011-2012, was the third time the Defense Forces Technical Research Centre provided a lecture package in operational analysis methods. Previously, however, the length of the lecture package was three days. Expanding it to five days allowed for more time for hands-on exercises and team work.

The workshop contained lectures, exercises and results considerations. The aim of the workshop was to familiarize the students with operational analysis research methods and tools. The tactical level combat simulation software Sandis was used as a teaching tool and was distributed to all students, along with documentation and supplementary material.

The program of the workshop was as follows.

- Introduction to operational analysis
- Combat modeling
- Presentation of Sandis
- Themes and models
- Indirect fire
- Creating, modifying and adding troops and equipment
- Direct fire
- Field medicine
- Electronic warfare (EW)
- Data farming and comparison of alternatives
- Team exercises and presentations

Two assignments were performed as team exercises and the workshop concluded with presentations by the teams.

3. Tactical simulations and modelling software Sandis 1

The Sandis 1 software was designed and implemented at the Finnish Defence Forces Technical Research Centre. Implementation began at 2002 due to a need to study effectiveness of different combat equipment and tactical alternatives at tactical level. After two years of part time working with trial and error, the customer needs were analysed and basic mathematical structure and methods for fast tactical level simulation were ready in 2004. The methods were initially programmed in the mathematical language MATLAB and as Excel spreadsheets and they were used in classified projects during 2005-2006. The first version of the Sandis 1 software was released in 2006. It was implemented in the Java programming language and combined a map user interface with combat models.

The basic features of the Sandis 1 software were published in (Lappi 2006). However, the model development has continued and a number of customer requirements have been met every year. The international cooperation started at the 1st Nordic Military Operational Research Symposium in 2004, where initial mathematical concepts were published (Lappi 2004, Kangas & Lappi 2004). The software has been used in international co-operation and case studies are documented in (Heath et al. 2009), (Bruun et al. 2011), (Åkesson et al. 2012) and (Lappi et al 2012a). The international partners have learned to use the software and given their contribution to the model development.

3.1 Modelling communication connections in Sandis 1

Sandis 1 includes models for communications, which may be disturbed by electronic or information warfare. The Sandis education in the workshop included examples of modelling EW effects on communication and outcome (losses) in combat modelling. Figure 1 shows the visualization of communication connections in Sandis (Pajukanta et al. 2008).

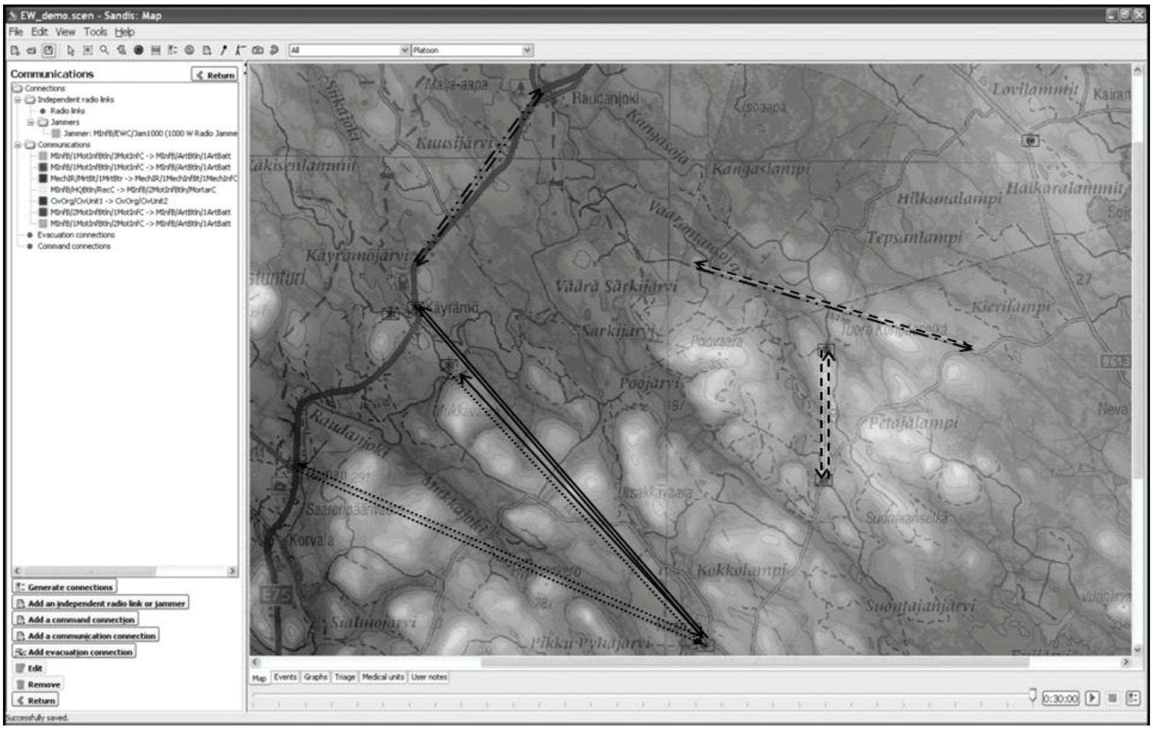

EW_demo.scen - Sandis: Map
File Edit View Tools Help
All
Platoon
Communications
Return
Connections
Independent radio links
Radio links
Jammers
Jammer: MInfB/EWC/Jam1000 (1000 W Radio Jamme
Communications
MInfB/1MotInfBtn/3MotInfC -> MInfB/ArtBtn/1ArtBatt
MInfB/1MotInfBtn/1MotInfC -> MInfB/ArtBtn/1ArtBatt
MechIR/MrtBt/1MrtBtr -> MechIR/1MechInfBt/1MechInfC
MInfB/HQBtn/RecC -> MInfB/2MotInfBtn/MortarC
CivOrg/CivUnit1 -> CivOrg/CivUnit2
MInfB/2MotInfBtn/1MotInfC -> MInfB/ArtBtn/1ArtBatt
MInfB/1MotInfBtn/2MotInfC -> MInfB/ArtBtn/1ArtBatt
Evacuation connections
Command connections
Generate connections
Add an independent radio link or jammer
Add a command connection
Add a communication connection
Add evacuation connection
Edit
Remove
Return
Successfully saved.
Map
Events
Graphs
Triage
Medical units
User notes
0:30:00
Raudanjoki
Kuusijärvi
Kangaslampi
Hilkunalampi
Tepsanlampi
Kieriälampi
Käyrämojärvi
Väärä Särkijärvi
Särkijärvi
Poojärvi
Petäjälampi
Kokkolampi
Suontajanjärvi
Pikku-Pyhäjärvi
Lovilammit
Haikaralammit
Neva
E75

Figure 1: Appearance of active radio links in Sandis 1 window

Lines are as follows: solid line means that the connection is functioning, long dash double dot line (−··) means jammed connections, dotted line (····) means too low signal-to-noise ratio connections and dashed line (−−−) to interfered connections (Pajukanta et al. 2008). Note that connection may have different status depending on the direction. In the software, connections are emphasized by using colour plots.

3.2 Previous Sandis workshops

Sandis educational workshops have been held before. The first and second workshops were co-operation workshops within the development of the software and were targeted to more specific questions. The third and fourth workshops were focused to modelling and research specialists with educational theme. The fourth one was particularly focused on studies of alternative tactical situations. The documentations of these workshops are available in (Hämäläinen et al. 2009) and (Hämäläinen 2011). As we have educated Sandis before in ATMAC, these two events combined to form the 5th Sandis workshop held in 2012. The Sandis workshops have been targeted at a specific audience and have not been open to the general public.

4. Research observations

During the workshop, two different research cases were under study with separate teams. One teams considered the effectiveness of indirect fire in rugged terrain. Another team worked with field medicine modelling. The initial version of the indirect fire model is presented and published in (Heininen 2006). Development, implementation and validation are presented in (Lappi 2012). The improved model, which considers terrain shapes, was now used with high-resolution elevation data obtained by laser scanning (National Land Survey of Finland 2012). The team compared the differences in losses caused by indirect fire in flat and rugged terrain. The team also studied the effects of forest on the losses. The results and detailed description of the model are reported in (Lappi et al. 2012b). The field medicine modelling team used and tested Sandis 2 within wide evacuation scenarios.

New results were obtained during the team works. Improvement of the indirect fire model to consider real terrain data showed the sensitivity of the results with respect to the parameters. These results imply that errors due to the ignoring of terrain effect are bigger than expected (Lappi et al. 2012b).

It was decided not to publish any proceedings from this workshop, due to having only two research teams. Instead, the teams were free to publish their reports at suitable forums.

4.1 Feedback of the workshop

The participants were asked for feedback at the end of the workshop. The experiences of the workshop were collected with a questionnaire, where the participants were asked to evaluate the workshop on eight questions on a scale of 1 to 5 (best), with an added free word field. The questionnaire and the average values of the answers are given in table 1.

Table 1: Feedback questionnaire of the workshop with average values of answers

Overall grade for the workshop, scale 1-5 (1 weak- 5 excellent)	Average grade
The content lived up to my expectations	4,14
The overall arrangements were successful	3,63
The teaching was competent	4,25
The IT arrangements worked well	4,63
The teaching methods were balanced (The proportion of exercises to theory was suitable)	4,20
The size of the teams was suitable	4,07
I can use Sandis in my work	3,38
The venue was successfully chosen	4,00

4.2 Testing Sandis 2 in the workshop

During the workshop, software developers, modellers, subject matter experts and users tested a new version of the software. This broad group represented different needs for the software development. In combination with the application to realistic education and research cases, this proved to be fruitful for the software testing process.

4.3 Lessons learnt

The combination of research, education and software testing was demonstrated in various ways. The education focused on training in using Sandis with selected topics, while research groups made further studies on similar topics by exploring new techniques in modelling and simultaneously testing a new version of the software. The end result was students learning combat modelling, validation and verification of simulation models, and new scientific observations, which are reported elsewhere (Lappi et al. 2012b).

Compared to traditional classroom education, the workshop enabled more guidance for students, since there were more than a dozen experts around to support and help with detailed questions. The educational exercises provided a number of suggestions for improvement of the software. Since there were programmers present, the viability of the suggestions was easily analysed.

The common presentation of the results from the case studies and exercises motivates and increases knowledge exchange between the participants. As there were researchers and military personnel in the audience during the presentations, the feedback had high quality.

5. Conclusions

The workshop contained team exercises for education and more challenging research case studies. Discussions between participants lead to self-learning and consideration of different aspects of the problems. The educational goal was to give a comprehensive introduction to the Sandis software in one week. The time restriction meant that large research problems could not be studied in detail. Instead, the teaching goal was to show how such problems can be treated, e.g. by solving separate smaller questions. These smaller case studies improve the understanding of the problem, and can be used in teaching and show the amount of work needed to completely study the research problem.

The workshop was focused on research and education in modelling and simulation, but several linked areas were identified. These areas are illustrated in figure 2.

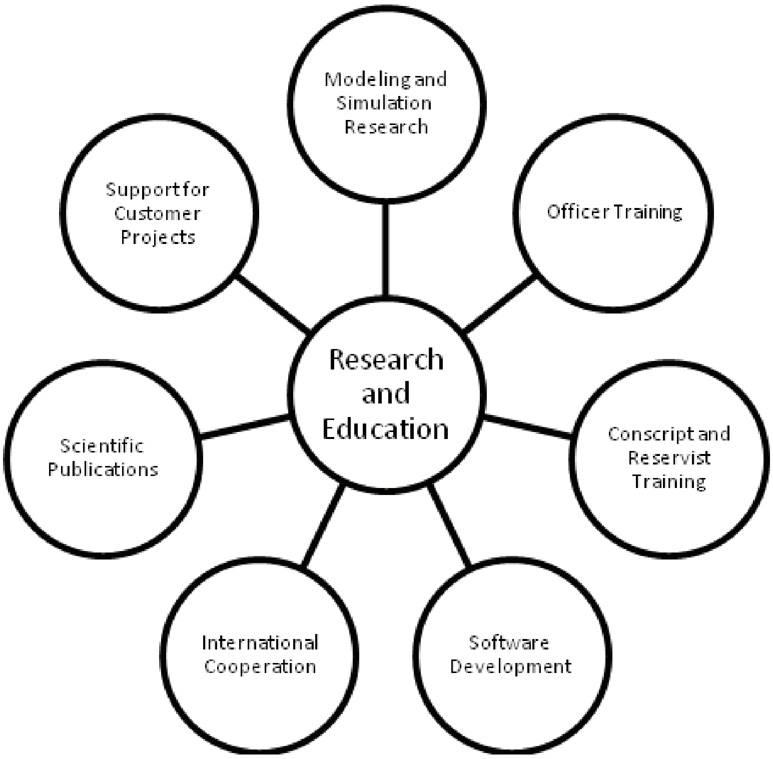

Modeling and Simulation Research
Support for Customer Projects
Officer Training
Research and Education
Scientific Publications
Conscript and Reservist Training
International Cooperation
Software Development

Figure 2: Illustration of the topics connected in the workshop.

The topics for the case studies and team exercises can be selected to support customer projects, as long as customer confidentiality is not breached. The target audience for the education may be quite broad, including besides officers also conscripts and reservists, as well as civilian researchers. Moreover, the results from research carried out during the workshop may merit scientific publication.

This kind of workshop supports research with increased resources. However, with a compact schedule it requires careful advance planning and timing, in order to simultaneously achieve research goals. In our case it proved to be efficient to combine research, software development and education.

Acknowledgements

Authors wish to thank all the participants of the workshop and Päivölän Kansalaisopisto for providing well functioning facilities and positive atmosphere for the workshop.

References

Bruun R.S., Hämäläinen J.S., Lappi E.I., Lesnowicz Jr E.J., (2011) "Data farming with SANDIS software applied to mortar vehicle support for convoys", *Proceedings and Bulletin of the International Data Farming Community*, Issue 9, pp. 18-21.

Heath, S., Dolk D., Lappi, E., Sheldon, B., Yu L. (2009) "Investigating the use of simulation tools for mass casualty disaster response", *Proceedings and Bulletin of the International Data Farming Community* Issue 6, pp. 22-25.

Heininen, T. (2006) "A method to calculate the lethality of fragmenting ammunition", *Lanchester and Beyond - A Workshop on operational analysis methodology*, J.S. Hämäläinen (editor). PVTT Publications No. 11, pp. 19-30.

Hämäläinen, J.S., Lappi, E. I., and Åkesson B. M. (Editors). (2009) *Third International Sandis Workshop*, PVTT Publications No. 19.

Hämäläinen, J.S. (Editor) (2011) *Fourth International Sandis Workshop*, PVTT Publications No. 23,.

Kangas, L. and Lappi, E. (2004) "An example of Markovian combat modeling", *Proceedings Nordic Military Operational Research Symposium*, Sirpa Korpela and Tapio Heininen, (editors), Defence Forces Technical Research Centre.

Lappi, E. (2004) "A method to evaluate operational performance of electronic warfare systems". *Proceedings Nordic Military Operational Research Symposium*, Sirpa Korpela and Tapio Heininen, (editors)., Defence Forces Technical Research Centre.

Lappi, E. (2006) "SANDIS -ohjelmisto - todennäköisyysjakaumilla laskeva taistelumalli tutkimuskäyttöön", *Insinööriupseeri 2006 - Insinööriupseeriliiton juhlajulkaisu*, pp. 92-96.

Lappi, E. (2012) *Computational methods for tactical simulations*, Doctor of Military Science thesis, National Defence University, Department of Tactics and Operations Art, Publications 1, N:o 1/2012.

Lappi, E., Heininen, T., Yüksel Ergün, I., Hörling, P., Hinshaw F., Lappi, T., Ots, K. (2012a) "Future indirect fire cost effectiveness", *Proceedings and Bulletin of the International Data Farming Community,* Issue 11, pp. 7-10.

Lappi, E., Sysikaski, M., Åkesson, B., Yildirim, U.Z. (2012b) "Effects of terrain in computational methods for indirect fire*",* *Proceedings of the 2012 Winter Simulation Conference*, C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. M. Uhrmacher, (editors). Berlin.

Lehtonen J., Kanerva-Lehto, H. ja Koivisto, J.. (2006) ”Tutkimuspaja mahdollisuutena yhdistää opetus ja T&K” *Turun ammattikorkeakoulun puheenvuoroja* 24, Turun ammattikorkeakoulu, Turku.

National land survey of Finland 2012. "Laser scanning data", http:/www.maanmittauslaitos/en/node/3177.

Pajukanta, S., Åsen, W., Sainio, J., Åkesson, B., & Lappi, E. (2008) ”The electronic warfare model in operational analysis tool Sandis”, *The 2nd Nordic military OA symposium 2008,* Stockholm: FOI.

Poranen, T. (2009) ”Projektiopetuksen ja tutkimuksen yhdistäminen” http://www.cs.tut.fi/tapahtumat/projektiopetus09/TimoPoranen_11082009.pdf, valid 1.3.2013.

Åkesson B., Horne G., Mässeli K., Narayanan F., Pakkanen M., Shockley J., Upton S., Yildirim Z., Yigit A. (2012) "MSG-088 Data farming case study on humanitarian assistance / disaster relief", *Proceedings and Bulletin of the International Data Farming Community* Issue 11, pp. 3-6.

From Auftragstaktik to Comprehensive Approach: Key Leader Engagement in Strategic Communication

Arto Hirvelä[1], Aki-Mauri Huhtinen[1] and Tommi Kangasmaa[2]
[1]National Defence University, Helsinki, Finland
[2]Defence Command, Helsinki, Finland
arto.hirvela@mil.fi
aki.huhtinen@mil.fi
tommi.kangasmaa@mil.fi

Abstract: There is a huge variety of perception management tools in the cyber security environment. War no longer seems to be about international politics and trade; instead, it is now perceived as a struggle within an omnipresent network which nobody is able to escape from. The different networks have enabled multiple actors – from states to non-governmental organizations, from individuals to the mafia – to organize themselves and cooperate. The globalization of the economy and the multiplication of information networks have also enabled and increased the probability of cyber attacks and propaganda. Encountering these kinds of problems demands comprehensive planning. The Western military today is not sufficiently organized, trained, or equipped to analyze, plan, coordinate, and integrate the full spectrum of capabilities available to promote national or global interests. The Western military is not the asset it should be in the security policy and crisis management. Security policy development and crisis management are not determined primarily by the military or by military interests, but by considering a variety of factors including social, economic, ecological, and cultural conditions. It is therefore normally not possible to guarantee security through unilateral national action, or with armed forces only. Given the complexity and interdependency of different actors and nations in crisis management, it is necessary to achieve greater harmonization among all appropriate actors in the analysis, planning, management, and evaluation of interventions in complex contingencies and emergencies. From the military perspective, there is a need for clear political guidance in crisis management, but also a need for operational freedom at every level. According to the German Army Command and Control Regulation, *Auftragstaktik* is based on mutual trust and demands from each soldier that, in addition to the conscientious performance of duty and willingness to achieve the objectives ordered, he or she is prepared to accept responsibility, to cooperate, and act independently and resourcefully in accordance with the overall mission. The command grants subordinated commanders freedom in the way they execute the mission. In crisis management, the military tries to turn political objectives into actions in the operational area. The military needs the freedom to communicate – in accordance with political guidance – with different audiences during the operation. Because there are many discrepancies in how communication takes place in different networks and because we increasingly use social media tools, we need synchronizing themes, messages, images, and actions that contribute to Strategic Communication across the joint force. These can be critical to mission accomplishment. Key Leader Engagement can help the military command avoid the crisis of perception management in the form of a "say-do gap" in a critical media environment. As a tool for implementing a communication strategy program, the employment of Key Leader Engagement cells has ensured that whenever commanders meet with leaders, they deliver an effective, consistent message that supports the command's goals. In this paper the Comprehensive Approach (CA) model is discussed as a continuum of the classical *Auftragstaktik* and the basic ideas of the Comprehensive Approach concept, such as Strategic Communication (SC), Key Leader Engagement (KLE), and Combat Camera (CC).

Keywords: key leader engagement (KLE), key leader enhancement (KLEN), strategic communication, core narrative, comprehensive approach

1. Introduction

The nature of war is based on several contradictions and paradoxes. The principle of continuity suggests that success must be exploited relentlessly at every stage. At the same time, the culminating point of attack will lose momentum even as it succeeds. The commander must therefore know when to stop his advance. (Handel, 1992: 181-182) The basic paradox of war is that an actor wishing to terminate a war should not appear too eager or too weak. Also intelligence may be reliable on the strategic level but not on the operational; surprise may be readily achieved on the tactical level but not on the strategic; and defence may be stronger on the tactical and operational levels but not necessarily on the strategic. (Handel, 1992: 8)

Conceptualizations of war were based largely on traditional Clausewitzian definitions. Political and/or military power was to be used for compelling the enemy to consent to 'our' will. After the Cold War, the key conceptualization in the US became *Revolution in Military Affairs*, which was based on a non-idealistic or non-normative way of war. (Raitasalo & Sipilä, 2006) Instead, it used war as a test framework for the western technological revolution. The new cyber war was a promise of systematic and controlled use of information (bytes, messages, etc.) At the same time, the new military organization had to become agile and resilient (Liblicki 2011).

”Scientific explanations in the natural sciences are always of the as-if kind”, said Csikszentmihalyi (2000: 9). They do not presume to stand for the reality explained. They are just models that help to simplify the behaviour of the reality under study; they do not claim to represent the wholeness of that reality in the abstract form. This construction can also be seen between a psychological operation and Key Leader Engagement. Many times in PSYOPs only leaflets have been distributed without measurement – but with the hope that the abstract message would be understandable for the target audiences. In KLE, however, the actors have to step forward and face their real adversaries. Body language and nonverbal behaviour are more important than purely textual messages. Everyone can immediately see the reactions to the communication without any doubt.

The core questions of the assessment in CA are the following: are the operations progressing according to plan and are they approaching the anticipated outcomes? If not, what needs to be adjusted? Traditionally, these questions have been addressed by reports through the chain of command and these reports have usually been of a qualitative nature and based on the subordinate commanders' skill and experience. In rational western military culture, an action has been based on quantitative measures of materiel military power and assessments. With evidence-based quantitative methods, one gets an abundance of data, but there have been great difficulties in linking the data to the high level goals. This means CA is not understood as the key to the parallax view. The old rational military leadership as a management style saying that everything that counts cannot be counted and everything that is counted does not count still holds true. The question is to find a balance between experience-based (leadership) and evidence-based (management) methodology.

2. Comprehensive approach

In general, there is no unanimous definition of Comprehensive Approach (Johnson, 2010: 8). For example, it has been identified, as a philosophy, a network of actors, a means to an end, conflict resolution, or as a network of key drivers or actors in the information system. Moreover, it has been designated as the model for crisis management, counter-insurgency, or the prevention of conflicts. In this paper, CA is understood as a planning model metaphor that is used before, during, and after conflicts. In addition, this paper comments only on the US and NATO planning model, which emphasizes the role of the armed forces as the leading actor in operations and leaves out the comprehensive approaches of, for example, the UN, the EU, or other national governments.

The essential foundation of any comprehensive approach is recognizing the need to combine the efforts of multiple actors (both civilian and military) operating in contemporary, complex conflict environments in order to produce a coherent and effective response to the situation (Johnson, 2009: 8–9). This combination ought to take place on several, overlapping levels, for example, at the inter-organization, intra-organization, and national levels, or at the politico-strategic, military-strategic, operational, and tactical levels (Rintakoski & Autti, 2008: 25; Simon & Duzenli, 2009).

Security policy development and crisis management are not determined primarily by the military or by military interests, but by considering a variety of factors including social, economic, ecological, and cultural conditions. It is therefore normally not possible to guarantee security through unilateral national action, or with armed forces only (MNIOE, 2009: 11). Given the complexity and interdependency of different actors and nations in crisis management, it is necessary to achieve greater harmonization among all appropriate actors in the analysis, planning, management, and evaluation of interventions in complex contingencies and emergencies (MNIOE, 2009: 12).

The comprehensive approach seeks the mutual engagement of governmental, non-governmental, and multinational actors from strategic to functional level (MNIOE, 2009: 3). It incorporates governmental and non-governmental actors in order to shape a regional environment and create stability. It seeks to broaden the context of pre-crisis, crisis, and post-crisis management by comprehensively engaging all relevant ministries, agencies, and organizations with a shared assessment and common goals in an interagency framework (MNIOE, 2009: 11). The comprehensive approach includes clear and achievable strategic-political guidance to ensure that the desirable early cooperation or collaboration between interagency actors works toward common aims. It requires compatible approaches to planning and implementation across organizations. In addition, flexibility is necessary, and approaches to analysis, planning, execution/management, and assessment/evaluation need to be tailored based on the organizations involved and the situation on the ground (MNIOE, 2009: 12).

The comprehensive approach could be considered an extension of all military *Auftragstaktik* that require flexibility and guidance. According to the German Army Command and Control Regulation, "*Auftragstaktik* is based on mutual trust and demands from each soldier that, in addition to the conscientious performance of duty and willingness to achieve the objectives ordered, he or she is prepared to accept responsibility, to cooperate, and act independently and resourcefully in accordance with the overall mission. The command grants subordinated commanders' freedom in the way they execute the mission" (Wittmann, 2012: 5)

The style of *Auftragstaktik* leadership is based on the strong sense of personal responsibility. Also the phenomenon of delegation, self-regulation and trust, freedom of action, participation, and flexibility are linked to the concept of *Auftragstaktik* as a comprehensive leadership and management. (Wittmann, 2012: 15) *Auftragstaktik* is heuristic and favors the use of a decentralized command system, because there is a limited time as to how long the general military plan can survive contact with the enemy. (Wittmann, 2012: 22) This is one reason why the high technological military systems try to avoid open contact with the enemy today. You cannot base your way of manoeuvre as planning any more. You lose you mathematical-rational paradigm of waging war. Like many great strategists have said (see Heuser 2010 and Hill 2010), a plan would just endure until the operations have begun, then the plan had to be adapted by ad-hoc decisions and up-dates based on real-time events following the core intent of the strategic commander. The bottom-up feedback process by the tactical and operational commanders, including the strategic plan if necessary, could be revised and re-orientated. (Wittmann, 2012: 22)

3. Strategic communication

Global communication and the 24/7 flow of information no longer stops at the national borders. The networked society does not stop for a moment, which blurs the boundary between the beginning and the end of issues. Strategic Communication is an example of how states and military organizations aim at continuous, simultaneous, and parallel participation at all operational levels, ranging from the strategic-political to the tactical. (Hirvelä & Huhtinen, 2012: 9)

Strategic Communication means that the sharing of information is organized in a manner that minimizes the possibility of a possible media spectacle as the result of a military operation. With their actions, the armed forces represent the domain of meaning, mediatisation and over-interpretation. SC organizes and coordinates the entire field of information flows. (Hirvelä & Huhtinen, 2012: 9)

Strategic communication is a concept that unites efforts of governmental organizations to influence intended key audiences in support of national interests. The concept tries to answer the challenges posed by changes in the information environment; the increased flow of information; the increased number of networks and reach of media; the increased value assigned to information, and the greater impact of e-media.

The military needs to have the operational freedom to execute operations. From operational to tactical level, there is a need for *Auftragstaktik* to enable flexible execution of a mission. At the same time, there is a need to maintain our reputation as a credible and trustworthy actor. Therefore, because there are many discrepancies in how we communicate in different networks and because we increasingly use social media tools, we need synchronizing themes, messages, images, and actions that contribute to SC across the joint force and that can be critical to mission accomplishment.

Some basic concerns of the art of war are the value of intelligence, the utility of deception, the feasibility of surprise attack, and the possibility of reliably forecasting and controlling the course of events on the battlefield. The key choice for military command is between being more rationale or using a genius's artistic intuition. The need to preserve the professional autonomy of military actions is equally a crucial issue. The military command and control means the balance between political control and military free operational manoeuvre actions. (Handel, 1992: 177)

Strategic Communication is also an important tool for managing organizational or national reputation. Reputation is something that is talked about, a story that spreads. The centre of gravity for a reputation is what the story is, whether the viewpoint is favourable or not. Between corporations and companies the important thing is a positive distinction between one another (Aula ja Heinonen, 2002: 36). The difference between the identity and reputation of an organization is that identity is the organization's presentation of itself to its various stakeholders and the means by which it distinguishes itself from all other organizations. Identity must be communicated for it to become reputation.

Reputation matters because people make decisions based not only on reality, but also on their perception of reality (Fombrun & Van Riel, 2004: 2). The importance of reputation is emphasized by the change in operational environment of military and non-military actors alike. Global networks have made possible the rapid spread of information but also the spreading of rumours. Organizations have to take more responsibility for their actions because they affect a larger crowd than before and are made transparent by uncontrolled informers. Even organizations that do not operate globally may have an international audience observing their actions (Aula & Heinonen, 2002: 277). Organizations need new tools to effectively deal with the 24/7 reality of new media. Inaccurate information often defines stories about organizations before they have a chance to shape the debate.

Reputation management is not just about communication, it is also about actions. Reputation can be effectively managed by meeting with stakeholders and organizations (Aula & Heinonen, 2002: 90–91). This makes KLE an important tool for reputation management.

3.1 Key leader engagement – new solutions?

Key Leader Engagement (KLE) can help the military command avoid the crisis of perception management in the form of a "say-do gap" in a critical media environment. The problem with KLE has been that these engagements frequently take place in an ad-hoc fashion and are rarely incorporated into other operations and strategies. In military organizations the KLE cell oversees KLE processes and includes representatives from Public Affairs, the Planning Division (J5), Information Operations, and Civil Affairs. KLEs are designed to support a menu of Communication Strategy, Information Operations, Public Affairs, Psychological Operations, and Defence Support to Public Diplomacy objectives. The cell develops a detailed background briefing on each key leader, and then suggests specific approaches to convey the command's overall theme for encouraging support for stability and reconstruction activities. As a tool for implementing a communication strategy program, the employment of KLE cells has ensured that whenever commanders meet with leaders, they deliver an effective, consistent message that supports the command's goals. In Afghanistan, the Force Strategic Engagement Cell (FSEC) used the weekly Engagement Synchronization Meeting to synchronize its KLE targets with engagements conducted by other directorates, diplomatic missions in the country, and other agencies who had interests in those targets. Representatives of these organizations were invited to attend the meeting and discuss their own engagements as FSEC personnel introduced their plans for KLE. (Hull, 2009: 20)

KLE is not just any meeting with influential persons but a planned activity with decided objectives. Still, KLE is something that any person from a commanding staff should be able to conduct. The important thing is that it needs to be planned, prepared, have clear objectives, and documented. A meeting plan should include proposed messages and desired effects, issues to be avoided, information gathering, and an assessment of the meeting (NATO Bi-SC, 2010: 114). The person conducting the meeting needs to have an Influence Briefing Package that may include key participants, desired effects, key message, considerations and information to gather. Persons conducting KLE should understand that effects achieved will not probably be immediate, but develop in the longer run. Commanders cannot look at a single engagement with a single effect as a possibility; instead, they should view an engagement target as one with which a series of engagements might occur as the relationship develops (Hull, 2009: 22).

A KLE matrix is a visual representation of planned KLEs at the each level of command. At a minimum, the matrix should include who should conduct the engagement and the level of engagement, and possibly how this engagement should be conducted (face to face, via telephone, VTC, etc.). It is also recommended that the engagement activities planned by other levels of command be included in order to provide a comprehensive view of the KLE activities at all levels. (NATO Bi-SC, 2010: 53)

Dialogue during KLE can be useful, but must be filtered through an understanding of the perspective, position, and/or agenda of the leader engaged. Information is most useful when it comes from sources that can be trusted to tell the truth, such as people with whom we have built genuine relationships over time. Social media sites (i.e. blogs, Facebook, Twitter, etc.) offer another means of dialogue and should be considered when developing communication strategies. However, an understanding of the operational environment is essential as access to these new sites is limited in many non-western or underdeveloped countries. This would limit the dialogue to populations that are typically urban, with a good income, etc., and using social media would only give you dialogue with a small upper segment of the society in question.

There has been confusion as to whether key leaders are commanders from our organization or targeted influential persons from other organizations. For example, in the Combined Joint Staff Exercise in Sweden 2012 there was a lot of discussion in the Land Component Command (LCC) whether KLE means Sweden's own commanders or leaders from other organizations that they need to meet. In the end it was decided that the key leader was the commander LCC. That limited possibilities for active engagement and influence.

Militaries are trained to be effects-oriented, and often the expectation is that those effects will be immediate. These expectations tend to result in impatience with establishing and building the relationships necessary to uphold useful dialogue. All too often, commanders expect that the desired outcome of a relationship or KLE can be achieved in a single engagement or with few engagements. It is also a mistake to think that it is always direct engagement with an insurgent leader, or important members of the government, that is the best approach since it is the most direct approach. (Hull, 2009: 35)

The KLE concept fails if mutual respect is not achieved between communicating leaders. To engage effectively, one has to have a solid reputation that is valid and respectful within the cultural environment where the action takes place. This reputation is easily lessened or even destroyed by means of negative information dissemination.

3.2 Key leader enhancement

In order to protect the reputation of one's own key leaders, attention should turn to a concept called Key Leader Enhancement (KLEN). Key Leader Enhancement is a preventive and proactive process where the aim is to lessen the vulnerabilities and emphasize the strengths of your own leaders before their reputation is attacked. The KLEN strategy aims to educate one's own key persons to communicate within the respective cultural environment in a way that shields their reputation and shows respect.

NATO considers this to be an important element of the main message. The NATO Information Operations Reference book (2010: 17) states that: The analysis has to identify how the Alliance is perceived as a whole, and also how the individual NATO member or partner nations are seen by the local population. It is fundamental for the credibility of an actor that he or she is respected by the opposing key leader.

In the first stages of operational planning, the info-ops staff has to analyze and prepare the key leaders who best upholds friendly forces morale and who will be deployed to negotiate with the opposing force's key leaders. The NATO Public Diplomacy Handbook emphasizes the importance of cultural awareness. There is a need to give IO officers' and commanders' comprehensive cultural training so they can tailor the right message (Kelton, 2006: 2).

The important idea is still missing in the handbook; the inherent message that is woven into the person from the very moment he or she is raised into the high position. Personal history is inescapable. Everything important said or done in a high position is archived on the internet. Those are the building blocks for our credibility. When preparing for first meeting with the opposing key leader, there is surety that personal history is checked and the first and most important picture of the participants is already built. The worst case scenario is that leaders are not taken as a credible and respected communicator which leads to negative consequences or even to a situation where influencing leader will be a persona non-grata.

Reputation can and will be attacked. Personal reputation is easily attacked through stories, pictures, and videos. This method is widely used in western media to win elections, push back opposition, or to create business opportunities.

Cultural differences are the easiest way to attack. As an example, the present operation in Afghanistan makes huge demands on cultural awareness. Alcohol, sexuality, religious and political views, and the attitude to women's rights are easy targets for defaming leaders. Mistakes from the past can be taken back into the discussion. Cultural aspects that are condemned in the local culture are those that will form the opponent's key messages.

The easiest, and very convincing method, is a visual attack. A still picture can deceive more than a thousand words. Videos are just still pictures multiplied. Framing, choosing the right place and time, or even manipulating pictures and videos are effective ways of giving a totally negative impression of the person who is targeted. The results can be so severe that no one wants to make any official or unofficial contact with that person. The only way to defend one self against such attacks is proactive counter information.

To counter these threats, the decided story needs to be told first. The first telling the story has the right and opportunity to tell the story in the right context. The story that has already been told is the strongest one. It gives possibility to refer to something that has already been big news.

A credible defence can be built. Today's media is full of stories. In this environment, the stories that are short, interesting, and delivered in small amounts are the winning ones. The KLEN approach is a new key tool for proactive information protection. A steady flow of stories related to the key leaders, building a narrative that counts, and continuous reputation management ensures that the leader who engages the key leaders of future areas of operation can rely on their personal reputation. These stories are easily built and disseminated through visual short stories distributed via global media.

Coordination between the IO staff in the operation and planners at home is essential. Those who are in the operation can start to build the reputation of the future commander by telling the story in advance. Those at home start to build story blocks that can be seen by the key leaders who will be engaged. The effect of coherent stories told by the internet and by mouth is very convincing.

The weakness is that leaders who will be deployed to be in key positions are not accustomed to the life-story building method. It takes a lot of time and cultural knowledge to build that kind of a media campaign. It is also hard to admit one's own weaknesses that will be told in the story.

If the story is told right, the engaged key leader will already have respect for the commander because the commander is known for his or her respect for the local culture. The story tells about a credible and just person whose aims to strengthen relations between two sides. It tells a story that is mentally written in the key leader's own cultural language.

3.3 Combat camera

One of the new tools in the Finnish Defence Forces IO toolbox is Combat Camera. Combat Camera consists of several professional soldiers who are also media professionals and conscripted media trained personnel. The new capability supports information operations, public affairs, intelligence, and legal advisors. Combat Camera troops are specialized infantry who are capable of going to the front line and delivering material quickly via mobile internet, satellite, or C3-systems.

Combat Camera conscripts are recruited from all Services – the Army, the Navy, and the Air Force. The Combat Camera unit takes recruits through intensive combat-related media training. Training includes interview techniques, tactical training, IO training, new media training, crime-scene documentation training, special environment training in all Services and aerial shooting training. After 12 weeks of special training, the Combat Camera troops are deployed to the field. When they finish their military service, the troops are transferred to the reserve. In the reserve, they can apply for international operations as Combat Camera.

Active Combat Camera personnel are taught a deeper understanding of psychological operations. They train with other officials, international colleagues, learn special infantry tactics, and they are given warrant officer training in three separate phases. They deploy regularly to international operations.

Today you cannot get away with lying. Facts come up faster than you can deny. Today you have to choose between truths. This is where you can use this tool effectively. Produced material can and will be delivered quickly to domestic and international audiences. The present media environment requires professionals who are capable of strengthening the strategic message from the moment they push the record button.

4. Conclusion

Kinetic effect remains the main task of armed forces, because the largest portion of resources is used for kinetic weapon systems. It is for the non-kinetic effect to intensify, support, and duplicate the kinetic effect. Strategic communication includes the means and tools for non-kinetic effect. In operational planning and the definition of kinetic targets, it is possible to make use of the non-kinetic concept that allows for the evaluation of the effectiveness of a kinetic target in relation to strategic communication. This kind of activity requires thinking that adheres to the comprehensive approach model.

The Comprehensive Approach and *Auftragstaktik* are concepts that try to bring flexibility and the capability to react and fulfil both military and political objectives. Even though the military needs the freedom to operate, this must be a controlled freedom from the perspective of reputation management and influencing. The comprehensive approach includes clear and achievable strategic-political guidance to ensure that the desirable early cooperation or collaboration between interagency actors works towards common aims. Strategic Communication also gives the political and strategic boundaries within which there is a freedom to operate. Key Leader Engagement needs to consider both the restriction set by Strategic Communication and take advantage of the flexibility provided by *Auftragstaktik*.

Key Leader Engagement is an effective tool to gather information, manage reputation, and influence selected audiences. It is important to enlarge the toolbox for influencing. The Finnish Defence Forces are building a capability called Combat Camera. This type of capability is used increasingly by western military organizations today. The structure and goals vary from one country to another, but in essence the strengths and weaknesses are the same. An inherent visual production capability gives a good possibility to build credible stories of those leaders who will be sent to operations.

As a conclusion it can be argued that Strategic Communication tool, KLE transforms speech or communication into immediate effects. Interaction with the target is instantaneous and takes place on all sensory levels, including non-verbal body language. In PSYOPS or emails, for instance, this is not the case, and the verification of comprehensive effects is difficult. In addition, a KLE event can make use of cultural symbols more effectively than PSYOPS products.

The role of CC is to extend the KLE event as well as to duplicate and distribute the message. At the moment CC has not yet been integrated into KLE. The organization and planning of CC as an element of KLE should be subject for further research.

References

Aula, P & Heinonen, J (2011) *M2 Maineen uusi aalto*. Helsinki: Talentum Media Oy.

Borg, J (2007) *The Art of Influencing People*. FT Press, New Jersey.

Csikszentmihalyi, M (2000) *Beyond Boredom and Anxiety*. Jossey-Bass Publishers, San Francisco.

Fombrun, CJ & Van Riel, CBM (2004) *Fame and Fortune: how successful companies build winning reputations*. Upper Saddle River: Pearson Education, Inc.

Handel, MI (1992) *Masters of War. Classical Strategic Thought.* Frank Cass, London.

Heuser, B (2010) *The Evolution of Strategy. Thinking War from Antiquity to the Present.* Cambridge University Press, London.

Hill, C (2010) *Grand Strategies. Literature, Statecraft, and World Order.* Yale University Press, London.

Hirvelä, A & Huhtinen, A-M (2012) Strateginen Kommunikaatio informaatioyhteiskunnan kansallisten tavoitteiden edistämisessä - realismista sosiaaliseen konstruktioon, *Tiede ja Ase No. 70*, Helsinki: Hakapaino Oy

Hull, JF (2009) *Iraq: Strategic reconciliation, targeting, and key leader engagement*, article available http://www.strategicstudiesinstitute.army.mil/pubs/display.cfm?pubID=938 accessed 9 Jun 2009

Johnson, T.F. (2010) “The Comprehensive Approach and the Death of the Term ‘EBAO’” *The Three Swords Magazine,* Issue No. 17, pp 9–13, http://www.jwc.nato.int/files/17_10_Magazine.pdf, accessed 11 Jan 2012.

Libicki, MC **(**2011) “Cyberwar as a Confidence Game” *Strategic Studies Quarterly*, Spring 2011, pp. 132-146.

Kelton, R (2006) The Culture Variable in the Influence Equations, Draft 3.0.

MNIOE Applied Concept (2009) *The Military Information Operations Function within a Comprehensive and Effects-Based Approach*, Final Draft Version 3.0, Bonn: MNIOE.

NATO Bi-SC (2010) *Information Operations Reference Book*, Version 1

Raitasalo, J. and Sipilä, J. (2006) "Reconstructing war after the Cold War". In Jeppsson, T. and Mikkola, E. (eds.) *Perspectives on the Evolving Nature of Military Power*, Finnish National Defence University, Department of Strategic and Defence Studies, Helsinki. Series 2, Research Reports No. 36, pp 1–24.

Rintakoski, K & Autti, M (2008) eds *Comprehensive Approach. Trends, Challenges and Possibilities for Cooperation in Crisis Prevention and Management. Seminar Publication*. Crisis Management Initative. Helsinki: Edita Prima.

Simon, G. and Duzenli, M. (2009) "The Comprehensive Operations Planning Directive" *NRDC-ITA Magazine*, Issue No. 14, pp 16–19, http://www.nato.int/nrdc-it/magazine/home.html accessed 12 Jan 2012.

Wittmann, J (2012) *Auftragstaktik – Just a command technique or the core pillar of mastering the military operational art?* Carola Hartmann Miles – Verlag, Berlin.

Preparing for Cyber Threats in Companies With Information Security Policies

Ilona Ilvonen and Pasi Virtanen
Tampere University of Technology, Tampere, Finland
ilona.ilvonen@tut.fi
pasi.virtanen@tut.fi

Abstract Contemporary companies in any industry are increasingly dependent on information systems. The information systems may have legacy elements, even from the time when an internet connection was not common for every company. Today most companies are online all the time, and their internal systems are used in environments that are already or easily connected to the internet. The amount of devices connected to the internet is approximated to grow fivefold by the end of the decade (Evans 2012). The internet population is estimated to over two billion individuals at the moment (James 2012). Cyber threats are usually discussed from the perspective of national infrastructures and national safety. Although national infrastructure comprises organizational infrastructures, the operations of companies many times do not follow national borders. Companies operate in a truly global environment. Hence, the threats that companies face in their operations are not national, they are global. The weakest link in a global network may prove to be dramatic for the whole entity. Companies use information security policies and procedures as a tool to manage their information security. The policies usually address multiple threats that the information of the company is facing. Most of the threats addressed are direct threats to the company. Another type of threat is the use of one organization as a passage to another. The problem is to recognize the possible indirect effects and to prepare for them as well. The paper analyses threats and their potential effect on the operations of different companies with the use of scenario analysis. The scenarios are built based on a literature review on cyber threats. One outcome of the analysis is that to a company it is irrelevant where a cyber threat originates from and who it is targeted for. If the threat is specifically targeted to the company or if the threat is collateral in nature is not important; preparing for the threat is important in both cases. The paper discusses the pressures that the cyber threats pose for information security policies and raises a question of whether information security policies are used in companies to prepare also for cyber threats.

Keywords: security policy, cyber threats, companies, information security management tools

1. Introduction

Contemporary companies in any industry are increasingly dependent on information systems and connections between them (ICT). The used information systems may have legacy elements, sometimes even dating back to the time when an internet connection was not common feature. Today most companies are online all the time, and their internal systems are used in environments that are already or easily connected to the internet. The internet population is estimated to over two billion individuals at the moment (James 2012). The amount of devices connected to the internet is approximated to grow fivefold by the end of the decade (Evans 2012). Some of these users are not there with good intentions.

According to a definition cyber threats are Internet-borne activities that may harm or have the potential to harm a computer or network and compromise the confidentiality, integrity, or availability of network data or systems (CCIP 2013). In public cyber threats are often discussed from the national infrastructure's and national safety's perspectives. The operations of companies do not always follow national borders even though their organizational infrastructures are subjected to one national infrastructure. Companies sometimes operate in a truly global environment. Hence, the threats companies face in their operations are not national, they are global. In the national cyber strategies it however seems, that the operation of companies is assumed to abide by national boundaries.

Companies should use internally confirmed information security policies and procedures as tools to manage their information security. The policies usually address multiple threats that the information of the company is facing. Most of the threats addressed are direct threats to the company. However, it is important to understand the complex nature of the cyber dimension and not to be short-sighted in this regard. The problem is to recognize also the possible indirect effects and to prepare for them as well. "If it runs on computers and computer networks, it's a potential target" says the chairman of the U. S. government's subcommittee on Cybersecurity, Infrastructure Protection, and Security Technologies, Patrick Meehan (2013). The paranoia-arousing question is: what does not run on computers? Companies are increasingly reliant on computer systems for all activities

The purpose of this paper is to explore the cyber threats that businesses may face and how they can prepare for the risk in advance in their information security policies. The research question in this paper is ‘What challenges does the cyber dimension of threats present for the companies and their information security policies?’ The paper presents theoretical background on information security policies and the cyber threat phenomenon in section two. In section three the paper analyzes the threats and their potential effect on the operations of companies with the use of scenario analysis. In the last section the paper assesses a company’s possibilities to take the cyber threats into account in its information security policy.

2. Theoretical background on information security policy and cyber threats

2.1 Information security policy

Information security policy is a document that companies use for preparing for and fighting against information security threats (Tipton & Krause 2004; Peltier et al. 2005; von Solms & von Solms 2004b). In literature the policy has been analyzed, for example, from the perspective of how it should be written (Peltier et al. 2005; Barman 2001; Höne & Eloff 2002), how to get employees to comply with the policy (von Solms & von Solms 2004b; Boss et al. 2009; Bulcuru et al. 2010; Herath & Rao 2009) and how the policy affects the information security culture of an organization (von Solms & von Solms 2004b; Lacey 2010; Van Niekerk & Von Solms 2010). The information security policy is considered an important tool for managing information security in organizations (Peltier et al. 2005; Ilvonen 2009).

The approach to documenting an information security policy can be technical or managerial (Baskerville & Siponen 2002). The technical approach means that the information security policy is used for guiding the use of information systems, and the way the systems communicate securely with each other. The managerial approach to information security policy refers to the policy document being targeted to all employees, and the scope of the policy being not just on the use of information systems. (Baskerville & Siponen 2002; Lopes & Sá-Soares 2012) In this paper these approaches are seen as complementary, and one approach is not chosen over the other. The information security policy is considered a key document directing both the information systems security and the secure behavior of employees.

Information security policy is defined differently by various authors (Ilvonen 2009). In this context information security policy is considered a document that follows the guidelines introduced by international standards (Höne & Eloff 2002), and thus it:

- defines the scope and objective of information security in the company, as well as states the purpose of the policy itself
- communicates management commitment and approval
- states the information security principles
- communicates the roles and responsibilities of employees regarding information security
- states how information security is monitored and reviewed
- is directly connected with other policies, procedures and strategies of the company

This list presented by Höne and Eloff (2002) is an exhaustive set of requirements for the policy statement. Other requirements for the content of the policy are that it must be based on a risk analysis conducted in the company (von Solms & von Solms 2004a) and that it is written in a short and clear manner (Barman 2001).

Identification of cyber threats is important to information security policy. Information security, as presented above, should by definition be based on identified risks. Identifying threats is the first step of risk analysis (Peltier et al. 2005) and thus an essential part of preparing an information security policy.

2.2 Cyber threats

According to a definition cyber threats are Internet-borne activities that may harm or have the potential to harm a computer or network and compromise the confidentiality, integrity, or availability of network data or systems (CCIP 2013). According to another definition “Cyber threats to a control system refer to persons who attempt unauthorized access to a control system device and/or network using a data communications pathway.” (ICS-CERT 2013)

The threat is not necessarily just harm done upon a computer or a network. It may also be harm introduced through the computers and networks to other kind of infrastructure. For example a self-extracting malware may come from internet or via a physically inserted USB-stick and contaminate software and further cause physical malfunctions in a power plant or water supply having further effects onto another type of infrastructure. Some sort of access to the target's network is required to happen in order for the intrusion to take place. This paper claims that the definition of the threat having to be internet-borne is not required. The term *cyber threat* may include features that are programmed and introduced to a system from physical media, thus not necessitating the use of the internet. Still they are featured by their nature of being originating from and affecting in the information and communications technology (ICT). However, a network on some level is needed to the malware to take effect. It needs a way to its goal. Thus, cyber threat in this paper means an activity or a group of activities affecting the computer system(s), or disseminating themselves via computer systems to other systems in order to do harm.

The actual activities that are regarded as cyber threats include a large variety of operations. Their unifying factor is their medium: the ICT. Weapons are there to inflict harm or damage (MOT Oxford Dictionary of English 2013). According to this definition all the threats fall under this category. However, in literature the cyber dimension is further divided into cyber weapons and non-weapons (Rid & McBurney 2012). The weapon-like approach is defined to mean the situation in which the means, computer code, is used or designed to be used to cause "physical, functional, or mental harm to structures, systems or living beings" or used to threatening to do so.

A counterpart for the weaponized class of cyber threats is logically the non-weapons. These are not used to cause harm or to damage but most often to gather information, to spy (Choo 2011; Malgeri 2009; Rid and McBurney 2012). This means that their effects are less severe and thus less punishable. The legislation in many countries is having hard time coping with this phenomenon (Foggetti 2008; Jahankhani 2007; Jie-miao 2008; Marion 2010). To draw the line separating the weapons and non-weapons might prove to be extremely difficult. However, both weapons and non-weapons are to be dealt with in one's information security policy as their methods of operation resemble one another.

One of the upcoming cyber threats according to one of information security company Kaspersky's senior researchers is malicious software introduced to mobile appliances (Emm in Waters 2012). What these malicious programs aim at remains unknown. They may be just data gathering or they may be pure malice with some kind of earning logic attached to them. (Choo 2011; Dunham 2009; Hypponen 2006)

Another type of cyber threat that the experts expect to emerge is the accessing to large data repositories and stealing data from these for various purposes (Schneider and Levin in Waters 2012). The increasing number of offshoring offers in this regard is also responsible for new risks (Kshetri 2010). There are numerous applications using cloud services for the data storage but also the more traditional data bases (such as medical records) are high-risk areas in the sense of possible target for cybercrime (Armbrust et al. 2010; Pearson and Benameur 2010; Pearson 2009). Also the number of more targeted cyber threats are expected to grow (Kshetri 2005). This may be partly due to the openness of companies and the ease of finding information of the possible targets and partly to the grown capability of the aggressors (Ottis 2008). A newer angle to mischief is 'datnapping'. The data thefts are used for demanding ransom payments in return for returning the data (Keizer 2011). It is done either by encrypting the data or locking the computer, both ways making the data inaccessible for the rightful user. Both obstacles are to be removed against payment, a ransom.

One of the main features of a cyber threat is that the actual perpetrator cannot be guaranteed to be successfully identified. The intruder or the actual objectives of the intrusion or attempts of intrusion remain more or less blurred. The parties involved as listed by the U.S. homeland security (ICS-CERT 2013) include:

- hackers
- hacktivists
- industrial spies
- organized crime groups
- terrorists

- national governments.

When the actions of these groups are considered, one may intuitively come to a conclusion that their objectives vary greatly from one another.

Hackers are thought to be intruding unauthorized domain almost for the fun of it. The mischief may be even left on the level of gaining access. They may break into networks for the thrill of the challenge and boast for their peers about their skills for doing so. Tools and advice for prospective hackers are readily available in the internet. Hacktivists as a category appeared first in 2008. They have some sort of political or social agenda in their hacker actions (Waters 2012). The agenda may be directed against a government, its policies, a public sector operator or equivalent. A hactivist's goal may be also just to draw attention to a point they are trying to make. The activity of hacktivists is estimated to grow with the time (Waters 2012).

Industrial spies and organized crime groups operate with profit-based objectives. Their function is based on making money by using illegal activities of some sort (intrusion related information and knowledge gathering and theft, extortion related to the previous, etc.). Their secondary objectives may include aggressions against the target organization's infrastructure to make profit to competitors or other groups listed here.

National governments and terrorists form a group of their own based on their motivation and interests. Terrorists are individuals or groups that carry a deep-rooted grudge against their target. Terrorism means by definition the "unofficial or unauthorized use of violence and intimidation in the pursuit of political aims" (MOT Oxford Dictionary of English 2013). Terrorists may be supported by national governments hostile towards the target with similar objectives. The goals of their actions (from espionage to attacks) may be to gain technological advances from the target, or disrupt the functions of the infrastructure of the target in order to attack the economy and everyday life thus creating disturbance.

3. Cyber threat scenarios for a company

In the previous section different kinds of threats and threat sources are mentioned as causing cyber threats. In this section three more specific scenarios are built based on the above mentioned theoretical perspectives. The point in introducing these scenarios is to analyze what kind of challenges they pose for the information security policy, and how companies could prepare for this kind of threat scenarios in their information security policies.

Scenario analysis is a method not uncommon in information security field. Especially in business continuity planning (BCP) this method is used to find out the possible effects a realization of a threat has for the business of a company (Lam 2002). Within the length limitations of this paper we will conduct a small scale typical-scenario analysis. In a company context the analysis should include discussions on how to react to threat scenarios and the actual recovery plans from the realization of the treat (the BCP). In the scope of this paper we concentrate on introducing the threats and the consequences of them to the operations of a company. We then discuss in the last section how the companies could use their information security policies to prepare for the threats.

The scenarios could be built from the perspective of origin, or from the perspective of end result. In this paper the scenarios are described from the perspectives of the end result, i.e. what the cause or the threat is to the company. Two of the scenarios are caused by malware. Malware can penetrate the company network either through downloads from a wrong source by an unaware employee, or by the careless use of portable storage media. Thus scenarios 2 and 3 can be caused by different ways of malware entering the internal systems of a company.

Scenario 1: Unavailability of web services

The denial of service (DoS) attack has been used lately as one way to harm the operations of not only companies, but entire nations (Czosseck et al. 2011). The logic of the attack is simple: the webservers of the target are flooded with traffic, and consequently they crash. The DoS attack may be a way to seek access to the systems of a company, or a way to simply cause harm. From the point of view of a company, however, the reason for the attack is not important. Countering the attack and recovering from the situation quickly is.

The reason why DoS attacks or more generally the threat of losing availability of the web-services of a company can be devastating is that many companies rely on their web-pages for business operations. One case is that the services of the company are directly distributed via internet, such as in media companies or companies that sell products directly from their web-stores. If the web services are not available, customers cannot buy the products of the company, and thus revenue is stopped. This seize of revenue can also be caused indirectly if the revenue of a company is reliant on the amount of page loads, i.e. amount of viewers for advertisements. On the other hand, the company may be reliant on the web services for internal communications and operations. If the domain of a company is crashed due to excess inbound traffic, the outbound traffic may also be interrupted. To recover from the crash of web services the excess traffic has to be stopped somehow, and the web-service re-launched as soon as possible.

Scenario 2: “Datnapping”

In the case of the malware “datnapping” the data of the company the harm for the company is caused by making data and possibly the use of ICT tools impossible. The recovery from the situation in the worst case requires both time and money, if the company cannot survive the situation without outside help. In a better case it requires just time and effort to re-install computers and databases from backups. The logic of datnapping data and computers is that the alternative solution to recover the data back can be so time-consuming that the company may choose to pay the “ransom” the datnapper, since it is cheaper than taking the effort of recovery. Sometimes full recovery would not be possible due to bad management of backups, and thus paying the ransom is the only way to get the data back.

Scenario 3: Information leaks

In the case malware works for example by keylogging or somehow otherwise by sending information outside the company, the harm for the company can be both, direct or indirect. Immediate harm may be caused by the information leaking outside the company, and the company suffering from image loss. If the company loses the trust of customers, it may prove very costly for the company. Indirectly the loss of information may harm the operations of the company by loss of competitive advantage, if its competitors are able to close the gap between them and the company. Loss of sensitive research and development information may reveal the plans of a new product to a competitor early, which could turn out devastating for the launch of that product.

Sometimes malware may be introduced into a company without anyone noticing, and without any damage. The Stuxnet virus is one example of this kind of malware that was widely spread, but it only did damage for a limited number of targets (Rid & McBurney). This kind of malware may wait for a long time to activate, but if it is targeted for the operations of a company, it may potentially do physical damage in addition to intangible damage. This is why computer systems that are used for overseeing physical processes should be very carefully isolated; one carelessly used, innocent looking usb-datastick may be all it takes to destroy the physical equipment of a company.

4. Discussion and conclusions

The above-described scenarios are general, but we can argue that they also are typical. Companies have announced that they have fallen victim to such scenarios (Waters 2012), and intuitively many more will do the same. How can the scenarios then be avoided, and what challenges to they pose for the information security policy?

The information security policy should define the roles and responsibilities of people, which helps a company in scenario 1. The unavailability of web services may cause havoc inside and outside the company, and in such a situation it is vital that everyone knows their role and responsibility in the event. Technical information security policies should be in place to limit the consequences of a DoS attack or a similar event, so that recovery of it is made fast, if the event cannot be avoided altogether. The role of a policy here is thus to prepare the company to face the cyber threat and react in an organized manner to it.

In scenarios 2 and 3 the directions given in the information security policy are the key to avoid the events. If all employees

- avoid opening untrusted links

- scan their email for viruses
- do not use USB-storage devices for moving files between computers unless absolutely necessary and under strict precautions

i.e. comply with information security regulations, the scenarios can be avoided to a large extent. With wider and wider use of ICT tools for both work and leisure this becomes difficult: links are not always what they seem to be, and a seemingly innocuous file may turn out to be something else completely (e.g. Hypponen 2006). The training of employees so that they are aware of the threats is thus as important as instructing them with the information security policy. The policy and training, however, need to be in line with each other (Bulcuru et al. 2010). Of course, there is only so much that the policy can do: some studies have shown that the policy does not help at all in reducing information security incidents (Doherty & Fulford 2005).

The challenge with all the threats is the global nature of many companies. Although different countries have different legislation, the companies that operate across borders should both comply with the legislation and protect their information at the same time. If they can track excessive logs of their network communications in one country and not do the same in another country, they have a dilemma: The policy should be the same across the company, but the legislative grounds for implementing the policy are different. A threat that could be tackled in one country may be left unaddressed in another, for example because the company cannot log and monitor email-communications. Building the kind of scenarios presented in the previous section may help a company find these challenges, which is the first step toward solving them.

The cyber threat perspective to information security policy raises a question that could be asked about information security policies in general: Are the policies addressing identified threats? And most importantly, are the employees of companies aware of the threats that the information security policy is supposed to address? Today the challenges to information security are changing fast, and the information security processes of a company should answer to this change. Approaching the policy from the point of view of threats and scenarios might help companies to reconceptualise their information security policy, and formulate the policies from a threat perspective. This way they could more effectively be used in managing the information security of the company.

References

Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., and Stoica, I. (2010) “A view of cloud computing,” *Communications of the ACM,* Vol.53, No.4, pp. 50–58.

Barman, S. (2001) *Writing Information Security Policies*. New Riders, Indianapolis. 216 p.

Baskerville, R. & Siponen, M. (2002) "An information security meta-policy for emergent organizations". *Logistics Information Management*. Vol. 15, No. 5. pp. 337-346.

Boss, S.R., Kirsch, L.J., Angermeier, I., Shingler, R.A. & Boss, R.W. (2009) "If someone is watching, I'll do what I'm asked: mandatoriness, control, and information security". *European Journal of Information Systems*. Vol. 18, No. 2 Special Issue: Behavioral and Policy Issues in Information. pp. 151-164.

Bulcuru, B., Cavusoglu, H. & Benbasat, I. (2010) "Information Security Policy Compliance: An Empirical Study of Rationality-Based Beliefs An Information Security Awareness". *MIS Quarterly*. Vol. 34, No. 3. pp. 523-548.

Czosseck, C., Ottis, R. & Talihärm, A. (2011) "Estonia after the 2007 cyber attacks: legal, strategic and organisational changes in cyber security", Proceedings of the 10th European Conference on Information Warfare and Security, Tallinn, Estonia, ed. R. Ottis, Academic Conferences International

Choo, K.-K. R. 2011. *Cyber threat landscape faced by financial and insurance industry*, Australian Institute of Criminology.

Doherty,N.F. & Fulford,H. (2005) *Do Information Security Policies Reduce the Incidence of Security Breaches: An Exploratory Analysis.* IGI Global.

Dunham, K. (2009) *Mobile malware attacks and defense*, Syngress Publishing.

Foggetti, N. (2008) “Transnation Cyber Crime, Differences between National Laws and Development of European Legislation: By Repression,” *Masaryk UJL & Tech*. Vol.2, pp. 31.

Evans, B. (2012) Big Data Set to Explode as 40 billion New Devices Connect to the Internet. [http://www.forbes.com/sites/oracle/2012/11/06/big-data-set-to-explode-as-40-billion-new-devices-connect-to-internet/] Accessed 20.12.2012

Herath, T. & Rao, H.R. (2009) "Protection motivation and deterrence: a framework for security policy compliance in organisations". *European Journal of Information Systems*. Vol. 18, No. 2 Special Issue: Behavioral and Policy Issues in Information. pp. 106-125.

Hypponen, M. (2006) “Malware goes mobile,” *Scientific American* Vol.295, pp. 70–77.

Höne, K. & Eloff, J.H.P. (2002) "Information security policy — what do international information security standards say?". *Computers & Security*. Vol. 21, No. 5. pp. 402-409.

Ilvonen, I. (2009) "Information security policies in small Finnish companies", Proceedings of the 8th European Conference on Information Warfare and Security, Lisbon, Portugal, 6-7 July 2009.

Information Services and Technology (IST) (2013). Viruses, Spyware, and Malware. Massachusetts Institute of Technology. http://ist.mit.edu/security/malware. 20.02.2013

Jahankhani, H. (2007) "Evaluation of cyber legislations: trading in the global cyber village," *International Journal of Electronic Security and Digital Forensics* Vol.1, No.1, pp. 1–11.

James, J. (2012) How much data is created every minute? A DOMO infographic. [http://www.domo.com/blog/2012/06/how-much-data-is-created-every-minute/?dkw=socf3] Accessed 20.12.2012.

Jie-miao, C. (2008) "China's Legislation on Criminal Jurisdiction over Cyber Crimes", *Modern Law Science* Vol. 3, pp. 012.

Keizer, G. (2011). Ransomware squeezes users with bogus Windows activation demand

But F-Secure sniffed out unlock code to stymie extortion scheme. ComputerWorld. Apr 11th, 2011. http://www.computerworld.com/s/article/9215711/Ransomware_squeezes_users_with_bogus_Windows_activation_demand 20.02.2013

Kshetri, N. (2005) "Pattern of global cyber war and crime: A conceptual framework," *Journal of International Management,* Vol.11, No.4, pp. 541–562.

Kshetri, N. (2010) "Cloud computing in developing economies," *Computer* Vol.43, No.10, pp. 47–55.

Lacey, D. (2010) "Understanding and transforming organizational security culture". *Information Management & Computer Security*. Vol. 18, No. 1. pp. 4-13.

Lam, W. (2002) "Ensuring business continuity". *IT Professional*. Vol. 4, No. 3. pp. 19-25.

Lopes, I. & Sá-Soares, F. 2012, "Information security policies: a content analysis", PACIS - The Pacific Asia Conference on Information Systems. Hochiminh, Vietnam.

Malgeri, J. 2009. "Cyber security: a national effort to improve," In 2009 Information Security Curriculum Development Conference, pp. 107–113.

Marion, N. E. 2010. "The council of Europe's cyber crime treaty: An exercise in symbolic legislation," *International Journal of Cyber Criminology* Vol.4, No.1&2.

Meehan, P. (2013). Cyber threats hit close to home. Philly.com, the Inquirerer on Feb. 4th, 2013. http://www.philly.com/philly/opinion/inquirer/20130204_Cyber_threats_hit_close_to_home.html.Read 06.02.2013

MOT Oxford Dictionary of English (2013). mot.kielikone.fi/mot/ttkk/netmot.exe

Ottis, R. (2008) "Analysis of the 2007 cyber attacks against estonia from the information warfare perspective," In Proceedings of the 7th European Conference on Information Warfare.

Pearson, S. (2009) "Taking account of privacy when designing cloud computing services," In Software Engineering Challenges of Cloud Computing, 2009. CLOUD'09. ICSE Workshop on, pp. 44–52.

Pearson, S., and Benameur, A. (2010) "Privacy, security and trust issues arising from cloud computing," In Cloud Computing Technology and Science (CloudCom), 2010 IEEE Second International Conference on, pp. 693–702.

Peltier, T.R., Peltier, J. & Blackley, J. (2005) *Information security fundamentals*. Auerbach Publications, Boca Raton, Fla.

Rid, T., and McBurney, P. 2012. "Cyber-Weapons," *The RUSI Journal* Vol. 157, No.1, pp. 6–13.

The Centre for Critical Infrastructure Protection (CCIP). What Are the Cyber Threats?(2013). New Zealand's Government Communications Security Bureau (GCSB). http://www.ccip.govt.nz/ 20.02.2013

The Industrial Control Systems Cyber Emergency Response Team (ICS-CERT). (2013). Cyber Threat Source Descriptions. Department of Homeland Security. Found: http://ics-cert.us-cert.gov/csthreats.html, 20.02.2013

Tipton, H. & Krause, M. (eds) 2004, *Information security management handbook*, 5th edn, CRC Press, Boca Raton.

Van Niekerk, J.F. & Von Solms, R. (2010) "Information security culture: A management perspective". *Computers & Security*. Vol. 29, No. 4. pp. 476-486.

von Solms, B. & von Solms, R. (2004a) "The 10 deadly sins of information security management". *Computers & Security*. Vol. 23, No. 5. pp. 371-376.

von Solms, R. & von Solms, B. (2004b) "From policies to culture". *Computers & Security*. Vol. 23, No. 4. pp. 275-279.

Waters, J. The New Year's Biggest Cyberthreats. The Wall Street Journal. Europe Edition, on Dec. 29th 2012. http://online.wsj.com/article/SB10001424127887323277504578193833434470690.html, Read 20.02.2013

Strategic Communication as a Communication Model: Functions, Characteristics and Criticism

Saara Jantunen
Department of Military Pedagogy and Leadership, National Defence University, Helsinki, Finland
sijantunen@gmail.com

Abstract: This article addresses the internal challenges of Strategic Communication by discussing the basic principles presented in the 2009 *Strategic Communication Joint Integrating Concept* (SC JIC) and the 2010 *Commander's Handbook for Strategic Communication and Communication Strategy* (CHSCCS) and relating them to the challenges of current communication landscape. Whereas the SC JIC is a concept, the CHSCCS is a pre-doctrinal document that aims to bridge between current practices and the migration into doctrine. These two documents thus cover the aspects of both theory and practice. Although Strategic Communication is no longer used as a concept, the successor, "communication synchronization", continues to follow the principles of it's predecessor. The discussion addresses the core aims and objectives of the doctrine (i.e. the concept of legitimacy), the elements essential in its execution (messages, themes and narratives) and the concept's interpretation of communication as a process. The analysis is carried out by discussing the doctrine in the context of asymmetrical and symmetrical communication models and by comparing and contrasting it to the tradition of Mass Communication Research (MCR). This discussion is relevant in the analysis of Strategic Communication, because it reveals the strong theoretical parallels between Strategic Communication and MCR in terms of the outlook on the mechanics of the communication process. The history and challenges of MCR explicates the problems and challenges of the 21st century Strategic Communication. The ultimate conclusion is that since Lasswell, little has happened in the field of military communication. Despite all efforts to evolve, Strategic Communication is still essentially a propaganda model.

Keywords: strategic communication, communication theory, mass communication research, asymmetrical doctrine

1. Introduction

Strategic communication is a key tool in strategic leadership. This article focuses on Strategic Communication in the context of military communication. It will argue that the current doctrine of Strategic Communication contains a number of inherent problems and inconsistencies that result from the failure to abandon the effects-centric paradigm that is based on the framework of the early 20th century communication theory. These inconsistencies hamper the communication process and therefore complicate strategic leadership.

U.S. Strategic Communication, according to the 2009 Strategic Communication Joint Integrating Concept (hereafter SC JIC), is a communication model the purpose of which is to create legitimacy. This paper discusses the US concept of Strategic Communication as laid out by the SC JIC and the 2010 Commander's Handbook for Strategic Communication and Communication Strategy (hereafter CHSCCS). The concept is then compared and contrasted to the general theoretical framework of communication theory. The following sections discuss the theoretical understanding of communication which the operational model of Strategic Communication is based on. The aim is to locate Strategic Communication in the field of communication theory.

The SC JIC, as the title states, is a concept, whereas the CHSCCS is a pre-doctrinal document that aims to serve "as a bridge between current practices in the field and the migration into doctrine" (p. i). These two documents thus cover the aspects of both theory and practice. The SC JIC determines the strategic function and objective of the doctrine, as well as the understanding of what communication is about, whereas the CHSCCS presents a much more detailed account of the previous with an emphasis on the practical implementation of communication.

2. The mechanics of communication

The SC JIC as well as CHSCCS both present an outlook on the operational model of communication. The SC JIC cites David K. Perlo's (p. 5-6) *The Process of Communication: An Introduction to Theory and Practice* from 1960:

> *Communication works this way: A source puts out a signal intended to convey a meaning. The receiver recognizes and selects the signal, if he chooses, from among the various signals available to him: he interprets the signal based on his own frame of reference and interests to create meaning. While the source may have an intended meaning in mind, it is the receiver who actually*

> *provides the ultimate meaning, which may or may not be the meaning the source intended. The challenge in effective communication is to anticipate what signal will trigger the desired interpretation.*

This model demonstrates that the Stimulus-Object-Response (S>O>R) model is still the basis of modeling communication. Further, the CHSCCS (p. II-10) outlines effective communication as a hierarchical model of themes, messages and narratives:

- Theme: an overarching concept or intention, designed for broad application to achieve specific objectives.
- Message: a narrowly focused communication directed at a specific audience to create a specific effect while supporting a theme.
- Narrative: enduring strategic communication with context, reason/motive, and goal/end state.

Strategic documents, according to the CHSCCS, produce narratives. Messages, instead, should support the themes, "themes should support [..] the next higher level themes, and themes at all levels should support strategic themes and the enduring national narrative" (p. xiii). This results in "consistent communications to global audiences" (p. xiii). Consistent communication, also addressed as “one voice” or synchronized communication, is seen as the key requirement of all communication efforts. The failure to produce synchronized, unified and harmonized narratives undermines all communication efforts. The importance of synchronization was further highlighted in December 2012, when Strategic Communication as a concept was officially killed (Department of Defense, 2012) and replaced with *communication synchronization*, which thus was one of the elements of Strategic Communication.

Whether addressing Strategic Communication or “communication synchronization”, it can be argued that the perception of communication these concepts propose view communication as a mechanical, logical and very straight-forward process. But, if we address the primary objectives of strategic communication, namely the production of legitimacy and coherence, what does the communication process proposed by the concept look like? In order to discuss the communication practices as instructed by the concept, it is important to bear in mind the context in which communication is produced: persuasion and influence. Added to the behaviorist understanding of communication, this context marks the conflict between (U.S.) military Strategic Communication and the strategic communication as it is understood in the civilian/corporate world. This conflict becomes apparent on the first pages of the SC JIC, where the motivations and ethics of military communication efforts are specified: while it treats communication as a vehicle to the hearts, minds and behavioral choices of the audience, it also recognizes the negative connotation of the term 'influence' and makes, in fact, an effort to renegotiate its definition (p. iii):

> *The term influence sometimes carries negative connotations because the term is often associated with deceptive manipulation or exploitation. Influence will not have that connotation in this concept. Influence is a pervasive and fundamental form of any social interaction, as essential to cooperation as it is to competition or conflict.*

However, this seems to be at odds with what is further stated about communication in the doctrine. Renouncing the attempts to practice negative influence is one of the attempts the concept makes to approach the idea of sharing meaning, also argued to be an element in the concept. Although the terms manipulation has officially been removed from doctrine (p. v), the influence spectrum still recognizes the communicative practices of urging and coercion:

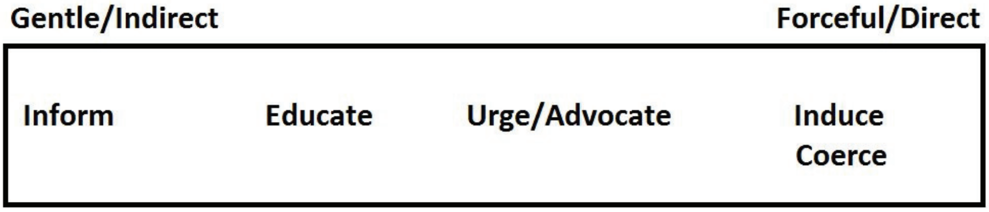


Figure 1: The influence spectrum according to the 2009 strategic communication joint integrating concept

The question is whether urging, advocating, inducing or coercing can be practiced without manipulation in the context of (information) warfare. Is it possible to differentiate between persuasion (legitimate) and manipulation (illegitimate)? According to Van Dijk (2008: 212), manipulation is "illegitimate influence" that produces inequality and that is used to serve the interests of the manipulator against the interests of the manipulated. Strategic Communication, in turn, aims at communicative asymmetry, where competing narratives become irrelevant.

Herein lies thus a paradox. The influence spectrum presented in Figure 1 suggests that influence as understood in the concept of Strategic Communication is synonymous to one-way flow of communication rather than an interactive relationship between the communicator and the target. Communication is practiced for the sole reason of influence, but to create a desired effect in the audience and to make the audience think and behave in a desired way. This obviously links Strategic Communication to the tradition of effects-centric kinetic doctrines such as Effects Based Operations and the Comprehensive Approach. To make a comparison it could thus be said that much like there are no kinetic operations for the sake of "just operating", there are no communication operations for the sake of "just communicating". Strategic Communication should be understood as a non-kinetic equivalent of effects-based strategy and tactics.

As argued in the following chapters, with the references to the 1960s models of communication, the Pentagon is attempting to maintain the theoretical framework of MCR, which, during the 1960s, was marked by efforts to establish a theoretical framework (Pietilä, 2005: 105-109). Part of these efforts was Lasswell's model, which presented communication as a transmission process. In Lasswell's theory the central analytical concepts were the contents, channels and effects of communication (Pietilä, 2005:108; Lasswell & Blumenstock, 1939/1970; Lasswell, 1960). It was the effects that were intended to be measurable and to generate the much needed empiricism in the new field of research. What was in focus was the behavior of the audience: "in effects analysis the characteristics of media content were seen as causal factors that may produce specific behavioral consequences" (Pietilä, 2005: 109). Lasswell (1946a: 80) himself established that of audience responses (attention, comprehension, enjoyment, evaluation and action) all responses (reactions) above the level of attention are effects (1946b: 97).

In order to discuss the paradigmatic properties of Strategic Communication, three key parallels between early 20th century communication theory (MCR) and Strategic Communication as defined in SC JIC and CHSCCS are considered next. These include the focus on creating effects, cognitive persuasion and communication as a transmission process, discussed next.

2.1 Focus on causing and measuring effects

The first key parallel relates the effects-centric paradigm of Strategic Communication to the Lasswellian approach to communication as the producer of effects. Strategic Communication is a "continuous function" that integrates Joint Force actions "to maximize the desired effect on selected audiences" (SC JIC p. ii). Theoretically this means that the foundations of Strategic Communication never evolved beyond Lasswell's propaganda model. As stated, Lasswell's (1946a: 80) categories of effects resonate with the influence spectrum presented in Figure 1. In the 21st century, 'effect' is one of the key terms in strategic and operational planning: the effects-centric doctrines, whether kinetic or non-kinetic, aim at "desired effects".

MCR underwent a methodological crisis during the 1960s, which led to the adaptation of the 'communication as interpretation' paradigm and the transmission model was abandoned (Pietilä, 2005: 126). Yet, the so-called Lasswellian magic bullet theory and the emphasis of the significance of measurable effects are still what the American doctrine writers aspire to. All in all, the claim made here is that due to the doctrinized requirement of "causing effects", the Pentagon's approach to the process of communication has not left the theoretical framework of MCR. The effects-centric doctrines of warfare require communication is seen as a causal process that results in measurable effects. The crisis of Strategic Communication as well as its predecessors, such as perception management, is largely due to this inability to modernize and develop paradigmatically.

2.2 Cognitive persuasion

The commitment to the behaviorist tradition of conceptualizing communication leads us to the second parallel between MCR and the concept of Strategic Communication. The Yale school neo-behaviorists treated the effects of communication-as-learning much like the SC JIC refers to the influence of communication as "education". Further, the Yale school researchers (see Hovland, Janis & Kelley, 1953) argue that the individual may be persuaded by offering a more rewarding experience to replace the old one and that this, in fact, is a learning process:

> *We assume that opinions, like other habits, will tend to persist unless the individual undergoes some new learning experiences. Exposure to a persuasive communication which successfully induces the individual to accept a new opinion constitutes a learning experience in which a new verbal habit is acquired. That is to say, when presented with a given question, the individual now thinks of and prefers the new answer suggested by the communication to the old one held prior to exposure to the communication. (Hovland, Janis & Kelley, 1953)*

The CHSCCS repeats the line of reason:

> *The Battle of the Narrative is often thought of as a battle for the local audience to "buy" our "story" and push out the enemy's "story," such as "we are the good guys, we are here to help you and bring you a better quality of life." This perspective on the "Battle of the Narrative" is incorrect. The battle is not merely to push aside, defeat or gain superiority over the enemy's narrative; it is to completely supplant it. In fact, upon our winning the battle of the narrative, the enemy narrative doesn't just diminish in appeal or followership, it becomes irrelevant. (p. II-13)*

The Battle of the Narrative is thus a battle over persuasion: as argued by Hovland, Janis & Kelley (1953), the best story wins.

This approach to the persuasion process understands communication as a cognitive process rather than conditioning. Echoing the challenge of measuring the effects of communication, the SC JIC states that "[a]ssessing the cognitive impact of a signal is much more difficult, for example, than assessing the physical impact of an air strike on a target" (p. 6). This analogy is another example of the methodological problem of reducing communication research into the study of cause and effect. The problem arises when the organic society is subjugated to causal theory. As the analogy above suggests, this has been experienced not only by MCR scholars, but by the Pentagon as well as the U.S. Joint Forces Command. The difference between the theorists here is that the veterans of MCR admitted the methodological crisis in the 1960s (see Pietilä, 2005: 126, Kunelius, 2010: 141-142), whereas Strategic Communication keeps aspiring to the past. Also, it was the cognitive theory of learning that challenged the behaviorist learning theory and established the theory of cognitive consistency (Pietilä, 2005: 116). As the two theories are mutually exclusive - influence cannot be both a conditioning process and a cognitive process at the same time - the concept of Strategic Communication does not recognize the paradigmatic problem in applying both behaviorist and cognitive frameworks at the same time.

2.3 Communication as transmission

The third parallel between 21st century Strategic Communication and the 1960s MCR is the understanding of communication as a transmission process. This parallel has already been addressed above, but returned to here because the paradox between the theoretical modeling (namely communication as a S>O>R relationship) and the practical implementation of Strategic Communication should be highlighted.

Despite defining the communication process as a form of behaviorist conditioning, the objective that is set is "shared meaning":

> *Strategic communication essentially means sharing meaning (i.e., communicating) in support of national objectives (i.e., strategically). This involves listening as much as transmitting and applies not only to information, but also physical communication - action that conveys meaning. (SC JIC p. ii)*

As argued here, the repeated and specific references to communication as "signals" and "transmission" link Strategic Communication to the tradition of the early propaganda models, where the targets tend to be seen as individuals rather than an audience (Kunelius, 2010: 153). At the same time, the doctrine clearly indicates that it aims to influence collectives: sharing is specifically understood as an attempt to create a collective. It can be asked whether meanings can be shared by simply transmitting them, as sharing suggests a two-way flow of information (Coombs & Holladay, 2007).

Ultimately, what Strategic Communication and the MCR tradition have in common is the social context of their emergence. What marked the turn of the 20thcentury were the World Wars and the age when the media were seen as a factor that may change and possibly even threaten the social order of society. Correspondingly, Strategic Communication is modeled to serve the purpose of maintaining and legitimizing a certain (global) political order.

3. The symmetry and asymmetry of communication

As argued this far, despite its attempts to renounce negative manipulation, Strategic Communication fails to define itself in terms other than propaganda and behaviorism. This, however, does not mean that it lacks the attempt to build oneness and create shared experience.

In addition to being a communication model, the concept of Strategic Communication also contains elements of Public Affairs, as its function is to ensure "conditions favorable" for the national interest of the U.S. government. Therefore it is worth considering Strategic Communication as a doctrine of Public Relations.

The first obvious problem brings us back to the question of dialogicality and interactiveness.

> *Dialogue requires listening. Listening represents the "two" in two-way communication while "one" is simply sending a message via one-way communication. (Coombs & Holladay, 2007: 31)*

At least in theory, the SC JIC includes "listening" and "sharing" in its operative model. As argued here, the practical application of listening and sharing may prove difficult. The paradox of listening and persuading culminates in the question of dialogue. In order to persuade, a certain attentiveness to the audience's reactions, i.e. the effects of communication, is necessary. However, does this type of listening count as dialogue? According to Coombs & Holladay (2007), it does not. According to the concept of Strategic Communication, listening is a matter of persuasion rather than aiming at a communicative balance between the organization and the audience. This two-way asymmetric model of public relations has been defended as an organization's way to look after its interests and is characterized as "advocacy" (see Grunig, 2001; also Grunig, 1992). A PR-strategy may appear as naturally asymmetrical and to favor either the interests of the organization or the public, but according to Grunig (2001: 25), the middle of the continuum of these interests contains "a symmetrical win-win zone" where both the organization and the public can "engage in mixed-motive communication". Unlike the original symmetrical model, the mixed-motive model would would not force the organization to sacrifice its interests, and would ideally lead to practices of collaboration instead of advocacy (Grunig, 2001).

Obviously, collaboration is an unrealistic request to an organization that specifically aims at asymmetry. However, the public and stakeholders have the tendency to communicate, whether their input was called for or not. The asymmetric methods force the counterparty to engage in respectively asymmetric activities in order to pursue its interests. This turns us to the question of audiences. The problem with Strategic Communication is its undefined audiences and the matrix of their interests. It is practically impossible to predict their expectations, which makes proactive communication strategies virtually impossible. In the asymmetric model, stakeholders can force the organization into dialogue (Coombs & Holladay, 2007: 54), which has also been the case with the ongoing military operations. The Abu Ghraib and WikiLeaks scandals have forced the organization to comment and evaluate its practices in public. In these cases the strategic stakeholders have included U.S. soldiers who have been able, intentionally or not, to reveal something about the practices from within the organization - something that is difficult for an outside stakeholder to do. In this sense WikiLeaks has been a significant element of asymmetry, and in fact a dire result of the asymmetric nature of U.S. Strategic Communication. As the scandals prove, without symmetry and mutual influence, communication becomes unpredictable and uncontrollable. This means that communication will turn from flows into bursts, each causing the organization significant damage and further undermining their credibility and legitimacy.

Further, it seems that one of the major shortcoming of Strategic Communication is its failure to recognize its stakeholders. The CHSCCS makes numerous references to the importance of the analysis and maintenance of stakeholder relations, but keeps referring to governmental and non-governmental organizations, interagency representatives, intergovernmental and international organization representatives, media, etc. Considering the recent scandals and their impact on public opinion, the focus should perhaps be on the global audience and the individual soldier. Failing to treat the soldiers as stakeholders has in this case led to major reputational damage. During the wars in Iraq and Afghanistan, most PR scandals have been caused by individual soldiers. These "strategic corporals" have had different motivations to communicate their perspective, some by leaking information on purpose, as was the case with Bradley Manning and the WikiLeaks scandals, and some by recording their war experience in photos and videos, never intending the outside world to see them, as happened with the photos taken in Abu Ghraib and the video of U.S. soldiers urinating on dead Taleban fighters. When the soldier's experience in the field does not correspond with the official strategic communication, they will continue to produce micronarratives that undermine the Grand Narrative. These narratives will be produced regardless of the outlet. What the concept of Strategic Communication ignores is the need for a feedback channel that would ensure that in addition to making the individual soldier responsible for generating "synchronized messages" where words correspond with actions, there should be a way for the soldiers to engage in a bottom-up dialogue with their organization. However, the concept of Strategic Communication recognizes communication as a top-down process. While the doctrine of Strategic Communication focuses the practices of persuasive influence only on the outside stakeholders and the adversary, the approach to the needs of the inside stakeholders, i.e. the soldiers, is underdeveloped. This will continue to pose a threat to the U.S. communication policy.

But stakeholders and audiences will continue to communicate. Ultimately this means that the organization faces the options of either evading transparency, or engaging in a dialogue with its members and stakeholders. This dialogue, or the lack of it, determines both the ethics and the efficiency of communication: the absence of it maintains distrust both in the minds of the soldiers participating in the operation, and in the minds of the global audiences and other stakeholders. This, in turn, complicates the legitimatory function that is the very objective of the organization.

4. Implementation of SC: Practice in the field as instructed

The previous section introduced the theoretical principles and paradoxes of Strategic Communication. This section provides an example of the core problem Strategic Communication: the execution of communication as a controlled and managed one-way process, namely the use of the so-called themes and messages cards, which was one of the steps towards the current model of communication synchronization.

The theory-to-practice manual CHSCCS goes as far as providing examples of themes and words that should be used to influence the perceptions about the enemy (p. IV-22). The purpose of these cards is to prevent the "say-do gap". The inconsistencies between the words and deeds of the military force troubles the organization not only on the level of decision making, but at the grass-root level of individual soldiers:

> *[W]ith actions at all levels sends conflicting messages and significantly inhibits the creation of desired outcomes. Many refer to this as a "say-do gap." To help solve this problem in the CENTCOM and OUTHCOM AORs, units issued each soldier a card with key themes and messages to carry with them at all times. This approach was designed to synchronize words and activities all the way down to the individual level. This card helped soldiers and activity participants consistently communicate the desired message and guided their actions during unanticipated circumstances. (CHSCCS, p. III-10)*

The themes and messages card presents an understanding of the practical implementation of Strategic Communication: how strategic level objectives should be transformed into dialogue. Individual soldiers are seen as operators who negotiate with the stakeholders, and in order to equip them for these exchanges, the card consists of lists of themes and messages that should be enforced:

Here, it is the individual soldier who is eventually made responsible for the generating of strategically applicable verbal and physical communication. Unlike the culture smart card that provides the basic information about the cultural environment, the themes and messages card is not intended to inform and educate, but to instruct the individual soldier the correct way to practice conditioning. This means that Strategic Communication is not truly a model for communication, but a means of internal organizational control. It determines the roles within the organization and the correct usage of language. The instructions on the practical application of SC emphasize dialogue (CHSCCS, p. A-2): "Effective communication requires a multi-faceted dialogue among parties. It involves active listening, engagement, and the pursuit of mutual understanding, which leads to trust." This instruction would be in favor of symmetrical communication, but the themes and messages card does little to promote deep understanding or dialogue. Instead, it models communication as an asymmetrical process that aims at persuasion rather than dialogue and mutual attunement. Moral agency may be manuscripted into the smart cards that promote democracy and partnership, but instead of treating soldiers as moral agents, the doctrine reduces the soldier into an instrument of asymmetric capability. With the themes and messages of the official Grand Narrative determined in the smart card, there is little space for "multifaceted exchange of ideas" or "deep comprehension of others". The objectives and methods of Strategic Communication are once again in dissonance.

SC Standing Themes and Messages	Themes to Stress (through actions and/or words)
Related Themes and Messages	Themes to Avoid

Figure 2: The outline of the themes and messages card (CHSCCS, p. K-1)

5. Discussion

This article has attempted to demonstrate that maintaining the old paradigm of communication will not take military communication to the 21st century. Instead, it will increase the inter-organizational conflict between the organization and its primary stakeholders, the soldiers. In order to increase its resilience against reputational damage, the organization cannot expect to be able to control or manage communication or information flows, as the current technological development as well as the modern communication culture both support the individualistic, 24/7 participation in communication. The communication doctrine that aspires to the early propaganda model will face increasing difficulties in the operational environment that Strategic Communication is intended to cover.

However, the perception of communication as a transfer process and conditioning is inseparable from the tradition of military communication. Strategic Communication is a parallel, non-kinetic doctrine to the kinetic doctrines of the 21st century, all of which are based on the principle of measurability of effects. The methodological crisis of the tradition Mass Communication Research suggests that the effects-centric military communication paradigm will be plagued with the same, inherent problems. The nature of the organization complicates the adoption of symmetric communication, because from the perspective of strategy and tactics, communication is seen as an asymmetric capacity. This keeps it tied to the problematic tradition of costly and burdensome management, which essentially makes Strategic Communication a control model. Its function is not merely to produce narratives, but to provide control within the organization. The question remains: is it possible for military communication to escape its purpose? If symmetry and transparency are unrealistic methods of communication, Strategic Communication will remain essentially a propaganda model.

Finally, soldiers are Agents and stakeholders. They react, make mistakes, and generate communication that expresses their experience. Now that the modern communication landscape is based on the principle of access, Strategic Communication cannot legitimately control the flows of information. This has led to the concept of information age 'strategic corporals', who may either support the efforts of the military operation - or undermine them. In fact, the communicative actions of the soldiers, voluntary or involuntary, may just be true 'strategic communication'. Political, strategic level statements made in the Pentagon press room or the politically motivated key words memorized from a smart card are easily challenged by the depictions of the soldiers' reality. Compared to the official strategic narrative, the information value of the unofficial micronarratives is far greater. What makes the soldiers' narratives strategic is their truthfulness and authenticity - which is something Strategic Communication cannot offer. The problem may thus be that the Department of Defense fails to recognize one of its key groups of stakeholders: the soldiers, those actors that literally live the reality of strategic decisions but whose experience the doctrine effectively attempts to silence. The introduction of communication synchronization means yet another communication doctrine is doomed to fail. In practice, nothing changed much. Instead, communication synchronization is a step towards a mode controlled and managed communication model. With the refusal to evolve paradigmatically, the problems of the previous concept are bound to be inherited. In this case, communication synchronization will increase the probability of bursts of asymmetry against the U.S. armed forces both inside and outside the organization.

References

Cioppa, T.M., 2009. Operation Iraqi Freedom strategic communication analysis and assessment. *Media, War & Conflict*, Vol 2: 25-45.

Coombs, W.T. & Holladay, S.J., 2007. *It's not just PR: Public Relations in society*. Oxford: Blackwell.

Department of Defense, 2009. *Strategic Communication Joint Integrating Concept* (Version 1.0, 7 October 2009).

Department of Defense, 2012. *Memorandum for commanders of the combatant commands.* November 28, 2012.

Festinger, L., 1962. *A theory of cognitive dissonance*. Stanford: Stanford University Press.

Grunig, E.J., 1992. Symmetrical systems of internal communication. In James E. Grunig (ed.), *Excellence in Public Relations and communication management*. Hillsdale: L. Erlbaum Associates.

Grunig, E.J., 2001. Two-way symmetrical Public Relations: Past, present and future. In Robert L. Heath, *Handbook of Public Relations*. Thousand Oaks: Sage.

Hovland, C. I., Janis I. J. & Kelley, H. H., 1953. *Communication and persuasion: Psychological studies of opinion change.* Westport: Greenwood Press.

Kunelius, Risto (2010): *Viestinnän vallassa: Johdatus joukkoviestinnän kysymyksiin.* Helsinki: WSOY.

Lasswell, H.D., 1946a. Describing the contents of communications. In Smith, B.L., Lasswell, H.D. & Casey, R.D., *Propaganda, communication, and Public Opinion*. New Jersey: Princeton University Press.

Lasswell, H.D., 1946b. Describing the effects of communications. In Smith, B.L., Lasswell, H.D. & Casey, R.D., Propaganda, communication, and Public Opinion. New Jersey: Princeton University Press.

Lasswell, H.D., 1960. The structure and function of communication in society. In Wilbur Schramm, *Mass Communications* (2nd edition). Chicago: University of Illinois Press.

Lasswell, H. D. & Blumenstock, D., 1939/1970. W*orld Revolutionary propaganda: A Chicago study*. New York: Books for Libraries Press.

Malrieu, J.P., 2002. *Evaluative Semantics: Cognition, language and ideology.* London: Routledge.

Martin, J.R., & White, P.R.R., 2005. *The language of evaluation: Appraisal in English.* New York: Palgrave Macmillan.

Pietilä, V., 2005. *On the highway of mass communication studies*. Cresskill: Hampton Press.

Snow, N. & Taylor, P., 2006. The Revival of the Propaganda State: US Propaganda at Home and Abroadsince 9/11. *International Communication Gazette*, Vol 68, p. 389-407.

US Joint Forces Command, 2010. *Commander's Handbook for strategic communication and communication strategy* (Version 3.0, 24 June 2010).

Van Dijk, T.A., 2008. *Discourse & power.* New York: Palgrave Macmillan.

Van Leeuwen, T., 2007. Legitimation in discourse and communication. *Discourse & Communication*, Vol 1, p.91-112.

Improving Cyber Defence of Tactical Networks by Using Cognitive Service Configuration

Anssi Kärkkäinen
Defence Command Finland, Helsinki, Finland
anssi.karkkainen@mil.fi

Abstract: Nowadays, cyber threats are growing due to a large number and complexity of interconnected computers, networks and microprocessors. These new threats require new approaches for security of military networking. This paper proposes a cognitive service configuration model for tactical military networks. In the model, the user and other services are established and maintained dynamically by adapting the service configuration continuously. The dynamic environment of ICT services creates a moving target for an attacker causing the pre-prepared cyber attacks to be useless. The paper presents a functional architecture of the cognitive service configuration, and uses a distributed denial-of-service (DDoS) attack as an example to demonstrate the features of the proposed cognitive service configuration. Some implementation challenges are also been discussed. The paper shows that the cognitive service establishment and maintenance is a promising approach to provide more protection on military computer networks and services, although there are still many details such as decision-making, and control mechanisms that require more research in the future.

Keywords: phrases: cyber threat, cognitive network, threat management

1. Introduction

Information sharing and networked command and control systems bring advantages to modern warfighters. Increasing amount of information and its accuracy helps decision-makers to make faster and more precise decisions. One of the critical requirements for the military networks is cyber security. Traditionally, information security means that information could not be stolen or modified, networks function as desired, and information availability is guaranteed. Information security sets controls to protect information and data, but cyber security also covers information and communication systems and networks themselves, and it includes all the means required to protect the systems and networks for misusage. This means that information is not an only protected object. Through cyber security the information system should be protected against all imaginable attacks so that the system is able to provide services continuously.

The expansion of cyberspace has generated new and increasing threats to military communications. Cyberspace can be seen as an electronic medium of computer networks in which online communication and information sharing takes place. Thus, all military equipment with a programmable micro circuit or computer creates a new attack surface to be exploited by an adversary. Growing cyber threat requires new approaches to protect information systems and avoid damages caused by successful cyber attacks. This paper proposes a cognitive information service structure to create an information system as a moving target. The cyber attacks are not defended by placing packet filters or other intrusion prevention systems on the edge of the information system. The damages are avoided by using a dynamic system configuration that is able to adjust its current configuration by using a cognitive process.

The structure of the paper is following. In Section 2 related work is presented. Section 3 discusses on military networking and the implementation of information services. Section 4 presents cyber threats on the information services and networking. Section 5 introduces the functional architecture and demonstrates the cognitive service configuration using a DDoS attack as an example case. Section 6 presents some implementation challenges. Finally, section 7 concludes the paper.

2. Related work

Cyber or information security has been researched a lot during the recent years. Security threats of military communication networks are widely discussed in various papers, and new methods against the threats have been developed. The security research has focused more on a certain piece of security in cyberspace instead of considering cyber security as an overall issue. For example, threat detection, cyber security threats and mission assurance are discussed in (Buford et al 2008), (Kawano et al 2005), (Morris et al 2011) and (Bodeau et al 2010). These studies still have quite narrow scope and propose techniques to detect and react cyber threats. Morris et al focus on the description of the system on cyber mission information needs, whereby collection, processing, management and mission model updates are based on cyber-related information from a variety of resources including commercial news, blogs, wikis, and social media sources. The result is a dynamic capability for cyber mission management that provides proactive, on demand cyber information to analysts, professionals, policy makers, and support personnel.

Only a few research papers of cognitive information services have been published. In (Zheng et al 2008) the next generation of Service Delivery Platform (NGSDP) is presented. The paper proposes a new intelligent and cognitive service delivery platform model, which is a SDP model integrated with the concept of cognitive networks. The new model helps the Service Delivery Platform turns to an autonomous platform without unnecessary manual interventions. The cognitive features provide a dynamic system configuration for easier service creation and maintenance, but cyber security benefits are not discussed.

Dzmitry Kliazovich et al (2009) propose a novel concept in cognitive network management and protocol configuration in which any protocol of the TCP/IP protocol reference model can be extended to dynamically tune its configuration parameters based on history performance. The approach is focused on cognitive adaptation between network nodes. The adaptation is provided by the Cognitive Information Service (CIS). The paper also presents performance evaluation results which are obtained for cognitive adaptation of main TCP congestion window parameters.

In (Jimenez-Molina and In-Young 2011), a novel cognitive engineering mechanism to optimize service functionality coordination is presented. The paper bases this resource aware approach for service coordination optimization on two theories from cognitive psychology; the human-processing system theory of Navon and the multiple resource theory of Wickens. The purpose of the study is to show how to manage service configurations of time-shared mobility activities. However, again, cognitive features are not used throughout a whole information system configuration, and not for cyber security purposes.

3. Military networks and information services

3.1 Military networks

The military networks function under the extreme circumstances. The networks are deployed in harsh environments where temperature, weather and other factors set high requirements for functioning. For wireless communication, the movement of troops brings a challenge with the mobility of the networks. In addition to that, the military networks are located in a hostile environment, where an active adversary is always present. Hence, for network availability and usability, it is important that the networks are well secured from external and internal attacks. The military networks are more and more based on commercial technologies and protocols. Military specific technology is costly, and it requires special knowledge for maintenance and configuration. In a case where commercial technologies do not fulfill high military requirements, some modifications and development are always needed. However, commercial hardware, applications and protocols are widely used in tactical networking systems.

For communications, the tactical environment or level of military operations is the most challenging. Figure 1 presents some fundamental characteristics of military tactical networks. On the tactical level, basic communication infrastructure is based on combat net radio, mobile (IP) nodes and long-haul radio relays. The priority features of the tactical network may cause traffic delay because data traffic of these networks may have to compete for bandwidth with other services at higher priority. For example, data traffic may have to unexpectedly wait several seconds or more while high-priority voice traffic is carried on the same underlying links. The tactical networks also may also have especially strong infrastructure protection requirements (Fall 2003).

3.2 Information services

Military information services are typically built using commercial technologies and products. However, there are some special services that are based on military-specified protocols, applications and interfaces. Figure 2 illustrates a typical information service structure. The figure shows a simplified view of the many protocols that impact network communications. It is to be noted this is just a very small portion of the all available protocols and applications in use. More important is to realize that each of these protocols could create cyber security problems because they are capable of being abused by potential adversaries (Joshi et al 2008).

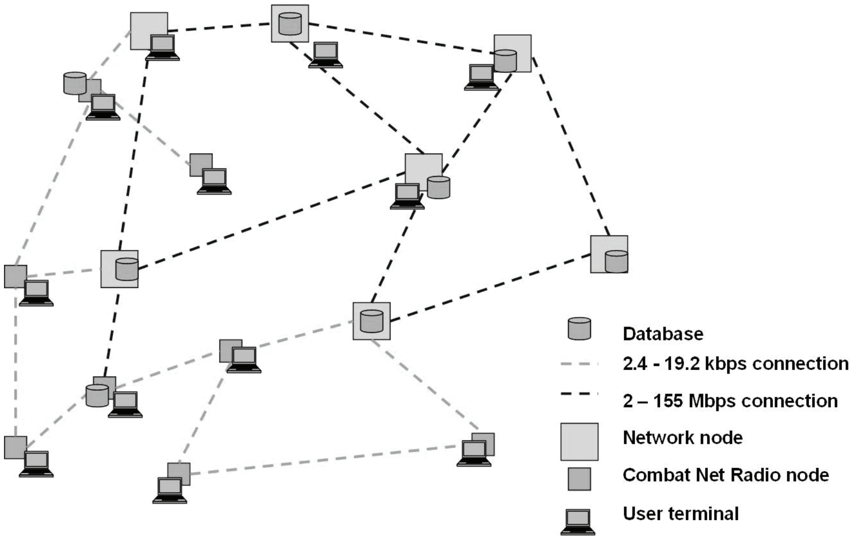


Figure 1: Characteristics of military tactical networks

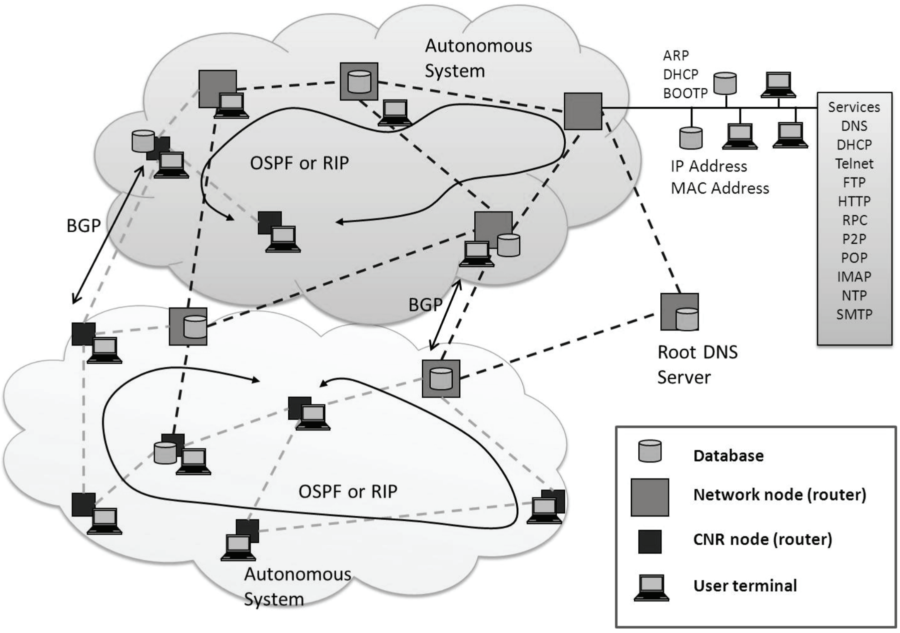


Figure 2: Many protocols are implemented in network communications and information systems.

In a typical scenario, a client application in host *A* wants to connect a server or service on host *B* on network *Q*. The client and server applications run as processes on the respective hosts, and the client application generates data that is sent to a lower level protocol for transportation. At the transport layer, a port number (P_A, P_B) identifies the process in the hosts. The hosts also have IP addresses (IP_A, IP_B) that belong to their networks. A connection between the client and server can thus be uniquely identified through the tuple < P_A, IP_A, P_B, IP_B >.

Network interface cards only recognize the MAC address. When host *A* creates a MAC frame, source and destination MAC addresses are added into the frame. While host *B* belongs to another network, a mapping between MAC addresses and IP addresses must be provided at a network gateway. The mapping can be obtained using the address resolution protocol (ARP). The gateway is responsible for routing IP datagrams to another router in the networks. The IP packets are routed in the network using routing information through routing protocols like the routing information protocol (RIP), open shortest path first (OSPF) and border gateway protocols (BGP). Except of knowing an exact IP addresses, the applications and services of host *A* typically use domain name service (DNS) to determine the IP address of host *B*. The DNS uses a hierarchical structure in which a local name server that is known to every host in the network. If the address is not known by the local name server, it contacts a root name server (13 servers worldwide). (Joshi et al 2008)

Cyber security problems take place for a variety of reasons, but one common reason is that servers listening at known ports have bugs in implementations. For example, a hostile adversary may create packets that can be sent to buggy services. When the service is compromised, it can enable the attacker to take control over the host, which means that the attacker can install malicious code or steal data files.

4. Cyber threat

4.1 Cyber threat overview

The emergency of large-scale cyber operations has moved network security attacks from the realm of hactivists to criminal organizations and states what makes threat more dangerous with potential for great economic and political harm. At the same time, organizations, business, governments and other actors are globally networked which provides attackers new and even better opportunities to conduct cyber attacks. There are numbers of techniques and methods to attack in cyberspace, and thus it is not possible to list all these methods of attack and their variations. In the following, some attack methods (Schiller 2010) are listed (in order of mildest to severe):

- Cyber espionage is the act or practice of obtaining secrets (sensitive, proprietary or classified information) from enemies using illegal exploitation methods on networks, software and or computers.
- Web vandalism includes e.g. defacing web pages or denial-of-service attacks. This is normally shifty combated and of little harm.
- Propaganda is used to spread political messages through networks.
- Gathering data means that classified data that is not securely stored or handled is intercepted and even modified, making espionage possible from the other corner of the world.
- Distributed denial-of-services attacks are generated using a large number of computers controlled by an attacker launching a DoS attack against target systems. A huge amount of traffic prevents normal users to access the service.
- Equipment disruption may put soldiers and troops in high dangerous. In military operations communication equipment is vital. Orders and communications can be replaced or interrupted by using different exploitation methods.
- Attacking critical infrastructure leads to damage in a real world. The purpose is not to steal or modify secret data but to affect critical infrastructure systems (power, water, fuel, communications, commercial, and transportation), and make them malfunctioning.
- Compromised fake hardware may include a back door for the attacker. Common hardware used in computers and network devices may contain malicious code hidden inside the software, firmware or even the microprocessors.

4.2 Cyber threat on military information services

Cyber attacks usually do not take place in one shot. Typically, the attacker first engages in mapping out the opponent's networks, resources, IP addresses, open or vulnerable services, and so on. This is called reconnaissance or cyber intelligence, and the attacker may try to get information that appears to be harmless if discovered, but may have some impact on cyber security later. The reconnaissance phase is followed by exploitation of vulnerabilities, information thief, taking over of hosts etc. In the following, some examples of cyber security attacks also relevant to military information systems are described in more detail.

At the transport layer, TCP is the most common protocol. It is used by many applications layer protocols such as HTTP and FTP. A TCP SYN flood attack exploits the three-way handshake of opening a TCP connection in which SYN and ACK flags are exchanged between hosts A and B. During the attack, an adversary sends a flood of crafted SYN segments to a server with spoofed source IP addresses. Since only a set number of TCP connections are accepted, services are denied to valid users. Lacking of authentication of the source IP address makes it difficult to block such attacks because separating the valid users from malicious requests is hard.

Address spoofing and sequence number guessing attacks are methods to get access to the services since the services use the IP address or host name to provide access. Spoofing IP addresses and host names is trivial, and there have been some attacks where root access to a host has been obtained by sending crafted packets with spoofed IP addresses. Sequence number guessing is used to exploit the protocols carried inside IP packets (such as TCP or DNS). Sequence numbers are provided to separate handshakes. If the IP address is spoofed and the attacker wants to fool the server into believing that a valid user has been connected, the attacker must guess the sequence number generated by the server. The sequence number is supposed to be random and hard to guess, but poor implementations of the protocols have allowed malicious entities to easily guess the sequence number.

Worms are self-replicating, malicious software pieces that crash hosts or services, open back doors to perform hostile operations. A worm probes other hosts for bugs and vulnerabilities for exploitation. The worm typically sends crafted packets to certain port numbers and IP addresses. Recent worms are intelligent. They may use search engines to discover vulnerable hosts. They are also able to use email addresses and address book entries. It is common to include zero-day exploits in worms to make it almost impossible to patch the exploit in time, enabling the worm spread extremely fast. Phishing is a growing social engineering attack method. In phishing, legitimate users are fooled into revealing information such as passwords, credit card numbers and so on by making them open legitimate looking email attachments, or click a malicious link on a web site. Even more dangerous attack is pharming in which DNS caches are poisoned with fake entries so that a user is directed to a fake web page although a valid URL was typed in. DNS cache poisoning is possible when servers use old, vulnerable versions of software.

Wireless military communication systems such as tactical mobile ad hoc networks provide even more attack surfaces for the enemy. Naturally, the security threats challenged by the tactical networks are the extension and expansion of that are confronted by wired networks, but added with the threats caused by wireless channels and interfaces. Table 2 lists some major cyber threats on the tactical wireless systems. The table also shows the exploited protocols on each threat of cyber attack. (Shi-Chang et al 2010)

Table 2: Passive and active cyber threats on the military information service infrastructure

Threat	Description	Security risks
Passive		
Wiretapping	Interception of transmitted data packets.	Access to confidential information
Traffic analysis	Analysis of characteristics of packet frequency, length and etc.	Disclosure of the structure of network and communications
Active		
Packet replay	A data packet is re-transmitted.	Unauthenticated functions in the target system (malfunctioning)
Fraud counterfeiting	A network entity behaves as another entity to carry out network activities,	Communication with unauthenticated network entities.
Packet tampering	Data is modified, or deliberately delayed transmission, or a passive change in the order.	Integrity failure. Injection of malicious code.
Denial of service	An authorized entity cannot access to the services	Service availability failure. Critical systems and services are not available when required.

5. Cognitive information service configuration

5.1 Cognitive networks and cognitive process

In the cognitive information service configuration, a cognitive process is required to provide intelligent, self-learning controlling and maintenance for the information services. The cognitive process is introduced by the cognitive networks of which basic idea is discussed in (Mahmoud 2007), (Thomas et al 2005) and (Thomas et al 2006). Cognitive network is a network that can dynamically adapt its operational parameters in response to user and service needs or changing environmental conditions. The network can learn from the adaptations and exploit knowledge to make upcoming decisions. Originally, the applications of cognitive networks enabled the vision of pervasive computing, seamless mobility, ad hoc networks, and dynamic spectrum allocation, among others. The most important element of the cognitive network is a cognitive engine or process that includes all learning and decision-making features needed to reach service level goals. The cognitive process tries to exactly perceive the current network situation and plan and decide to meet the end-to-end goals in an entire network aspect. The process could be viewed as the commonly known OODA loop in which a network or service observes, orients, decides and acts. Figure 3 shows the phases of the loop in context of cognitive processing.

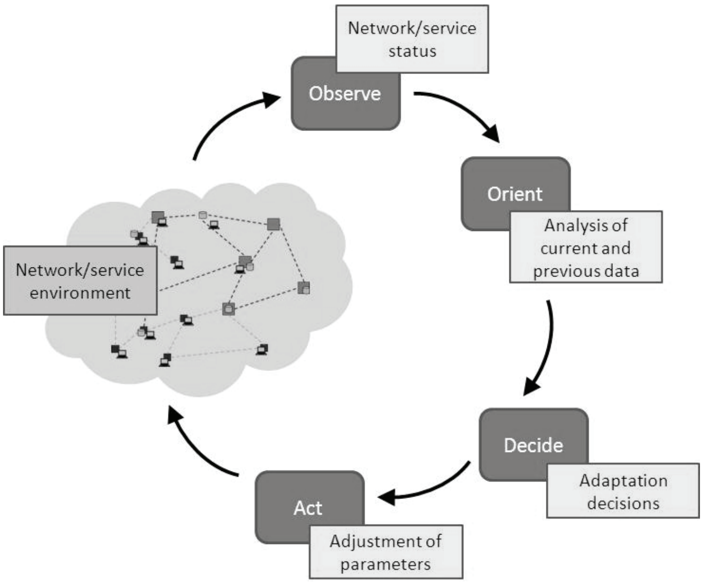


Figure 3: Cognitive process follows the OODA loop.

The observation phase is critical because the effect of a cognitive network's decisions on the network performance depends on how much network state information is available. If a cognitive network has knowledge of the entire network's state, cognitive decisions should be more "correct" than those made in ignorance. For a large, complicated system such as mobile ad hoc networks, it is unlikely that the cognitive layer of the network would know the total system state. It could be very high costly to communicate status information beyond those network elements requiring it, meaning the cognitive network will have to work with less than a complete picture of the network resource status. The orientation phase plays also an important role in the cognitive process. In this phase, all observed information and the previous knowledge are added together and analyzed. Filters and weighting are examples of methods used in the orientation phase. In the decision phase, the best decision for the desired end-to-end information security goal is made. Finally, the parameters of the cognitive network elements are adjusted and modified in the acting phase. The cognitive elements are allowed to act selfishly and independently (in the context of the entire network) to achieve local goals.

5.2 Functional architecture

Figure 4 depicts the functional architecture of the cognitive information service configuration. The architecture consists of three layers that are the target layer, cognitive layer and reconfiguration layer. The target layer is guided by security policies, security situation awareness (vulnerability libraries) and service level agreements. A security policy defines the security goals and elements of an organization's computer systems. It declares what security means are used in different situations, and how and who are authorized to the system (Joshi et al 2008). The vulnerability libraries include information about current software and hardware vulnerabilities. The libraries are important to keep the system updated, and to avoid vulnerable system configurations. A service level agreement (Hiles 2002) sets minimum requirements for information services. Service level agreements can contain numerous service performance metrics with corresponding service level objectives The purpose of the target layer is to define an end-to-end service target for each information service. A service goal can be for example an availability level or sub services available for dedicated users. The goal is fed to the cognitive layer as a set of individual goals (service element targets) for each cognitive service elements. The main task of the cognitive service element is to decide how network and server parameters are reconfigured in network and service elements at the reconfiguration layer.

Each cognitive service element uses the previously defined cognitive process (Figure 3) to make a decision for adaptation. The decision-making is based on learning from history data, current situation and the desired service element (SE) target. The optimization of decisions is processed at several levels depending in the service structure. The optimization may occur at the service element level, server level or entire service system level. At the SE level, decisions are made with information from a single element. The server level optimization means that all the elements in a server provide information for decision-making. At the entire system level, information from all servers is used. The control channel is used for status information sharing between SEs to guarantee the most optimal decisions. The software adaptive service and network element (SASNE) is required for adjusting service configuration and networking parameters. The cognitive behavior needs software-adaptable hardware and devices. SASNE transforms the higher level decisions to actual configuration orders at the configuration level. A configuration order may include e.g. new IP addresses, or a change of transport protocol. Table 3 lists some reconfigurable protocols or attributes.

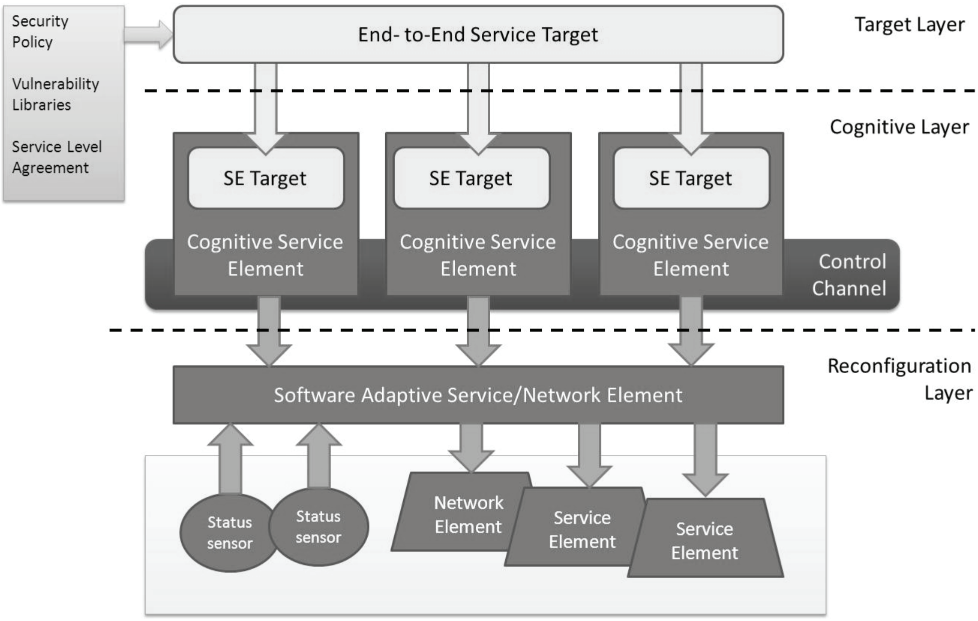


Figure 4: The overview of the cognitive information service structure

Table 3: Reconfigurable protocols or attributes by the cognitive process

Protocol/attribute	Alternatives/options
Application protocol	HTTP, HTTPS, FTP,
Transport protocol	TCP, UDP, others
IP address	Changing IP addresses
Port number	Changing standard ports, changing port numbers randomly
Encryption key	Key lengths
Encryption algorithm	AES, DES, Blowfish
Firewall filters	Blocking certain sources and destinations, dropping desired protocols

The status sensors monitor current situation. Sensor information is used for decision-making, and the sensor information may launch a new adaptation phase. The adaptation phase is also initialized when the end-to-end target is changed according to vulnerabilities, policies or service level modifications.

5.3 Protection against cyber threats: An example of DDoS defence

The protection against cyber attacks is based on the ability to change the system configuration according to status sensors, target goals, current vulnerabilities, or security policies. For enemy's intelligence the information system is a moving target which may require a lot more intelligence and data gathering than traditional, static service configurations. The purpose of this section is to demonstrate the protection mechanism of the cognitive information service configuration. A distributed denial-of-service (DDoS) attack is used as an example for the demonstration. In a DDoS attack, a large number of computers are used to launch a coordinated attack against a target system or service (Douligeris and Serpanos 2007). The effectiveness of the DoS significantly is multiplied by using the resources of multiple compromised computers, which serve as attack platforms. Minimum information required to launch a DoS attack is just the IP address and the port number of the target service. Depending on the attacking method more information is needed.

DDoS defense mechanisms include various activities to prevent, detect, response and tolerate intrusions. The best mitigation strategy is to completely prevent the attacks. The DDoS attacks should be stopped in the first possible point after launching. The preventive DDoS defense mechanisms include (Douligeris and Serpanos 2007):

- Globally coordinated filters that drop attack packets before they cause serious damage.
- Disabling unused services
- Applying security patches can shield the hosts against attacks. Computers should be updated according to the latest security patches.
- Changing the IP address is a simple way to guard against a DDoS attack.
- Disabling IP broadcasts prevents the use of host computers as reflectors in flood attacks. This intrusion prevention mechanism can be effective only if all the neighboring networks have also disabled IP broadcasts.
- Load balancing is a simple approach enabling network providers to increase the connections' bandwidth and prevent their crash in case a DoS attack.
- Honeypots can be used to avoid DDoS attacks. Honeypots can be used to deceive the enemy to attack the honeypot instead of the system being protected.

Figure 5 depicts the architecture of DDoS attacks and cognitive service reconfiguring as a prevention method. In the DDoS attack, an attacker controls a few handler computers that use a large number of agents to generate flooding traffic to the target system. While attacked the target server reconfigures system parameters so that the DDoS attack cannot cause any damage. The cognitive process changes the system parameters related to the defensive mechanisms described above. The process adjusts filter settings so that incoming packets from illegitimate origins are blocked out. Unused services are disabled (unused services should be disabled by default).

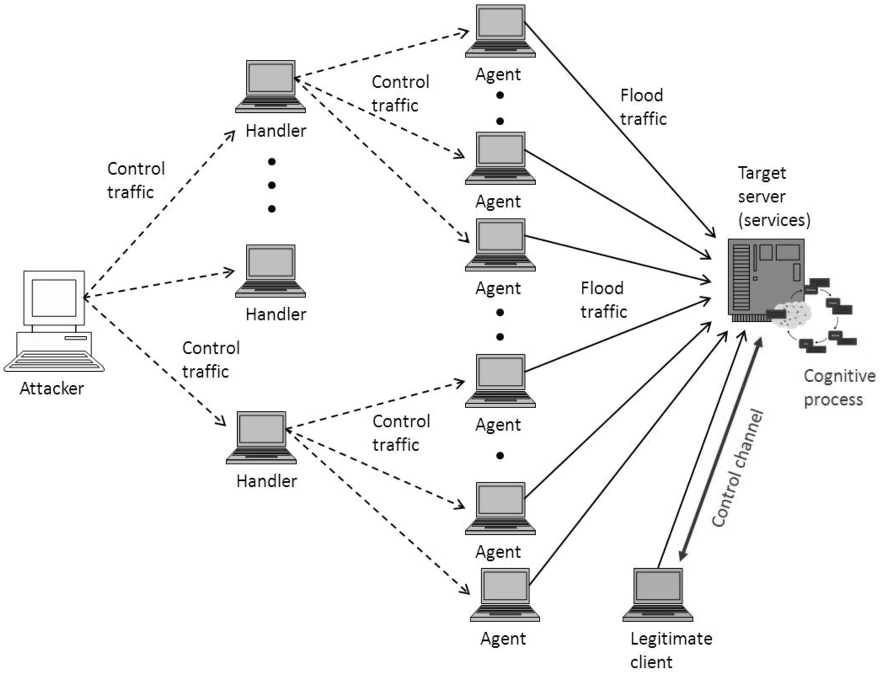


Figure 5: Architecture of DDoS attacks and the cognitive service reconfiguration

The cognitive process takes care of continuous patching. According to vulnerability information, the system patching is guaranteed automatically. The IP addresses and port numbers of services are changed immediately after the attack. This requires a feedback channel to the users so that the user knows correct service addresses. In Figure 5, the control channel between the target server and the client is used to for these purposes. In addition, the cognitive server may build honeypot services that deceive the enemy. The easiest way is to use original settings such as original IP addresses and port numbers making the adversary believe the target services are still up and running.

6. Implementation challenges

Cognitive network research has already been carried on years, but there are still few prototypes to demonstrate factual cognitive features. Thus, this idea of cognitive information service configuration is just taking initial steps, and many details are not solved yet. One of the major challenges with the cognitive process is a decision-making process and learning functionality. It is possible to teach servers to act in a certain way for a certain cyber attacking scenario, but it is not clear how the systems could function in an optimal way in those situations with no prior data. To keep services available for legitimate users the control channel is required. This provides a new attack surface for an enemy. Control channel attacks may paralyze the entire information system leaving the user without any service. The control channel adds traffic between the client and the server, and it requires high- level security protocols that also consume the limited bandwidth of the tactical networks. A challenge is also how end-to-end targets are formulated so that lower level elements are able to make the most optimal decisions. It could be difficult to formalize vulnerability information into an understandable format. This could be provided by using standardized format although a risk is that some information is lost.

7. Conclusion

Modern cyber threats and attack methods require new approaches for information service protection. This paper proposed a cognitive service configuration model for the military networks. In the model, the cognitive process controls all the service configuration parameters. The parameters and protocols are configured automatically according to the monitored situations. The paper presents the functional elements of the cognitive information service configuration structure. The purpose was to show an initial idea of making a military information service system as a moving target for the enemy attacks. The paper illustrates that the cognitive service establishment and maintenance could be a promising approach to provide stronger protection on the computer networks and services. The DDoS example demonstrates that using an automatic reconfiguration the damages of attack may be avoided. The future work includes detailed research on the architectural elements. The cognitive process and decision-making needs to be studied in more details. Also, the control channel issues require more research. This paper is also missing initial performance and functionality calculations to demonstrate and analyze the overall maturity of the proposed model.

References

Bodeau, D.J., Graubart, R. and Fabius-Greene, J. (2010) “Improving Cyber Security and Mission Assurance via Cyber Preparedness (Cyber Prep) Levels”, *Proceedings of the IEEE Second International Conference on Social Computing (SocialCom)*, pp 1147 – 1152.

Buford, J.F., Lewis, L. and Jakobson, G. (2008) “Insider Threat Detection Using Situation-Aware MAS”, *Proceedings of the 11th International Conference on Information Fusion*, pp 1 – 8.

Douligeris, C. and Serpanos, D. N. (2007) *Network Security: Current Status and Future Directions*, John Wiley & Sons, Inc.

Fall K. (2003) “A Delay-Tolerant Network Architecture for Challenged Internets”, *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications,* pp 27-34, New York, NY.

Hiles, A. (2002) *The Complete Guide to I.T. Service Level Agreements: Aligning It Services to Business Needs*, Rothstein Associates Inc, 2002.

Jimenez-Molina, A. and In-Young Ko (2011) “Cognitive Resource Aware Service Provisioning”, *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, Lyon.

Joshi et al (2008) *Network Security: Know It All*, Elsevier Inc.

Kliazovich, D.et al (2009) “Cognitive Information Service: Basic Principles and Implementation of a Cognitive Inter-Node Protocol Optimization Scheme”, *Proceedings of the IEEE Global Telecommunications Conference, 2009*, Honolulu, HI.

Mahmoud, Q. (2007), *Cognitive Networks: Towards Self-Aware Networks*, Wiley-Interscience.

Morris et all (2011) “A perceptually-relevant model-based cyber threat prediction method for enterprise mission assurance”, *Proceedings of the IEEE First International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support*, pp 60 – 65.

Schiller, J. (2010) *Cyber Attacks & Protection*, CreateSpace, Paramount, CA.

Shi-Chang, L. et al (2010) “Research on MANET Security Architecture Design”, *Proceedings of the* International Conference on Signal Acquisition and Processing, pp 90 - 93.

Thomas, R., DaSilva, L. and MacKenzie, A. (2005) “Cognitive networks”, *First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp 352 – 360.

Thomas, R., Friend, D., DaSilva, L. and MacKenzie, A. (2006) “Cognitive networks: adaptation and learning to achieve end-to-end performance objectives”, *Proceedings of the IEEE Communications Magazine*, Vol 44, Issue 12, pp 51 - 57.

Zheng, Y., Lu, H. and Sun, Y. (2008) ”An Intelligent and Cognitive Service Delivery Platform Model”, *Proceedings of the IEEE Second International Symposium on Intelligent Information Technology Application, Shanghai*.

Efficient Remote Authentication

Michael Kiperberg and Nezer Zaidenberg
University of Jyvaskyla, Jyvaskyla, Finland
michael@truly-protect.com
nezer@truly-protect.com

Abstract: Kennel, R. and Jamieson, H. (2003) describe a method of a remote machine authentication. By authentication the authors mean that the remote machine is non-virtual and the operating system on the remote machine is not malicious. The described method does not consider the variety of versions of each operating system. The description completely ignores the existence of modules that can be plugged into the operating system. In this paper we adapt the method described by Kennel, R. and Jamieson, H. (2003) to the real world, so that the method can be applied without prior knowledge of the operating system or the modules on the remote machine.

Keywords: virtual machine, digital rights management, remote authentication, database

1. Introduction

Averbuch, A., Kiperberg, M. and Zaidenberg, N. (2011) show how a content distributor can verify that a remote machine is non-virtual and that the operating system is authentic. This verification procedure allows the distributor to execute her game on the remote machine while keeping the game's code hidden. The verification procedure consists of two interleaved processes.

The first process, whose goal is to verify that the operating system is authentic, computes a checksum of the memory locations that hold the code of the operating system. The idea of computing and sending the checksum of some memory region is not new. For example the AOL Instant Messenger was sending to a server a hash of a requested memory region, and the client was blocked if this hash was not the expected one (AOL, 2012; PyxisSystemsTechnologies, 2002).

The second process, whose goal is to verify that the remote machine is non-virtual, is executed in parallel with the first process. The second process combines the checksum computed so far with a value of some performance counter. The last idea is based on the assumption that performance counters cannot be easily simulated in a timely manner (Bedichek, R., 1990; Magnusson, S. P. and Werner, B., 1994; Witchel, E. and Rosenblum, M., 1996).

Therefore, if the remote machine is a virtual machine, the result produced by the verification procedure will be either erroneous or computed after a long period of time. The distributor rejects all the erroneous and the late results, thus guaranteeing the two properties of the remote machine: the machine is non-virtual and the operating system is authentic.

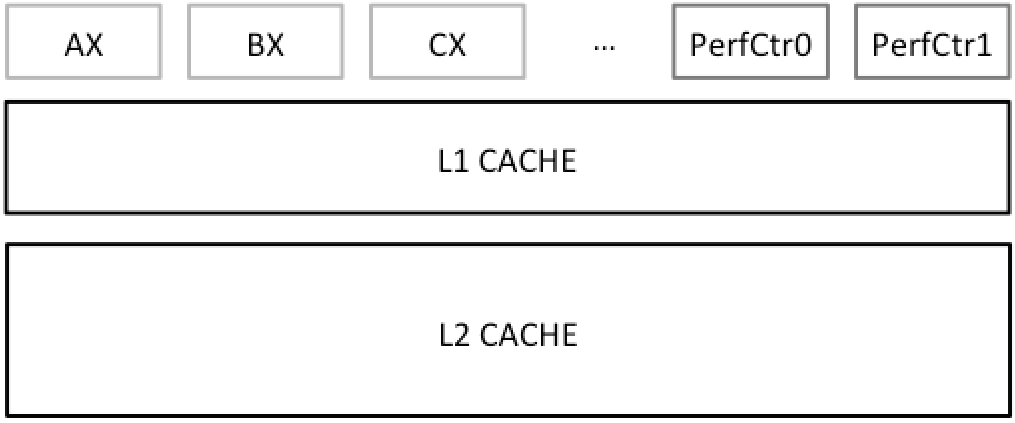


Figure 1: CPU internal memory. The registers AX, BX and CX can be accessed in non-kernel mode. The registers PerfCtr0 and PerfCtr1 can be accessed only in kernel mode. The L1 and L2 caches access policy is determined by access policy of the underlying memory

In this paper we concentrate on the second property, i.e. the authenticity of an operating system. By "authenticity of an operating system", we mean that the code that is executed in kernel mode is trusted by the distributor. The reason behind the importance of authenticity lies in the way the game is executed on the remote machine. In order to be able to execute the game on a remote machine without revealing the game's code, the distributor supplies the game in an encrypted form. The decryption key is transferred to the remote machine that passes the verification procedure. The transferred key is stored in one of the CPU registers that can be accessed only in kernel mode. That is why it is so important to be sure that no malicious code gets executed in kernel mode. Figure 1 depicts the internal structure of a CPU memory.

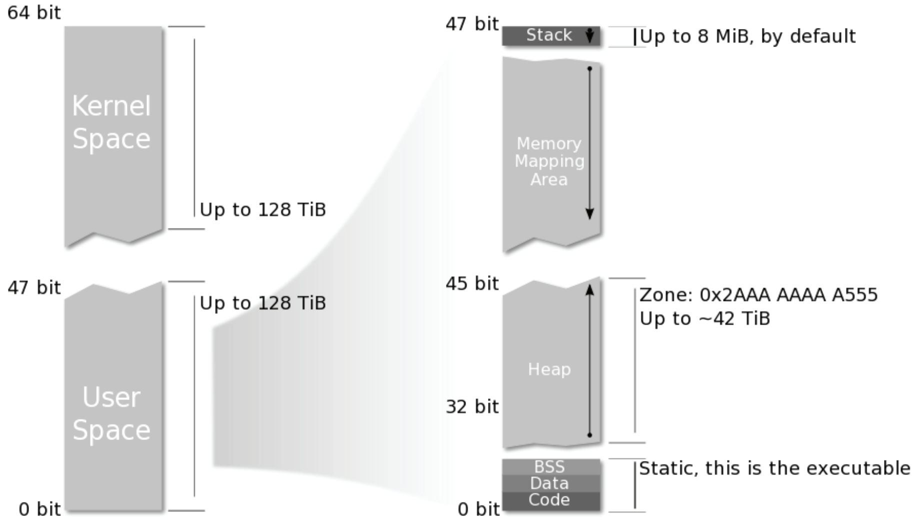


Figure 2: Memory layout of a process on 64-bit machine running Linux

We should note that not only the code of the operating system is executed in kernel mode, but also the code of device drivers and other kernel modules is executed in kernel mode. Figure 2 depicts the layout of the process virtual memory on 64-bit machine running Linux. The kernel and kernel modules are located in the "Kernel Space" segment depicted in Figure 2. However, the exact location in memory of each kernel module or even its presence can differ from machine to machine, thus making it difficult for the distributor to reproduce the result of a verification procedure.

In this paper, we try to overcome the above difficulties by pre-calculating the results.

The paper is organized as follows. Section 2 gives a detailed description of the problems that are then solved in the section 3. Section 4 summarizes the paper and outlines the future work.

2. Meaning

Since CPUs today cannot execute program in an encrypted form as was proposed by Best, M. R. (1980), the instruction of any encrypted program should be decrypted prior to their actual execution in a CPU. Therefore, the decryption key should be held in the remote machine that executes the program. Since we do not trust any component on the remote machine, in the sense that we believe that a malicious user can analyze externally the content of any component, it seems to be impossible to allow program execution on a remote machine without revealing the code of this program. That is why we refine our trust policy to include a single component – the CPU.

Our approach to executing a program without revealing its code is as follows. Suppose the user wishes to play some game. The user asks the distributor to send him this game. The distributor requests the user to validate his machine; we call the user's machine the "remote machine". The distributor generates a test, which is a code to be executed on the remote machine. This code computes a checksum of some memory region. The distributor sends the test to the remote machine and waits for the result. If the result is correct, the distributor encrypts the game using some random key and sends the key and the game to the remote machine. The remote machine stores the key in a CPU register that can be accessed only in kernel mode (see Figure 1). This key is used to decrypt the game function-by-function on-demand. Since only small portions of the game are decrypted at any time, there is a small probability that they will be evicted from the CPU cache.

The only problem with this approach is that any code running in kernel mode has access to the register that contains the decryption key. This code can read the content of the register and print it to the screen, which will be enough to decrypt the entire program. That is why we should be able to guarantee that no such code is executed in kernel mode. On Linux, there are only two types of code that is executed in kernel mode. The first type is the kernel itself and the second type is the code of kernel modules that are plugged into the kernel.

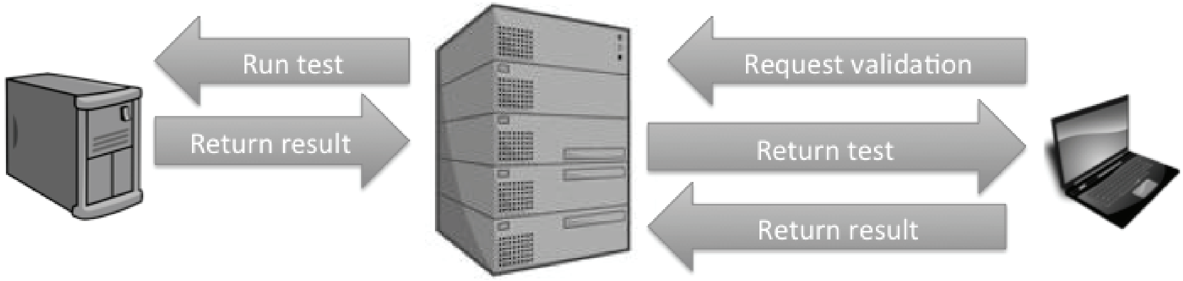


Figure 3: Naïve verification procedure. The remote machine (on the right) requests validation. The distributor (in the middle) generates a test and runs it on a local machine (on the left). The local machine returns the result to the distributor. The distributor sends the same test to the remote machine. The remote machine runs the test and returns the result to the distributor. The distributor compares the results received from the local and remote machines. If the results are equal the validation succeeded

In order to understand the difficulties that arise from the variety of kernels, modules and their combinations, we should explain first, how the distributor verifies the result of the test that it sent to the remote machine. The distributor verifies the correctness of the result by executing exactly the same test on some of its local machines. Then, it compares the result on the local machine with the result on the remote machine. The test is qualified as correct if and only if the results are equal. Figure 3 depicts the verification procedure. Obviously the results will differ if the two machines have different CPUs or contain different content in the memory region under test (see Figure 2). Note, that the later does not imply that the kernel or the device drivers are malicious; the memory may, as well, contain exactly the same drivers that for some reason were loaded in a different order and thus residing in different locations of the memory.

3. Solution

Kennel, R. and Jamieson, H. (2003) noted that the verification protocol should be CPU aware. To accommodate this, their protocol starts by sending a message containing the CPU information from the remote machine to the distributor. We propose to include in this message additional info about the version of the kernel, the list of kernel modules, their layout in memory, etc.

Upon receiving this information, the distributor can setup the same arrangement on one of his local machines and afterwards run the test. The problem with the above approach is that it is very time and resource consuming. The setup can take tens of minutes. For each such setup the distributor has to allocate a machine. This machine cannot be used for other tests. Therefore, in the real world, it will not be feasible to perform such a setup.

In order to solve this problem, we propose to prepare tests for many different configurations of CPU model, kernel version and kernel modules layout. Each such test and the corresponding result will be stored in the "test database". This database will be constantly augmented by new tests. Upon request from a user the distributor will pick one of the tests that suits this user and send the test to the user. Figure 4 shows the proposed verification procedure.

It is possible, however, that no tests are available for the configuration of a particular user. There are many reasons for this. It is possible that the user uses a new version of the kernel that was not known to the distributor. It is possible that one of the kernel modules was not known to the distributor. It is also possible that all the components were known to the distributor, however it is the first time the distributor encounters the combination of these components together.

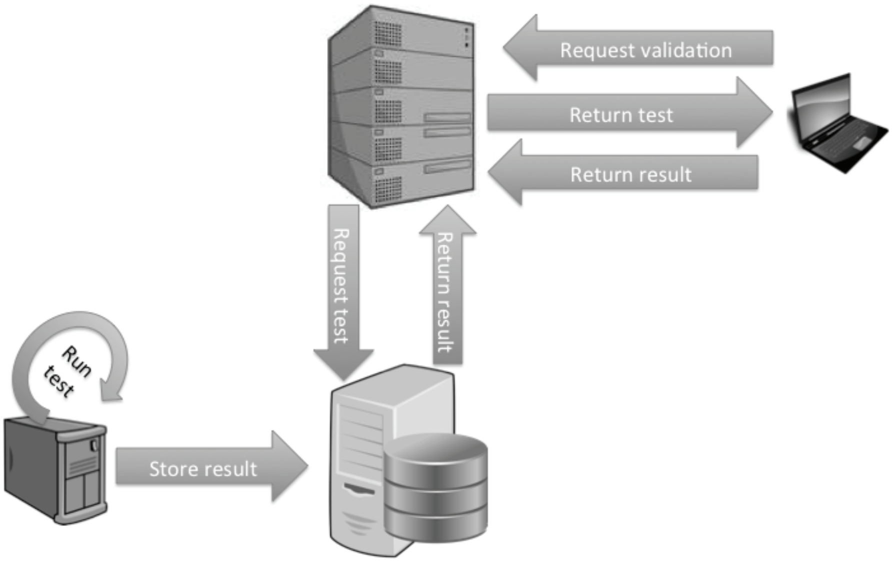


Figure 4: New verification procedure. The local machine (on the left) generates and stores test results for various tests in the database (in the middle-bottom). The remote machine (on the right) requests validation. The distributor (in the middle-top) fetches some test result from the database and sends the corresponding test to the remote machine. The remote machine runs the test and returns the result to the distributor. The distributor compares the result of the remote machine to the result fetched from the database. If the results are equal the validation succeeded

In our opinion, the first two cases are relatively rare. Although it is time consuming to augment the database with a new version of kernel, new distributions of Linux are released only once or twice a year. A large percentage of kernel modules are actually device drivers. In spite of the fact that there are many device drivers, an average user uses only a few of them. Figure 5 demonstrates that there are less than a thousand device drivers. Most of them are present in all versions of the kernel. Therefore, we believe that drivers are not released very often. The remote machine will send every new kernel component (either the kernel itself or one of its modules) to the distributor for investigation. The distributor will assure that the component is not malicious and create a test for the configuration of the remote machine.

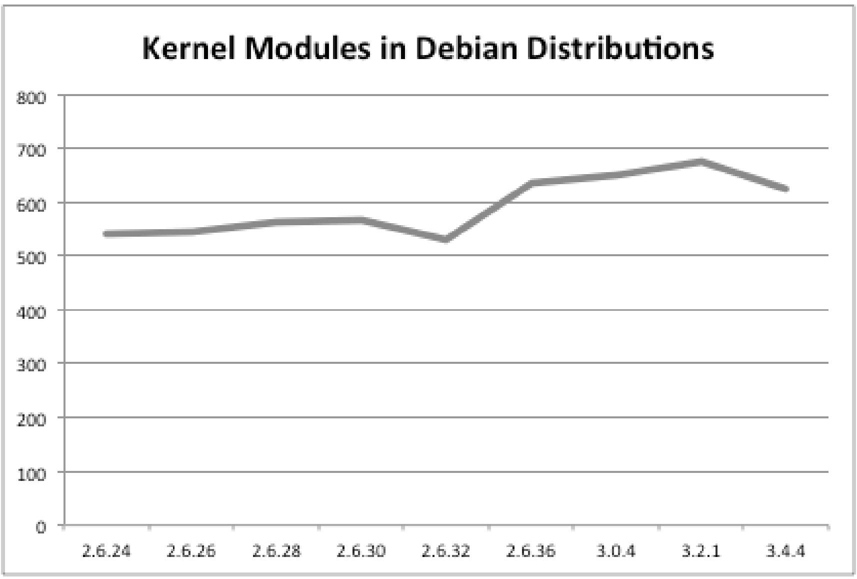


Figure 5: Kernel Modules in Debian Distributions. The X-axis represents the various versions of the Linux kernel that were used in the Debian distributions. The Y-axis represents the number of kernel modules in the directory /lib/modules/A.B.C/kernel/drivers/ where “A.B.C” is the kernel version

The third case is much simpler since it does not involve a, possibly manual, verification of a kernel component. The test for the new configuration can be generated automatically.

During the construction of tests and results for new, the user having one of these configurations will not be able to start the verification procedure. In this case, the user will be informed of the situation and asked to try again later.

4. Procedure entities role

This section gives a detailed description of the verification procedure. The description covers all the entities participating in the procedure (see Figure 4).

The procedure begins when the remote machine requests to perform verification. In the request message, which the remote machine sends to the distributor, it indicates the configuration of its system. The configuration consists of the following information:

- CPU model,
- Kernel version and location in memory,
- List of kernel modules, their version and locations in memory.

The verification procedure supports only relatively new CPU models, since newer models can efficiently simulate the older ones. Therefore, after receiving the configuration C, the verification procedure fails immediately if the CPU model is not supported. Otherwise, if the configuration was not known previously, the distributor adds it to the list of known configurations and fails the verification procedure. If the configuration was known previously, the distributor fetches the set S of tests and results for C. Then a pair (T, R) of a test and its result is picked uniformly at random from S. The distributor sends T to the remote machine and sets a timer. The remote machine runs the test T and returns the computed result R' to the distributor. When the distributor receives the result R', it checks whether R' was received within an allowable interval of time. If not, the verification procedure fails. Otherwise, R and R' are compared and the remote machine is informed accordingly.

Each local machine produces new tests in an infinite loop as follows. It scans the list of all known configurations in the database, and fetches only those configurations that correspond to the CPU of the local machine. It then chooses a configuration with the smallest number of tests and produces a test for it. Note that there is at least one local machine for every supported CPU model.

5. Conclusion

This paper describes a feasible method that allows verifying that a remote machine is non-virtual and its operating system is authentic. The described method allows accommodating for bursts of verification requests, thus making the entire system more scalable. The method breaks the dependency between the test generation procedure and the remote machine verification procedure. The later makes it possible to share the generated tests between multiple verification procedures, thus allowing even greater scalability.

In our future work we will concentrate on porting the ideas of this paper to other operating systems, specifically Windows and Mac OS X. Likewise, we will try to automate the verification of kernel components.

References

AOL (2002) *The America Online Instant Messenger Application*, http://www.aol.com/.

Averbuch, A., Kiperberg, M. and Zaidenberg, N. (2011) *An Efficient VM-Based Software Protection*, NSS.

Bedichek, R. (1990) *Some efficient architecture simulation techniques*, USENIX Technical Conference.

Best, M. R. (1980) *Preventing software piracy with crypto-microprocessors*, COMPCON.

Kennel, R. and Jamieson, H. (2003) *Establishing the genuinity of remote computer systems*, 12th USENIX Security Symposium.

Magnusson, P. S. and Werner, B. (1994) *Some efficient techniques for simulating memory*, Technical Report R94-16, Swedish Institute of Computer Science.

PyxisSystemsTechnologies (2002) *AIM/oscar protocol specification: Section 3: Connection management*, http://aimdoc.sourceforge.net/faim/protocol/ section3.html.

Witchel, E. and Rosenblum, E. (1996) *Embra: Fast and flexible machine simulation*, Measurement and Modeling of Computer Systems.

Retrospective Evaluation of Cyber Security Strategic Reports for the Next two Decades: From 2000 to 2030

Ahmet Koltuksuz
Yasar University, College Of Engineering, Dept Of Computer Engineering, Turkey
ahmet.koltuksuz@yasar.edu.tr

Abstract: Retrospective taxonomical evaluation of the strategic cyber security reports of the last decade is very much important if we are to understand and to appreciate how the very concept of Cyberspace has evolved along with other and equally important concepts and, definitions like cyber security, cyber terrorism, cyber warfare, information warfare for which we seem to struggle every day. For that reason, seven cyber security strategy reports; covering a decade from 2002 to 2011, evaluated, compared and contrasted. Armed with the strategic cyber security reports of the last decade, as the second aim of this paper, we will try to examine and understand if the near future cyber security strategies will be valid with our current understandings. In order for us to do that two next decade projective papers will be put under scrutiny. Our gatherings from all of those examined, plus experiences gathered in the field over the years will be provided as conclusions

Keywords: cyberspace, cyber warfare, information warfare, strategic cyberspace analyses, strategic reports and recommendations for cyber security

1. Introduction

Ever since Gibson has coiled the term Cyberspace, in his celebrated science fiction novel Neuromancer in 1984 as “a consensual hallucination experienced daily by billions of legitimate operators, in every nation, by children being taught mathematical concepts. A graphical representation of data abstracted from the banks of every computer in the human system. Unthinkable complexity.” (Gibson, 1984), it has been defined many times over and, has been extensively used by many researchers. Much later, another definition of Cyberspace has surfaced as “a time-dependent set of interconnected information systems and the human users that interact with these systems.” (Ottis, 2010).

Furthermore; Ottis also attributed any cyber conflict due to the control of cyberspace with different aims such as from obtaining illegal revenue, information gathering, and disruption of the operations of the enemy to the fights over domain name ownership (Ottis, 2010).

Nye approaches cyberspace architecturally and, delineates many layers of activities, starting with a unique hybrid regime of physical and virtual properties. “The physical infrastructure layer follows the economic laws of rival resources and increasing marginal costs, and the political laws of sovereign jurisdiction and control. The virtual or informational layer has economic network characteristics of increasing returns to scale, and political practices that make jurisdictional control difficult. Attacks from the informational realm where costs are low can be launched against the physical domain where resources are scarce and expensive. But conversely, control of the physical layer can have both territorial and extraterritorial effects on the informational layer” (Nye, 2010).

While one cannot imagine a daily life without it, the cyberspace is itself actually very new entity. Started with the initiation of US Department of Defense, the ARPANET, in 1969 with only a few connected computers, we reach to billions of users constantly connected in only four decades. And now, cyberspace is nothing but the life itself.

However having realized the importance of cyberspace, many universities, research institutes, think-tanks and other establishments & private corporations have come forward with strategic analyses and reports each of which got many findings and recommendations for further actions for governments and businesses to operate & to control in cyberspace.

Starting as early as 2000s, those strategic cyber security reports have tried to combine national security issues with that of cyber security which is essentially borderless and as a concept and as a reality out of reach of the power of national governments.

Therefore it is imperative to understand the evolution of the concept of cyber security by carefully examining and classifying those reports that came to life within last decade if we are to operate in cyberspace in the next two decades.

In this paper, some of the important cyberspace strategy reports and/or papers; which were prepared by different organizations for the President of the United States or by the office of the President of the United States to the various governmental bodies or to organizations covering a whole decade starting from the year of 2002 up to 2011, were selected in order to illuminate how the concepts of cyber security, cyber warfare and information warfare; in terms of both conceptually and practically, have been evolved.

As the aforementioned approach constitutes the first aim of this paper, the second one is; after establishing a firm taxonomical base through reports coming from the last decade, to evaluate and to discuss the strategic approaches of some of the recent reports in order to conclude and, to recommend some operational approaches for the next decade.

2. A decade of cyberspace strategies: 2002-2011

2.1 Report #1 - 2002: (Lewis, 2002).

In this report Lewis defines cyber terrorism as “the use of complex network tools to shut down critical national infrastructures” and delineates the steps of cyber terrorism as

- Exploit vulnerabilities,
- Penetrate computer networks and,
- Disrupt and shut down critical functions.

This paper also brings out a new definition to cyber-attacks as “Weapons of Mass Annoyance” and, claims that unless a cyber-attack is accompanied by a physical one it may be classed as more of an annoyance than a real threat. It concludes that the nations are more robust and resilient than early theories of cyber terror assumed.

2.2 Report #2 - 2003: (President, 2003).

This strategy document is actually one of the three key strategy documents. The other two complementing documents are

- The National Strategy for Homeland Security and,
- National Strategy for the Physical Protection of Critical Infrastructures and Key Assets.

The National Strategy to Secure Cyberspace is a call for rising cyber security awareness as well as putting processes in place to continuously identify and remedy vulnerabilities. It states three strategic objectives as “Prevent cyber-attacks against America’s critical infrastructures”, “Reduce national vulnerability to cyber-attacks”; and “Minimize damage and recovery time from cyber-attacks that do occur”. In order to achieve these strategic goals, five priority areas were defined and, those are as follows

- A National Cyberspace Security Response System,
- A National Cyberspace Security Threat and Vulnerability Reduction Program,
- A National Cyberspace Security Awareness and Training Program,
- Securing Governments’ Cyberspace and,
- National Security and International Cyberspace Security Cooperation.

Before providing a road map in the form of Actions and Recommendations which are quite lengthy and are still valid today, this report concludes with the following determinations:

- Reliance on cyberspace will only continue to grow in the years ahead.
- There is a need to work continuously to secure the cyber systems that control our infrastructures.
- This is a complex and evolving challenge, as new vulnerabilities and techniques surface each day.
- As it is, this is only a first step in a long-term effort to secure our information systems.

- Within the federal government Department of Homeland Security (DHS) will play a central role in implementing the National Strategy to Secure Cyberspace.

2.3 Report #3 2005: (PITAC, 2005).

This report was prepared by PITAC and, they provide the President independent expert advice in IT. As such, PITAC members are IT leaders in industry and academia with expertise relevant to critical national IT infrastructure. The report determines that:

- The Information Technology Infrastructure is 'Critical',
- Ubiquitous Interconnectivity means Widespread Vulnerability,
- Software is a Major Vulnerability,
- Attacks and Vulnerabilities are Growing Rapidly,
- Endless Patching is not the Answer,
- Fundamentally New Security Models, Methods Needed and,
- Central Role for Federal R&D needed.

The PITAC committee goes on with some key findings and related recommendations listed below as table 1

Table 1: PITAC committee findings & recommendations

#	Findings	Recommendations
1	Federal R&D budget provides inadequate funding for fundamental research in civilian cyber security.	The NSF budget for fundamental research in civilian cyber security should be increased by $90 million annually
2	The Nation's cyber security research community is too small to adequately support the cyber security research and education programs necessary to protect the United States.	The Federal government should intensify its efforts to promote recruitment and retention of cyber security researchers and students at research universities.
3	Current cyber security technology transfer efforts are not adequate to successfully transition Federal research investments into civilian sector best practices and products.	The Federal government should strengthen its cyber security technology transfer partnership with the private sector.
4	The overall Federal cyber security R&D effort is currently unfocused and inefficient because of inadequate coordination and oversight.	The Interagency Working Group on Critical Information Infrastructure Protection should be focal point for coordinating R&D. This group should be strengthened and integrated under the NITRD Program

2.4 Report #4 2008: (Lewis, 2008).

This report which is created by a CSIS commission opens with three major findings and they are as follows:

- Cyber security is now a major national security problem for United States.
- Decisions and actions must respect privacy and civil liberties.
- Only a comprehensive national security strategy that embraces both the domestic and international aspects of cyber security will make us more secure.

Report goes on with recommendations such as listed in table 2

2.5 Report #5 2009: (DHS, 2009).

Prepared by Department of Homeland Security, this report is one of a kind among the others for it specifically provides a multi-layered strategy for securing control systems. It identifies coordinating mechanisms such as

- Federal Partners Working Group
- Government and private Sector Coordination
- Planning
- Research and Development

Table 2: CSIS 2008 report findings & recommendations

#	Findings	Recommendations
1	Create a comprehensive national security strategy for cyberspace.	U.S. must protect cyberspace using all instruments of national power in order to ensuring national security, public safety and economic prosperity. President must unite all agencies and open discussion of how best to secure cyberspace and present issues of deterrence and national strategy to the broad national community of experts and stakeholders. In order to prevent disorder in management of the cyber security, major agencies must be coordinated from White House.
2	Lead from the White House	Creating a new agency for the cyberspace in the Executive Office of the President
3	Reinvent the public-private partnership.	Public and private sector entities can collaborate and share information on critical cyber security in a trusted environment.
4	Regulate cyber space.	The president should work with appropriate regulatory agencies to develop and issue standards and guidance for securing critical cyber infrastructure.
5	Modernize laws.	Laws for cyberspace are decades old written for the technologies of a less-connected era. New administration should update these laws. These laws are not effective in fighting against cybercrime.
6	Modernize digital identities and security	Creating a strong authentication system for access to critical infrastructure. Taking steps to increase the use of secure Internet protocols & increase the use of firewalls. Working with industry to develop and implement security guidelines for procurement of IT products. Creating guidelines and standards for secure products and services. The United States should make strong authentication of identity, based on robust in-person proofing and thorough verification of devices, a mandatory requirement for critical cyber infrastructures (ICT, finance, government services).Biometric is the most secure and convenient authentication tool.
7	Conduct training for cyber education and research and development for cyber security.	The president should direct the relevant agencies as To create training programs and career paths for the federal cyber workforce and to work with the National Science Foundation to develop national education program. To provide overall coordination of cyber security research and development.

The efforts towards the creation & maintenance of cyber security awareness within sectors that constitute the critical infrastructures such as energy, water & wastewater, nuclear, chemical, transportation, banking & finance, information technology and communications, postal & shipping, agriculture and food, emergency services, healthcare & public health, defense industrial base, commercial facilities, critical manufacturing, national monuments & icons and, government facilities were explicitly defined. The recommendations that this report provides are

- Provide leadership in development of control systems security principles.
- Assume full engagement in the NIPP partnership for control systems security.
- Maintain a high level of outreach and awareness within the CIKR stakeholder community.
- Coordinate and participate in the identification and analysis of gaps in control systems security technologies, policies, and planning.

2.6 Report #6 2010: (President, 2010).

This report, again from the White House, reflects the views of Obama Presidency over the matters of security. Although it addresses the security matters in general, it does have one whole section devoted solely for securing the cyberspace. Obama administration clearly underlines the fact that the Cyber security threats represent one of the most serious national security, public safety, and economic challenges as “The very technologies that empower us to lead and create also empower those who would disrupt and destroy.” While the report specifically mentions about cyber criminals it also delineates the problems caused by individual criminal hackers as well as organized criminal groups and terrorist networks. As for the solutions, the report states that “We will deter, prevent, detect, defend against, and quickly recover from cyber intrusions and attacks by: Investing in People and Technology and by Strengthening Partnerships”

2.7 Report #7 2011: (DoD, 2011).

Although coming from DoD, this report is by no means offering offensive strategies but only defensive ones. It should also be noted that the report is also the first evidence that the USA admits their system is vulnerable to cyber-attacks. DoD differentiates five strategic initiative areas as follows

Table 3: DoD strategic initiatives

#	Strategic Initiative	Details
1	DoD will treat cyberspace as an operational domain to organize, train and equip so that DoD can take full advantage of cyberspace's potential.	To organize, train and equip for cyberspace to support national security interests. Coordinate training for operations in a "degraded" environment, including the use of red teams in war games, operating with presumption of a security breach, and development of secure networks for redundancy purposes.
2	Employ new defense operating concepts to protect DoD networks and systems.	This includes enhancing best practices and *"cyber hygiene"* featuring updated software and better configuration management. The steps to strengthen workforce communications, accountability, internal monitoring, and information management capabilities to mitigate insider threats. The DoD will also focus on maintaining an active cyber defense to prevent intrusions. The DoD will develop new defense operating concepts and computing architectures including mobile media and secure cloud computing to embrace evolutionary and rapid change.
3	Partner with other U.S. Government departments and agencies and the private sector to enable a whole-of-government cyber security strategy.	Commercial assets such as Internet Service Providers and global supply chains, constituting a vulnerability that DoD and DHS will work together to mitigate. Their joint planning will increase effectiveness of cyber needs while respecting privacy and civil liberties and will conserve budget resources. The DoD is also establishing pilot public-private partnership to enhance information sharing. A Whole-of-government approach will lead to continue to support interagency cooperation to analyze and mitigate supply chain threats to government and private sector technology.
4	Build robust relationships with U.S. Allies and international partners to strengthen collective cyber security.	The DoD will seek "robust" relationships to develop international shared situational awareness and warning capabilities for self-defence and collective deterrence. The DoD will assist US efforts to help develop international cyberspace norms and principles, dissuade and deter malicious actors, reserve the right to defend vital national assets as necessary and appropriate. Cooperation with allies to defend allied interests in cyberspace, work to develop shared warning capabilities, build capacity, conduct joint training, share best practices and develop burden sharing arrangements.
5	Leverage the nation's ingenuity through an exceptional cyber workforce and rapid technological innovation.	Catalyze US scientific, academic, and economic resources to build a pool of talented civilian and military personnel to operate in cyberspace. Adopt differing levels of oversight based on prioritization of critical systems. Improved security measures for hardware and software. Development and retention of cyber workforce is central to strategic success outlined in this strategy.

3. Comparative evaluation of the reports of the first decade of 2000s

- The report in **2002** evaluates the events in somewhat not-so-serious manner like defining the cyber threats as "nuisances" or "Weapons of Mass Annoyance". This attitude of the report towards the issues at hand can be classifies as a lightweight approach as best. It seems that the cyber threats were known by the authors nevertheless undervalued.
- The **2003** White House report clearly sets the beginning point of cyber security strategies by being very detailed and quite organizational towards to the solution of cyber problems. Actually this report is one of its kinds when approaching the cyber security issues in an organizational manner put forward by a nation state.

Because of the structure and of the origin of those two reports, making of a comparative evaluation would be totally pointless.

- In **2005**, The Advisory Committee report to the White House evaluates the cyber security as a prioritization problem and in that, provides a very through research initiatives to counter and to remedy the cyber security issues. This report is important for the fact that the committee clearly puts forward the importance of research & cooperation among many actors.

Actually, combining of the reports of 2003 & 2005 creates a one whole report for the fact that while the former one defines the structural organization, the later one puts forward as to what & how those structured bodies would be doing against the cyberspace threats in terms of research & development.

- **2008** CSIS report clearly indicates the increased awareness of the think tanks to the issue of which the CSIS; as the owner of this report, is one of them. Perhaps the most important and differentiating part of this report from its contemporaries is the call for the regulation of cyberspace by laws which would be enforced by the president of US of America. So, the judicial players were called into the game as the effort to control the cyberspace at last.

This effort of trying to control the cyberspace through laws probably stems from feeling helpless from a managerial point of view, but actually it was a demonstration of the will to exercise a judicial and/or managerial power over a field; which by definition, has no geographical and/or political borders and also free from four dimensions. Hence, the futile calls at best.

This call also clearly indicates that how the understanding of the limits of cyberspace was different than of today.

- Like the 9/11 event which paved the way of the formation of DHS, the **2009** report of this very same governmental body defined the methodology for the securing of the Control Systems. Although there may be other contemporary documents on the same topic; notably the Scada documentation, this report is surely a milestone in itself in this regard.

The findings and especially the recommendations of this report are still valid today, and since the other aforementioned reports do not specifically deal with the security of Control Systems, it would not be a fair comparison if one compares this report with the others. However, Control Systems security is one of the major issues of Cyberspace at the moment as it was by then and thus, it would be a fatal mistake to omit this very crucial report in historical timeline when analyzing & evaluating cyberspace security reports.

- **2010** White House Security report was the first actual demonstration of the understanding of the many dimensions of cyberspace security. This report did mention the strategic and critical importance to stop the cyber terrorists and cyber organized criminal groups. This endeavor, fight against cybercrime and cyber terrorism, earned its place in a governmental strategic paper for the first time; which also reflects the determination & resolution of the government of a nation state to overcome of this one of a kind problem.
- From this paper's point of view, the end of the first decade of 2000 was clearly marked by a report **coming from DoD of US of America in 2011. Actually, this was a first written declaration of the armed** forces about a cyber-warfare & about information warfare. Although there were five declared initiatives altogether in this report, the fourth one; which underlined the importance of international cooperation among the nations, was a unique one among all other reports detailed above.

This call of DoD's is also unique for it lists an extensive list of recommendations in a format of a defensive methodology in strategic initiatives. Thus, this report also shows that how the new threats originating from a strange entity; known as a cyberspace, for a military are understood & evaluated by a very military organization.

4. The future projections

Actually, the making of the future starts today or better put it, has started yesterday. The tactical and/or operational level of any action might be based on the strategies developed way before or sometimes years before than the action itself. Therefore, some studies; although published couple of years back, might be playing a key role to appreciate the importance of yet some other strategic reports concerning the very near future. A good example for that kind of a report is by NATO.

4.1 NATO 2010: (NATO, 2010).

A strategic report of NATO which will be in effect until **2020**, clearly defines, as the very first time within NATO, the cyberspace issues as findings plus, provides recommendations to remedy them. One major finding in that report openly points to the serious gaps in NATO's cyber defense capabilities.

The importance of this report lays in the fact that, not only it openly admits the existence of gaps & vulnerabilities in cyber defense capabilities of NATO itself, but it also classifies the cyberspace threats as both "unacceptable and increasingly dangerous".

In this strategic report, NATO recommends some actions all of which will cause some movements in cyberspace for a decade to come, such as

- The cyber security will be one of the requirements for homeland security,
- A new classification of asymmetrical & unconventional threats including the cyber assaults and,
- To develop an array of cyber defense capabilities throughout its member states. That action calls for new formations that will cover possibly all areas from education & training to the establishment of cyber command posts.

4.2 The office of the director of the national intelligence 2012: (ODNI, 2012)

The series of strategic and midterm & long term forward looking reports from National Intelligence Council (NIC) of US of America comes in every 5 years under the name of "Global Trends". The very recent Global Trends report of NIC tries to approach the cyberspace and its issues from the perspective of a next decade and a half.

In "Global Trends 2030" report, the NIC anticipates that

- Even small groups will be able to have cyber instruments as weapons,
- As the power of the individual will rise, the chance of nonstate actors conducting a cyber-attack will also be rising,
- Individual experts in cyber systems will be selling their expertise to criminal groups.

As the new technologies will arrive so will new issues associated with those technologies such as

- Establishing and maintaining the security of large data sets like that of zetta bytes (270) or even yotta bytes (280) of data in a Cloud.
- Controlling the cyber arms-race which will likely to occur as the technology will always be getting cheaper and cheaper.

5. Conclusions

1. From many aforementioned reports it is quite clear that, we are still very far from the solutions to the problems of cyberspace. And even, it can be stated boldly that, we have just begun to understand what the cyberspace is all about; let alone providing solutions.

2. As far as the cybercrime goes: Norton Cybercrime report (Norton, 2012) clearly states

- *The scale of consumer cybercrime:556 million victims per year which is more than the entire population of the European Union* 1.5+ million Victims per day, 18 victims per second. 2 of every 3 online adults have been victims of cybercrime in their lifetime.
- The Global Price Tag of Consumer Cybercrime: USD $110 billion. Average cost per victim is USD $197.

All these figures, especially when compared with the recent years', have a tendency to grow by at least 50% which in turn are correlated & consistent with the projections for the next decade and a half to come (ODNI, 2012).

3. Moreover, we are not so sure whether there will be at least some solutions to the problems of cyber security as long as we stick to Turing Machine model with Shannon Information Theoretical approach to the information (Koltuksuz, 2006).

4. It is quite clear that we will need new hardware & software to deal with large data sets in yotta bytes. For that, the new operating systems coded in new programming languages with associated new data structures to run on a large scale multicore architecture will surely be in demand. With that kind of a new and very advanced hardware and software some very strong cyber-attack types will definitely emerge.

5. So far as we understand, there is no way to control the cyberspace, nor will ever be. As stated by US Air Force, Cyber Command, "The network is complex and cannot be completely secured", (Shugg, 2011) and, by US Navy Fleet, Cyber Command "Unlike the physical domain, achieving dominance may be impossible. Cyber warfare necessitates considerable demand on intelligence and resources. We need to know our targets and vulnerabilities, and understand the relationship between them", (Leigher, 2011). Therefore, it is quite safe to assume that the future conflict will be dominated by new cyber weapons.

6. For the solution to the cyber terrorism, the much needed cooperation in between public and private sectors plus intergovernmental cooperation still have not been accomplished and it looks like that it will be even harder in years to come.

7. It seems that Gibson's prophecy about cyberspace has finally come true.

Acknowledgements

The author wishes to thank to all grad students of the course titled *CENG504 Advanced Information Security, (Fall 2012, Grad. School, Yasar University)* for their much valued discussions. Thanks are also to the anonymous referees for their much valued critical reviews upon which this paper gained its current content.

References

DHS, 2009. Strategy for Securing Control Systems, Coordinating and Guiding Federal, State and Private Sector Initiatives, Washington, DC., USA: Department of Homeland Security.

DoD, 2011. Strategy for Operating in Cyberspace, Washington, DC., USA: Department of Defense.

Gibson, W., 1984. Neuromancer. New York: Ace Books.

Koltuksuz, A. T. S., 2006. Intelligence Analysis Modelling, International Conference on Hybrid Information Technology-ICHIT 2006. s.l., IEEE, DOI: http://doi.ieeecomputersociety.org/10.1109/ICHIT.2006.157, pp. 146-151.

Leigher, R. A. W., 2011. Defense Systems. [Online] Available at: http://www.defensesystems.com/Articles/2011/01/27/AFCEA-West-cyber-warfare-panel.aspx [Accessed 17 February 2011].

Lewis, J., 2002. Assessing the Risks of Cyber Terrorism, Cyber War and Other Cyber Threats, Center for Strategic and International Studies, Washington, DC., USA: CSIS.

Lewis, J., 2008. Securing Cyber Space for the 44th Presidency, A Report of the CSIS Commission on Cyber security for the 44th Presidency, Washington, DC., USA.: CSIS.

NATO, 2010. NATO 2020-Assured Security; Dynamic Engagement, Brussels, Belgium.: NATO.

Norton, S., 2012. 2012 Norton Cybercrime Report, USA: Norton.

Nye, J. S., 2010. Cyber Power. Cambridge, USA: Belfer Center for Science and International Affairs.

ODNI, 2012. Global Trends 2030, Washington, DC., USA.: The Office of the Director of the National Intelligence (ODNI) of the U.S. of National Intelligence Council.

Ottis, R. L. P., 2010. Cyberspace: Definition and Implications. Dayton, Ohio, USA, Academic Publishing Ltd., pp. 267-270.

PITAC, 2005. Cyber Security: A Crisis of Prioritization, President's Information Technology Advisory Committee, Washington, DC., USA.: The White House.

President, 2003. The National Strategy to Secure Cyberspace, The President's Office, Washington, DC., USA.: The White House.

President, 2010. National Security Strategy,The President's Office, Washington, DC., USA: The White House.

Shugg, B. G. C., 2011. Defense Systems. [Online] Available at: http://defensesystems.com/articles/2011/01/31/air-force-cyber-command-ready-for-operations.aspx [Accessed 17 February 2011].

Alerting Security Authorities and Civilians with Smartphones in Acute Situations

Jaana Kuula[1], Olli Kauppinen[1], Vili Auvinen[1], Santtu Viitanen[1], Pauli Kettunen[1] and Tuomo Korhonen[2]
[1]University of Jyväskylä, Department of Mathematical Information Technology, Finland
[2]Central Finland Police Department, Jyväskylä, Finland
Jaana.Kuula@jyu.fi
Olli.Kauppinen@jyu.fi
Vili.Auvinen@jyu.fi
Santtu.Viitanen@jyu.fi
Pauli.Kettunen@jyu.fi
tuomo.korhonen@poliisi.fi

Abstract: The speed of communication and the recognition of emergency notifications are key issues in alerting people in acute situations. This article describes case studies which The University of Jyväskylä has made for studying how well smartphones can be used for emergency alerting in different situations. In the study empirical tests were carried out with three different test groups. The first test was carried out with security professionals within the internal organization of the Finnish police. In the second test police sent emergency alerts to the private citizens' personal smartphones. The third test was carried out in a school of 500 pupils. All tests were carried out by using a smartphone based alerting system that has been developed at the University of Jyväskylä. The alerting system utilizes multiple technical features of smartphones for ensuring that the alerts will get through and become noticed in all circumstances. It also operates independently from commercial telecommunication operators, and if open WLAN is available, emergency alerts can be sent even if mobile phone base stations are down. The tests show that for ensuring the perception of safety alerts in all situations smartphone alerts must be given in different forms at the same time. For example, in noisy environments voice alerts need to be intensified with the vibration feature of the phone. Voice alerts may also be supplemented with a visual and textual message. Alerts also need to be given in a different form for different users if the users are security professionals, adult aged private citizens or if they involve underage children. The tested smartphone based alerting system may be used both in normal times and during the state of emergency.

Keywords: smartphones, mobile alerting systems, public warnings, crisis communication

1. Introduction and background

Mobile alerting systems are becoming more common and important at the side of broadcasting type of notification systems. For example, Aloudat et al. (2011) consider them as an established part of mobile government strategies worldwide. Meissen and Voisard (2008) see the effective implementation of early warning systems as one of the best investments for disaster prevention and mitigation. As an example, during the Tohoku Earthquake in 2011 in Japan warnings were sent to circa 52 million people by SMS and CBS. First warning was given in 8.6 seconds from first wave and to Tokyo warnings arrived 65 seconds before earthquake (Yamasaki, 2012).

At the same time as new emergency alerting systems are taken in use, in the modern world people are already stressed up with many other kinds of alarms and signals. There might even be a risk that the individuals' overload of digital signals is so heavy that people are not able to recognize critical emergency signals, even on their personal mobile phones.

Häkkinen has studied the failure of alarms in his doctoral thesis (Häkkinen, 2010) and notes that many people ignore alarms even if they can be crucial for life. He sees many causes for the failure. One reason is that human alarms are often implemented through a variety of technical means and presented as abstract signals. Häkkinen suggests that multimodal alarms should be used for addressing the sensory and cognitive factors that impact alarm detection and understanding. According to Coombs (2007) fast reaction to crisis is important. The first reaction should also be logical and correct. This causes pressure for authorities and for crisis management. Ready-made notification templates will help making correct response actions, and wide usage of communication channels will help reaching a high adoption rate.

Häkkinen (2010) sees the alarm process as detection, perception, recognizing and responding. All of these phases must success for taking an appropriate action. Due to the environment for example voice alarms may not be heard at a crowded railway station and a vibration alarm may not be noticed in the pocket of a motorcycle rider. Perceiving alerts means that the detected alarm should receive attention and be processed. If attention towards it cannot be maintained, the alarm can be lost among other signals or noise. After perceiving the alert, existing knowledge will be used for recognizing the signal. People will then use the received information and earlier known protection procedures for surviving the situation. According to Sillem and Wiersma (2006) people are very capable in receiving information and the successful design of warning messages maximizes the probability of having each step of the alerting process to be completed.

According to Häkkinen (2010) many physical, environmental, sensory or cognitive disabilities do not match with the alarm processing. People's sensory systems can also be temporarily weakened by fatigue, acute illness or injury, stress, foreign language or methods, background noise and intoxicants. Also false and pointless alerts decrease the motivation to respond alarm.

A research group at the University of Jyväskylä made an empirical study for testing the usage of smartphones for emergency alerting purposes in three different user groups. The first test was performed within the internal organization of the Finnish police in order to measure how well smartphone alerts will be noticed and reacted by security professionals. The second test was addressed to civilian people in order to find out how well similar alerts will be noticed by private citizens and in which way authorities should structure the alerts which are directed to the private people. The third test was carried out in a public school of 500 children. The purpose of this test was to find out issues which authorities have to take in notice when considering mobile alerts in schools.

All tests were carried out in November 2012–February 2013 and they were made by using a smartphone based alerting system which was developed at the University of Jyväskylä in 2010–2012.

2. The system design of the tested alerting system

The tested mobile alerting system is based on a combined server and smartphone application which also operates with other kinds of media. The server end includes an integration interface with several external systems like with emergency response centers and other security authorities. It also contains a web-based application for the input of manually given notifications. The server also collects and stores information. According to Amailef and Lu (2011) alerting systems are highly data dependent, for which reason databases play a significant role in these systems.

The mobile application is connected to the server and it is the primary communication channel in the system. So far it runs on the Android platform. The Android platform was chosen as the operating system because it is the most widely used mobile platform in the World and can perform various multiform alerting operations. (Kuula et al. 2013) For others than Android phones alerts are delivered as SMS messages with the assistance of a third party SMS gateway service.

The system can also exchange information with other communication channels. In many situations multichannel messaging is needed and more efficient than delivering alerts through one channel only. In addition to smartphones, messages may be delivered to PC's, laptops and tablets, social media, electronic bulletin boards and public media. According to Vihalemm et al. (2012) the use of alternative sources and channels of warning messages will help people to emotionally and cognitively cope with crises. Also Hughes and Palen (2009) suggest that emergency management could use Twitter and similar micro-blogging technology as a way of getting information to the public. They expect that this would also fuel personal technology adoption and instruct operation in emergency warning, response and recovery situations. Muralidharan et al. (2011) have noticed that nonprofit and media organizations use information dissemination and disclosure effectively, but fail to capitalize the two-way communication nature of social media. The tested smartphone phased system has all these abilities. The operating concept of the system is presented in Figure 1.

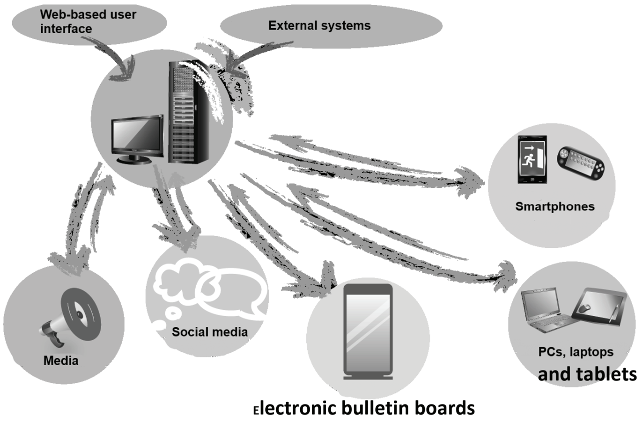


Figure 1: The operating concept of the tested smartphone based alerting system

With the system emergency alerts can be sent to selected location-based areas or for selected groups. The location-based alerting area will be defined for each emergency individually by giving the geographical coordinates and range of the influence area of the emergency to the system. The range of the alerting area can vary from one building into the whole city or wider. In group-based alerts notifications will be sent for selected user groups regardless of their geographical location.

The system also includes a GIS-based graphical user interface and it supports multilingual and multimodal presentation. Multimodality is the system's main advantage compared with Common Alerting Protocol (CAP) and it is based on Häkkinen's (2010) Multi-Modal Alarm Specification Language (MMASL).

Figure 2 illustrates the selection of the alerting area through the GIS-based graphical user interface and some visual images which can be used as a part of MMASL based alerting messages.

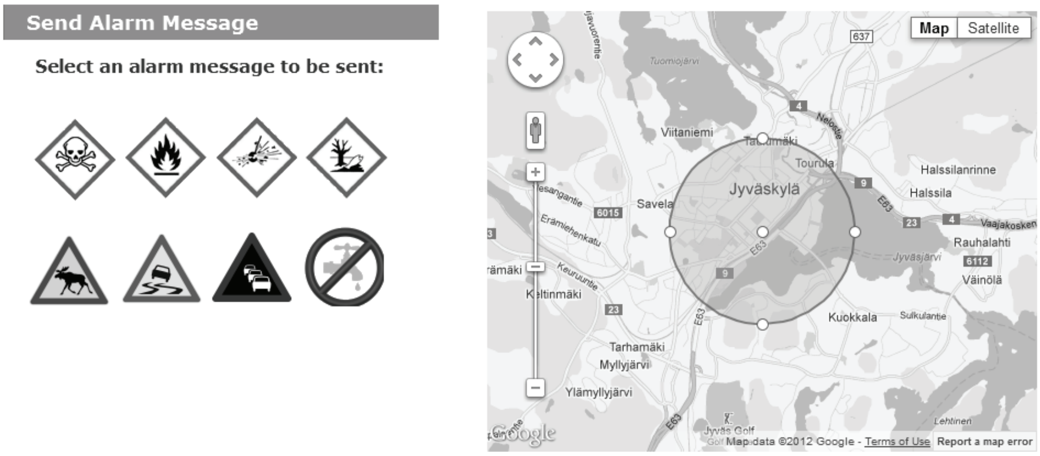


Figure 2: Visual alerting icons and the Web and GIS-based user interface of the tested system

The mobile end of the system operates on the Android platform. Requirements are Android version 2.2 or newer and Google account. The application runs at the background of other operations and sends the user's location information to the server timely. The native Android application enables using all available functionalities of the smartphone. It's main advantage is that the emergency notifications can be forced through to the user over other services. The user will be notified even if the ringtone is muted.

The Android platform handles the location information and network connection without the user's or application's need to think about it. Localization methods depend on the user settings and features of the device. Normally the system uses GPS-, WLAN- and mobile network-based localization.

The system operates on ordinary mobile networks (2G/GPRS/EDGE, 3G/UMTS/HSPA, 4G/LTE) and WLAN (IEEE 802.11). The usage of WLAN and web-based connections is a big advantage, because they are not dependent on mobile telecommunication operators. This is critical in emergencies where the base stations of mobile networks are down. That may happen in ordinary storms and because of heavy loads of snow on trees in winter, which both cut trees and break off the electricity supply on base stations. Also other problems may occur in telecommunications both in normal times and during crises (Kuula et al., 2012). For avoiding the vulnerability of mobile networks, also the device-to-device communication and mobile ad-hoc networks (MANET) could be used. That would help avoiding communication overload on base stations and enable communication without mobile networks.

In the emergency the user will receive a notification with an alert (alerting cue) and a message. Both parts can be customized with different auditory, visual and tactile effects. Customization enables sending notifications with a different priority and with selected effects which give the users fast a truthful understanding of the situation. Figure 3 illustrates how the alert (cue) is received on the smartphone and how textual message will be shown on the screen after that. The latter view includes also a question to user. By answering to that question the user may ask for help.

A big challenge in the real life is how to get people to use these systems. Users often have concerns about their privacy while using real time location-based systems and therefore the privacy mode should be enabled in such systems (Aloudat et. al., 2009, Al-Akkad and Zimmermann, 2011). According to Al-Akkad and Zimmermann (2011) some people fear for creating a surveillance society while gathering masses of data from mobile phones.

Wu (2009) sees that the usefulness of SMS-based alerting systems has multiple levels of meanings to the users. The ease of use is more about the users' ability to control the system behavior. It is also a subjective norm which needs to be examined with relation to its originating source. According to Kaasinen (2005) user acceptance and intention to use are built on the perceived value of the service, perceived ease of use, trust and on the perceived ease of adaption in the actual usage phase.

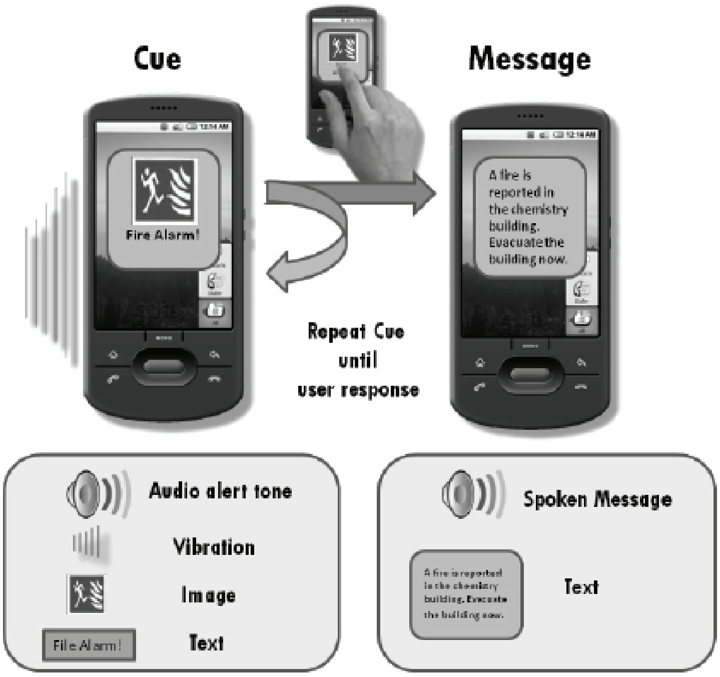
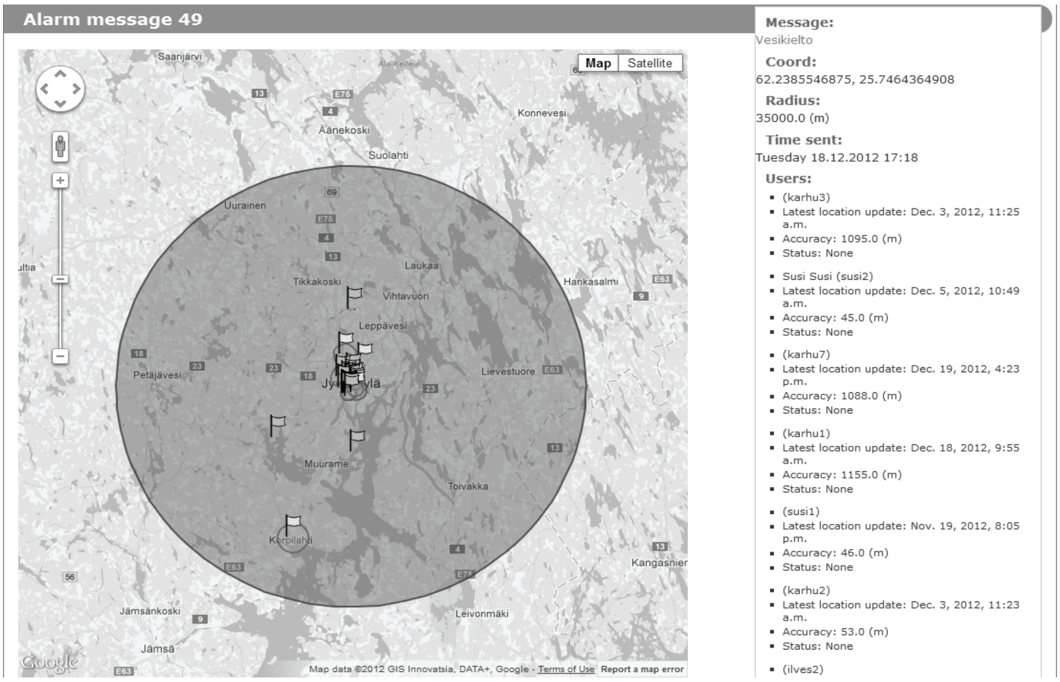

Cue
Message
Fire Alarm!
A fire is reported in the chemistry building. Evacuate the building now.
Repeat Cue until user response
Audio alert tone
Vibration
Image
File Alarm!
Text
Spoken Message
A fire is reported in the chemistry building. Evacuate the building now.
Text

Figure 3: Notification receiving process on smartphone (Häkkinen, 2010, p. 78)

The two-way communication of the tested system enables to gain information from the people in the hazard zone. Figure 4 illustrates the map view of the real time situation in the emergency area after giving a location based emergency alert. Users who need help are indicated with a red flag (not in the picture) whereas the green flag indicates that the user is all right. A yellow flag means that the user hasn't signed the question, for example because of getting hurt and not being able to use the phone. The phone may also be lost or out of order.

Table 1 contains a summary of some central features of the tested system. The table has been modified from the original table of Lee at al. (2011) and it concludes the major differences between the tested smartphone system and other alerting systems which operate on ordinary mobile phones.



Alarm message 49

Message:
Vesikielto

Coord:
62.2385546875, 25.7464364908

Radius:
35000.0 (m)

Time sent:
Tuesday 18.12.2012 17:18

Users:

- (karhu3)
- Latest location update: Dec. 3, 2012, 11:25 a.m.
- Accuracy: 1095.0 (m)
- Status: None

- Susi Susi (susi2)
- Latest location update: Dec. 5, 2012, 10:49 a.m.
- Accuracy: 45.0 (m)
- Status: None

- (karhu7)
- Latest location update: Dec. 19, 2012, 4:23 p.m.
- Accuracy: 1088.0 (m)
- Status: None

- (karhu1)
- Latest location update: Dec. 18, 2012, 9:55 a.m.
- Accuracy: 1155.0 (m)
- Status: None

- (susi1)
- Latest location update: Nov. 19, 2012, 8:05 p.m.
- Accuracy: 46.0 (m)
- Status: None

- (karhu2)
- Latest location update: Dec. 3, 2012, 11:23 a.m.
- Accuracy: 53.0 (m)
- Status: None

- (ilves2)

Figure 4: Overview of a given notification

Table 1: Comparison of older and smartphone based services, modified from Lee et al. (2011)

	Base station based mobile emergency alerts	**Smartphone based emergency alerts**
Basic device	Cellular phone	Smartphone
Network	3G, 4G	3G, 4G, WLAN
Service Type	SMS, CBS, LBS	Push messages, mobile client application
Time of being used	During the emergency	All the time
Type of message	Text	Voice signal, image, text, vibration, map
Need for client application	Not needed	Installation needed

3. The research design

It is difficult to define how public warnings should be given. The research problem was therefore split and divided into smaller problems. As the police and the city are in charge of the public safety, a close cooperation was started with the Central Finland Police Department and with the City of Jyväskylä for testing the university's alerting system.

With the police, two pilot studies were designed. One was organized within the internal organization of the Finnish Police and another with civilians. In the first pilot emergency alerts were given by the police to the test users of two preparedness groups of the police around the country. In the second test police gave security alerts to the private citizens' personal mobile phones. The third pilot was carried out in a 500 pupils' school in the City of Jyväskylä, because in recent years there have been some unfortunate shootings, bomb threats and other violence at schools in many countries. The police were observing also the school pilot.

The participation of the Central Finland Police Department and the City of Jyväskylä in the empirical tests was extremely valuable. According to Aloudat et al. (2009) it is important to investigate the perspectives of the crucial stakeholders, like of the prospective users and the government. Aloudat et al. (2009) noticed that these users believe that location-based services have the potential to aid people in emergencies, but there are several major disagreements for example about the privacy, cost, specification and management issues of these systems.

3.1 Case study with the police

The first pilot was organized with the Central Finland Police Department and with the Finnish Police. Detailed results of the pilot have been published in the ISCRAM 2013 Conference (Kuula et al, 2013).

The purpose of the study was to evaluate the performance and usability of a smartphone based mobile alerting system for the alerting, command and communication purposes of the police. The study should also give information of the smartphones' usability for alerting civilian people. The test was carried out with the standardized smartphones which were given to the police officers' use.

For the study a test group of ten policemen was formed from two different preparedness groups of the police and introduced to the use of the system in a video conference. The leading police officer then sent in the situation room test events of simulated real-like incidents to the test users' around the country. Users were obligated to sign all messages immediately regardless of where they were and what time of the day it was. Immediately after sending the alerts all users' location and status appeared on a real-time map on the screen in the situation room. Each user's position and status were indicated with a green, red or yellow flag.

The overall performance and usability of the alerting system was evaluated by the police officers' and civilians' ability to notice, understand and react on the incoming alerts. Users' ability to notice and understand alerts was evaluated from the auditory, sensory, visual and cognitive senses' point of view. The policemen's readiness to take action was measured by their reaction time on commands and alerts which were sent in the working hours and at the free time and night.

In a numeric evaluation in scale 1-5 (1=poor, 5=excellent) more than 70 % of the police users evaluated the sound and volume of the alerts as excellent or good. Sometimes the background noise was so loud that alerts could not be heard well. In the written comments users mentioned that in a real situation the vibration and silent alerts would be useful. Vibration helped to notice alerts in noisy situations. The alerting sound and icon gave to the users the first information about the incident, and the textual message gave additional information about it.

Table 2 presents the users' reaction times to all alerts. Data has been taken from the log from the server and it indicates the time between sending the alert and getting a receipt about it from the user. The table shows average reaction times for all alerts without separating them into categories. When day and night time alerts as well as voice and silent alerts are viewed separately, reaction times vary.

More than 60 % of users signed all alerts in less than two minutes and more than 70 % in less than five minutes. Some of the 25.1 % of users whose reaction time was more than ten minutes or who did not sign some alerts at all have according to the log obviously not carried the test phone with them after the office hours. In that case they have not been able to sign alerts in the evening or at night.

The visual map view gave to the leading police officer a good comprehension of the course of the action after giving the alert to the other police officers. The map showed instantly where geographically the police officers were and how fast they could be got on duty to handle the crisis.

Table 2: Shares of different reaction times in the test with the police (N=195)

Less than 2 minutes	64.1 %
More than 2 and less than 5 minutes	8.7 %
More than 5 and less than 10 minutes	2.1 %
More than 10 minutes or no signature at all	25.1 %

3.2 Case study with the police and civilian people

In the second test police send mobile emergency alerts to the private citizens. The test group was formed from twenty-five volunteers who were searched through various mailing lists and a short introduction about the system was given to them by a written document and email. The testers' age varied from 20 into 60 years or older. People were very eager to participate in the test with the police. Sillem and Wiersma (2006) had also noticed that earlier in their own studies and say that people are open for new technologies and very keen to participate in a research about citizens warning.

The police sent location based emergency notifications to the test group around the City of Jyväskylä and wider. Most notifications were given with a tight geographical limitation and directed into different sub areas of the city. The range of the widest alerts was 500 kilometers and they reached the City of Helsinki in the South, Arctic Circle in the North, Russian border in the East and Sweden in the West.

The purpose of the study was to gain empirical knowledge about how smartphone alerts should be given by the security authority to the population. The test simulated real emergencies and after the study information was collected from the users with an internet survey. Test alerts warned users for traffic accidents and jams, armed and dangerous persons, intruders, escaped criminals, missing persons, industrial fires and for spoiled drinking water. The police were also prepared for giving authentic alerts to the test users and on 21.12.2012 at 8:11 the Central Finland Police Department gave its first authentic smartphone alert by warning people for a severe traffic accident on the Vaajakoski motor highway. This alert was at the same time the first real smartphone based warning message from the Finnish Police to the civilian people ever.

The range of the alerting area, visualization and the alerting tone of the message were decided individually each time depending on the type of the incident. Also the two-way communication was tested in order to help the work of the police. For example, users were informed about a dangerous person in their neighborhood and asked if they had seen the person.

The civilian testers were active technology users and almost all used in the test their private phones. When the users were asked to answer the internet survey after the test, the response rate was 90%.

One of the questions enquired about the users' experiences of severe and dangerous situations and about the tested system's usefulness in those situations. A half of the users had experienced dangerous situations and evaluated that a smartphone based warning system would have been of help in those cases. Some results of the questionnaire are presented in Table 3.

Table 3: Civil users' evaluations concerning the tested alarm system (1 = poor, 5 = excellent) (N=20)

Question	Average	Variance
Overall usability of the alerting system	3.95	0.47
Easiness of the interpretation of the messages	4.25	0.72
Usefulness of the messages	3.42	1.26
Personal relevance of the messages to the user	2.90	1.36
Superiority of the application-based alerting compared to SMS messaging	4.39	0.49
Possible constraints in taking the system in real use (1 = very much, 5 = none)	3.26	0.94
The system's ability to improve the user's personal feeling of safety	3.65	0.87
Gained benefits compared to the required effort of using the system	4.10	1.25
Recommendation of taking a finalized system in real use	4.65	0.34

3.3 Case study at the school

The third pilot was carried out in a school because the school violence has been discussed a lot lately and because the school differs from the other tests with the police. Test users were teachers and other personnel of the school but as the environment was full of underage children that caused special requirements for the test. Users were introduced to the use of the system in a face to face group meeting at school.

Arranging physical protection in schools is a separate issue whereas this study focused on giving security alerts in schools. This study was arranged in Kilpinen school which represents an average school of the ca. 50 public schools in the city. The school building was built at the end of 1960's and there are nearly 500 children of 13–16 years old in that school.

The limitations of the building became obvious in the beginning of the study when the signal strength of the 3G mobile network appeared to be so weak inside the stony walls that the test could not be run on it. The 3G network was then strengthened so that the signal could be reached better inside the building but even that left some shadow areas inside the school. These were caused by the structure of inner walls and by some heavy objects which the school needed. The representatives of the city told that the signal strength of the 3G network may be weak even inside the newest schools because of the 3–4 layered energy saver windows. According to Waitinen's (2011) studies of the physical safety in primary and secondary schools in Helsinki good safety cultures include good safety management practices, well-developed understanding of safety hazards and the requirements of basic safety, open and communal safety-related work and an appreciation of safety evidenced through everyday practices.

Another obstacle in schools is that the employer may not have provided mobile phones for teachers or that the teachers are forbidden to use mobile phones in the class. Teachers may also not want to install security systems on their personal phones. Security alerts can be received also with tablet computers, PCs, smartboards etc. and from the technical point of view smartphones are not necessarily needed in class. For preventing panic and chaos amongst the children, it might however be better to give security alerts discretely to the teachers first. That is possible only with the teacher's personal devices which children cannot see or hear.

The personnel were asked about their attitude towards school safety before starting the study. According to the survey personnel wanted panic buttons into the school. Existing systems did however not enable installing more than two buttons in the building and a private company was ordered to install more buttons into the school. Also mobile buttons were given to the personnel. The call buttons were then integrated into the tested smartphone system so that all alerts which were made with the buttons would trigger an alarm on the test users' mobile phones.

During the test alerts were given with fixed and mobile panic buttons, call buttons which were built on the smartphones and through the alert system's user interface on the web. All alerts were directed to the school personnel's smartphones and the most serious alerts to the private security company. In the most serious incidents a 112 emergency call would be made to the emergency response center.

Table 4 shows the personnel's attitudes towards the school safety. Answers show that the teachers and other personnel are interested in security issues. Most interest was paid to panic buttons and some teachers, student counsellors and school nurses had asked for getting them in their offices and class already earlier. Especially the student counsellors and nurses work alone and sometimes they may need to call their co-workers for help to relieve the threatening situation. The tested smartphone based alerting system was not familiar in the school in advance but the personnel's attitude towards it was quite positive. The employer has provided mobile phones only for a small number of teachers, but if all teachers would have them they would be willing to install the security system on their phones. Many of them would however not want to install security systems on their private phones at school.

The survey showed also that even if many of the interviewed persons do not need a personal security system, they still want to install it for helping their colleagues. Attitude towards external security professionals is mixed. It looks like the personnel would like to handle threatening situations by themselves or that they see it as a police issue. It also looks like they are unwilling to call the private security company to the school because the school needs to pay extra each time when the security guard comes to the school. The services of the police are free but the police will be called for help only in serious incidents. All schools also have a named school police who visits the school regularly and gives security education for the pupils. When the researchers interviewed the school police he said that the personnel talks only about is the pupil's safety and not about themselves.

Table 4: Some examples of the personnel's attitudes towards school safety before starting

Statements	Average	Variance
Built-in panic buttons inside the building are a good enhancement in the school's safety	4.15	0.88
Panic buttons on the personnel's mobile phones are good enhancement in the school's safety	4.12	1.41
Mobile devices would suite well as the personnel's internal communication channel at school	3.53	1.39
I would take a mobile communication and security application in use on my phone at work	4.03	2.20
I would take a mobile communication and security application in use on my private mobile phone	3.50	2.26

4. Summary and conclusions

The study shows that even if the emergency is the same, smartphone based alerts need to be given in a different way in different environments and user situations. In the study user tests were made with police officers, adult aged private citizens and in a school of underage children.

One common feature with the three cases is that the society is not yet fully prepared for a wide implementation of smartphone based emergency systems. For example the open air mobile communication infrastructures are inadequate in some places which might cause problems for the full exploitation of mobile emergency systems. Mobile communication infrastructures are also vulnerable. If for example base stations will lose their energy supply or if they are destroyed, all mobile phones in the area will be muted. If however, there is an open WLAN available the tested kinds of smartphone based alerting systems will be able operate even if all base stations are down.

The study in the school pointed out that wireless infrastructures can be inadequate also inside the buildings. Although this was proved only in one building it is quite evident that in every country there are plenty of old or technically tricky buildings in which there are hundreds or thousands of people inside daily and that inside those buildings networks may not operate properly. Wireless networks are also vulnerable for overload, and when something bad happens the risk for overload will increase.

When public smartphone based emergency alerting applications will be installed widely also the peripheral devices may cause some problems. In this study all tests were made with standardized smartphones which operate on the Android platform. In a real situation the variety of mobile devices among the population is much bigger and not all of them are able to receive smartphone messages. With the tested alerting system these devices may be though reached with SMS. All people do also not have mobile phones at all or they are not allowed to use them at work.

When people receive smartphone based alerting messages, according to this study more than 60 % of them would react on them within the first two minutes. If the reaction time will be counted for voice alerts or night time alerts only the average reaction time will be even shorter. During the daytime voice alerts are distracted in many places by background noise and by many other tones and signals which disturb the detection and perception of emergency alerts. The vibration of the phone will improve the detection and it will be useful also in situations where voice alerts cannot be used. In the study best reaction times were received at night when 80 % of the test users signed alerts within the first minute. At the night time most of the users were at home in bed and there was no background noise which would have drowned the alerting sound.

Finally, even if the technology runs smoothly, the success of emergency alerting is dependent on the user him/herself. If the user does not care about the alerting message, or if he/she does not receive it in time because the alerting device is not in hands or if it is not working, the alert will not be noticed. In this study for 25 % of the test users it took more than 10 minutes to sign the alert and some did not sign all alerts at all. In these cases the users may not have carried the test device with them all the time. In a real situation, part of the people might be careless with their phones as well, and for that reason all people would not receive emergency alerts even if they would have a personal phone and even if the alert would reach their personal device.

The study concludes that smartphones are an appropriate and flexible alerting technology for many situations and in many ways they are better than several other technologies on the market. Alerting needs however be planned carefully according to the environment, user group and situation. One must also consider each time whether the benefit of the intended message would be bigger that the harm which it would cause. This decision has to be made also now when most of the public alerts are given by broadcasting methods, but when emergency alerts will be transferred to the private people's personal devices that decision will become even more important.

Acknowledgements

Authors wish to thank Tekes – The Finnish Funding Agency for Technology and Innovation, University of Jyväskylä, Magister Solutions, Parco Group, Jämsän Apteekki, and the European Regional Development Fund for funding this research. Authors also thank Central Finland Police Department, National Police Board and the City of Jyväskylä for their support and participation in this study.

References

Al-Akkad, A. and Zimmermann, A. (2011). User Study: Involving Civilians by Smart Phones During Emergency Situations. *Proceedings of the 8th International ISCRAM Conference*, Lisbon, Portugal.

Aloudat, A., Michael, K. and Abbas, R. (2009). Location-Based Services for Emergency Management: A Multi-Stakeholder Perspective. *Eighth International Conference on Mobile Business*, pp 143–148. ICMB.

Aloudat, A., Michael, K. and Abbas, R. (2011). Recommendations for Australia's Implementation of the National Emergency Warning System Using Location-Based Services. *Journal of Ubiquitous Systems & Pervaisive Networks,* Vol. 3, No. 2, pp 59–66.

Amailef, K. and Lu, J. (2011). A mobile-based emergency response system for intelligent m-government services. *Journal of Enterprise Information Management*, Vol. 24, No.4, pp 338–359.

Coombs, T. (2007). *Ongoing Crisis Communication: Planning, Managing, and Responding*. SAGE Publications, New York, NY.

Hughes, A., L. and Palen, L. (2009). Twitter adoption and use in mass convergence and emergency events. *International Journal of Emergency Management*, Vol. 6, No. 3-4, pp 248–260.

Häkkinen, M. (2010). *Why alarms fail: a cognitive explanatory model*. Doctoral thesis. Jyväskylä studies in computing, 127, University of Jyväskylä, Finland.

Kaasinen, E. (2005). *User acceptance of mobile services - value, ease of use, trust and ease of adoption.* VTT Publications.

Kuula, J., Räsänen, J., Kettunen, P., Kauppinen, O. and Panasenko, V. (2012). Mobile Emergency Messaging and the Vulnerability of Crisis Communication. *Proceedings of the NBC 2012 – 8th Symposium on CBRNE threats: How does society scope,* Turku, Finland, June, pp 11–14.

Kuula, J., Auvinen, V., Kauppinen, O, Kettunen, P., Viitanen, S. and Korhonen, T. (2013). Smartphones as an Alerting, Command and Control System for the Preparedness Groups and Civilians: Results of Preliminary Test with the Finnish Police. *Proceedings of the 10th International ISCRAM Conference,* Baden-Baden, Germany, May.

Lee, J., Niko, D., Hwang, H., Park, M. and Kim, C. (2011). A GIS-based Design for a Smartphone Disaster Information Service Application. *First ACIS/JNU International Conference on Computers, Networks, Systems, and Industrial Engineering. IEEE*

Meissen, U., Voisard, A. (2008) Increasing the Effectiveness of Early Warning via Context-aware Alerting. In: Fiedrich, F.; Van de Walle, B. *Proceedings of the 5th International ISCRAM Conference*, pp 431–440.

Muralidharan, S., Rasmussen, L., Patterson, D. and Shin, J.-H. (2011) Hope for Haiti: An analysis of facebook and twitter usage during the earthquake relief efforts. *Public Relations Review*, Vol. 37, No. 2, pp 175–177.

Sillem, S. and Wiersma, E. (2006). Comparing Cell Broadcast and Text Messaging for Citizens Warning. *Proceedings of the 3rd International ISCRAM Conference*, Newark, NJ (USA), May.

Vihalemm, T., Kiisel, M. and Harro-Loit, H. (2012). Citizens' Response Patterns to Warning Messages. *Journal of Contingencies and Crisis Management*, Vol. 20, No. 1.

Waitinen, M. (2011). *Turvallinen koulu?: Helsinkiläisten peruskoulujen turvallisuuskulttuurista ja siihen vaikuttavista tekijöistä*. Helsingin yliopisto, Unigrafia.

Wu, P. F. (2009). User Acceptance of Emergency Alert Technology: A Case Study. *Proceedings of the 6th International ISCRAM Conference* – Gothenburg, Sweden.

Yamasaki, E.(2012).What we can learn from Japan's early earthquake warning system.*Momentum*, 1.

Strategic Communication for Cyber Security Leadership

Rauno Kuusisto[1, 2, 3] and Tuija Kuusisto[3, 4]
[1]Finnish Defence Forces Technical Research Center, Riihimäki, Finland
[2]University of Jyväskylä, Department of Mathematical Information Technology, Jyväskylä, Finland
[3]National Defence University, Department of Tactics and Operations Art, Helsinki, Finland
[4]Ministry of Finance, Helsinki, Finland
rauno.kuusisto@mil.fi
tuija.kuusisto@vm.fi

Abstract: This paper describes a preunderstanding of phenomena and characteristics that need to be considered when studying decision-making systems and strategic communication in physical-cyber world. The purpose of this paper is to form a preliminary hypothesis about how to identify characteristics that a leader needs to focus when aiming cyber security leadership. First, the paper studies the key concepts and terms of cyber security. Then the paper outlines a structure and activities based framework for increasing understanding of required leadership in cyber context. The framework is derived from research on organizations and information technology. The framework models that activities are executed either in the physical or cyber world and they are occurring in one of those worlds. Recently, the activities executed in the cyber world and occurring in the physical world have gained most interest. An example is malware affecting on manufacturing systems or controlling systems. The paper analyzes the results of two limited media surveys about cyber related newspaper articles. Newspaper articles were published on one of the main newspapers in Finland during the years 2011 and 2012. The key findings of the first media survey showed that organization structures as well as norms and rules were not discussed. This means that it was not known or under general interest that what were the responsibilities and who was responsible of what. In addition, the internal integration of the community had no commonly agreed departure point. The second media survey showed that the internal discussions in society about the norms and rules of cyber activities and behavior were started. People want to what kind of norms will guide the world and how this cyber-physical world will be perceived in futures. A strong need to organize cyber world seems to be in front of us. However, the ways of integration of the various communities is still unclear. This means that the basic question is: Whom we want to let to lead us?

Keywords: cybersecurity, cyber security theory, information security, governance

1. Introduction

Cyber security has gained increasing interest among the management, the users and the producers of information systems and e-services. One of the research problems of cyber world is to identify the issues that a leader needs to communicate when developing cyber security leaderhips. This paper describes a preunderstanding of phenomena and characteristics that are needed to be considered when studying decision-making systems and strategic communication in physical-cyber world. The purpose of this paper is to form a preliminary hypothesis about how to identify characteristics that a leader needs to focus when aiming cyber security leadership.

Cyber security or cybersecurity is typically defined as ‘measures taken to protect a computer or computer system (as on the Internet) against unauthorized access or attack’. First known use of cybersecurity is the year of 1994. (Merriam-Webster 2012) According to von Solms (2010) cyber security is the fifth wave of information security. Its main interest is on the security of Internet-based systems. The previous waves being technical, management, institutional and governance all exist in parallel with each other. (von Solms 2010) On the other hand, cyber security issues are discussed and activities are taken on all those waves or levels of organizations. At the technical level focus is more on data and information networks and system security issues like software vulnerabilities or denial-of-service attacks. At the management level interest lays on administrative and policy issues and often on cyber strategies. At the institutional and governance levels organizations are focusing on shared values, norms and goals as well as good governance practices. Deeper and wider understanding of cyber security principles is needed for a society to gain benefits of the digitalization of its vital functions without over controlling or loosing of cyber security. This balanced cyber security in a society means that organizations are able to efficiently and securely commit their tasks with the support or in cyber world.

Cyber is defined as ‘of, relating to, or involving computers or computer networks’. Cyber originates from cybernetics. Cybernetics interest lay on organizations, patterns, and communication in entities. It is derived from Greek world ‘kybernnētēs’ meaning ‘pilot, governor’ or ‘to steer, govern’. (Merriam-Webster 2012) So, the basic meaning of cyber is to pilot, steer or govern. A narrow definition of cyber says cyber is piloting, steering and governing of computers and on data networks. A threat is ‘an expression of intention to inflict evil, injury, or damage’ (Merriam-Webster 2012). Following these definitions cyber threats are defined as expressions of intentions to damage piloting, steering or governing of computers and on data networks. A broader definition of cyber extends it to piloting, steering and governing of systems that are connected by information and communication technology and data networks. Accordingly, cyber threats are described as expressions of intentions to inflict evil, injury, or damage systems that are connected by information and communication technology and data networks. These broader definitions allow the studying of vital functions and structures of society without locking the study in the structural restrictions of any technology.

Based on the broader definition of cyber, the concept of cyber world (Fig. 1) includes not only the computers and data and information networks, but also the complete and comprehensive system of human existence in those networks. This kind of interpretation of the concept ‘cyber world’ makes it possible to deal with the essential issues and phenomena that emerge from this novel domain. Those issues include human social behavior supported by information technical solutions. Information technology is set as a position of enabler rather than dominator of human existence.

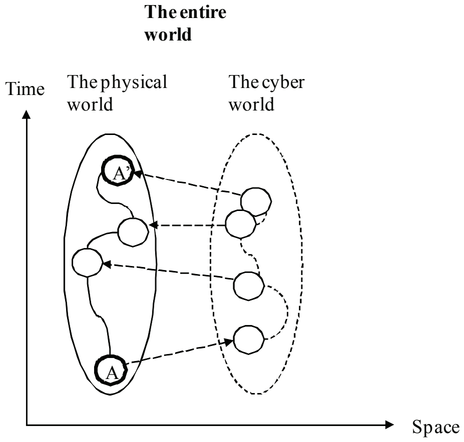


Figure 1: The interaction between the physical world and the cyber world

Leadership is a popular term typically defined as ‘the office or position of a leader’. Other definitions of leadership include ‘capacity to lead’ or ‘the act or an instance of leading’. One of the definitions of leading is ‘providing direction or guidance’ (Merriam-Webster 2012) So, from an action point of view cyber security leadership is defined as providing direction of guidance to cyber security activities. However, cyber security leadership includes capacity to lead. It has to meet the capacity to leed needs of a society or an organization performing the technical, management, institutional and governance activities of cyber security.

Some of the definitions of the term strategic communication include ‘purposeful use of communication by an organization to fulfill its mission’ (Hallahan et al. 2007) and ‘focused United States Government efforts to understand and engage key audiences to create, strengthen, or preserve conditions favourable for the advancement of United States Government interests, policies, and objectives through the use of coordinated programs, plans, themes, messages, and products synchronized with the actions of all instruments of national power’ (DOD 2012). Strategic communication is close to effecting by information approaches known as information warfare. It includes collaborative and participating encouragement approaches to communication and the use of modern technology like social media in cyber world. Strategic communication and cyber security have several relations. From cyber security point of view, communication including strategic communication is considered as an enabler of cyber security. Cyber security leadership includes the collecting and distributing of effective information about cyber security situation, resources, means and goals. In addition, strategic communication supports cyber security on long term by actively refining shared values and norms of a society or an organization. On the other hand, cyber security is needed for strategic communication to reach its goals with the support of cyber world and in the cyber world.

In cyber world communication including strategic communication between organizations and individuals is different than in physical world, because cyber world does not support all senses. The most important missing sense is touching. This changes the way of communicating. The missing power of could be compensated in various ways. These ways may transform ways of self-expression in physical world, as well. There is lots of information in the cyber world that is most essential for the functionality of the cyber world itself, but which does not reveal itself to the users of the cyber world. There are archives of essential information about and for individuals and organizations in cyber world. The physical location of those archives is not necessarily in a good control or even known by the users. The duration of archiving information is neither necessarily user controlled. Such information that is nonrecurring thus vanishing for ever after expressed in the physical world may be archived in the cyber world. This may have consequences to the way of humans will be acting during time.

Information in both cyber and physical world is vulnerable for various security threats. Cyber world as an endlessly expanding domain offers splendid opportunities for misuse of information. There are lots of new kinds or in novel way manifesting structures and activities in the cyber world that are waiting for relevant and acceptable regulations that does not yet exist. Regulations and norms form a complex system. Some norms are acceptable and good for someone's point of view, but vain or even disastrous for someone else. Interpretation of norms and their value vary for several reasons. Cyber world is more or less open. Information travels at the speed of light and borders are in different places than in the physical world.

2. The physical world and the cyber world framework

A physical and cyber world framework is outlined for increasing the understanding of required leadership in cyber context. The framework is derived from research on organizations and information technology. The framework consists of two dimensions: physical world and cyber world as presented in Figure 2. The framework models that activities are executed either in physical or cyber world and they occur in one of those worlds. For example, traveling by a car is an activity that is executed and occurs in the physical world. If the drivers needs a map update for his navigation system and pays the maps by a credit card he would perform an activity that is executed in the physical world but occurs in the cyber world. The purchased maps update to the navigation system would be an activity that is executed and occurs in the cyber world. The driver receiving route instructions is an activity executed in the cyber world and occurring in the physical world.

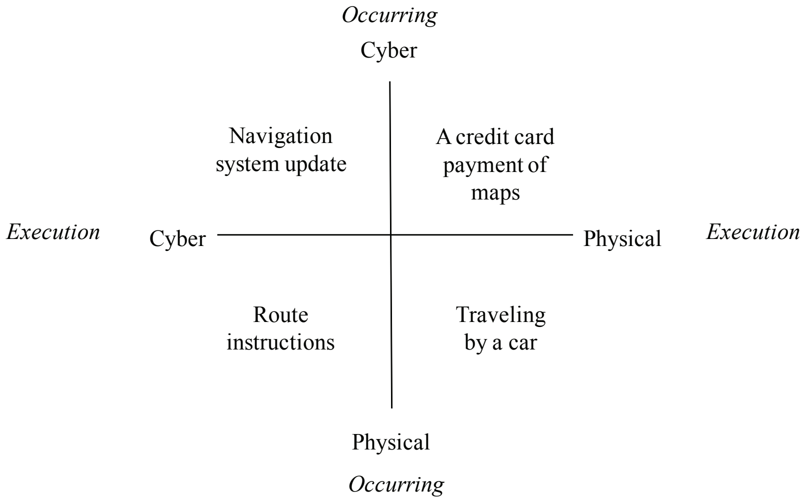


Figure 2: The physical world and the cyber world framework

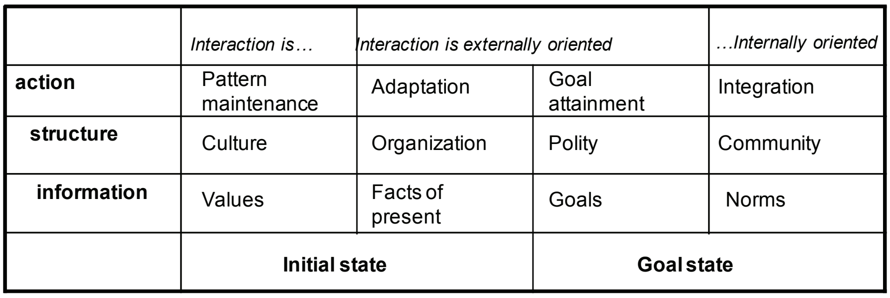
Closer example to cyber security is the explosion of a datacenter that is an activity executed in physical world and occurring in cyber world. A typical denial-of-service attack is executed and occurs in the cyber world. Recently, those activities executed in the cyber world and occurring in the physical world have gained most interest. An example is malware affecting on manufacturing systems or controlling systems. Further on, the impacts of the activities occurred can be outlined like malware affecting on manufacturing system would cause unpredictable behavior of the physical components of manufacturing systems.

3. A system model of a society or an organization

All systems as well as social systems can be considered as information-driven activity cycles in a structure. This very early discovered feature (Aristotle) says that systems consist of a structure, actions and information. They will produce activity, when right kind of information is fed into their structures. The produced activity will acts as input information for the system to produce more activity. Habermas (1984, 1989) combines theories of the social sciences and system thinking. He states that a social system contains time and space dimensions. It has initial state and goal state. Its communication orientation is both internal and external. Habermas (1989) argues referring to Talcott Parsons (1951) thinking that information directing activities of an actor contains four basic classes: Values, norms, goals and external facts. Actor is, e.g., a state, an organisation, a team, or even an individual. In the information refining process of an actor the values have effects on the norms, which both have effects on the goals and further on all those have effects on the exploiting of external facts. Activities using external facts to change values are adaptation, goal attainment, integration and pattern maintenance functions. The structural phenomena of social systems contain culture, community, polity and institutions. Cultural systems are more solid than communities, which are again more solid than a polity structures and institutions (Habermas 1989). Information fed in a structure produces various actions based on information categories. Values affect on ways to act and maintaining patterns. Norms urge forward the integration to the community. Goals will guide to reach objects and external facts produce adaptation to the requirements of the surroundings.

Facts of present, including means and resources are used for putting such an activity in practice, which will lead the actor to fulfil its goals as optimally as possible. Originally in Habermas' theory, the user of resources, i.e., institution is economy. However, the concept of a resource using structure can be applied to other entities. For example, marketing, production or research and development departments are potential resource using structures in an enterprise. (Kuusisto 2004) Cyber security and strategic communication are potential resource using structures in a society and in an organization. Leadership aspect on cyber security considers cyber security as a resource using structure. Leadership on cyber security is required to reach the goals of a society or an enterprise, i.e., to secure the units and vital functions of a society. Especially, leadership needs strategic communication as a means to provide and deliver information for directing the resources of cyber security.

A general model that contains the aspects on a social system described in previous paragraphs is presented in Figure 3. The model is derived from a more detailed figure of organization dynamics presented in (Kuusisto & Kuusisto 2009). Each field contains a certain kind of action, structure and information. Time dimension contains initial state and futures´ state and space dimension contains internal interaction and external interaction. Interactivity relationships exist between a field and the fields above or below that field. In addition, interaction exists across neighbouring action and information fields. Pattern maintenance interacts with norms and facts of present, adaptation interacts with values and goals, goal attainment interacts with facts of present and norms and integration interacts with goals and values. So, information of different functional parts of the system is a combination of the influence of neighbour parts of the system and external input of each subsystem of the comprehensive system. It can be easily recognized that this kind of system is complex thus being emergent.



	Interaction is…	*Interaction is externally oriented*		*…Internally oriented*
action	Pattern maintenance	Adaptation	Goal attainment	Integration
structure	Culture	Organization	Polity	Community
information	Values	Facts of present	Goals	Norms
	Initial state		**Goal state**	

Figure 3: A system model of a society or an organization

People are acting in the structures of a social system guided by the structure itself and obeying more or less the internal norms. People acting in a social system are producing information both inside the system and between other, neighboring systems. This kind of information flowing and continuous emergence of new kinds of interpretations forms a complex system that may be difficult to figure out and that is practically impossible to control. However, the model can help to create understanding about the complex nature of the ever interacting and dynamically evoluting system of various subparts and phenomena of the comprehensive system thus helping us to figure out relevant enough acts to make it more convenient to live in this kind of new surroundings.

4. A media survey on cyber related newspaper articles

The first author of this paper conducted limited media surveys on cyber related newspaper articles in autumns 2011 and 2012. He collected cyber related news published between 6th of September and 17th of November in 2011 and between 7th of August and 15th of November in 2012. The news was published mostly on one of the main newspapers in Finland, Helsingin Sanomat (2011 & 2012). The number of news collected was 83 in 2011 and 136 in 2012. Two groups of graduate students categorized the news according to the model described in Figure 3 as a part of their studies. The first author of this paper conducted a content analysis of the categorized news according to Krippendorff's (1980) method. The results of the analysis are presented in Tables 1 and 2.

Table 1: Reported cyber-related news (%) in one main newspaper in Finland in autumn 2011

	Interaction is…	*Interaction is externally oriented*		*…Internally oriented*
action	Pattern maintenance	Adaptation	Goal attainment	Integration
	16	17	20	4
structure	Culture	Organization	Polity	Community
	7	0	6	3
information	Values	Facts of present	Goals	Norms
	9	10	7	1
	Initial state		Goal state	

Table 2: Reported cyber-related news (%) in one main newspaper in Finland in autumn 2012

	Interaction is…	*Interaction is externally oriented*		*…Internally oriented*
action	Pattern maintenance	Adaptation	Goal attainment	Integration
	13	13	9	3
structure	Culture	Organization	Polity	Community
	7	5	6	3
information	Values	Facts of present	Goals	Norms
	10	17	5	8
	Initial state		Goal state	

The media surveys showed that cyber-physical world modelled in Figure 2 are discussed on public. The key findings of the first media survey include that organization structures were not discussed. This means that it was not known or under general interest that what were the responsibilities and who was responsible of what. Norms and rules were not discussed either. This means that the internal integration of the community had no commonly agreed departure point. That led the society to a situation where the adaptation to the current situation alone was likely to be the driving force of decision-making. The key findings of the second media survey showed that the internal discussions in society about the norms and rules of cyber activities and behavior were started. In addition, the external discussions in society about the facts of cyber activities and organization were increased compared to the results in 2011. This means that people want to what kind of norms will guide the world and how this cyber-physical world will be perceived in futures. A strong need to organize cyber world seems to be in front of us. However, the ways of integration of the various communities is still unclear. This means that the basic question is: Whom we want to let to lead us? On the other hand, is should be noted that the presented results of the media surveys are one interpretation about current status of cyber world in society.

The key findings of the media surveys show that in the current cyber-era the structures of the changing world are unclear, the future does not reveal as understandable patterns and the future is undetermined and open. Activities perceived from the world are changing and cultural changes effect on maintaining activity patterns. Decision-making apparatus is changing and goals reveal themselves in a different way than before and seemingly out of own control. It is not necessary clear and well enough known which are the mutually understood norms and what is and will be the safe community to belong to and which are the other communities to integrate with. The future of "me" reveals uncertain, because the concept of "us" is tottering. Adaptation to the new situation takes place under the control of somebody else. Feeling is that the future will be reached in an uncontrolled way.

5. Conclusions

The paper presented concepts, models and the results of media surveys for increasing understanding of those phenomena that may be important, when studying the complexity of cyber security. The media surveys show that the proposed models presented in Figure 2 and 3 are plausible for studying issues that a leader needs to communicate when developing cyber security leaderhips. In addition, the media surveys pointed out that the focus of discussion about cyber-physical world in a society is changing over time. Further on this means that a leader needs to change his strategic communication profile to be able to effectively lead and develop cyber security. Continuos collecting of empirical data are required for optimizing strategic communication. On the other hand, wider and deeper sets of empirical data are needed to verify and validate the models and conclusions based on the models.

However, there is a need for cyber security leadership. This means providing of directions and guidance to cyber security activities in the comprehensive cyber-physical world. Strategic communication serves as a means to provide and deliver information for planning and directing of the resources of cyber security.

References

DOD (2012) *DOD Dictionary of Military Terms*, http://www.dtic.mil/doctrine/dod_dictionary/, visited Dec 19, 2012

Habermas, J. (1984). *The Theory of Communicative Action, Volume 1: Reason and the Rationalization of Society*. Boston, MA: Beacon Press.

Habermas, J. (1989). *The Theory of Communicative Action, Volume 2: Lifeworld and System: A Critique of Functionalist Reason*. Boston, MA: Beacon Press.

Hallahan, K., Holtzhausen, D., van Ruler, B., Verčič, D. & Sriramesh, K. (2007) "Defining Strategic Communication" in *International Journal of Strategic Communication*, Vol 1, No. 1, pp 3-35.

Helsingin Sanomat (2011 & 2012) The main daily published newspaper printed in Finland.

Krippendorff, K. (1980) Content Analysis: An Introduction to Its Methodology, Newbury Park, CA, Sage.

Kuusisto, R. (2004). *Aspects on Availability*. Helsinki, Finland: Edita Prima Oy.

Kuusisto, R., Kuusisto (2009) T "Information Security Culture as a Social System" in Gupta, M & Sharman, R. *Social and Human Elements of Information Security,* Information Science Reference, IGI Global, Hershey, New York, pp. 77-97.

Merriam-Webster (2012) *Merriam-Webster Online Dictionary*, http://www.merriam-webster.com/, visited Dec 17, 2012

Parsons, T. (1951) *The Social System*, Glencoe.

von Solms (2010) *The 5 Waves of Information Security – From Kristian Beckman to the Present*,

in Rannenberg, K, Varadhajaran, V and Weber, C. (Eds.) SEC2010, IFIP Advances in Information and Communication Technology, Vol 330, pp 1-8

Dangers of Social Networking Sites- the Propagation of Malware

William Aubrey Labuschagne and Namosha Veerasamy
Council for Scientific and Industrial Research, Pretoria, South Africa
wlabuschagne@csir.co.za
nveerasamy@csir.co.za

Abstract: Users sometimes lack the security knowledge to protect themselves whilst carrying out activities online. One of the most popular tools used online is social networking tools. The popularity of Facebook and Twitter has become exponential with users making regular posts and updates. Due to the popularity of these sites, users easily engage and communicate with each other. Users may place personal details, hobbies and preferences in posts- all of which may look legitimate. Catchy phrases, controversial words and emotive language are all ways of enticing users into clicking on links. However, social networking site users may currently be unaware of the dangers, threats, attacks and malware that can stem from these popular forums. Malware, phishing attacks and digital attacks are emerging from these popular forums. The aim of this paper is to help uses protect themselves against malware on the social networking platforms. Various shifts in malware have taken place which include piggy- backing off files, email, spamming and now the instant messaging capabilities of social media sites provides an ideal avenue from which to dispense the next generation of malware which includes psychological tactics to influence users to perform undesired actions. Users may seemingly be unaware that a simple click on a spam message or obfuscated uniform resource locator (URL) can be triggering the download of malicious malware that will command their computer to form part of botnets. It is therefore essential to create some awareness of these dangers and explore how users can protect themselves. The paper will illustrate the dangers of social networking malware through examples. In addition, the paper will discuss propagation techniques used in social networking malware. The aim of the paper is to create user awareness to minimise the risk of falling prey to malware in popular social networking platforms. The paper will recommend best practices to users to guard against falling prey to social networking malware. In addition, the design of a high-level system to identify potential social network media malware will be proposed. Through this paper, users can better identify potential malware before they infect themselves.

Keywords: awareness, social networking sites, malware, propagation, social engineering

1. Introduction

Social networking sites have increasingly been adopted by users to conduct various activities during the last few years. These sites allow like-minded people to connect and collaborate with other users. Each social networking platform is developed around a theme, for example LinkedIn is used for business relationships and to conduct job searches while Facebook is used to stay in touch with friends as well as making new friends. These platforms allow users to communicate with other users with communication features that include chatting and posting of messages. In 2012, the five major social networking sites were Facebook, Pinterest, Twitter, Google+ and LinkedIn (Larson 2012).

Although the uses of social networking sites have been advantageous from a social connectivity point of view, many disadvantages have been highlighted. For example, the Nielsen's survey in 2012 reported on the negative effects that social networking has on productivity within the workspace (McGarry 2012). They found that within the United States of America (USA) 121 billion minutes were spent on social networking sites between July 2011 and July 2012.

With the high uptake of social networking site by users around the globe another negative effect has taken form, the use of these platforms for nefarious purposes by cyber criminals. They use the laws of economics within the digital environment to make a profit by unleashing cyber attacks on potential victims on these popular platforms. Faghani describes social networking as consisting of high clusters of smaller networks with a degree distribution which could follow power law distribution, hence friends of a user on Facebook could be very influential while others might not (Faghani, Saidi 2009). This implies the influential friends could help to propagate the malicious payload of cyber criminals quicker that other less influential friends. The sheer volume of users on social networking platforms would make money making endeavours worthwhile. Even a success rate of 1% within a target population of 1 million users is lucrative. In addition, the effort to reach the target population is minimal and cyber criminals have a wide choice of different vectors to use to attack these unsuspecting user. The attacks that could be used on social networking sites include but are not limited to SPAM, phishing, malware and identity theft.

An example of the propagation of malware through social networking sites took place in 2008. Facebook users started receiving messages on their walls instructing them to update the Adobe Flash plugin after clicking on the link to watch a movie in July 2008. The users who complied with the instruction were inadvertently infected with malware that resulted in unfavourable behaviours by their computers, including prevention of the infected computers to connect to Anti Virus (AV) vendors, as well as the modification of search results from search engines. The malware prevented AV's to update virus signatures that could be used to remove the malware from the infected system furthermore the users unknowingly visited web sites controlled by the attackers. More than 400 000 personal computers were infected by the social networking malware called Koobface (Protalinski 2012).

Users may seemingly be unaware that a simple click on a spam message or obfuscated uniform resource locator (URL) can be triggering the download of malicious malware that will control their computer from powerful Command and Control Centres forming part of botnets. It is therefore essential to create some awareness of these dangers and how users can protect themselves.

This paper addresses the various techniques of social network propogation, mitigation techniques and identification methods. A high-level system is proposed in order to help identify potential social network malware. The next section will describe social networking malware that has been encountered. The common propagation techniques will be identified by determining how the malware was spread between users.

2. Social networking malware

Facebook is a social networking platform where malware can be used to propagate and infect unsuspecting users. Bradbury listed attack vectors that could be used from Facebook. These attacks include but are not limited to clickjacking, click fraud, survey scams and rogue apps (Bradbury 2012). The following social networking malware will be analysed to determine what techniques were used by cyber criminals to ensure they became viral and spread amongst users:

- Koobface: Koobface is a well-known bot that utilised social networking sites including Facebook, Twitter and MySpace to propagate (Villeneuve, Deibert & Rohozinski 2010). The main mode of operation was to automatically spam numerous users with a catchy phrase and therefore entice users to click on the embedded link. Cyber criminals would hide the malicious looking URL with the use of short URL services that could include bit.ly (Baltazar, Costoya & Flores 2009). For example, *www.malcioussite.com* would become *http://bit.ly/YemU4W,* hence the user would not know where the final destination of the link would be. The obfuscated link would prompt users to upgrade their Flash Player or Adobe Acrobat. If users clicked on the malicious link, malware would be installed on the user's computer.
- **Most Hilarious Video Ever:** On May 29 2010, posts were made prompting users to click on a link to the "Most Hilarious Video Ever" (Runald 2010). Due to the title of the posts, various users were convinced to click on the link. When the user clicked on the link it took them to a spoofed Facebook page. Eager to view the video, users entered their credentials on the fake Facebook login page where their credentials were captured. Victims of the attack had their credentials captures and many of their accounts were compromised (Arthur 2010) .
- **Strauss-Kahn Video:** In June 2011, a new attack surfaced on Facebook. The headline contained carefully constructed words to draw the attention of the reader and it also included a provocative image to lure users to click on the link to view the x-rated video clip (McMillan 2011). The video clip allegedly contained content depicting disgraced Strauss-Kahn, a former International Monetary Fund (IMF) Managing Director, and a hotel maid. The headline and graphical photo appeared as part of a news feed on Facebook. It was in the format of a video link that a friend had liked (Smith 2011).
- **Rihanna and Hayden Panettiere:** The Rihanna and Hayden Panettiere video was another version of the Strauss-Kahn Video whereby celebrities were used as part of a ploy to draw users' attention to the post containing explicit content (Cluley 2011). This attack started in June 2011. The attackers created their attack using two components to draw users to the malicious content. First using two well-known television celebrities Rihanna and Hayden Panettiere (who are well-known to most users) and secondly adding a sexual context.
- **Syrian Activists Phishing Attack:** Various Syrian opposition activists have been targeted by phishing attacks that attempt to steal their YouTube and Facebook login credentials. Some of the attacks install

malicious surveillance software on the victims' devices. One of the attacks stemmed from a link posed in the comments section of Facebook pages of leading Syranian opposition members including Burhan Ghalioun, Chairman of the Syrian Opposition Transitional Council. After clicking on the link in the comments section, a page was displayed that had the appearance of a Facebook security download application. If the user clicked on the download, they were actually installing a malicious keylogger that collected key strokes made by the user (Galperin, Marquis-Boire 2012).

This section discussed the various methods that social network malware can be propagated. The next section addresses the identification of social networking malware components.

3. Social networking malware components

Table1 lists the findings of the analysis of social networking malware. The analysis was categorised using posts, links and social engineering. These three components as depicted in **Figure 1** are highly effective on social networking sites as these platforms uses these to ensure a social experience.

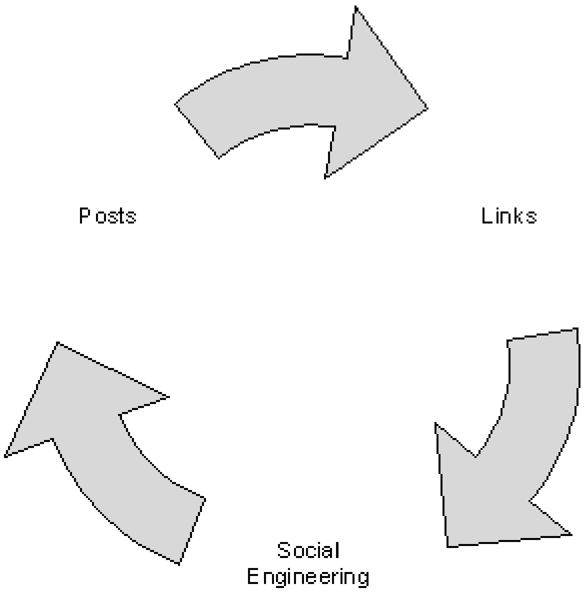


Figure 1: Effective components for propagation

The cyber criminals used the message-posting feature on Facebook to spread the dangerous payload. In Facebook, all users have a timeline where messages are posted. These messages are visible to all the other users who have access to the victim's timeline. The newsfeed will typically broadcast the message to all the users.

Malware cannot be deployed directly on a social networking site. Cyber criminals need to deploy the malware on external locations and subsequently create a link from the social networking site to the location of the malware. With the malware deployed, the next step is to ensure that users are lured to the message; this is achievable using attention grabbing content that include enticing images and provocative text. Social networking malware cannot execute by themselves and require the user to activate it (Provos et al. 2007). For example, the user needs to click on a link to view the video clip. Social engineering is used by cyber criminals, which influences users to perform actions that they would not under normal circumstances have performed (Hadnagy 2010). Grandjean noted common social engineering ploys used by malware writers including pornographic links and images, fake emails from financial services, threatening emails, urgent news and celebrity misbehaviour (Grandjean 2008). All the analysed social networking malware have similarities in that each of them posted a message on the timeline of unsuspecting users. In addition, a user had to click on the link to be directed to the malicious site with the malware and finally all of them used social engineering tactics.

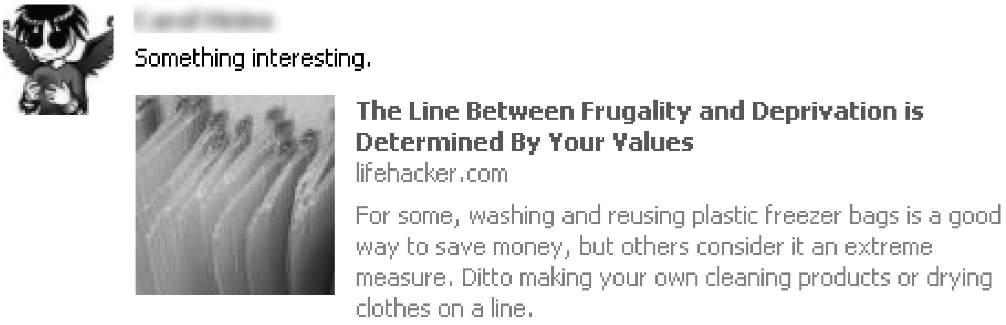
Table1: Social networking malware propagation techniques

Name	Post	Links	Social Engineering
Koobface	Yes	Yes	Yes
Most Hilarious Video Ever	Yes	Yes	Yes
Strauss-Kahn Video	Yes	Yes	Yes
Rihanna and Hayden Panettiere Video	Yes	Yes	Yes
Syrian Activists Phishing Attack	Yes	Yes	Yes

As a result, from the analysis of the mentioned social networking malware the common three components used are posts, links and social engineering. Kritzinger and von Solms state the vulnerability of personal Internet users is due to the fact that they lack the information security knowledge to understand and protect their personal computers and therefore also their personal information (Kritzinger, von Solms 2010). Next, the three components will be described within the context of messaging on Facebook.

4. Facebook messaging

A Facebook message contains at most three elements which can be described as the text, an image and a link to an external web resource for example a web page as depicted in Figure **2**. **Figure 3** illustrates clearly the different components found within a Facebook message.



[redacted]

Something interesting.

The Line Between Frugality and Deprivation is Determined By Your Values

lifehacker.com

For some, washing and reusing plastic freezer bags is a good way to save money, but others consider it an extreme measure. Ditto making your own cleaning products or drying clothes on a line.

Figure 2: Facebook message

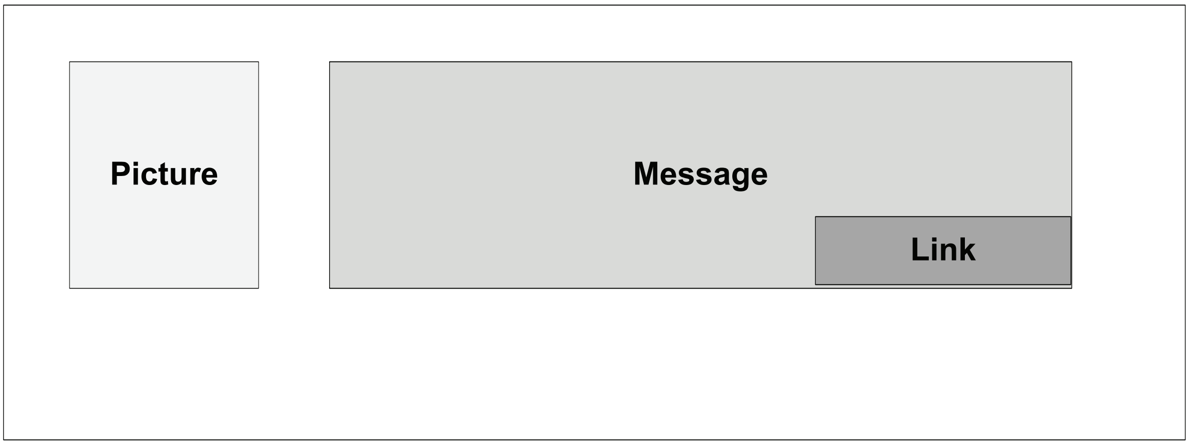
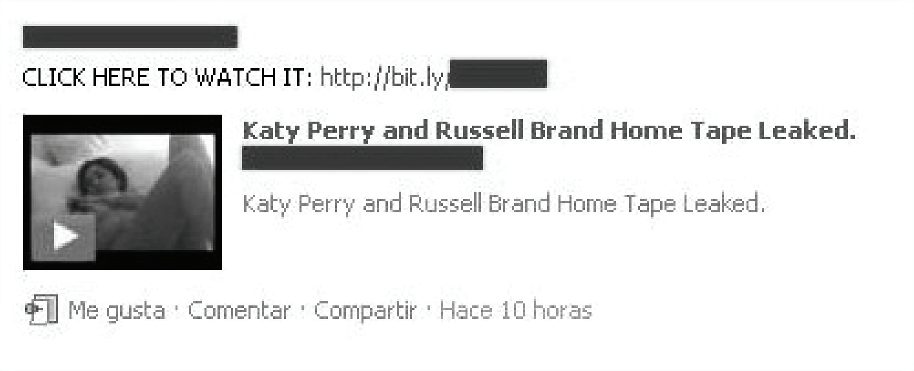


Figure 3: Facebook message high level design

This implies a Facebook message will always have at least a text element. The other elements can be added to draw the reader's attention with the use of images and also be more useful with the use of links to web sites which contain further information. The textual data provides no visual cues but does describe the context of the textual data. An image provides a graphical cue that users can use to infer the context of the message. The link provides the user with a mechanism to visit an external web site, which could contain more information. For example in normal circumstances a message (such as "*Samsung released new Galaxy phones with new design and features*") might be posted by a Facebook friend about an interesting article that was read about the new smart phones that are currently rolled out. . A user who reads this message should be able to infer the context of the posted message. The timeline where the message will be visible will compete with other posted messages and could be missed since it is competing for the user's attention. A user who posted the message could add an image of the new design of the afore-mentioned smart phone to draw the attention of the user. This will make the message more prominent and increase the chance of it being spotted by other reading users. The posting user could then also add a link to an external web site that provides more information on the new smart phones.

Figure 4 is an example of a malicious message posted on Facebook. All three components are present. Provocative images together with enticing text are used within a post to draw the attention on the user, an obfuscated external link to watch a video clip of a well-known celebrity. This message will also be posted on the timeline hence all the other unsuspecting users will see the message. If the user clicks on the link, they would be instructed to download software to watch the video clip. The software is malware and infects the web browser to ultimately spread the worm to other Facebook users (Correns 2012).



[redacted]

CLICK HERE TO WATCH IT: http://bit.ly/[redacted]

Katy Perry and Russell Brand Home Tape Leaked.

[redacted]

Katy Perry and Russell Brand Home Tape Leaked.

Me gusta · Comentar · Compartir · Hace 10 horas

Figure 4: Malicious message

The next section addresses how security vendors are addressing the threat originating from social networking malware that uses links to propagate and infect other computer systems.

5. Security measures

End users who do not have the required security knowledge have a high possibility to fall victim to social networking malware attacks. Cyber criminals are aware that users on social networking sites trust these sites implicitly and many examples in recent times have shown that users befriend strangers on these platforms without verifying the identity of the other users (Labuschagne et al, 2012). The Internet is faceless which implies that a user does not know with whom he is really communicating. Merely receiving a message from a friend on a social networking site does not imply that one can trust the content of the message. Humans are social beings and cyber criminals exploit these traits by designing attacks that focus on their curiosity. For example, the death of a celebrity drives users on the Internet to read articles covering the story. Cyber criminals have used celebrities to lure unsuspecting users to malicious websites and subsequently infect their systems with malware. Users receiving a message on social networking sites should ask the following questions before proceeding and clicking on the link:

- Does the message evoke an emotion (curiosity)?
- Does the user trust the origin of the message?
- What action does the message require the user to complete?
- Is the message a hoax, this can be determined by visiting sites which lists the latest hoaxes for example *www.hoax-slayer.com*?

These questions are used to assist in the decision making process before the user proceeds by clicking on the link. The users have control over the situation by positively answering these questions.

Social networking malware is innocuous and only becomes a threat once the user activates it. The major threat resides with the payload that the victim needs to trigger by executing it after downloading the malware from the malicious website. Security vendors have implemented mechanisms to warn and prevent users from accessing the malicious websites. For example, website reputation services can be used to scan URLs for malicious content.

Angai analysed the effectiveness of Safe Browsing Services in 2010 (Angai et al. 2010). They used Norton Safe Web, McAfee SiteAdvisor and Google Safe Browsing services to determine how effective website reputation services are in identifying malicious websites. They found that these services are inadequate in protecting normal users when only one service is use;, they suggested a more effective solution be the use of more than one website reputation services.

These results are supported by another study conducted by StopTheHacker.com (Jaal LLC) (Anonymous2010). This involved the testing of 721 URLs to determine the accuracy of website reputation service and identify the convergence in terms of safe/unsafe website. These tests included using the following website reputation services: McAfee SiteAdvisor, Norton Safe Web, Google Safe Browsing, Microsoft Bing and Comodo SiteInspector. The study concluded that the effective detection rate is not sufficient to allow users to implicitly trust these services. These services utilise different methods for analysing URLs in the identification of malicious sites. Due to different implementations, these website reputation services vendors use different databases to stores the data about the malicious sites, thus causing inconsistency between different sets of data from the different website reputation services. In addition, the size of the Internet makes the process of scanning all the available URLs within the Internet a cumbersome process. Furthermore cyber criminals create new malicious URLs every minute (Anonymous2011) hence it is impossible to have a list of all the malicious URLs on the Internet since these newly created URLs must be reported as suspicious and then analysed for malware.

In some cases some URLs might be wrongly identified as malicious. The website reputation services will only be evoked when the user clicks on the link. This clearly shows the need for a more comprehensive solution that would prevent the user from clicking on the link. Next, a system is proposed to provide a holistic solution in addressing the propagation of social networking malware using social engineering tactics.

6. Proposed system description

The proposed system consists of different components that will be used to look at each of the potential vectors individually and present the user with an early warning system. The system will analyse the context of the posted message by predicting if it is a social engineering attack tricking users to click on a link. Each of the individual components will be analysed to determine if they together form part of a social engineering attack. The different possible combinations are listed in Table 2**.**

Table 2: Warning system table

Text (Message)	Picture (Image)	Link	Threat Rating
No	No	Yes	Medium
No	Yes	Yes	High
Yes	Yes	Yes	High
Yes	No	No	Low
Yes	Yes	No	Low
Yes	No	Yes	High

In the event that only the link is available then the link will be analyzed in isolation. This would also be categorised as a medium threat since social engineering attacks would use all three components to be most effective. If the message contains a link then the system will also need to analyse the other components, if they are present. The text in the message will be tested to determine if it contains words which could draw attention for example the use of attention words like "*OMG*", "*Shocking*", "*WOW*", "*Revealing*", "*Latest news*", "*sex*", "*pornography*", "*naked*", etc. The link will also be analysed to determine if it has been identified as a malicious link using a website reputation service called Google Safe Browsing (GSB). A message containing a link and text containing attention words would be classified as a high potential threat. If the message contains a picture, then analyses of the picture is required to identify the theme of the picture. If the theme of the picture is classified as enticing or provocative together with the presence of a link in the message, then the message is classified as a high potential threat.

It is a possibility that a user posted a message that contains an enticing picture and a link to an site containing adult content. Even so, many of these adults' sites as reported by Wondracek, contain malware which could attack unsuspecting users visiting these sites and these users should be warned about the potential danger

(Wondracek et al. 2010). Therefore, if the message contains an enticing picture, attention words and a link then users would be made aware of the potential danger of a social engineering attack. The system would provide the user with appropriate notifications to prevent them from performing undesired actions, which could put the users' security at risk for example downloading software to watch an online video. A message containing no links is considered as safe or no threat even if the picture is marked as enticing and it the text in the message contains attention words. The threat exists when a link is present which could link to malware.

At an implementation level, different technologies could be implemented to analyse each of the different components as depicted in **Figure 5**. The textual analysis would be conducted using keyword frequency analysis to determine if certain attention words are present within the message.GSB could be used to analyze the links to determine if they have been classified as malicious. The implementation of GSB does have effects, which should be considered. For example all the black listed web sites are stored in a database. The location of the database determines how quickly the analyses will occur. Also, the location will also affect how up to date the database data will be. Two choices are available when using GSB, download the entire database with all the blacklisted websites or use the application programming interface (API) to connect to the online database (Kuo et al. 2005). Furthermore, analysis is required to determine what theme is depicted in the image (Yee et al. 2003). The result should consist of a threat rating which could identify potential social engineering attacks and inform the user of a suggested action before clicking on the link.

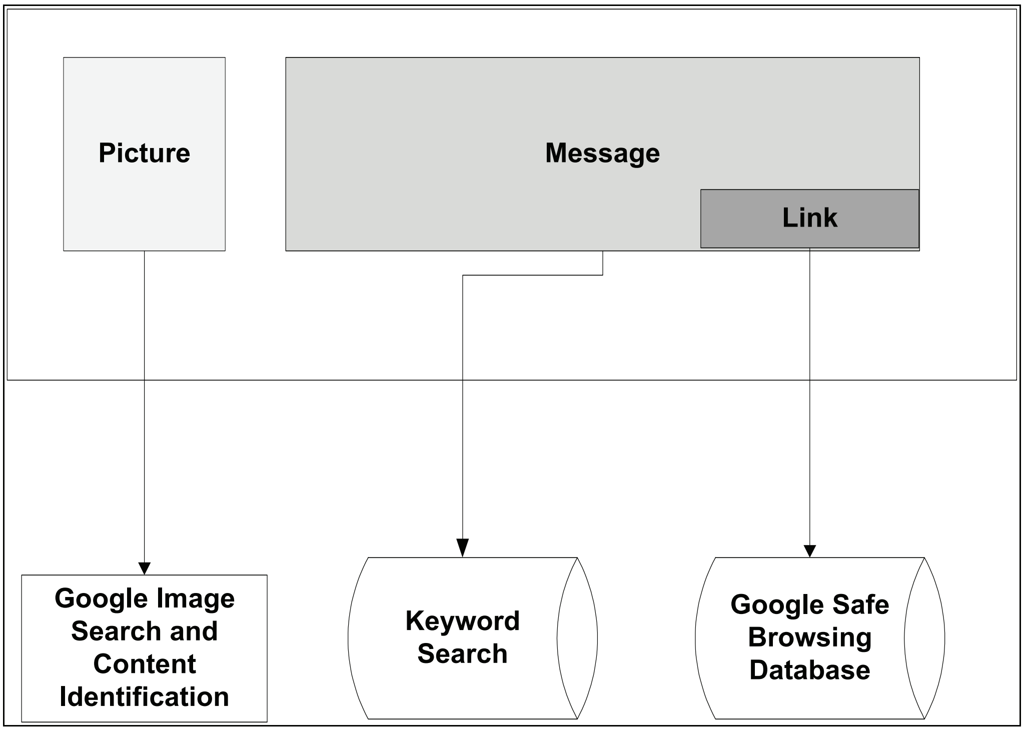

Picture
Message
Link
Google Image Search and Content Identification
Keyword Search
Google Safe Browsing Database

Figure 5: High level system description

7. Future work

The development and deployment of the proposed system will be pursued as future work. The complete system will be developed first to test the conceptual idea subsequent work will be focused on the optimization and effectiveness of each individual component. In addition, security awareness content will be created which addresses the issues of social engineering malware and how the propagation of these could be mitigated on social networking sites.

8. Conclusion

Social networking sites have been widely adopted as part of everyday life, which include staying in contact with friends but also conducting business. Duality also exists within the social networking domain whereby cyber criminals are using these platforms to attack users for nefarious purposes. Due to the nature of social networking sites that promotes sociability, many cyber attacks have incorporated social engineering into their arsenals. Many examples of successful social networking malware has been reported during the past few years where Koobface was the most profitable for cyber criminals but also very effective. These social networking malware uses specially crafted messages posted on the users' profile. The messages entice users to perform undesired actions for example clicking on a link to view a video and then unknowingly install malware that infects their computer systems and opens themselves to many other attacks. These messages contain three components: a link to an external web site, pictures and text to draw the attention of the inquisitive user.

Website reputation services do exist which can be used to analyze the link but studies have shown that at the current time these are not as effective-hence the user should be prevented from clicking on the link. User's attention should be drawn to these links with the use of text and images. These components should also be analyzed to determine if the message contains social engineering tactics and inform the user about the potential threat. A high-level system design is proposed to address each of these components as a social engineering prevention system to mitigate the threat of social engineering malware on social networking sites.

References

Angai, F., Ching, C., Ng, I. & Smith, C. (2010), "Analysis on the Effectiveness of Safe Browsing Services", .

Arthur, C. (2010), *Twitter 'funniest video' link hides malware threat* [Homepage of guardian.co.uk], [Online]. Available: **http://www.guardian.co.uk/technology/blog/2010/may/20/twitter-funniest-video-security-threat-malware** [2013, 01/27].

Baltazar, J., Costoya, J. & Flores, R. (2009), "The real face of Koobface: The largest web 2.0 botnet explained", *Trend Micro Research,* vol. 5, no. 9, pp. 10.

Bradbury, D. (2012), "Spreading fear on Facebook", *Network Security,* vol. 2012, no. 10, pp. 15-17.

Cluley, G. 2011, (2011), June 1-last update*, Rihanna and Hayden Panettiere sex video spreads Mac malware on Facebook* [Homepage of Sophos], [Online]. Available: **http://nakedsecurity.sophos.com/2011/06/01/rihanna-hayden-panettiere-lesbian-sex-video-mac-malware-facebook/** [2013, 01/17].

Corrons, L. (2012), *Katy Perry and Russell Brand baits to spread a new Facebook worm* [Homepage of Pandasecurity], [Online]. Available: **http://pandalabs.pandasecurity.com/katy-perry-and-russell-brand-baits-to-spread-a-new-facebook-worm/** [2013, 01/27].

Faghani, M.R. & Saidi, H. (2009), "Malware propagation in online social networks", *Malicious and Unwanted Software (MALWARE), 2009 4th International Conference on* IEEE, pp. 8.

Galperin, E. & Marquis-Boire, M. (2012), 2012, March 29-last update*, Syrian Activists Targeted With Facebook Phishing Attack* [Homepage of Electronic Frontier Foundation], [Online]. Available: **https://www.eff.org/deeplinks/2012/03/pro-syrian-government-hackers-target-syrian-activists-facebook-phishing-attack** [2013, 01/27].

Grandjean, E. (2008), "A prime target for social engineering malware", *'Mcafee security journal'Debuts,* , pp. 16.

Hadnagy , C. (2010), *Social Engineering: The Art of Human Hacking,* 1st edn, Wiley.

Kritzinger, E. & von Solms, S.H. (2010), "Cyber security for home users: A new way of protection through awareness enforcement", *Computers & Security,* vol. 29, no. 8, pp. 840-847.

Kuo, C., Schneider, F., Jackson, C., Mountain, D. & Winograd, T. (2005), "Google Safe Browsing. Project at Google", *Inc., June–August,* .

Labuschagne, W.A., Eloff, M.M. & Veerasamy, N. (2012), "The dark side of Web 2.0", *IFIP Advances in Information and Communication Technology,* vol. 386/2012, no. ICT Critical Infrastructure and Society, pp. 237-249.

Larson, D. (2012), *Infographic: Spring 2012 Social Media User Statistics* [Homepage of Tweetmaster.com], [Online]. Available: **http://blog.tweetsmarter.com/social-media/spring-2012-social-media-user-statistics/** [2013, 01/25].

McGarry, C. (2012)*, Nielsen survey: Social media sucking up most of our time* [Homepage of PCWorld.com], [Online]. Available: **http://www.pcworld.com/article/2019194/nielsen-survey-social-media-sucking-up-most-of-our-time.html** [2013, 01/25].

McMillan, R. (2011)*, Facebook video scam puts malware on Mac and Windows* [Homepage of computerworld.com], [Online]. Available: **http://www.computerworld.com/s/article/9217229/Facebook_video_scam_puts_malware_on_Mac_and_Windows** [2013, 01/27].

Protalinski, E. (2012), *Facebook to expose hackers behind Koobface worm* [Homepage of ZDNet], [Online]. Available: **http://www.zdnet.com/blog/facebook/facebook-to-expose-hackers-behind-koobface-worm/7462** [2012, 09/19].

Provos, N., McNamee, D., Mavrommatis, P., Wang, K. & Modadugu, N. (2007), "The ghost in the browser analysis of web-based malware", *Proceedings of the first conference on First Workshop on Hot Topics in Understanding Botnets*USENIX Association, Berkeley, CA, USA, pp. 4.

Runald, P. (2010)*, Most Hilarious Video attack on Facebook* [Homepage of Websense], [Online]. Available: **http://community.websense.com/blogs/securitylabs/archive/2010/05/28/most-hilarious-video-attack-on-facebook.aspx** [2013, 01/27].

Security Threat Report (2011). Available: http://www.sophos.com/en-us/medialibrary/Gated%20Assets/white%20papers/sophossecuritythreatreport2011wpna.pdf [2013, 02/15].

Smith, C. (2011), *Facebook Malware Attack: Fake Strauss-Kahn Video Infects Mac And PC Users (UPDATE)* [Homepage of The Huffington Post], [Online]. Available: **http://www.huffingtonpost.com/2011/06/01/facebook-malware-strauss-kahn-video_n_869576.html** [2013, 01/27].

Villeneuve, N., Deibert, R. & Rohozinski, R. (2010), *Koobface: Inside a crimeware network,* Munk School of Global Affairs.

Website-Reputation Services Agree to Disagree (2010), 2010, 17 January-last update [Homepage of StoptheHacker], [Online]. Available: **https://www.stopthehacker.com/2010/01/17/website-reputation-services-agree-to-disagree/** [2013, 01/15].

Wondracek, G., Holz, T., Platzer, C., Kirda, E. & Kruegel, C. (2010), "Is the Internet for porn? An insight into the online adult industry", *Proceedings (online) of the 9th Workshop on Economics of Information Security, Cambridge, MA*.

Yee, K.P., Swearingen, K., Li, K. & Hearst, M. (2003), "Faceted metadata for image search and browsing", *Proceedings of the SIGCHI conference on Human factors in computing systems,* ACM, , pp. 401.

The Ways, Means and Ends in Cyber Security Strategies

Martti Lehto
Department of Mathematical Information Technology, Faculty of Information Technology, University of Jyväskylä, Jyväskylä, Finland
lehto.martti@kolumbus.fi

Abstract: For all nations, the information technology revolution quietly changed the way business and government operate, as well as the daily life of citizens. The asymmetrical threat posed by cyber-attacks and the inherent vulnerabilities of cyberspace constitute a serious security risk confronting all nations. In recent years attacks against critical infrastructures, critical information infrastructures and the Internet have become ever more frequent and complex because perpetrators have become more professional. Threats in cyberspace can be classified in many ways. This is evident when you look at cyber security on a multinational level. One of the most common models is a classification based on motivational factors. Most nations use this model as a foundation when creating a strategy to handle cyber security threats as it pertains to them. This paper will use the five level model: cyber activism, cybercrime, cyber espionage, cyber terrorism and cyber warfare. The National Cyber Security Strategy articulates the overall aim and objectives of the nation's cyber security policy and sets out the strategic priorities that the national government will pursue to achieve these objectives. The Cyber Security Strategy also describes the key objectives that will be undertaken through a comprehensive body of work across the nation to achieve these strategic priorities. Cyberspace underpins almost every facet of the national functions vital to society and provides critical support for areas like critical infrastructure, economy, public safety, and national security. National governments aim at making a substantial contribution to secure cyberspace and they have different focus areas in the cyber space. In this paper the cyber security priority areas are in three categories: the public sector, the private sector and citizens. In this context the level of cyber security reached is the sum of all national and international measures taken to protect all activities in the cyber space. This paper will analyze the definitions, threats and objectives of the Cyber Security Strategies made by Australia, Canada, Estonia, Finland, Germany, the Netherlands, the United Kingdom and the United States.

Keywords: cyber security strategy, cyber definition, cyber threats

1. Introduction

The global community continues to experience an increase in the scale, sophistication and successful perpetration of cyber-attacks. As the quantity and value of electronic information has increased so too have the efforts of criminals and other malicious actors who have embraced the Internet as a more anonymous, convenient and profitable way of carrying out their activities. Of primary concern is the threat of organized cyber-attacks capable of causing debilitating disruption to a nation's critical infrastructures, functions vital to society, economy, or national security.

A national cyber security strategy is a tool to improve the security and resilience of national information infrastructures and services. It is a high-level, top-down approach to cyber security that establishes a range of national objectives and priorities that should be achieved in a specific timeframe. As such, it provides a strategic framework for a nation's approach to cyber security. (ENISA 2012)

Cyber security strategies are analyzed through comparative research methods. This research strategy illustrates the similarities and differences between selected cases. Comparative research methods have been used for a long time in cross-cultural research when it comes to identifying, analyzing and explaining similarities and differences between cultures. This research uses the normative perspective: comparisons facilitate a class analysis of findings, and enable the creation of explanatory models for the findings. A comparison that conforms to the classification generates an analytical framework for the purpose of studying the differences and similarities between different cyber security strategies. The findings are then contextualized and amalgamated into the entirety.

2. Cyberspace definition

Many countries are defining what they mean by cyber world or cyber security in their national strategy documents. The common theme from all of these varying definitions, however, is that cyber security is fundamental to both protecting government secrets and enabling national defence, in addition to protecting the critical infrastructures that permeate and drive the 21st century global economy. (NATO 2012)

The Australian cyber security strategy defines cyberspace on the foundation of Australia's digital economy and the importance and benefits of ICT (Information and Communications Technology) to the entire national economy. In accordance with the strategy "Australia's national security, economic prosperity and social well-being are critically dependent upon the availability, integrity and confidentiality of a range of information and communications technologies. This includes desktop computers, the Internet, mobile communications devices and other computer systems and networks." In short, it is all about the world of networks and terminals.

The Canadian cyber security strategy starts out with the definition of cyberspace: "Cyberspace is the electronic world created by interconnected networks of information technology and the information on those networks. It is a global commons where more than 1.7 billion people are linked together to exchange ideas, services and friendship." Cyberspace is not only limited to physical networks; rather, it is a world consisting of the exchange of information, communication and different services.

Strictly speaking, Estonia's cyber security strategy does not define cyberspace. A definition of a kind can be found in the descriptions of the implementation of cyber security. "The security of cyberspace acquires a global dimension and the protection of critical information systems must be elevated, in terms of national security, on a par with traditional defence interests." Furthermore, the strategy states that "the security of the Internet is vital to ensuring cyber security, since most of cyberspace is Internet-based." The strategy introduces networks, users and the information content into cyberspace.

Finland's cyber security strategy succinctly states: "an international term for this interdependent, multipurpose electronic data processing environment is the cyber domain."

In Germany's cyber security strategy "Cyberspace is the virtual space of all IT systems linked at data level on a global scale. The basis for cyberspace is the Internet as a universal and publicly accessible connection and transport network which can be complemented and further expanded by any number of additional data networks. Cyberspace includes all information infrastructures accessible via the Internet beyond all territorial boundaries." This global network (infrastructure + services) is invaluable for the state, the business community and the everyday activities of individual citizens alike.

The Netherlands' cyber security strategy does not specifically define cyberspace. The section that details cyber security also includes a definition of some kind over cyberspace. "Cyber security is freedom from danger or damage due to the disruption, breakdown, or misuse of ICT." This description follows the traditional data security perspective.

The United Kingdom clearly defines cyberspace: "Cyberspace is an interactive domain made up of digital networks that is used to store, modify and communicate information. It includes the Internet, but also the other information systems that support our businesses, infrastructure and services." The strategy illustrates the critical infrastructure which is necessary for society's everyday activities. "In the UK digital networks already underpin the supply of electricity and water to homes, help organise the delivery of food and other goods to shops, and act as an essential tool for businesses across the UK." In addition to this the network connects the citizens with various services.

According to the U.S. viewpoint "Cyberspace is their [critical infrastructures] nervous system — the control system of our country. Cyberspace is composed of hundreds of thousands of interconnected computers, servers, routers, switches, and fiber optic cables that allow our critical infrastructures to work. Thus, the healthy functioning of cyberspace is essential to our economy and our national security." The definition highlights the critical infrastructure rather than network services, information contents or service users.

There are terms and concepts associated with cyberspace which are difficult to define due to the very nature of cyberspace and different phenomena therein. Cyberspace is a man-made ecosystem. While land, air, sea and space domains exist without any human presence, cyberspace requires continuous human attendance and activities. Cyberspace fuses all ICT networks, databases and sources of information into a global virtual system. Cyberspace structures include the economy, politics, armed forces, psychology and information (Grobler et.al 2011). Some researchers also include societal and infrastructure domains in cyberspace. Nonetheless, the Internet is an integral and elemental part of this new world.

3. Cyber threat landscape

Threats in cyberspace are difficult to define as it is hard to identify the source of attacks and the motives that drive them, or even to foresee the course of an attack as it unfolds. The identification of cyber threats is further complicated by the difficulty in defining the boundaries between national, international, public and private interests. Because threats in cyberspace are global in nature and involve rapid technological developments, the struggle to meet them is ever-changing and increasingly complicated.

Threats in cyberspace can be classified in many ways. The threat landscape is a list of threats containing information about threat agents and attack vectors. By exploiting weaknesses/vulnerabilities, threats may lead to a loss or takeover of assets. (ENISA 2012b)

The European Network and Information Security Agency (ENISA) uses a cyber threat model consisting of threats. The threats include different forms of attacks and techniques as well as malware and physical threats.

Table 1: Cyber threats (ENISA 2012b)

Cyber-attacks and techniques	Malwares	Physical threats
Drive-by Exploits Code Injection Attacks Botnets Denial of service Phishing Compromising confidential information Targeted Attacks Identity Theft Abuse of Information Leakage Search Engine Poisoning	Exploit Kits Worms/Trojans Rogueware/Scareware Spam	Physical Theft/Loss/Damage Rogue certificates

In the ENISA-model “a threat agent is any person or thing that acts (or has the power to act) to cause, carry, transmit, or support a threat”. Some of the major threat agents in cyberspace are corporations, cybercriminals, employees, hacktivists, nation states, and terrorists. (ENISA 2012b)

One of the common threat models is a fivefold classification based on motivational factors: cyber activism, cybercrime, cyber espionage, cyber terrorism and cyber warfare. With a typology such as this motives can reduced to their very essence: egoism, anarchy, money, destruction and power. This fivefold model is derived from Myriam Dunn Cavelty’s structural model (Dunn Cavelty 2010, Ashenden 2011).

Level 1 consists of cyber activism which encompasses cyber vandalism, hacking and hacktivism. For a single company or an individual their activities can cause significant economic losses. The recent activities of the Anonymous hackers have been more effective than in the past.

Level 2 consists of cybercrime. The Commission of the European Communities defines cybercrime as “criminal acts committed using electronic communications networks and information systems or against such networks and systems".

Level 3 consists of cyber espionage. This can be defined as action aimed at obtaining secret information (sensitive, proprietary or classified) from individuals, competitors, groups, governments and adversaries for the purpose of accruing political, military or economic gain by employing illicit techniques in the Internet, networks, programs or computers. (Liaropoulos 2010)

Level 4 consists of cyber terrorism which utilizes networks in attacks against critical ICT systems and their controls. The purpose of the attacks is to cause damage and raise fear among the general public, and to force the political leadership to give into the terrorists’ demands. (Beggs 2006)

Level 5 cyber warfare consists of three separate entities: strategic cyber warfare, tactical/operational cyber warfare and cyber warfare in low-intensity conflicts. No universally accepted definition for cyber warfare exists; it is quite liberally being used to describe the operations of state-actors in cyberspace. Cyber warfare *per se* requires a state of war between states, with cyber operations being but a part of other military operations.

Cyber threats can be assessed from the standpoint of cyber operations and cyber weapons. Modern cyber weapons are tailored to the task. For example, Stuxnet, Flame, Duqu and Gauss are all modular bot software whose desired impacts have been created from several components. These cyber weapons contain warheads which execute the actual function and transport bodies that take the warhead to their targets.

Australia's cyber security strategy identifies cybercrime as the primary cyber threat. "The global community continues to experience an increase in the scale, sophistication and successful perpetration of cybercrime. Just as we have seen the benefits of ICT in promoting legitimate economic activity, we now see cybercrime emerging on an unprecedented scale."

Canada's cyber security strategy includes a chapter dedicated to cyber threats. The strategy identifies three types of threats. "Cyber Crime: Once they have access to a computer, attackers can steal or distort the information stored on it, corrupt its operations and program it to attack other computers and the systems to which they are connected. Cyber Espionage: The most sophisticated cyber threats come from the intelligence and military services of foreign states. Cyber terrorism: Terrorist networks also are moving to incorporate cyber operations into their strategic doctrines. Among many activities, they are using the Internet to support their recruitment, fundraising and propaganda activities."

The Estonian cyber security strategy presents a three-tier cyber threat model. Its motivational factors encompass cybercrime, cyber terrorism and cyber warfare. Moreover, the strategy states that "advanced technologies and attack methods make it difficult to define with any certitude or clarity the motives impulsing an attack, threats can also be classified on the basis of methods employed and on the extent of damage inflicted."

Finland's cyber security strategy does not explicitly describe the threat. It only states that "threats against the cyber domain have increasingly serious repercussions for individuals, businesses and society in general. The perpetrators are more professional than before and today the threats even include state actors. Cyber-attacks can be used as a means of political and economic pressure; in a serious crisis pressure can be exerted as an instrument of influence alongside traditional means of military force."

When it comes to defining cyber threat, Germany's cyber security strategy parallels that of Finland, focusing on the cyber-attack. "Given the openness and extent of cyberspace it is possible to conduct covert attacks and misuse vulnerable systems as tools for an attack. In view of technologically sophisticated malware the possibilities of responding to and retracing an attack are rather limited. Often attacks give no clue as to the identity and the background of the attacker." In addition, the strategy lists actors that may commit the attacks: "Criminals, terrorists and spies use cyberspace as a place for their activities and do not stop at state borders. Military operations can also be behind such attacks."

The Netherlands' cyber security strategy also concentrates on a description of cyber-attacks. "When cyber-attacks occur, it is often difficult to identify the perpetrator, who may be a loner, an organization, a state, or a combination of all three. The nature of the cyber threat (cybercrime, cyber terrorism, cyber activism, cyber espionage, and cyber conflict) is also often unclear. But many cyber-attacks involve the same techniques and methods."

The cyber security strategy of the United Kingdom lists three threat agents: criminals, spies and terrorists. The strategy states that "criminals from all corners of the globe are already exploiting the internet to target the UK in a variety of ways. Some of the most sophisticated threats to the UK in cyberspace come from other states which seek to conduct espionage with the aim of spying on or compromising our government, military, industrial and economic assets, as well as monitoring opponents of their own regimes. Cyberspace is already used by terrorists to spread propaganda, radicalise potential supporters, raise funds, communicate and plan."

In turn, the U.S. cyber security strategy describes cyber-attacks as follows: "a spectrum of malicious actors can and do conduct attacks against our critical information infrastructures. Of primary concern is the threat of organized cyber-attacks capable of causing debilitating disruption to our Nation's critical infrastructures, economy, or national security. Cyber-attacks on U.S. information networks can have serious consequences such as disrupting critical operations, causing loss of revenue and intellectual property, or loss of life."

Cyber security strategies discuss cyber threats, vulnerabilities and the different forms of cyber conflict in a varying manner. Attacks against a country's critical infrastructure and ICT infrastructure are considered to be the most important threats. When their controls changed over to modern network-based systems their vulnerabilities increased. Consequently, this increased the threat of an attack. The strategies do not discuss cyber warfare; it is only mentioned in five strategies.

Cybercrime is felt to be constantly increasing. In particular, in the Internet world cybercrime causes enormous financial losses to individuals and companies on a global scale. The key security measures in the Internet environment are predominantly associated with the prevention of cybercrime.

Cyber espionage is the third most important threat. The targets of cyber espionage include the central government, armed forces, weapons industry and the business community in general as well as the academia. According to observations, cyber espionage amounts to a terabyte's worth of stolen information annually.

Most strategies mention cyber terrorism. It is not perceived as a separate entity; instead it is regarded as a functional area of terrorism. Nor is it described from the traditional viewpoint of the concept of terrorism. The Internet plays a significant role in cyber terrorism due to the fact that terrorist organizations can promulgate information over the Internet or recruit and train new members. Table 2 presents the prevalence of cyber threats in the strategies that were studied for this paper.

Table 2: Cyber threats in the cyber security strategies

Country	Cyber activism	Cyber-crime	Cyber espionage	Cyber terrorism	Cyber warfare
Australia	x	x	x	x	
Canada		x	x	x	
Estonia		x		x	x
Finland		x			
Germany		x	x	x	x
Netherlands	x	x	x	x	x
United Kingdom		x	x	x	
United States of America	x	x	x	x	x

4. Key objectives and focus in cyber security strategies

Cyber Security Strategies define a roadmap for the implementation of the strategy. It contains concrete activities that would meet the objectives of the strategy and a governance framework for the implementation, evaluation and maintenance of the strategy. The cyber security strategy also has a master plan for the implementation of the strategy and concrete action plans for each activity.

The Australian cyber security strategy (2009) clearly identifies the aims and the objectives of the cyber security policy. The citizens are put first in the Australian cyber security's objectives: "All Australians are aware of cyber risks, secure their computers and take steps to protect their identities, privacy and finances online." Next, the strategy's security measures focus on the business community, which "operates secure and resilient information and communications technologies to protect the integrity of their own operations and the identity and privacy of their customers." Third, "the Australian Government ensures its information and communications technologies are secure and resilient." The strategy lists 7 strategic priorities.

- Improve the detection, analysis, mitigation and response to sophisticated cyber threats.
- Educate and empower all Australians with the information, confidence and practical tools to protect themselves online.
- Partner with business to promote security and resilience in infrastructure, networks, products and services.
- Model best practice in the protection of government ICT systems.
- Promote a secure, resilient and trusted global electronic operating environment.
- Maintain an effective legal framework and enforcement capabilities to target and prosecute cyber-crime.

- Promote the development of a skilled cyber security workforce with access to research and development to develop innovative solutions.

Canada's Cyber Security Strategy prioritizes the objectives and aims in the opposite manner. "The Strategy (2010) is built on three pillars: [1] Securing Government systems, [2] Partnering to secure vital cyber systems outside the federal Government, and [3] Helping Canadians to be secure online. Canada's Cyber Security Strategy will strengthen the cyber systems and critical infrastructure sectors, support economic growth and protect Canadians as they connect to each other and to the world."

Securing Government Systems contains the following priorities:

- Establishing Clear Federal Roles and Responsibilities
- Strengthening the Security of Federal Cyber Systems
- Enhancing Cyber Security Awareness throughout Government

Partnering to Secure Vital Cyber Systems focuses on the following key areas:

- Partnering with the Provinces and Territories
- Partnering with the Private Sector and Critical Infrastructure Sectors

Helping Canadians to be Secure Online has two priorities:

- Combatting Cybercrime
- Protecting Canadians Online.

Estonia's cyber security strategy (2008) seeks primarily to reduce the inherent vulnerabilities of cyberspace in the nation as a whole. "Estonia's cyber security should be pursued through the coordinated efforts of all concerned stakeholders, of public and private sectors as well as of civil society." The Estonian cyber security strategy lists 5 strategic goals which also describe the measures used in achieving the goals. The cyber security goals are:

- Development and implementation of a system of security measures
- Increasing competence in information security
- Development of a legal framework for cyber security
- Development of international co-operation
- Raising awareness of cyber security.

There are three visions for cyber security in Finland's cyber security strategy (2013). The first one states: "Finland can secure its vital functions against cyber threats in all situations." The next vision is that "citizens, the authorities and businesses, can effectively utilise a safe cyber domain and the competence arising from cyber security measures, both nationally and internationally." The third vision, as per the Finnish Government Programme, is that "by 2016, Finland will be a global forerunner in cyber threat preparedness, and in managing the disturbances caused by these threats." The strategy lists the following 10 strategic guidelines:

- Create an efficient collaborative model between the authorities and other actors
- Improve comprehensive cyber security situation awareness among the key actors
- Maintain and improve the ability to detect and repel cyber threats and disturbances
- Ensure the police have sufficient capabilities to prevent, expose and solve cybercrime
- The Finnish Defence Forces will create a comprehensive cyber defence capability
- Participation in the activities of international organisations
- Improve the cyber expertise and awareness of all societal actors
- Provide national legislation for effective cyber security measures
- Assign cyber security related tasks to the authorities and actors in the business community
- Monitor the implementation of the Strategy and its completion.

The Cyber Security Strategy for Germany (2011) states that "the Federal Government aims at making a substantial contribution to a secure cyberspace, thus maintaining and promoting economic and social prosperity in Germany. Cyber security in Germany must be ensured at a level commensurate with the importance and protection required by interlinked information infrastructures, without hampering the opportunities and the utilization of the cyberspace." The strategy identifies the following 10 strategic areas:

- Protection of critical information infrastructures
- Secure IT systems in Germany
- Strengthening IT security in the public administration
- National Cyber Response Centre
- National Cyber Security Council
- Effective crime control also in cyberspace
- Effective coordinated action to ensure cyber security in Europe and worldwide
- Use of reliable and trustworthy information technology
- Personnel development in federal authorities
- Tools to respond to cyber attacks.

According to the cyber security strategy of the Netherlands (2011) the "goal is to strengthen the security of digital society in order to give individuals, businesses, and public bodies more confidence in the use of ICT. To this end, the responsible public bodies will work more effectively with other parties to ensure the safety and reliability of an open and free digital society." The strategy lists the following 6 priority activities:

- Setting up the Cyber Security Council and the National Cyber Security Centre
- Setting up threat and risk analyses
- Increasing the resilience of critical infrastructure
- Capacity for responding to ICT disruptions and cyber attacks
- Intensifying the investigation of cybercrime and the prosecution of its perpetrators
- Encouraging research and education.

The vision in the UK cyber security strategy is "for the UK in 2015 to derive huge economic and social value from a vibrant, resilient and secure cyberspace, where our actions, guided by our core values of liberty, fairness, transparency and the rule of law, enhance prosperity, national security and a strong society. Achieving this vision will require everybody, the private sector, individuals and government, to work together." The strategy lists the following 4 objectives for the purpose of achieving the vision:

- The UK to tackle cyber crime and be one of the most secure places in the world to do business in cyberspace
- The UK to be more resilient to cyber attacks and better able to protect our interests in cyberspace
- The UK to have helped shape an open, stable and vibrant cyberspace, which the UK public can use safely and that supports open societies
- The UK to have the cross-cutting knowledge, skills and capability it needs to underpin all our cyber security objectives.

The policy of the U.S. cyber security strategy (2003) is "to prevent or minimize disruptions to critical information infrastructures and thereby protect the people, the economy, the essential human and government services, and the national security of the United States." The Strategy articulates five national priorities. The first priority focuses on improving the ability to respond to cyber incidents and reduce the potential damage from such events. The second, third, and fourth priorities aim to reduce the numbers of cyber threats and overall vulnerability to cyber-attacks. The fifth priority focuses on preventing cyber-attacks with the potential to impact national security assets and improving international management of and response to such attacks.

The cyber security strategies identify citizens as well as the public and the private sector as focus areas of action. Australia's strategy clearly expresses the focus areas. In the other strategies they are often itemised in list form and it is impossible to establish whether they are deliberately presented in a prioritized order or through happenstance.

When it comes to the standpoint of public administration, some strategies emphasize the role of the government in relation to all administrative branches, and others highlight the role of the central government in seeing to the security of the nation.

From the individual citizen's perspective the strategies focus on the safe and secure use of the Internet as well as the protection of identity and privacy in the virtual realm and in using Internet-based services.

Based on a content analysis, Table 3, in which the numbers indicate the order of priority for each focus area, was generated.

Table 3: Cyber strategy focus

Country	Public sector	Private sector	Citizens
Australia	3	2	1
Canada	1	2	3
Estonia	1	2	3
Finland	1	2	3
Germany	3	2	1
Netherlands	3	2	1
United Kingdom	3	1	2
United States of America	3	2	1

Of all of the strategies researched six key priorities could be identified which appeared in almost every cyber security strategy. These priority areas are:

- Roles and responsibilities of cyber security
- Cyber security center / situation awareness
- Legislation and supervising the lawfulness of government actions
- Cyber security training and research
- Secure ICT products and services
- National and international cooperation

A number of various proposals for action were presented within the scope of these priority areas. The number of said proposals varied from 7 to 57, and there were marked differences as regards their level of detail. The bulk of the proposals related to various public sector-associated initiatives. The strategies do not attach these initiatives to any administrative branch, in other words, the strategies only present a number of different measures without assigning responsibilities to any organisation. Only the U.S. strategy lists the actors responsible for the different segments of critical infrastructure.

When it comes to cyber security management and responsibilities, the strategies present quite divergent approaches. Whereas the German strategy proposes the establishment of a National Cyber Security Council, the Finnish strategy states that "Each ministry is in its sector responsible for preparing cyber security related matters and appropriate arrangement of administrative matters". Many other strategies only make passing reference to the management of cyber security and the responsibilities of each actor.

5. Summary

This paper analyzed eight cyber security strategies. All of the strategies presented a focus, the cyber threat (at least in general terms), strategic objectives and cyber security proposals. Yet, significant variance was to be found in their scope and depth.

In their cyber security strategies states have selected different viewpoints, such as the standpoint of the public administration, the private sector (industry and the business community) and the citizen perspective. While each strategy identified the existence of these three points of view, there were differences in emphases and priorities.

The strategies’ cyber threat scenarios represent a classification based on the actors’ motives and the form in which the threat materializes, and its impact. Cyber security strategies varyingly address the different forms of cyber threats, vulnerabilities and cyber conflicts. Different attacks against the nation’s critical infrastructure and critical information infrastructure are considered to be the most serious threat. The next biggest threats include cybercrime and cyber espionage. While most strategies do mention cyber terrorism, it is perceived to be but one form of terrorism, rather than an individual function. Some strategies give mention to cyber warfare. However, no extensive analysis is given.

The creation and development of a national cyber security strategy requires close cooperation between all stakeholders. While many different definitions of cyber security exist, it is generally considered to be an instrument that helps governments manage security measures in controlling cyber security risks, and creates the kind of cyber resilience that serves the national goals.

References

Australian Cyber Security Strategy (2009)

Ashenden Debi (2011), Cyber Security: Time for Engagement and Debate, Proceedings of the 10th ECIW Conference, Tallinn, Estonia, pp. 15

Beggs Christopher (2006), Proposed Risk Minimization Measures for Cyber-Terrorism and SCADA Networks in Australia, Proceedings of the 5th ECIW Conference, Helsinki, Finland, pp. 9-18

Canada’s Cyber Security Strategy (2010)

Dunn Cavelty, Myriam (2010), The Reality and Future of Cyber war, Parliamentary Brief

ENISA (2012a), National Cyber Security Strategies, Practical Guide on Development and Execution, December 2012

ENISA (2012b), Threat Landscape, Responding to the Evolving Threat Environment, September 2012

Estonia Cyber Security Strategy (2008)

Finland´s Cyber Security Strategy (2013)

Germany Cyber Security Strategy (2011)

Grobler Marthie, van Vuuren Joey Jansen and Zaaiman Jannie, Evaluating Cyber Security Awareness in South Africa, Proceedings of the 10th ECIW Conference, Tallinn, Estonia, pp. 114-115

Liaropoulos Andrew (2010), War and Ethics in Cyberspace: Cyber-Conflict and Just War Theory, Proceedings of the 9th ECIW Conference, Thessaloniki Greece, pp. 177-182

NATO CCD COE (2012) National Cyber Security Framework Manual

Netherlands, The National Cyber Security Strategy (2011)

UK Cyber Security Strategy (2011)

United States of America, The National Strategy to Secure Cyberspace (2003)

Information Security – Military Standards Versus ISO 27001: A Case Study in a Portuguese Military Organization

José Martins[1], Henrique dos Santos[2], António Rosinha[1] and Agostinho Valente[3]
[1]Academia Militar – CINAMIL, Lisboa, Portugal
[2]Universidade do Minho – DSI, Guimarães, Portugal
[3]Instituto Geográfico do Exército, Lisboa, Portugal
jose.carloslm@gmail.com
hsantos@dsi.uminho.pt
antonio.rosinha@sapo.pt
agostinho.av@gmail.com

Abstract: The objective of this paper is to present a *Case Study* conducted in a Portuguese military organization, which seeks to answer the following research questions: (1) what are the most relevant dimensions and categories of information security controls applied in military organizations? (2) What are the main scenarios of information security incidents that are expected to occur? (3) What is the decision process used for planning and selection information security controls? Current trends in technological advances impose new security requirements to information systems. This is true for all application domains, including military and especially concerning their unavoidable links to cyberspace. However, most of the time these information systems are specified under rules adapted to a different environment, what can result in unexpected and dangerous security flows. This study aims to evaluate how the military doctrine of the Portuguese Army limits or promotes the implementation of the international standard ISO / IEC 27001 (information systems security management) and simultaneously to propose a formal method for the selection and management of information security controls, based on that standard and aligned with the military organization. This *Case Study* consists of three phases: the first phase involves the collection and analysis of key documentation of the organization; in a second phase, a questionnaire was applied in the military organization, to three distinct groups - decision-makers, information security specialists, and employees with functions specifically linked to information use; and finally, in a third phase, interviews with specialists are used to validate the results obtained from the other phases. This study reveals that (1) information security within the military organization is built on the basis of physical and human attack vectors, and targeting the infrastructure that supports the flow of information in the organization (i.e. Information Systems), (2) the information security controls applied in the military organization are included in ISO 27001; (3) planning and selection of applied information security controls are made by decision makers and information security specialists, aiming to protect mainly integrity of digital information. It appears that specialists impose their planning options essentially inferring knowledge from analogies (like following guidelines), or rather, seeking to select and retrieve past successful information security cases, i.e. similar scenarios concerning situations under planning and that may (likely) lead to the selection and implementation of the most efficient information security controls.

Keywords: information security management, information security method, case studies, military decision-making process, international standard ISO 27001

1. Introduction

Information security consists on the protection of information stored, processed or transmitted, against loss of confidentiality, integrity and availability, through the implementation of a diverse set of technical, administrative and physical controls. This type of security is fundamental to military organizations, which have, as one of its main objectives, to guarantee information superiority. These organizations operate in an environment of Information Warfare, which can be defined as a set of actions to preserve our Information Systems from exploitation, corruption or destruction, while simultaneously trying to exploit, corrupt and destroy opponents' Information Systems (IS) (Waltz, 1998). Currently the importance of information security to military organizations is increased by the development of doctrines of cyber war in several countries, more specifically the called, Computer Network Operations, which fall within the command and control warfare, and information operations.

The objective of this paper is to present a *Case Study* conducted in a Portuguese military organization, which seeks to answer the following research questions: (1) what are the most relevant dimensions and categories of information security controls applied in military organizations? (2) What are the main scenarios of information security incidents that are expected to occur? (3) What is the decision process used for planning and selection information security controls? This study seeks to examine both how the military doctrine concerning information security limit or enhance the implementation of the international standard ISO / IEC 27001 (2005) for information security management, in the Portuguese Army.

The military organization chosen to conduct the Case Study has as a well-defined mission, to support implementation of activities related to geographical science, cartographic technique and the promotion and development of scientific actions and technological research, in the field of geographic support and geomantic, directed to the Army, other branches of the Armed Forces and the civil community.

The selection of this military organization took into consideration that it is a level one organization, whose operation is vital to the fulfillment of the Army's mission, complying with the current military security regulations. In addition, it features an integrated system of quality, environment, health and safety at work (SIQAS), in accordance with international standards (i.e., NP ISO 9001: 2008, NP ISO 14001: 2004 and NP ISO 18001: 2007) and one of the most complete and technologically evolved Information Systems, among all Units, Establishments and Agencies (U/E/O) of the Portuguese Army.

In order to answer the research questions, the paper is divided into four sections: the first section presents the problem and sets out the main objectives of the study. The second section briefly presents the applied research plan. The third section presents the main results of the study. Finally to conclude the study, the fourth section will present the findings and the limitations of the study.

2. Research plan

In order to answer the indicated questions a Case Study takes place in a military organization, of the Portuguese Army. This *Case Study* consists of three phases: the first phase involves the collection and analysis of key documentation of the organization; in a second phase, a questionnaire was applied in the organization, to three distinct groups - decision-makers, information security specialists, and employees with functions specifically linked to information use; and finally, in a third phase, interviews with specialists are used to validate the results obtained from the other phases.

The orientation of the research in this study is predominantly interpretive. In this research only one Case Study is carried out. Considering that each military organization is a unique case, a similar study with more military U/E/O is required to allow the generalization of results. Despite, the option was to conduct a single Case Study with the greatest possible depth according with the limitations of military security imposed by the organization. Moreover, the difficulty in obtaining authorization to perform Case Studies like this and the required time for its completion also limits the application of the research method.

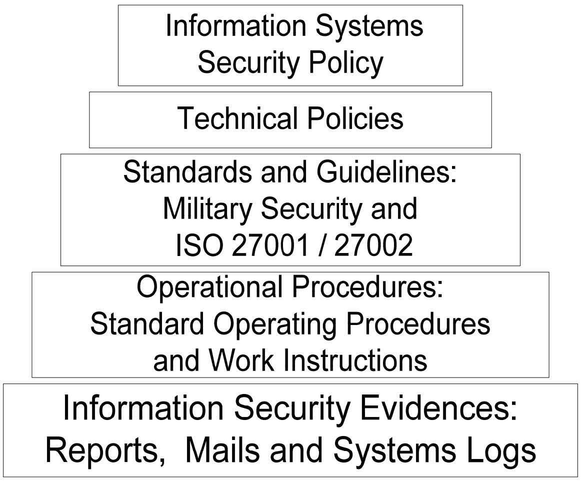
3. Research outcomes

The main results obtained from the Case Study reflect the three research phases applied to this study i.e. document analysis, the questionnaire, interviews and the participative observations.

Documental Analysis

In the documental analysis the main goal is to obtain and analyze the most relevant documentation of the organization to guarantee the fullest possible vision of the organization's business processes and of the information security. The main results in information security were obtained, in accordance to the hierarchy of the documents presented in Figure 1 and the key players of information security according to Figure 2.

The organization does not have an information security policy, but it does have an IS security policy, which integrates various technical policies. The elaboration of the policy and the implementation of the security controls have as its main guidelines security standards classified under the military Army (RAD 280-1, 2003; RAD 280-2, 2005; RAD-95, 2008; SEGMIL 1, 1986), complemented with the international standard for information security management ISO 27001. January 2013 marked begin of the critical controls' metric application for effective cyber defense, recommended by the SANS Institute (SANS – CSCCD 4.0, 2013). Moreover, in terms of organizational instruction, the organization combines work instructions in supporting the management model implemented, with grounds on SIQAS, with permanent military implementing rules (NEPs).



Information Systems Security Policy

Technical Policies

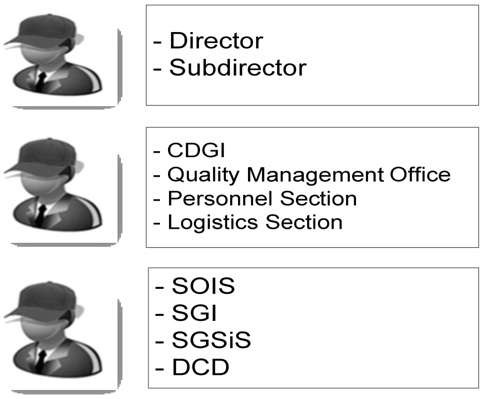
Standards and Guidelines: Military Security and ISO 27001 / 27002

Operational Procedures: Standard Operating Procedures and Work Instructions

Information Security Evidences: Reports, Mails and Systems Logs

Figure 1: Documental hierarchy for information security

The IS policy refers to the importance and the necessity of IS security, the structure of the security organization and define responsibilities for the entities involved. The main concerns are centered on the classification and control of organizational resources, the safety and security of personnel, physical and environmental security, management of operations and communication, access control to information, management and continuity of activities. This policy is complemented with technical standards for the use of technological equipment, passwords, software applications, the Internet and email. In terms of the primary responsible for security, actors can be identified in the three main levels of the organization, i.e. at the strategic level (e.g., director), the management level (e.g., quality assurance office) and at the operational level (e.g., system management section).



- Director
- Subdirector

- CDGI
- Quality Management Office
- Personnel Section
- Logistics Section

- SOIS
- SGI
- SGSiS
- DCD

Figure 2: Main actors of information security[1]

There are several entities related to information security, but there are no responsible for operational management coordination of the various security dimensions. However, the excellent working relationships, respect and obedience in the military chain of command, ensure that this management is carried out with the collaboration of the involved entities. In addition, the director of the organization (i.e. the military chief) is primarily responsible for all activities that are executed or left to run within the organization, and thus ultimately it is the director that is the operational manager for information security.

Questionnaire

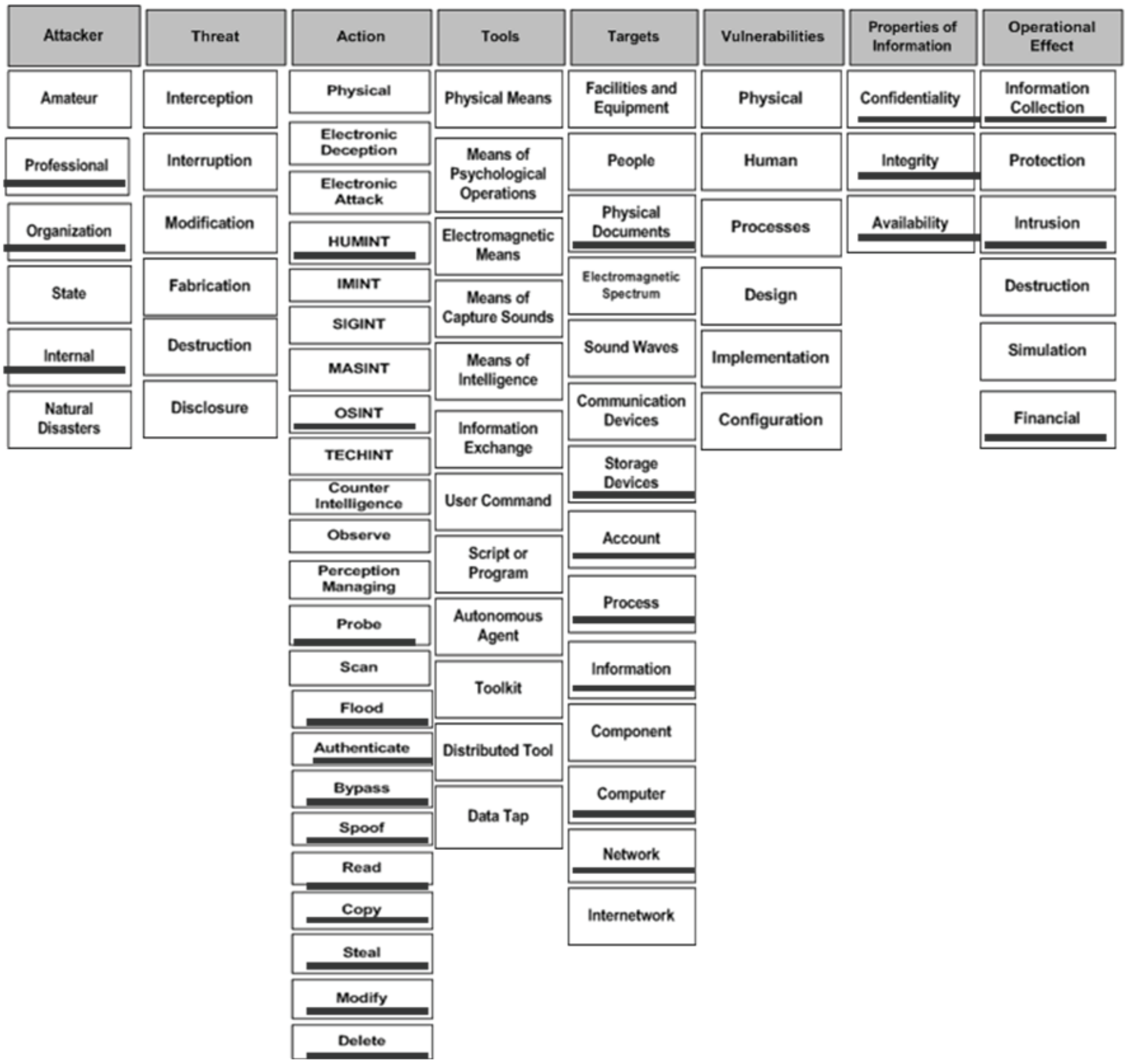
In a second phase of the study a questionnaire was completed by fifty employees from a total of one hundred and sixty employees. The main purpose of the questionnaire was to identify the major concerns in terms of

[1] Abbreviations: CDGI – Centre for Development and Information Management; SOIS – Operations Section, Information and Security; SGI – Management Information Section; SGSis – System Management Section and DCD – Department of Conception and Development.

vectors and attack methods, and simultaneously understand the decision process used for information security planning.

In terms of attack vectors directed to the organization, the main identified vectors (by priority order) are: the physical vector; the human vector; and the technological infrastructure vectors that support the processing, transmission and storage of information. Besides, using a generic model of information security incidents for military organizations in information warfare environment, developed by Martins, Santos, Nunes and Silva (2012a), it is possible to identify in Figure 3 the main concerns of the organization after analyzing some of the key variables in the model studied in the organization through the application of the questionnaire (i.e. attacker, action, targets, properties of information and operational effect).

The model suggests the variables, or the principle parameters of the problem, followed by the identification of the range of values or conditions that each parameter expresses. The variables of this model identify the possible attackers, threats and attack methods (i.e. the actions, tools/weapons) and targets that can be achieved to affect the fundamental properties of information security (i.e. confidentiality, integrity and availability) directly or indirectly by exploiting the vulnerabilities of the major components of the IS. Variables that have the conditions indicated in figure 3 (i.e. with an underline) allow to obtain some of the main scenarios of attack methods to the organization according to the perception of its employees, by identifying relationships between identified conditions of the proposed model.



Source: Adapted of Martins et al. (2012a)

Figure 3: Main scenarios of attack methods

Another main result obtained from the questionnaire, indicate that the fundamental entities for the planning of information security are the specialists in information security and that the main guidelines for security planning are primarily the experience of the specialists, the informal risk assessment done in association with the possible attack scenarios, and finally the laws, regulations and mandatory military standards. The opinion of the respondents is that this is a safe organisation, and information security is very important subject to the organisation.

Interviews

In the third and last phase of the Case Study, interviews were conducted simultaneously with observations of evidence, making the triangulation of data sources (i.e. documents, questionnaires, interviews and participant observation). From this interpretive and qualitative analysis, the main results obtained are summarized in the following points:

- The organisation applies the principle of military security: (1) the need to know, (2) the least privilege, (3) the defence in depth and (4) the responsibility, loyalty and trust of employees. It is the military chain of command that is responsible for defining what should or should not be known by each employee, in association with the principle of least privilege to ensure maximum security on classified information. On the other hand the organisation seeks to implement a set of security measures that ensure defence in depth against attack scenarios that are considered most prevalent. Within the chain of command there is also the perception that employees are responsible, loyal and trustworthy and that their permanence in office facilitates individual knowledge.
- The military organisation uses to plan, implement and audit information security according to a set of internal and external documentation. The external documentation is fundamental to the security standards in force in the military Army (SEGMIL 1, 1986), NATO military security standards and the ISO 27001 standard. The main internal documentary sources for security management are the SIQAS, the security IS policy, the internal emergency plan and NEPs of military security.
- The organisation meets the military security standards guidelines of the Army in three major areas: personnel safety and security, information security, material and facilities security. Information security is essentially concerned with the security of classified materials and the security of Communications and Information Systems (CSI).
- The key dimensions of information security in the organization are: the physical dimension, the responsibility of the SOIS; the technological dimension, which focuses primarily on the security of IS through the SGSis; in software development (if possible insurance), by the DCD and the management of geographic information, especially in control of its integrity and versions by SGI. The human security dimension has contributions from employees (i.e. responsible for human resource management) and a core support from the central command, which enables identifying “deviant” behaviour in employees. There is also a concern about drug and alcohol abuse in the organisation, according to the programme to prevent and combat drug and alcohol abuse in the military organizations. Finally the organisational dimension is centred on the directorate and deputy director, logistics section (i.e. responsible for purchasing, equipment maintenance), at the office of quality, which is responsible for monitoring and SIQAS and in SGSis.
- The lack of formally described information security requirements is basically obsolete by the experience and initiative of the elements responsible for SGSis, of the SOIS and the guidelines of the chain of command. The main concerns of the organisation in terms of information security is to ensure the integrity of the geographic information, the confidentiality of classified military information and availability of geographic information for different customers (i.e. Army). Although the organisation conducts an identification and qualitative risk assessment through the office of quality, the risk of information security is not done formally, but it is a constant concern for the chain of command.
- The selection of information security controls is made on the basis of the capacity to cope with a given method of attack, its ease of management and the intrinsic value of information that is intended to protect (although not qualified). Systems to handle classified information are certified by the National Security Office and installed by Transmissions Regiment of the Army. The handling of classified information conforms to what is regulated superiorly by the Army in SEGMIL 1. The organisation adopts

procedures that are defined superiorly to the military security of states, although they are mainly focused on the physical security of the organisation.

- To handle classified information every employee needs accreditation. In addition to the SOIS and the personnel section of the organisation, the sub-register of the Army Staff also operates, validating biographic data of the accredited collaborator and the National Security Office that conducts employees' security investigation process. The organisation has a manual of functions and a skills profile of all employees, which guides in training and awareness education.

From the study it appears that the organisation bases its information security in four main security dimensions: an organisational dimension, a physical dimension, a human dimension and a technological dimension. The categories of information security obtained in the organisation was based on an information security framework for military organisations within an Information Warfare environment, developed by Martins, Santos, Nunes and Silva (2012b).

In this study we highlight two aspects of the framework. First the security categories not yet implemented, but that the organisation acknowledges to be important for information security (marked in Figure 4 through 7 with a black bar) and secondly the specific military security that are not currently referenced in the international standard for information security management, i.e. ISO 27001.

In the *Organisational Security Dimension* (see Figure 4) there is a need for an information security plan that integrates all the scattered documents related to military security in diverse dimensions. There is too a lack of a model that allows the organisation to identify possible scenarios of information security incidents (i.e. identification and risk assessment information security) so as to plan and integrate the various security controls of various dimensions. The organisation need to formalise the acceptance criteria of risk and does not have an information security management across all processes organisation's business. Ultimately not currently have in-house expertise to enable it to perform computer forensics, requiring the support of higher ranked military in the area of forensic computer science.

The interviews allows us to identify differentiating aspects of this dimension and military organisation that are not addressed by ISO 27001, such as the importance of military values (i.e. loyalty, responsibility, trust) as unquantifiable contributions to information security. A fundamental aspect to consider is also the obligation to comply with security standards and military security principles.

In the U/E/O the military upper echelon audits to verify the operational status and to inform the need to change of the organization security status. On the other hand the organisation performs planned audits according to its SIQAS and focused on information security. The organisation seeks to integrate information security in implemented SIQAS, and the director has the primary responsibility for coordinating the information security management. The organisation promotes a security controls baseline approach, assuming it is aligned with best practises (hence the use of ISO 27001, in addition to the military security standards), and valuates the lessons learnt as a basis of knowledge about information security.

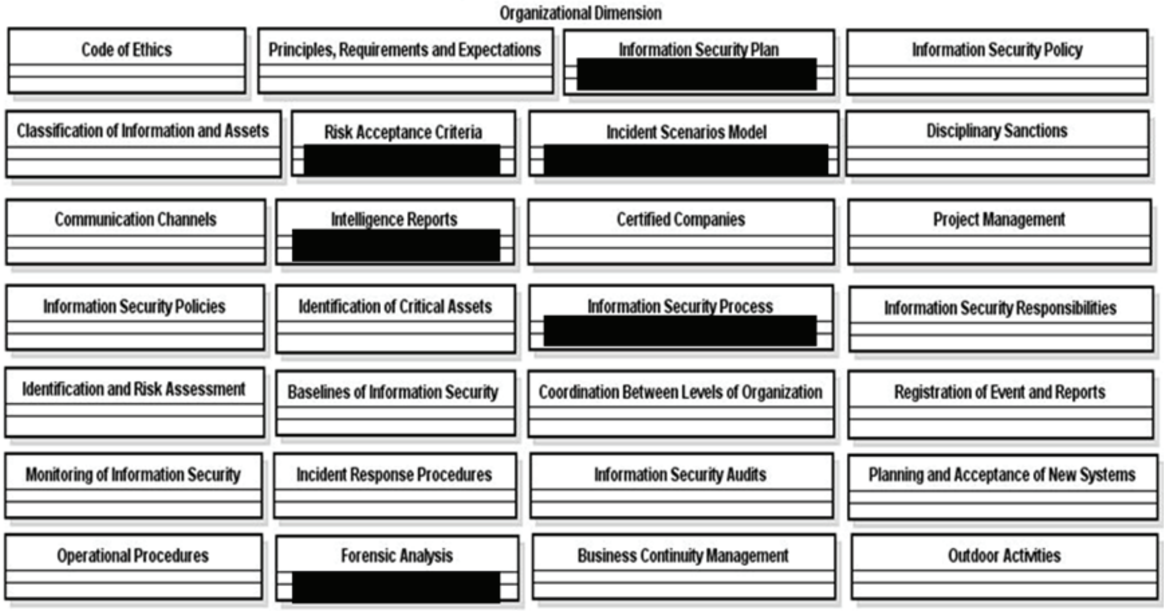
However, the essential aspect identified is the importance of the chain of command to establish the access control and information manipulation within the organisation. Simultaneously it ensures the implementation of coordination measures between the various organisation levels (i.e. weekly meetings and descending information, annual planning policy). In the case of classified military information, it is managed by the SOIS the director supervision (i.e. military chief), thereby reducing the stakeholders in the process.

The *Physical Security Dimension* categories are identified in Figure 5, which shows that this security dimension is one most covered. The main specifications of military organisations in this dimension are the existence of armed security personnel and military mentioned in order to be able to act by the use of physical force against an opponent. There is a requirement for security classification of physical areas in military organisations, to have a security plan and an internal emergency plan.

Another fundamental aspect is the possibility to conduct inspections to electrical and electronic materials before meeting where information shared in case of no guarantee of confidentiality may jeopardise the fulfilment of the mission of the Army. There is also a concern about the emissions protection of electromagnetic radiation equipment, in particular the system for exchanging classified information between

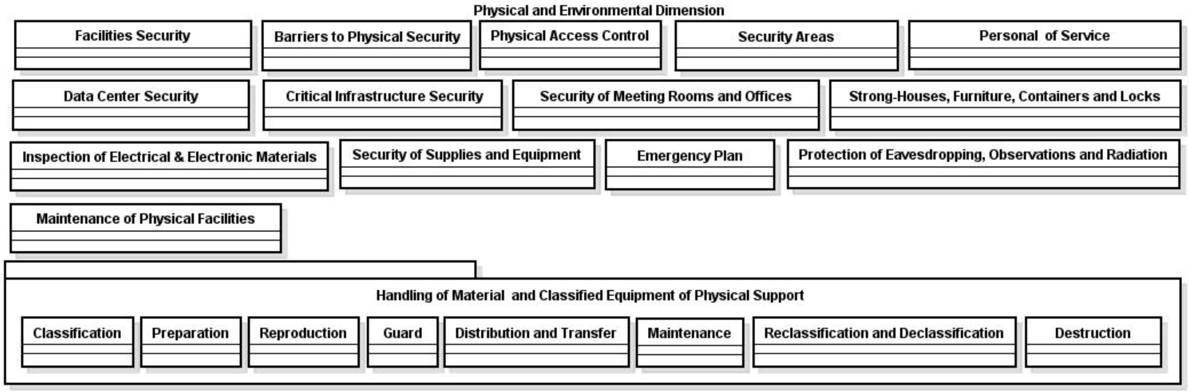
U/E/O military and the holding of classified materials and equipment on a physical medium (i.e. paper, The *Physical Security Dimension* categories are identified in Figure 5, which shows that this security dimension is one most covered. The main specifications of military organisations in this dimension are the existence of armed security personnel and military mentioned in order to be able to act by the use of physical force against an opponent. There is a requirement for security classification of physical areas in military organisations, to have a security plan and an internal emergency plan.

Another fundamental aspect is the possibility to conduct inspections to electrical and electronic materials before meeting where information shared in case of no guarantee of confidentiality may jeopardise the fulfilment of the mission of the Army. There is also a concern about the emissions protection of electromagnetic radiation equipment, in particular the system for exchanging classified information between U/E/O military and the holding of classified materials and equipment on a physical medium (i.e. paper, analogue photography).


Organizational Dimension
Code of Ethics
Principles, Requirements and Expectations
Information Security Plan
Information Security Policy
Classification of Information and Assets
Risk Acceptance Criteria
Incident Scenarios Model
Disciplinary Sanctions
Communication Channels
Intelligence Reports
Certified Companies
Project Management
Information Security Policies
Identification of Critical Assets
Information Security Process
Information Security Responsibilities
Identification and Risk Assessment
Baselines of Information Security
Coordination Between Levels of Organization
Registration of Event and Reports
Monitoring of Information Security
Incident Response Procedures
Information Security Audits
Planning and Acceptance of New Systems
Operational Procedures
Forensic Analysis
Business Continuity Management
Outdoor Activities

Source: Adapted of Martins et al. (2012b)

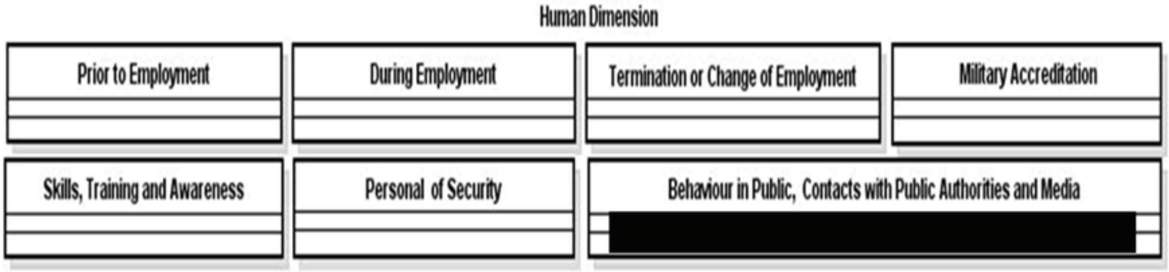
Figure 4: Information security categories in the organisational dimension


Physical and Environmental Dimension
Facilities Security
Barriers to Physical Security
Physical Access Control
Security Areas
Personal of Service
Data Center Security
Critical Infrastructure Security
Security of Meeting Rooms and Offices
Strong-Houses, Furniture, Containers and Locks
Inspection of Electrical & Electronic Materials
Security of Supplies and Equipment
Emergency Plan
Protection of Eavesdropping, Observations and Radiation
Maintenance of Physical Facilities
Handling of Material and Classified Equipment of Physical Support
Classification
Preparation
Reproduction
Guard
Distribution and Transfer
Maintenance
Reclassification and Declassification
Destruction

Source: Adapted of Martins et al. (2012b)

Figure 5: Categories of physical dimension of information security

In the *Human Security Dimension* the categories not yet implemented or operationally limited are indicated in Figure 6. The category addressing the behaviour of employees in public, contact with public authorities and the media, which also include several networks, is currently the least operationalised. In this organisation only the director is authorised to contact the media and is not authorised to have an official page on social networks (i.e. Facebook).

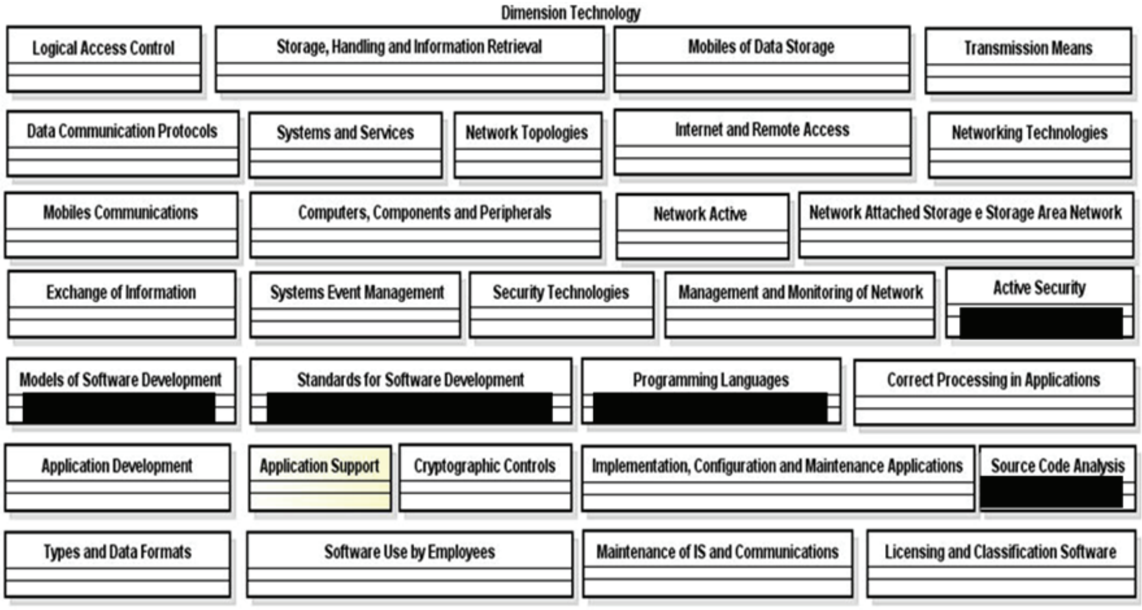

Human Dimension
Prior to Employment
During Employment
Termination or Change of Employment
Military Accreditation
Skills, Training and Awareness
Personal of Security
Behaviour in Public, Contacts with Public Authorities and Media

Source: Adapted of Martins et al. (2012b)

Figure 6: Categories of human dimension of information security

In this dimension the main specificities include the existence of a mandatory accreditation process for all employees who handle classified information, the provision of a host manual distributed to all employees, which integrates new employees and clarify various aspects of the inner working, including the information security. There is a core of support for the command, which allows detecting "deviant" employee's behaviour. There is an ongoing concern with aspects related to alcoholism and drug use by employees. Two important aspects that are particular to the organisation and contribute to information security are the organisation operating functions and the skills profile of employees, which allow each employee to have continuous update as to powers in accordance with employees' functions. An essential aspect is also a concern for the safety of all employees, including the organisation being certified in health and safety (i.e. the NP ISO 18001: 2007) and there is a concern to conduct daily briefings to ensure that military staff that are completing military service have a clear notion of their responsibilities and the procedures to be undertaken in the event of a security incident.

The unimplemented or of limited operation security categories in the *Technology Dimension* are indicated in Figure 7. Currently the categories that are less operationalised are those that have to do with internal software development. However, this organisation is a particular case since the majority of the U/E/O military do not develop software. Another limitation is the diminished ability to perform penetration tests with internal skills.
In this dimension what protrudes is the speciality of the transmission system of classified information between U/E/O military. This dimension is the one which the organisation more relies on international standards of information security controls and to implement security metrics, i.e. ISO 27001 and the critical controls for effective cyber defence recommended by the SANS Institute. It is also in this security dimension that there is a perception that passive security is not sufficient to ensure information security, highlighting the benefits of active security by means of penetration testing methodologies (Wilhelm, 2010).


Dimension Technology
Logical Access Control
Storage, Handling and Information Retrieval
Mobiles of Data Storage
Transmission Means
Data Communication Protocols
Systems and Services
Network Topologies
Internet and Remote Access
Networking Technologies
Mobiles Communications
Computers, Components and Peripherals
Network Active
Network Attached Storage e Storage Area Network
Exchange of Information
Systems Event Management
Security Technologies
Management and Monitoring of Network
Active Security
Models of Software Development
Standards for Software Development
Programming Languages
Correct Processing in Applications
Application Development
Application Support
Cryptographic Controls
Implementation, Configuration and Maintenance Applications
Source Code Analysis
Types and Data Formats
Software Use by Employees
Maintenance of IS and Communications
Licensing and Classification Software

Source: adapted of Martins et al. (2012b)

Figure 7: Categories of technology dimension of information security

The military organisation deploy information security management based on the framework obtained from the Case Study and shown in Figure 8. It can be seen that the military organisation focuses on the mission, i.e. operational requirements, study the opponent, the internal environment where it can act and is in conformity with the doctrine, law and military regulations. The security approach is primarily focused on the opponent modes of action and searching for the best combination of organisational, physical, human and technological security controls, to meet probable modes of action, i.e. methods of attacking an opponent.

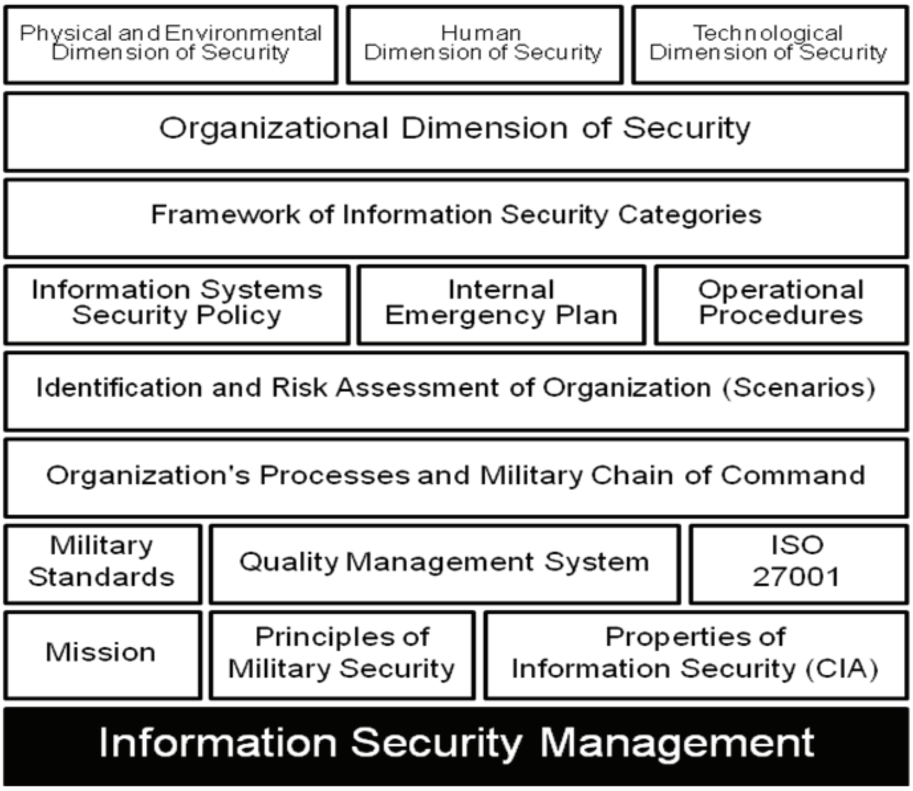

Physical and Environmental Dimension of Security
Human Dimension of Security
Technological Dimension of Security
Organizational Dimension of Security
Framework of Information Security Categories
Information Systems Security Policy
Internal Emergency Plan
Operational Procedures
Identification and Risk Assessment of Organization (Scenarios)
Organization's Processes and Military Chain of Command
Military Standards
Quality Management System
ISO 27001
Mission
Principles of Military Security
Properties of Information Security (CIA)
Information Security Management

Figure 8: Information security framework of the military organisation

In conclusion it appears that information security is oriented by possible attack methods, according to the vectors of a Physical, Human or Information Infrastructure attack. The information security controls are integrated into the main categories of security controls within the Organisational, Physical, Human and Technological Dimensions, with special concern about monitoring, allocation of responsibilities to employees and centred on the principles of war, as referenced in military doctrines, *Command Unit*.

The planning and decision-making about information security is based essentially in infer knowledge obtained from analogies got by experience and training (i.e. lessons learned), i.e. seek to select and retrieve past cases of information security incidents that resembles the situations under study and its subsequent adaptation to ensure the best combination of information security controls. Despite the excellence of the military organisation and its specialists, there is no formal model or methodological approach for planning information security.

4. Conclusions

This study reveals that (1) information security within the military organization is built on the basis of physical and human attack vectors, targeted to the infrastructure that supports the flow of information in the organization (i.e. Information Systems), (2) the information security controls applied in the military organization are included in ISO 27001; (3) planning and selection of information security controls are made by decision makers and information security specialists, aiming to ensure the integrity of digital information. It appears that specialists impose their planning options essentially inferring knowledge from analogies (like following guidelines), or rather, seek to select and retrieve past successful information security cases, i.e. scenarios similar to situations under planning and that may (likely) lead to the selection and implementation of the most efficient information security controls.

The present study does not allow the generalisation of the results obtained to all U/E/O military of the Portuguese Army, since only one single Case Study was carried out. However due to the characteristics of the organisation and its management model, this can be an excellent example (and easy) to replicate in other U/E/O military of the Portuguese Army, thus the study is intended for readers with managerial or leadership functions within a military organisation. In virtue of the inexistence in Portugal and especially within the Portuguese Army, of such studies, it is believed that the sharing of knowledge on this subject contribute to obtain an integrated view of the information security managing process in its multiple dimensions.

Acknowledgements

I would like to thank the Geographic Institute of the Army, for all of its support and availability during the completion of this study. The excellence of its management model and of its employees played a vital role in the investigation.

References

ISO / IEC 27001 (2005). Information technology-Security techniques-Information Security Management Systems-Requirements.

Martins, J., Santos, H., Nunes, P., & Silva, R. (2012a). Information Security Model to Military Organizations in Environment of Information Warfare, Paper presented at the 11th European Conference on Information Warfare and Security, Laval, France.

Martins, J., Santos, H., Nunes, P., & Silva, R. (2012b). Framework de Gestão de Segurança da Informação para Organizações Militares Orientada pelos Principais Vetores de Ataque, Conferência Anual da Associação Portuguesa de Sistemas de Informação, Universidade do Minho, Guimarães, Portugal, 7 de Setembro.

RAD 280-1 (2003). Segurança da Informação Armazenada, Processada ou Transmitida nos Sistemas de Informação e Comunicação do Exército (Reservado), Ministério da Defesa Nacional, Exército Português, Estado-maior do Exército, Portugal.

RAD 280-2 (2005). Orientação Gerais de Segurança para os Sistemas de Informação e Comunicação do Exército (Reservado), Ministério da Defesa Nacional, Exército Português, Estado-maior do Exército, Portugal.

RAD - 95 (2008). Regulamento para a Inspeção no Exército (Reservado), Ministério da Defesa Nacional, Exército Português, Estado-maior do Exército, Portugal.

SANS – CSCCD 4.0 (2013). Twenty Critical Security Controls (Version 4.0), retrieved from the Web March 1, 2013, http://www.sans.org/critical-security-controls/.

SEGMIL 1 (1986). Instruções para a Segurança Militar, Salvaguarda e Defesa de Matérias Classificadas (Reservado), EMGFA, Portugal.

Waltz, E. (1998). Information Warfare: Principles and Operations, Artech House.

Wilhelm, Thomas (2010). Professional Penetration Testing, Syngress.

Mobile Bullying in South Africa - Exploring its Nature, Influencing Factors and Implications

Grant Oosterwyk and Michael Kyobe
University of Cape Town, Cape Town, South Africa
grant@oosterwyk.za.net
michael.kyobe@uct.ac.za

Abstract: Due to an increase in mobile web adoption as well as active mobile users in South Africa, mobile phone bullying is escalating and has become a major concern in schools and communities. Many users of these technologies are often unaware of the severity of this problem and its consequences or simply ignore them. In addition, limited scientific research exists which examines the nature of mobile bullying in South Africa. The lack of awareness of mobile bullying, the negative effect thereof on the injured party, the potential legal liabilities and the lack of research in this field has prompted this study. This study examines the nature of mobile bullying and the key factors that influence mobile bullying. It aims to create awareness of this abuse and highlights the legal implications that could result from engaging in it. A conceptual model is developed that organisations can use to measure the influence of these factors on organisations and groups. Some propositions to test this model empirically are suggested.

Keywords: mobile bullying, influencing factors, legal implications, learners, South Africa, mobile technologies

1. Introduction and problem statement

Bullying is a serious social problem. It occurs globally and can occur during all stages of a person's life, from infancy through adolescence and even in the working environment. This advancement in communication and information technology has increased the potential for bullying via electronic communication devices (Li 2007).Mobile bullying is increasingly becoming a common issue in schools and communities. Often it goes unnoticed and victims are uninformed about which procedures to follow. Most research on bullying has focused on physical bullying and specifically amongst males. Indirect bullying and mobile bullying has not received much coverage. According to Badenhorst (2011:5) mobile bullying has not extensively been examined in South Africa but limited online articles do exist. Criminal law responses to mobile bullying include crimen injuria, assault, criminal defamation and exertion (physical assault). The public is expected to adhere to these laws to avoid potential liabilities (Sizwe 2009; Kyobe2010).

The consequences are also severe in terms of their effect on the injured or bullied party, for instance: depression, withdrawal from society, poor mental health and possible suicides (Cross *et al.* 2012). Cross argues that the danger in the online virtual world can be seen as a mere artefact of what could happen in the real world. The severity of this problem and lack of research in the field has driven the present study. This study examines the factors that lead to mobile bullying. A guide on the legal implications of this abuse is developed. In addition, a conceptual model that schools and other organisations can use to examine these influencing factors is developed. Propositions to test this model in future research are suggested. By understanding these influencing factors it is anticipated that awareness will be created and the invisible problems will be made visible.

The paper begins by reviewing the concept of bullying and distinguishing between different forms of bullying. Furthermore, the nature of the impact of mobile technologies are discussed, the underlying theoretical work and the factors influencing mobile bullying. Based on this analysis, a conceptual model is developed and propositions for testing this model are presented.

2. Traditional bullying, cyber-bullying and mobile bullying

2.1 Bullying: an overview

The most universal definition of bullying articulates that "a person is being bullied when they are exposed, repeatedly and over time, to negative actions on the part of one or more other persons" (Olweus 1987). These actions are purposefully inflicted to cause injury or discomfort to another person (Olweus1991). More specifically, bullying is at the foundation of a hierarchy of violent acts and is thus embedded in the broader picture of the increasing violence in South Africa (Morrison 2007). Furthermore, bullying is argued to be "abusive and is based on an imbalance of power" (Sullivan et al. 2004).

As all bullying is not predominantly physically violent in nature, it is important to distinguish between forms of bullying that is not always evident. Researchers propose that cyber-bullying constitutes a quarter to a third of traditional bullying. An article titled "New Bottle but old wine" written by Li (2005), states that there exists a digital divide between traditional bullying and mobile bullying. Thus the wide adoption of mobile phones amongst the youth is of great concern. While there are many benefits in using this technology, research also associates this usage to addiction, depression, lack of sleep, and cyber-bullying (Zulkefly *et al.* 2009). The following section will introduce and distinguish between the forms of bullying.

2.1.1 Traditional bullying

Traditional bullying has been defined as the misuse of a power acted on behalf of the aggressor to the target (Orpinas & Horne 2006). Furthermore, the aggressor is perceived as physically, socially, or psychologically more powerful than the target. Olweus (1993) argues this power imbalance exploited by the aggressor to control, inflict pain, and commit repetitious attacks over time, makes up the core of what constitutes bullying behaviour. Three forms of traditional bullying exist: social, verbal and physical (Crick *et al.* 2002). Based on literature these three forms of traditional bullying can be summarized as follows:

Table 1: Forms of traditional bullying

Social Bullying	An aggressive behaviour targeting an individual's mental state. This type of bullying also known as indirect bullying provides the ability of the perpetrator to be anonymous. This can include the spreading of false information, gossiping and the social segregation of groups (Smith & Rivers 1994). A major consequence of social bullying is falsely influencing the way in which other people see an individual's social status.
Verbal Bullying	Behaviours such as teasing, insulting or harassing someone face to face are a few examples of verbal bullying (Cole et al. 2006). Verbal bullying requires an individual to use cruel and foul language when communicating.
Physical Bullying	This form of bullying is conducted by physically attacking another person or group which could inflict physical pain. Victims can identify exactly who the perpetrator is with physical bullying, unlike with cyber-bullying.

2.1.2 Cyber-bullying

Many definitions exist on cyber-bullying, however for this study we have adopted a definition by author Bill Belsey, where he defines cyber-bullying as the "*use of information communication technology, known as ICT, to support deliberate, repeated, and hostile behaviour by an individual or group, with an intension to harm others"*. Forms of technologies used to cyber-bully include: text messages, emails, instant messaging, online social networking posts, blogs and mobile phones. According to Hines (2011:19), there are two major similarities between traditional and cyber-bullying. The first similarity is that it takes on an act of aggression by intentionally wanting to harm or hurt an individual. Secondly, both forms of bullying can often be repeated (Kowalski *et al.* 2008). Despite the similarities, anonymity is a huge exception as it allows the perpetrator to further torment their victim more than they normally would if it had been face to face.

According to Willard (2007), different online platforms can be used to cyber-bully, but it is important to note that cyber-bullying can also be grouped in different forms such as flaming, harassment, denigration, impersonation, exclusion and ostracism as seen in Table 2.

Table 2: Forms of cyber-bullying

Flaming	This method of cyber-bullying usually happens when two or more individuals exchange harmful emails (Friedman & Curral 2003). It also entails the sending of messages containing aggressive and unfriendly dialogues.
Denigration	According to Kowalski (2007), denigration occurs when an individual posts mean and hurtful things about someone online in the attempt to hurt the person or spread lies. The victim might not be able to report or delete the post due to the limit of rights to access accounts or the website which is not locally hosted.
Impersonation	This kind of bullying relates to the ability of being anonymous. The perpetrator can post false information online as if he or she was actually the legitimate user.
Exclusion and Ostracism	This occurs when an individual or group of people blocks someone from entering a specific group of friends online. Kowalski (2007) refers to it as "the buddy list".

2.1.3 Mobile bullying

Mobile bullying can be defined as a form of electronic online bullying through email, chat rooms, instant messaging and small text messages using mobile phones (Kowalski *et al.* 2007). From previous literature, it can be concluded that mobile bullying is using a mobile phone to perform the act of cyber-bullying. According to Priscillia (2011: 21), even though bullying is a phenomenon that existed well before the creation of the mobile phone and the World Wide Web, the two mediums have magnified the problem by creating a new avenue through which bullying can be executed. It is important to understand the aspects of mobile technology and how it creates value but also how it can contribute to abuse, damage, and bad practice.

3. Mobile bullying in SA

It must be noted that, although the vast majority of South Africans do not have access to running water and electricity, they do have access to cell phone technology. This combined with more affordable broadband prices, lends weight to the argument that although we lag behind the rest of world, the risk is similar to that evident in the US and Europe. Finn extends this thinking, by hypothesizing that South Africa, despite limitations to penetration, has experienced a rapid uptake of electronic media. This, coupled with the convergence of voice and data services, and the shift to Web 2.0 technologies, has created a fertile breeding ground for cyber violence, multiplying the risk exponentially. In addition, intended as a top-end communication device, the smart phone is now a standard offering with most pre-paid contracts in South Africa. All new smart phones typically include functionality that enable the user to access the internet, capture and display images and video, and can identify their GPS (Global Positioning System) location. Children are now able to communicate in ways that are completely foreign to both parents and educators.

South Africa has an undesirable reputation as one of the most violent countries worldwide (Burton & Mutongwizo 2009). South Africa's technological infrastructure is improving - especially mobile networks and internet service providers (ISP). Consequently, Burton argues that with the rise of newer technologies, new forms of violence emerge especially amongst those who are prone to get absorbed by it – our youth. A study conducted by the Centre for Justice and Crime Prevention indicate that nearly one half of adolescents has been victims of mobile bullying. The study reveals that 31% of the participants that has been interviewed experience some sort of mobile bullying whilst on schools premises, whereas 42.9% experienced it outside of school.

4. Mobile technologies values and limitations

The value creation role of mobile technologies has been presented very well in Pousttchi, Weizmann and Turowski's (2003:414) concept of mobile added values (MAV) theory. These authors claim that added value in the use of mobile phone technologies can be obtained through its ubiquity; context-sensitivity; identifying function and command and control functions. *Ubiquity* is the ability to send and receive data anytime and anywhere and thus, eliminate any restrictions whatsoever. *Context-sensitivity:* this attribute refers to the delivery of customized products or services, such as GPS, weather reports or local shops (Derballa & Pousttchi 2004). The attribute, *Identifying Function*, refers to the ability of an individual to authenticate with the specific mobile device. Further security measures can also be enabled on the actual device for better or further means of authentication. *Command and control functions* mean that mobile devices have the ability to make use of remote control applications that allows certain commands to control and manipulate network peripherals and devices.

These value creating attributes of mobile technology have however also facilitated its negative aspects. *Mobile technologies have four main features that facilitate mobile bullying*.

4.1 Mobile bullying via electronic mail (email)

Email messages contain various types of content such as typewritten content or files of text, images, audio, or video. Email can be used in various ways to inflict harm on an individual: sending harassing or threatening messages; attaching viruses to emails; or including personal information about a victim and sending it to many people.

4.2 Mobile bullying via chat rooms

Mobile chat rooms are web sites which allow for real-time communication between two or more users. Users enter a chat room under a "username", a name which they display to represent themselves, and can converse about any topic. Chat rooms can be especially dangerous for children and adolescents due to the added anonymity of usernames; mobile roaming capability. Bullies and sexual predators can therefore pose as a trustworthy friend and confidant (BBC 2002).

4.3 Mobile bullying via instant messaging

Instant messaging is a type of communication service that enables people to create a private chat room with another individual to communicate in real time over the Internet. This is analogous to a telephone conversation but uses text-based communication instead. In a survey of Canadian youth, 14% of users reported being threatened while using instant messaging.

4.4 Mobile bullying via small text messages (SMS, EMS, and MMS)

Commonly called "text messages", this service is generally provided by mobile phone distributors and can take the form of Short Message Service (SMS), Enhanced Messaging Service (EMS), and Multimedia Messaging Service (MMS) (Webopedia 2003). Bullying through text messaging appears to be a method of primary concern among researchers and adolescents alike.

5. The influential factors of mobile bullying

While technology can fuel bullying as discussed above, there are many other factors also contributing to mobile bullying which relate to behaviour, attitude, as well as the social environment of the adolescent. The consequences of a learner having a mobile phone on the school premises may lead to disruptions in class, dishonesty in tests as well as bullying other learners (Obringer & Coffey 2007:41). A feature of mobile phones is that it has video and photo capabilities which bring forth privacy concerns where learners can be photographed anywhere and shared (Obringer & Coffey 2007). This act may lead to the victim feeling isolated and excluded from society and too embarrassed to face fellow students in future (Srivastava 2005).

According to Li (2007), a difference between traditional and mobile bullying is that traditional bullying occurs at a specific time and place, while cyber victims may continue to receive SMSs, emails or see comments posted online. However, according to Zhang (2010:4), there is limited empirical research that shows whether the global nature of the internet has an influence on individuals to engage in mobile bullying activities.

A qualitative study was conducted on mobile bullying in 2009 targeting the youth and analysing the prevalence of mobile bullying and eagerness of the youth to inform their parents about this topic (Mishna, Saini & Solomon 2009). The general consensus amongst the youth was that they would not approach their parents, due to the fear of losing their rights to their mobile phones. Mishna *et al.* suggests that many bullying studies fail to address the thinking processes and reasoning which provoke adolescents to act the way they do. Recent studies by Thornberg (2007; 2010) have looked at moral reasoning and the "bystander effect" offline amongst secondary school learners.

Evidence of specific thinking processes like trivialization and dissociation that can lead to "bystanding" behaviour was found. Trivialization refers to people who are making the impact of mobile bullying seem insignificant. Dissociation refers to the act of bystanders not wanting to get involved and prefer being separated. Thornberg (2010) cites Kohlberg's theory of moral development, as well as the importance of emotions, attitude and social context on the development of morality among children. Kohlberg's (1984) theory argues that the adolescent thinking process develops at different stages and times, and this usually continues all the way through adolescence until adulthood is reached. The thinking process may affect the way the youth perceive certain socio-economic factors such as social networking, crime and education (Thornberg 2010).

In addition, the bio-ecological approach, by Bronfenbrenner (2005), provides an understanding of the development of individual relationships and how they are integrated in a social context. His model consists of four key concepts namely: person factors, process factors, contexts and time. Lerner (2005) states that this model is used to conceptualise the development of integration and to design research. Person factors involve the characteristics of how an individual would behave, for example the personality of the bully or victim. Secondly, process factors refer to the interaction between individuals, thus between the adolescent and the context, which refers to families, peers, teachers, schools and communities. Furthermore, Lerner (2005) suggests that the aspect of time is relevant to the concept of the theory because the maturation of the adolescents changes over time. The bio-ecological theory therefore helps to understand how adolescents relate to their social environments (Kruger 2011).

Rendering to Ajzen (1991), the theory of planned behaviour (TPB) is an empirically validated theoretical framework which can be used to investigate the variation in mobile phone usage. Furthermore, this theoretical framework verifies that behaviour results from a rational, systematic evaluation of salient information. According to Ajzen, behaviour is influenced by three constructs: *attitudes* (based on the individual's overall evaluations of the behaviour), *subjective norms* (peer pressure) and perceived behavioural control (the level of control an individual believes they have over behavioural performance). By using the TPB framework, one major advantage is the capability to link differences between high and low level behavioural performers allowing an understanding of fundamental behavioural influences.

Furthermore, a person's ability to efficiently use technology empowers them to potentially become a mobile bully. According to Ybarra & Mitchell (2004), people who claim that they are experts in internet knowledge environments, were found to be more aggressive than those who claim that they are not experts. Additionally, people who spend more time online on a weekly basis were 73% more likely to tend towards mobile bullying behaviours. Moreover, there is a possibility that online aggressive behaviour by adolescents could be as a result of spending an extensive amount of time on the internet.

Finn (2004) states that electronic communication tools decrease the chances of cyber-bullies showing any remorse towards the victim. Additionally, communication or messages in the virtual world can be misinterpreted and this can also promote a false sense of intimacy, which may lead to greater risk-taking and possible incidents of cyber-bullying. While Finn specifically looked at cyber-bullying, these factors apply to mobile bullying as well.

Moreover, according to Felson and Clarke's (1998) Routine Activity Theory (RAT), a reasonable explanation for victimisation could be provided. RAT states that offenders use anonymity as a weapon and is motivated to behave rebelliously. A study conducted by Christopherson (2006:3040) looked at the social impact of the internet on groups and individuals, with a focus point on the ability of being anonymous. The participants were aged between 13 and 19, using a mobile chat room. It was found that mobile chat rooms made it easier for the participants to be anonymous. Adolescents responded that being anonymous gave them the confidence to express themselves when chatting to another person and saying things they would not say in person.

In South African schools, a code of conduct is drawn up by the school governing body, after consulting educators, learners and parents. Any form of bullying or victimisation appears to be in conflict with this structure, which is implemented to enforce a disciplined school environment that aids quality learning. The phenomenon of mobile bullying is becoming an increasing problem which can present many problems for South African schools and educators. Many studies in education have been conducted in South Africa around traditional bullying and how to protect learners from it. However, with the new mobile bullying sensation, many researchers fail to study whether educators are adequately informed about it and how they will deal with it should any legal consequences arise (Moodley 2011). Although many people are aware of the act of mobile bullying through media reports, the lack of mobile bullying policies in schools however, is a great concern (Kowalski, Limber & Agatston 2008).

6. Mobile bullying and the South African Law

As mentioned before, there may not be a concrete law that regulates mobile bullying, but responses to civil law may accommodate for this. According to Prinsloo (2005:8), with reference to the South African Schools Act on Human Dignity of 1996, the existing legislations governing the use of electronic devices and new school rules strongly prohibit all forms of bullying.

In an effort to create awareness of the law and their implications, the authors have identified and developed a guide to assist researchers, educators and learners to understand good control practices recommended by codes of practice. Furthermore, the aim of the guide is to inform authorities about existing legislation which can be used to draw up policies that specifically govern mobile bullying in the school context. Schools and supporting organisations, such as law enforcement and government need to ensure accountability, transparency and measurability if they are to implement mobile bullying policies effectively.

Table 3: A guide to create awareness of mobile bullying legislations and implications

Regulations	Consequences
Higher Education Act	**The South African Schools Act 84(1996) upholds the rights of learners, teachers and parents and shows that these parties accept their responsibility towards the school as well as the governance thereof. Any form of bullying, is not tolerated (Gov. Gazette 1997).**
Electronic Communications & Transactions (ECT)	**Chapter 8 of this act governs legal criminal offenses relating to unauthorized access to data (e.g., through hacking), interception of data (e.g., tapping into data flows or denial of service attacks), interference with data (e.g., viruses) and computer related extortions, fraud and forgery. Perpetrators can be fined and/or imprisoned (Michalson & Hughes 2005).**
Common Law	**The common law provides principles such as the rule of natural justice as well as in loco parentis which play an important role in dealing with mobile bullying in schools. The rule of natural justice aims to ensure fairness and justice in all disciplinary actions. With the principle in loco parentis, it is the educators' duty to accept the role of the parent and see to it that the adolescent is protected from harm (Sizwe 2009).**
Children's Act 38 of 2005	**The Children's Act encourages all role players in a school environment to respect, promote and fulfil the rights of children as set out in the Bill of Rights. It is the responsibility of the school to ensure the child's best interest is considered (Government Gazette 2006).**
Regulations	Consequences
Protection of Personal Information Bill	**In the Promotion of Access to Information Act 2 (2000), the procedure that should be adopted to request access to the personal information or records of any individual, is detailed. In order for an individual to gain access to this information a written permission should be obtained from the consumer for the collection, collation, processing or disclosure of any personal information of the consumer (Michalson 2009).**
Civil Law	**In terms of Section 384 of the Criminal Procedure Act, the civil law can follow two distinct courses of action that can be employed. Firstly, an interdict can be brought before the High Court, restraining the perpetrator from continuing with the behaviour, and subsequently the applicant can sue for damages and defamation of character. Secondly, it may be the case that a perpetrator has provoked a breach of peace, by threatening a victim with injury to their person or property. The perpetrator may be sued (Badenhorst 2011).**
Criminal Law Act 2007	**Section 19 of the Criminal Law (Sexual Offences and Related Matters) Amendment Act 2007, provides that any person exposing or displaying, or causing exposure of child pornography to a child violates this act. A conviction will also result in the child's name being registered as a sex offender (Badenhorst 2011).**

7. Gaps identified in the literature

The literature review highlighted mobile bullying and the impact of mobile technologies amongst adolescents. The literature also looked at different forms of mobile bullying in general. One of the gaps that were identified in the literature was that although many incidents of mobile bullying exist - both locally and internationally, limited research exists exploring mobile bullying and the major influential factors thereof. Mobile bullying has not been researched amongst adolescents in South Africa even though mobile phones are the most common technological artefact used to communicate.

Currently there is limited information available on mobile bullying pertaining to the nature and extent of the problem within the school context. It was necessary to come up with a guide which can be used as a tool to develop policies specifically pertaining to mobile bullying as illustrated in Table 3.

8. The conceptual model

Figure 1 below represents the proposed conceptual model for examining factors influencing mobile bullying in South Africa. The model presents the constructs identified in literature which the authors consider to be key influences on the degree of mobile bullying. The theoretical work on which this model is premised are summarized as well. In the conceptual model, the dependent variable (degree of mobile bullying) consists of all the constructs measuring mobile bullying. The independent variable for this conceptual model consists of all the influential factors of mobile bullying. Firstly, anonymity influences the extent to which mobile bullying occurs because it empowers the perpetrator to inflict more harm than they would in a real life situation and this makes it difficult to trace the perpetrator (Kruger 2011). There is a lack of awareness regarding the risks and implications around mobile bullying in South Africa and this influences the way in which mobile bullying is perceived and understood. Furthermore, the existence of the law regulating mobile bullying and the knowledge of this law and the liabilities for non-compliance with it, will influence the extent to which mobile bullying occurs.

According to Thornberg (2010) the attitude and thinking processes of adolescents may affect the way adolescents perceive certain socio-economic factors, such as social networking, crime and education and influence their mobile bullying behaviour. An individual's ability to effectively use technology may impact their mobile bullying behaviour because individual's who spend more time online tend to be more aggressive. In the virtual world, messages can easily be misinterpreted and this leads to greater risk-taking, affecting the degree of mobile bullying. Lastly, the accessibility of mobile applications allows for anytime and easy access for potential mobile bullies and this influences the extent to which mobile bullying can occur.

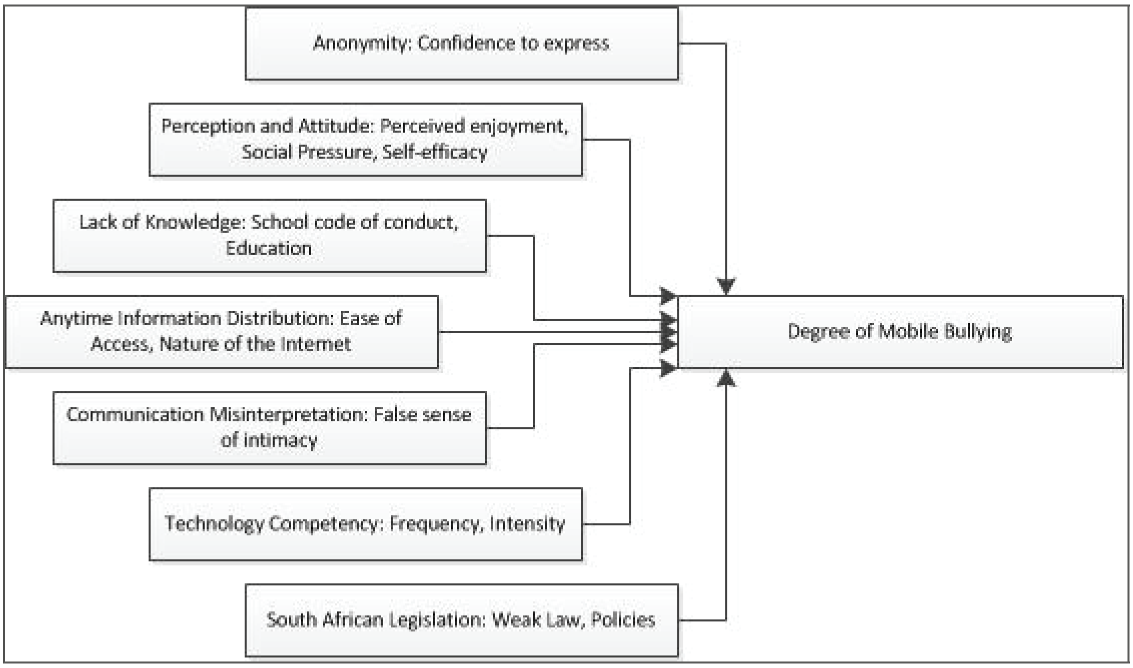

Anonymity: Confidence to express
Perception and Attitude: Perceived enjoyment, Social Pressure, Self-efficacy
Lack of Knowledge: School code of conduct, Education
Anytime Information Distribution: Ease of Access, Nature of the Internet
Communication Misinterpretation: False sense of intimacy
Technology Competency: Frequency, Intensity
South African Legislation: Weak Law, Policies
Degree of Mobile Bullying

Figure 1: Conceptual model

The following propositions have been formulated and can be used in future studies to test this model empirically.

> *Proposition 1: An individual's technology usage competence affects their mobile bullying behaviour. Existing literature supports the view that the higher the magnitude of expertise in using online communication tools the higher the likelihood of mobile bullying taking place. An example of measuring this relationship can be found in a study by Anandaraja et al. (2000), which analyses the usage of technology by using two indicators to determine if it has an impact on the probability of mobile bullying behaviour. The first indicator 'Intensity of use' measures the amount of time an individual spends using an electronic communication device. Secondly, 'Frequency of use' entails measuring the number of times an electronic communication device is used by an individual.*
>
> *Proposition 2: An individual's perception and attitude towards the internet affects their mobile bullying behaviour. An example of measuring this relationship can be found in a study by Cheung & Huang (2005), which explores three indicators on perceptions and attitudes. The first indicator 'Perceived enjoyment/fun', is represented by measuring the perception of how much an individual enjoys using electronic communication devices. Secondly, 'social pressure' is a*

representation of an individual's perception of peer pressure on themselves to conduct the act of mobile bullying (Cheung & Huang 2005). Lastly, 'Technology self-efficacy' is derived from the 'Internet self-efficacy' indicator which is defined by Anandaraja et al. (2000) as "an individual's beliefs about his/her ability to competently use the technology".

9. Conclusion

The objective of this paper was to explore the nature of mobile bullying in South Africa, to identify the key factors which influence it as well as the implications thereof. The key factors identified are: the power of being anonymous; the lack of knowledge amongst people and the judicial system of South Africa and the regulation regarding mobile bullying. Furthermore, the perception and attitude of adolescents towards mobile technology affects their mobile bullying behaviour. Mobile technology has provided a platform to access information at any time which improves the competency of learners to use technology to bully. In addition our review reveals gaps in research on mobile bullying; the prevalence of the problem in schools; and the awareness of legislation and implications. By identifying these gaps and the influencing factors, this paper has managed to better our understanding of what mobile bullying is and how it impacts society. To create awareness of this serious problem, we propose a guide that can be used to understand the South African legislation and how it can be utilised to protect individuals against mobile bullying. This study therefore proposes a conceptual model based on the premise that since there is an increase in the adoption of mobile technologies, emphasis needs to be placed on the process of identifying the influential factors of mobile bullying and therefore aim to minimise mobile bullying occurrences. This model can be empirically tested through propositions which measure the relationship between the influential factors and its impact on the degree of mobile bullying.

References

Ajzen, I. (1991)The Theory of Planned Behavior: Organizational Behavior and Human Decision Processes, Vol50,pp 179-211.

Badenhorst, C. (2011)Legal responses to cyberbullying and sexting in South Africa, Vol10, Online, Centre for Justice and Crime Prevention,http://www.childlinesa.org.za/index2.php?option=com_docman&task=doc_view&gid=221&Itemid=64

Bronfenbrenner, U. (2005) Making human beings human.BioecologicalPerspectives on human development. London: SAGE.

Burton, P. and Mutongwizo, T. (2009)Inescapable violence: Cyber bullying and electronic violence against young people in South Africa, Online, Vol8, Online, Centre for Justice and Crime Prevention,http://www.cjcp.org.za/admin/uploads/Issue%20Paper%208%20-%20Inescapable%20Violence%20-%20Cyber%20aggression.pdf

Cheung, Waiman; Huang, Wayne, (2005)"Proposing a framework to assess Internet usage in university education: an empirical investigation from a student's perspective", British Journal of Educational Technology, Vol. 36, No. 2,pp 237-253.

Christopherson, K.M. (2006)"The positive and negative implications of anonymity in Internet social interactions.Computers in Human Behavior", Vol. 23, pp3038-3056.

Cole, J. C. M., Cornell, D. G., andSheras, P. (2006)"Identification of school bulliesby survey methods",Professional School Counseling, Vol. 9, pp305-313.

Coyne, S. M., Archer, J., andEslea, M. (2006)''We're Not Friends Anymore! Unless.The frequency and harmfulness of indirect, relational, and social aggression", Vol. 32, pp294–307.

Crick, N. R., Grotpeter, J. K., andBigbee, M. A. (2002)"Relationally and physically aggressive children's intent attributions and feelings of distress for relational and instrumental peer provocations",Child Development, Vol73, pp1134–1142.

Cross, D. (2012)"CyberbullyingVersus Face-to-Face Bullying – A Theoretical and Conceptual Review".

Felson, M., and Clarke R.V. (1998)"Opportunity Makes the Thief: Practical theory for crime prevention",Police Research Series Paper 98, London: Home Office

Finn, J. (2004) "A Survey of Online Harassment at a University Campus",Journal of Interpersonal Violence,Vol19, No. 4,pp 468-483.

Friedman, R. A., andCurrall, S. C. (2003) "Conflict escalation: Dispute exacerbating elements of e-mail communication conflict",HumanRelations, Vol56, No. 11, pp1325–1347.

Government Gazette (1997) Republic of South Africa. Vol. 390,pp 22-24,Online, http://www.info.gov.za/view/DownloadFileAction?id=70759

Government Gazette (2006) Republic of South Africa. Vol. 492, pp 84-85, Online, http://www.info.gov.za/view/DownloadFileAction?id=67892

Hines, H.N. (2011)"Traditional bullying and cyber-bullying: are the impacts on self-concept the same?",Online,http://libres.uncg.edu/ir/wcu/f/Hines2011.pdf

Kohlberg, L. (1984)"Essays on Moral Development".San Francisco: Harper and Row.

Kowalski, R. (2007)“Electronic Bullying Among Middle School Students”,Journal of Adolescent Health, Vol. 41,pp22–30, Online, http://www.isb.sccoe.org/depts/csh/docs/Mar2011/Elec.Bullying.Middle.School.pdf

Kowalski, R. M., Limber, S. P., andAgatston, P. W. (2008)“Cyber bullying: Bullying in the digital age”, Oxford, UK: Blackwell.

Kruger, M.M. (2011)“Bullying in Secondary schools: teachers’ perspectives and experiences”,Online, http://scholar.sun.ac.za/handle/10019.1/17929

Kyobe, M. (2010)“Towards a framework to guide compliance with IS Security policies and Regulations in a university”,Online, http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=5588651&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxpls%2Fabs_all.jsp%3Farnumber%3D5588651

Kyobe, M.E. andShongwe, M.M., (2011)“Investigating the extent to which mobile phones reduce Knowledge Transfer barriers in Student Project Teams”,SA Journal of Information Management, Vol.13, No. 1, Art. # 424

Lerner, R.M.(2005)“Career contributions of the consummate developmental scientist”, Making human beings human: Bioecological perspectives on human development. London: SAGE Publications, pp11–26.

Li, Q. (2005)“New bottle but old wine: A research of cyber bullying in schools”Online, http://people.ucalgary.ca/~qinli/publication/cyber_chb2005.pdf

Li, Q. (2007)“Bullying in the new playground: Research into cyber bullying and cyber Victimization”, *Australasian Journal of Educational Technology*, Vol. 23, No.4, pp 435-454,Online, http://www.ascilite.org.au/ajet/ajet23/li.html

Michalson, L. (2009)“Protection of Personal Information Bill - the implications for you”,Online, http://www.michalsons.com/protection-of-personal-information-bill-the-implications-for-you/3041

Michalson, L and Hughes. B (2005)“Guide to the ECT Act”, MichalsonsAttorneys,Online, http://www.michalson.com

Mishna, F., Saini, M. and Solomon, S. (2009)“Ongoing and online: Children and youth’s perceptions of cyber bullying”,Children and Youth Services Review, Vol. 31, pp1222–1228.

Mollo, N.T. (2009)“A legal perspective on the establishment of anti-bullying policies in public schools”, Online, http://upetd.up.ac.za/thesis/available/etd-07212009-135205/unrestricted/dissertation.pdf

Moodley. (2011)Unknown, Online, http://www.inter-disciplinary.net/wp-content/uploads/2011/10/moodleybpaper.pdf

Morrison, B. (2007) “Restoring safe school communities: A whole school response to bullying, violence and alienation”,Sydney: The Federation Press.

Obringer, S.J. and Coffey, K. (2007) “Cell Phones in American High Schools: A National survey, Journal of Technology Studies”, Vol. 33, No.1/2,Online, http://www.proquest.com/

Olweus, D. (1987) “Schoolyard bullying: Grounds for intervention”,School Safety, Vol. 6, pp4-11.

Olweus, D. (1991)“Bully / victim problems among schoolchildren: Basic facts and effects of a school based intervention program”. In D. J. Pepler& K. H. Rubin (Eds.), The development and treatment of childhood aggression. Hillsdale: Erlbaum, pp 411-448.

Olweus, D.(1993) “Bullying at school. What we know and what we can do”, USA: Blackwell Publishers.

Orpinas, P.,and Horne, A. M. (2006) “Bullying prevention: Creating a positive school climate and developing social competence*”,Washington, D.C.: American Psychological Association*.

Pousttchi, K., Turowski, K. and Weizmann, M. (2003) “Added value-based approach to analyze electronic commerce and mobile commerce business models”, in R.A.E. Andrade, J.M. Gómez, C.

Prinsloo, I. J. (2005)“How safe are South African schools?”,South African Journal of Education, Vol. 25, No. 1, pp5-10.

Priscillia and Sinha, Akshay. (2011)“What’s your ASLR’ to ‘Do You Wanna Go Private?”,UNICEF,Online, http://www.unicef.org/southafrica/SAF_resources_MXitstudy.pdf

Rivers, I., and Smith, P. (1994)“Types of bullying behavior and their correlates: Aggressive Behavior”,Vol. 20, No.5, pp359-368.

Serra, S. Venter, H.S. (2011) “Mobile cyber-bullying and the children of South Africa: A proposal for a pre-emptive approach to risk mitigation by employing digital forensic readiness”, 2011 Information Security for South Africa (ISSA): Proceedings of the ISSA2011 Conference; Johannesburg, South Africa,ISBN 978-1-4577-1483-2.

Sizwe, S. (2009) “Cyber Crime in South Africa”,Online, http://www.hg.org/article.asp?id=5351

Srivastava, L. (2005)“Mobile phones and the evolution of social behaviour”, Behaviour & Information Technology,Vol. 24, No. 2,Online, http://www.ebscohost.com/

Sullivan K, Cleary M and Sullivan G.(2004)“Bullying in secondary schools. What it looks like and how to manage it”. London: Paul Chapman Publishing.

Thornberg, R. (2007)“A classmate in distress: schoolchildren as bystanders and their reasons for how they act”. Social Psychology of Education, Vol.10, pp5–28.

Turner, M., Love, S. and Howell, M. (2007)“Understanding emotions experienced using a mobile phone in public: The social usability of mobile (cellular) telephones, Telematics and Informatics”,Online, http://www.sciencedirect.com/

Webopedia.(2003) Online, http://www.webopedia.com/.

Willard, N. (2007)“Educator's Guide to Cyberbullying: Addressing the Harm Caused by Online Social Cruelty”.

Ybarra, M.L., and Mitchell, K.J. (2004)“Linkages between depressive symptomatology and Internet harassment among young regular Internet users”. *Cyberpsychology&Behavior*, Vol. 7,pp247-257.

Zhang, A.T. (2010)“Key influences of cyberbullying for universitystudents”,Online, http://www.pacis-net.org/file/2010/S01-01.pdf

Zulkefly, S. (2009)“Mobile Phone use Amongst Students in a University inMalaysia: Its Correlates and Relationship to Psychological Health”,Online,http://www.eurojournals.com/ejsr_37_2_03.pdf

The Strategic Communication of Russia, China and the USA in Latin America: War or Peace?

Evgeny Pashentsev
Lomonosov Moscow State University, Moscow, Russia
icspsc@mail.ru

Abstract: The new strategic goals and priorities, together with the growing reorientation of many Latin American countries in their political, military, business and cultural ties has a communication dimension. It is doubtful whether modern Russian weapons are much more competitive on Latin American markets than in the times of prosperity of the Soviet Union, or Chinese IT products are much better than their US equivalents. In the myriad of socio-economic, geopolitical factors of the mentioned changes the efficiency of the strategic communication of three countries in the region is one of the key factors in helping to define the region. In the present paper using the methods of comparative and content analysis we try to accentuate the differences and highlight the peculiarities of strategic communication of the USA, Russia and China in Latin America in the 21st century. How do these powers formulate their key objectives and strategies in the region, develop awareness of their national brands in the local audiences, ensure that those brands fulfill their promises and expectations? What are their main advantages and weaknesses in realization of strategic communication in the region? The efficiency of strategic communications as means of ensuring collaboration is negligible in case of strategic interests and goals mismatching drastically. In this case, strategic communication inevitably becomes a tool for information warfare. There should be no illusions cherished of the contrary situation. That is why compromises are essential, as well as searching for ways to combine interests. Strategic communication can be very fruitful for the creation of a climate that is useful for such a search, but it can aggravate the situation as well. To a certain extent, strategic communication itself is an important (and partly autonomous) factor of rapprochement or estrangement of the parties, and it is vital to ensure that it serves the accomplishment of the first task. Such a program of joint optimization of strategic communication in the region is completely impossible to implement in the circumstances of growing tension between Russia and China on the one hand, and the USA on the other. There is a need for a serious, revolutionary at its core economic, technological, social and political shifts in the three countries with consideration to their national peculiarities. This is the common interest in order to overcome the threat of a new world war and to provide conditions for a democratic and progressive development of mankind. The strategic communication of the countries will have to deal with ongoing theoretical and practical problems of further alignment of "words and deeds" in the real policy. It should make strategic communication overcome its largely propagandistic level and become not only an element of communication support of foreign policy, but even more of an equal part of foreign policy management, with the appropriate set of tools of communication management, i.e. of professional people management via communications.

Keywords: China, the USA, Russia, Latin America, strategic communication

1. Introduction

Most generally, strategic communication is the state's projection of certain strategic values, interests and goals into the conscience of domestic and foreign audiences. It is done by means of adequate synchronization of multifaceted activities in all the domains of social life with professional communication support. It is clear that such synchronization takes place in all three countries analyzed: Russia, China and the USA; reflecting the dynamics of the unique national symbiosis of the old and the new, of the local and the adopted aspects of administration forms and methods of influencing public consciousness.

In Russia, the term "strategic communication" is not used in official documents, unlike the USA. The latter have elevated the concept of strategic communication to a state policy at the highest level (White House, 2010). In spite of an abundance of state institutions, documents (Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics, 2004; Deputy Secretary of Defense, 2006; Strategic Communication and Public Diplomacy Policy Coordinating Committee (PCC), 2007; U.S. Senate, 2008; Department of Defense, 2010; Office of the Under Secretary of State for Public Diplomacy and Public Affairs, 2010 etc.) and scientific research (Paul, Ch., 2011; Corman S. R., Trethewey, A., Goodall, H. L., ed. 2008; Fisher, A. & Lucas, S., ed. 2011; Murphy, D. M., 2008; Stovichek B. E., 2007; Patterson S.J., Radtke J.M., 2009 etc.), it is still in its nascent stages of development there. Russia tends to use the term "state informational policy", which does not exclude, however, the need for strategic communication, as it is implicit in the term.

In 2011 the first Chinese book on strategic communication, entitled the *Essentials of Strategic Communication*, was published. The book was jointly published by two presses in Beijing: Press of Chinese Academy of Governance, and the Central Compilation and Translation Press. The book was highly evaluated by General Liu

Yongzhi, General Zhang Xusan and Chief Political Commentator of Phonex TV Mr. Ruan Cishan. They three respectively wrote prefaces to it. The term “strategic communication” becomes more and more popular in China on the official sites of different state structures.

Strategic communication in the realm of foreign policy combines synchronization of affecting an allied state, and non-state actors through friendly “deeds, words and images” and through a wide range of communications within the framework of information warfare addressing foes and enemies. However, separating one from another is extremely difficult for the following reasons: It is not easy to forge alliances in the contemporary international field, due to the conflicting interests of governing elites, who are often quite controversial. Such a phenomenon was evidence when the U.S. launched a large scale media campaign against their ally France due to its outspoken disapproval at the UN in 2003 with regard to America’s decision to invade Iraq.

The modern realities and interpretations of conflict deliberately blur boundaries between war and peace, between military and civilian systems and spaces, and between information and influence or manipulation (Armistead, E., ed., Arpagian N., 2010; 2004; Brunner E. M., Cavelty M. D., 2009). The extension of warfare into public consciousness has taken place before. However, today we can witness the evidence of a professional form of warfare being waged in the public consciousness using complex methods of communicational influence on a global level – pulling more people into the virtual world of the internet and social networking. Accordingly, it stimulates the further development and application of new kinds of informational impact and information weapons.

Next we try to show in what way synchronized deeds, words and images are being exploited in the current politics of Russia, China and the USA in Latin America. We shall start from Russia, which has made a significant breakthrough to Latin America over the last ten years.

2. Strategic communication of Russia in Latin America is on the rise

After the breakup of the Soviet Union the system of foreign policy propaganda fell into a badly coordinated, controversial, low-powered, inefficient and far from long-term Russian interests foreign policy communications of separate bodies without any strategic context. The new elite was deep in corruption, with Western backup, tried to hold on to power in the country with a declining economy, social sphere sliding into poverty, intensifying property stratification, science and education degradation, but with the richest natural resources and a powerful nuclear potential.

The situations has started to improve with the presidency of Vladimir V. Putin first, and then Dmitry A. Medvedev and now once again Vladimir Putin. There is a twist to more independent international course, although many strategic solutions of national development are delayed and the citizens’ well-being is connected to the oil prices and other resources which can’t but transmit negative images of Russia to Latin America.

In 2008 RIA Novosti opened its bureau in Cuba. It was the step towards a more complete presence of the leading Russian agency in the region. In 2010 further strengthening of the informational presence of Russia in the region took place as Spanish inserts of Rossijskaya Gazeta appeared in the biggest Latin American periodicals – “Clarín”, “Jornal do Brasil”, “Russia Today” also started broadcasting in Spanish.

“Mutually beneficial collaboration”, “equality”, “multipolar world” – the key messages of the modern Latin American Strategy of Russia, well accepted on the level of local elites and the broad strata of the people.

We can hardly ever argue the fact that the image of Russia is better in Cuba and Venezuela than in the former Baltic republics; and the image of the USA is better in the minds of the Latvians, the Estonians and the Lithuanians than in the Cubans or Venezuelans. The nature of foreign policy is not the main reason here. The matter is in the difficult history of small peoples neighboring world powers. No world power is persistent in respectable relations with small neighbors. Probably vice versa. Old offences are healing slowly and the strategic communication of Russia and the USA should not be directed at the warming up, at the same time this situation should not be considered from the point of view of the counterpart’s craft. The concrete analysis and concrete evaluation as well as mutual exchange of opinions and specialists are necessary.

Russia quickly and substantially returning to Latin America could go by:

- The evident faults in the US foreign policy during the last two decades in the region and in the world; evident unwillingness of Latin America to remain a "backyard" of the USA.
- The evident desire to increase the price of its services for those parts of Latin American elite which is interested in the USA; the desire to organize a fast modernization for those who don't believe in the US support.
- The absence of geographical contact and problematic past common with Latin America.
- As for cultural and interpersonal contacts, Russian culture is closer to multi-active Latin American culture than to linear - active or reactive ones.

Russia with its slowly recovering economic power and alternative resources of supply and production of arms (together with the EU, China and other countries) symbolizes an alternative for Latin Americans. At the same time, Russia can go by recollections about past Soviet power (historical heritage) but without fear of communism for the local elites. Neither those who like it, nor those who hate it will be able to find it in modern Russia.

The instability of strategic communication of modern Russia is mainly determined by the instability of its economic resource-based situation, scientific and technical inferiority (more serious than in the USSR) and hot social problems. But if the current model of the socio-economic and political development has changed radically in positive direction it could open new strategic perspectives and logically renovate strategic communication.

3. Competitive advantages of strategic communication of China in Latin America

The strategic communication of China, the second power analyzed in this paper, also plays a significant part in reinforcing its positions in the Latin American region. Regarding the changes in the presence of the three countries in Latin America after the end of the cold war, the USA have been demonstrating a gradual and in some parts an invisible decrease of the economical, military and political presence in the region. In the past decade this tendency acquired an obvious and in some countries, a dramatic character. Russia, who seemed to be a ridiculous political dwarf here in the 1990s, has managed not only to return to the region but to significantly expand its influence when compared with the Soviet period. China is the only one among the three powers who has been demonstrating during the two decades a steady expansion of its economical, military, cultural and informational presence in the region.

"In the present international climate, China considers it necessary to develop strategic partnership with the Latin American countries, "President of the People's Republic of China Hu Jintao said during his last official visit to Peru in 2008. "China is the largest developing country and Latin America is the largest developing continent. Our closer communication … is the demand of the present and the requirement for the development of the both sides," explained the president (Novaya Politika, 2008).

These words provide a key message to Latin America within the strategic communication of China. Another China's key message is formulated in the slogan of the so-called "peaceful rise". Through this message China presents its own system to the countries of Africa and Latin America as a model for the struggle against poverty.

A characteristic feature of the Chinese strategic communication is their well-considered and long- term nature. It is aimed at neutralizing the prospective threats of being rejected by the target audiences abroad for the reasons related to the rapid economic growth of China, as well as the growth of its military potential. The Chinese leaders in this or that way in their speeches stress that in terms of world development, revitalization of a country in the era of economic globalization can be well achieved through equal and orderly international competition and mutually beneficial cooperation. It's no longer necessary or possible to take the old path of challenging either the existing international order or other countries. The rise and fall of some big powers in the world tells China: Expansionism leads to nowhere; arms race leads to nowhere; seeking world domination leads to nowhere. As Deng Xiaoping once said, if one day China tries to seek hegemony in the world, people of the world should expose, oppose and overthrow it. The international community can hold us to account. China

Will never seek hegemony when it becomes more developed – this is the third key message to the countries of the world and Latin American countries can especially appreciate it. (Liu Huanxing 2011).

What competitive advantages can China count on in its strategic communication on the Latin American direction of its policy? Here, we will refrain from repeating the positions that have proved to be advantageous for Russia. They are mostly valid for China as well. However, China has its own extra advantages as well as weak points.

- The rate of economical development of China is incomparably higher than that of the USA and Russia and its current GDP, though two times smaller than the North American one, is three times larger than the Russian one. The real prospects of its further growth in the XXI century undoubtedly exceed both the Russian and the US potential.
- The population of China is four times larger than that of the USA and nine times larger than the Russian one. At the same time, it is more ethnically homogeneous than the populations of the two powers compared to China.
- China has the image of a developing country and a country whose status used to be close to colonial possession once in the early XX century. This makes the perception of China closer to a former colony and then for a long time politically and economically dependent Latin American states.
- The Chinese diasporas in the Latin American region is growing much faster than the North American and the Russian ones and is penetrating the political, economical and military structures of the Latin American countries on different social levels.
- A high political flexibility of China: the representatives of the Chinese diasporas are in the Latin American political establishment, the latter being not afraid (as compared to the USSR) of the "red threat" embodied by this country. At the same time, China is good at building perfect relations with most communist and left socialistic organizations of the continent.
- China possesses a higher dynamism and potential of informational and cultural influence than Russia, though it will still lag behind the USA for a long time.

Consequently, China is more than Russia associated with the symbol of alternative development for Latin Americans.

However the uncertainty of strategic communication of China comes from the ambiguity of the prospects of its further rapid economic development, the absence of its own raw materials supplies, an increasing (and successful) competition with the Latin American producers, a still large (though decreasing) lag in the scientific and technological field, and still unsolved social problems. Some of the advantages can transform (and partly have already transformed) into disadvantages, such as, for instance, the world largest population (the fact which on the mental level easily provokes still not panic but fear of the Chinese threat in the significant social strata of the region).

As for the theoretical development and the practical use of strategic communication, China, like Russia, is lagging behind the USA, being on the same technological, mostly propagandistic stage of development. A transition to the new "managerial" and integrative (not only from the point of view of administration) type of strategic communications in China is underway.

4. Strategic communication of the USA: The burden of bad decisions

The United States possess a whole set of competitive advantages allowing efficient strategic communication in Latin America. Among them are:

- the geographical proximity;
- a similarity of histories (both the USA and the Latin American countries emerged in the struggle against the colonialism of the European powers and together used to oppose (not once) the aggressive policy of the European powers in the Western hemisphere);
- the economical, military and scientific superiority of the USA in the world;
- the predominance of the USA in the trade, economical, military, scientific and educational relations with most countries of the region as compared to all other powers;

- the predominance of the US mass culture in the public conscience of the Latin Americans over the cultural influence of other powers;
- the superiority of the USA in the field of global media over any other country (the fact that clearly manifests itself in Latin America);
- a superiority in the theoretical and practical development of strategic communication on the Latin-American direction of the US foreign policy.

However, because of the vested and false pretences and decisions in the past and present, all these advantages have caused permanent and quite well-grounded perceptional stereotypes of the USA as “an aggressor-power”, “an external threat” and “the main basis of the reactionary forces in the region” in the mentality of a significant and sometimes even the bulk of the Latin-American population.

The rise to power in the USA of the new administration of B. Obama caused in Latin America a wave of great expectations however this is entirely a US domestic affair. The image of the new President was totally positive for most Latin Americans. The US Administration has outlined four pillars for the regional partnership with the Western Hemisphere: protecting citizen security; expanding economic opportunity and social inclusion; securing our clean energy future; and supporting democratic, transparent, and accountable institutions of governance. These were the key messages of the President Obama to the region many times repeated by him and the top officials of his administration and widespread in Latin America. The more repeated, the more effectively disseminated, the more practical implementation was expected among the public.

But very soon Obama's "change you can believe in" soon began to look like "more of the same." The recognition of the November 2009 elections in Honduras for a new president soon after June 28th coup in that country, provoked the negative reaction in many countries of Latin America. Also in June, reports began to surface about a secret agreement between the United States and Colombia to allow U.S. access to seven military bases in Colombia (Stephens 2010).

Systematic, repeated and often expensive attempts to provide a communication background for questionable and improper acts have already resulted in and, moreover, can result in future in strategic management expenses and a fall of trust of the USA. A gradual growth in the number of such expenses can have a cumulative effect and is a much more dangerous threat for the national security of the country than the questionable tactical successes

5. Conclusion

All the three countries actively develop their strategic communications in the region but, according to numerous expertise and public opinion polls, the “correspondence of words and actions” is perceived as best achieved by China. Among the reasons behind it is a more adequate and consistent Chinese strategy of developing the relations with the region and its communication support, a complex heritage of the relations between the USA and Latin America, and gross mistakes of the foreign policy of G. Bush administration, including the mistakes in the field of strategic communications, as well as an undecided and inconsistent way of their correction by the present B. Obama administration.

Once again we want to repeat that strategic communication is a projection in the mass consciousness of some strategic values, interests and goals in this or that way. And these goals can more or less coincide, coexist or compete, to be at enmity, to be in war or finally exclude each other. If strategic communication of Russia, China and the USA (desirably, that of other countries as well) projects a great number of coinciding basic values into the public conscience, this possibility will be the best option for most of the Earth’s population, as well as for those countries themselves and for global safety. If we want peace, our main goal is to obtain harmonious coincidence of interests, values and goals, though it does not depend on strategic communication completely. It doesn't mean defending immoral compromises, it means defending pluralism in the respect of the means and models of development based on dialectical unity of the main laws, as well as of national and regional peculiarities of human development.

We consider it possible to present our general recommendations for optimization of strategic communication of the three countries within the framework of the word and deed policy, the refusal of information war in favor of mutually beneficial cooperation.

- In order to decrease the tension between each other and improve the promotion of the images of the three nations in the continent, it is sensible to launch joint projects in the domains which are crucial for the increase of prosperity and liquidation of current arrearage of Latin American countries. For this is needed the increase of involvement of Latin American partners in these projects, especially in its hi-tech elements. The projects should be open for participation of other countries and unions, such as the EU, India, Japan etc..
- Global joint projects involving Latin American partners, such as search and development of alternative energy sources, life (especially its active period) prolongation, solution of alimentation and ecological problems etc. It is rather important to achieve the development of projects vital for all mankind by joint efforts. A priority interest in bi- and multilateral relations with Latin American countries to hi-tech branches wherever and as much as it is considered mutually beneficial.
- Maximal possible transparency on preparation and implementation of joint projects.
- Well-thought system of consultations and meetings of strategic communication experts in order to discuss the emerging problems on time. Just in 1996 Timothy L. Thomas, an analyst at the Foreign Military Studies Office, Fort Leavenworth, Kansas proposed:

One of the easiest ways for the West to begin joint talks on information warfare with Russia is through the medium of a conference among academics or through an unofficial organization or club…The academy could serve as a forum for broader discussions with the West and already appears oriented this way, having several foreigners on its membership roll. By starting this discussion soon, Russia and the West can prevent a new arms race over information systems and technologies from gaining momentum and spinning out of control. With the rate of progress in the realm of information technology, time really is of the essence (Thomas T. L., 1996, p.33).

A very good idea, but all of the opportunities of it are far from being realized yet.

- The efficiency of strategic communication as means of collaboration is negligible in case of strategic interests and goals mismatching drastically. In this case, strategic communication inevitably becomes a tool of information warfare. There should be no illusions cherished of the contrary.
- That is why compromises are essential, as well as searching for ways to combine interests. Strategic communication can be very fruitful for the creation of a climate useful for such a search, but it can aggravate the situation as well. To a certain extent, strategic communication itself is an important (and partly autonomous) factor of rapprochement or estrangement of the parties, and it is vital to procure that it serves to the accomplishment of the first task. We can fully agree with the point of view of Dennis M. Murphy, a professor of information operations and information in warfare at the U.S. Army War College: "Basic theory – you may not change someone's mind, but you can find areas of agreement where interests overlap" (Murphy, D. M., 2008).

Such a program of joint optimization of strategic communication in the region is completely impossible to implement in the circumstances of growing tension between Russia and China from one hand and the USA on the other. There is a need for a serious, revolutionary at its core economic, technological, social and political shifts in the three countries with the consideration of their national peculiarities. This is the common interest in order to overcome the threat of a new world war and to provide conditions for the democratic and progressive development of mankind.

The strategic communication of three countries will have to deal with ongoing theoretical and practical problems of the further alignment of "words and deeds" in the real policy. It should make strategic communication overcome its largely propagandistic level and become not only an element of communication support of foreign policy, but even more of an equal part of foreign policy management, with the appropriate set of tools of communication management, i.e. of professional people management via communications.

References

Abraham, A. J., 2004. The Strategic Communications Process-How to Get Our Message Out More Effectively, National War College Paper. National Defense University.

Armistead, E., ed., 2004. Information Operations: The Hard Reality of Soft Power. Washington.

Arpagian N., 2010. Internet et les resseaux sociaux: outils de contestations et vecteur d'influence. La Revue Internationale et strategique. №78.

Brunner E. M., Cavelty M. D., 2009. The formation of in-formation by the US military: articulation and enactment of information threat imaginaries on the immaterial battlefield of perception. Cambridge Review of International Affairs. P. 641 – 642.

Corman S. R., Trethewey, A., Goodall, H. L., ed. 2008. Weapons of Mass Persuasion: Strategic Communication to Combat Violent Extremism (Frontiers in Political Communication). Peter Lang Publishing.

Crandal, C.Russel, 2008. The United States and Latin America after the cold war; The New York: Cambridge University Press.

Department of Defense, 2009. Report on Strategic Communication. The Secretary of Defense, Washington, DC, December 2009, 20301-1000. Feb. 11 2010.

Department of Defense, 2010. Report on Strategic Communication. December 2009.The Secretary of Defense. Washington: Department of Defense DC, 20301-1000. February 11.

Deputy Secretary of Defense, 2006. Memorandum for Secretaries of the Military Departments. Quadrennial Defense Review (QDR), Strategic Communication (SC), 1010 Defense Pentagon, Execution Roadmap. Washington, DC 20301-1010. September 25, 2006.

Future of U.S. Public Diplomacy. *Testimony.Judith A. McHale Under Secretary for Public Diplomacy and Public Affairs. Before the SFRC Subcommittee on International Operations and Organizations, Human Rights, Democracy, and Global Women's Issues.*Washington, DC. March 10, 2010. Available from: http://www.state.gov/r/remarks/2010/138283.htm [Accessed 20 May 2011].

Haddock, E. K. , 2002. Winning with Words: Strategic Communications and the War on Terrorism: National Defense University.

Handricks, B., Wenner R., 2010. DIME is for integration: Strategic Communications as an Integrator of National Power. IO Sphere, May, pp.36-39.

Internet Journal, 2008. China is going to develop a strategic partnership with Latin American countries. Internet Journal, November 21. Available from: http://www.novopol.ru/text56287.html [Accessed 20 May 2011].

Murphy, D. M., 2008. The Trouble With Strategic Comunication(s), IO Sphere, Winter, pp.24- 26.

Novaya Politika, 2008. Available from: http://www.novopol.ru/text56287.html [Accessed 20 May 2011]

News of the Communist Party of China. Party Work. Available from: http://english.cpc.people.com.cn/66105/index.html [Accessed 20 May 2011].

Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics, 2004. Report of the Defense Science Board Task Force on Strategic Communication. September.

Office of the Under Secretary of State for Public Diplomacy and Public Affairs, 2010. Public Diplomacy. Strengthening U.S. Engagement with the World. A strategic approach for the 21stcentury.Washington: Office of the Under Secretary of State for Public Diplomacy and Public Affairs. 2/26/.

Pashentsev, Evg., 2012. Kommunikatsionny menedjment y strategisheskaya kommunikatsia. Moskva:Mejdunarodnij tsentr socialno-politisheskih issledovaniy.

Paul, Ch., 2011. Strategic Communication: Origins, Concepts, and Current Debates. Santa Barbara: Praeger.

Patterson S.J., Radtke J.M., 2009. Strategic Communications for Nonprofit Organization: Seven Steps to Creating a Successful Plan: Wiley.

Speech by H.E.Mr.Liu Huanxing, Chinese Ambassador to Botswana, at Botswana Defence Command & Staff .Available from: http://www.fmprc.gov.cn/eng/wjb/zwjg/z [Accessed 20 May 2011].wbd/t801574.htm[Accessed 20 May 2011].

Stavridis, James G., 2007. Strategic Communication and National Security, Joint Force Quarterly, 3rd Quarter.

Stephens , S.2010. Why Latin America is Disappointed with Barack Obama: Huffpost World. January 7. Available from:http://www.huffingtonpost.com/sarah-stephens/why-latin-america-is-disa_b_415341.html [Accessed 20 May 2011].

Stiglitz J.E. Freefall: America, Free Markets, and the Sinking of the World Economy. 2010. 1st ed. New York – London: W. W. Norton & Company.

Stovichek B. E., 2007. Strategic Communication. A Department of Defense Approach, USAWC Strategy Research Project. Carlisle Barracks, Pennsylvania: U.S. Army War College, 2007.

Strategic Communication and Public Diplomacy Policy Coordinating Committee (PCC), 2007. U.S. National Strategy for Public Diplomacy and Strategic Communication. Released June 2007.

The Foreign Policy Concept of the Russian Federation. Approved by Dmitry A. Medvedev, President of the Russian Federation, on 12 July 2008. Available from: http://rusembassy.in/ [Accessed 20 May 2011].

The News of the Communist Party of China website the Party Work. Available from: english.cpc.people.com.cn/66105/index.html [Accessed 20 May 2011].

The U.S. Department of State. Diplomacy in Action. Benefits From Existing Free Trade Agreements Available from: http://www.state.gov/e/eeb/tpp/bta/fta/c26474.htm [Accessed 20 May 2011].

Thomas T. L., 1996. Russian Views on Information-Based Warfare. Airpower Journal. Special Edition.

Vesti, 2010. Medvedev D. Russia comes back to Latin America. Vesti, April 15.

White House, 2010, Strategic Communications report to Congress “National Framework for Strategic Communication”, 16 Mar 2010, released 17 March 2010. Government Information Earl Gregg Swem Library.

National Security Auditing Criteria, KATAKRI: Leading Auditor Training and Auditing Process

Jyri Rajamäki[1] and Merja Rajamäki[2]
[1]Laurea University of Applied Sciences, Espoo, Finland
[2]Finnish Safety and Chemicals Agency (Tukes), Helsinki, Finland
jyri.rajamaki@laurea.fi
merja.rajamaki@tukes.fi

Abstract: The National Security Auditing Criteria, KATAKRI, were published in 2009, revised in 2011, and version III is currently under revision. The root of KATAKRI is to preserve the confidentiality of any confidential and classified information held by the organisation concerned. One of KATAKRI's aims is to combine the actions of authorities when verifying the security level of a company or other corporation by carrying out security auditing. From the enterprise operators’ point of view, the focus of security auditing is to eliminate unfair competition and maintain an equal opportunity field for operators. Another of KATAKRI’s aims is to improve national security when Finnish Defence Forces or other security authorities apply subcontracting. KATAKRI is also intended to help companies and corporations when they are developing their own security level. The purpose of this case study is to find out: what is expected from the security auditing process and from the leading auditor; what kind of competence the auditor should have; and how the security auditing training and qualification should be developed to correspond with the needs of the security field. The empirical research was conducted in the form of interviews, questionnaires and observations made as a student during the first KATAKRI leading auditor course executed 2/2/2012–12/12/2012. The combined results showed that deep knowledge of the security field and competence to manage overall security is required from security auditors. Furthermore, it was concluded that qualifications for security auditors should be created in accordance with ISO Standard 19011:2011, which provides a very strong competence model. In light of the above, it is recommended that the academic level, content and requirements of future audit and security auditing training should be clearly defined, and the quality of the training should be standardised and certified. The results also indicate that KATAKRI version II still has defects due to its inconsistency. One task of auditing processes should be collecting information about KATAKRI’s shortcomings, and they should be systematically analysed. Future leading auditor courses would be suitable scenes to analyse shortcomings and to propose improvements to KATAKRI. KATAKRI should be revised every second or third year.

Keywords: KATAKRI, national security auditing criteria, security auditing, security auditing training

1. Introduction

1.1 The National Security Auditing Criteria, KATAKRI

The root of the National Security Auditing Criteria, KATAKRI, is to preserve the confidentiality of any confidential and classified information held by the organisation concerned. KATAKRI was officially published in November 2009, and the first update (Ministry of Defence, *National Security Auditing Criteria, version II*) was published in mid-2011. Version III is currently under revision; the Internal Security Secretariat has appointed a working group to update KATAKRI by 31/12/2013.

According to the current version of the criteria, KATAKRI’s main goal is to harmonise official measures when an authority conducts an audit in a company or in another organisation to verify their security level. The National Security Authority (NSA) uses KATAKRI as its primary tool when checking the fulfilment of security requirements. The preface to the criteria states that the second important goal is to support companies and other organisations, as well as authorities and their service providers and subcontractors, in working on their own internal security. For that reason, the criteria contain recommendations for the industry that are separate and outside of the official requirements; it is hoped that useful security practices will be chosen and applied, thus progressing to the level of official requirements.

The Web page ‘Ministry of Defence of Finland – National Security Auditing criteria (KATAKRI)’ relates: ‘KATAKRI-criteria have been created from the perspective of absolute requirements and they do not include a marking system which is used in some criteria. The aim here is to make sure that at the end of an audit there would not be possibly unidentified but critical risks. The chosen approach means specific demands for the personnel conducting security audits and, as a result, high enough training level requirements are set to satisfy these demands.’

1.2 Auditing procedure

Many different types of audits exist, including financial audits, property assessments, supplier reviews, contractor evaluations, registration audits, equipment evaluations (ISO 19011 Expert), etc. Figure 1 illustrates internal (first-party) and external (second-party and third-party) auditing types. The common principle is that they compare applied procedures, as well as a set of collected information, against some established criteria.

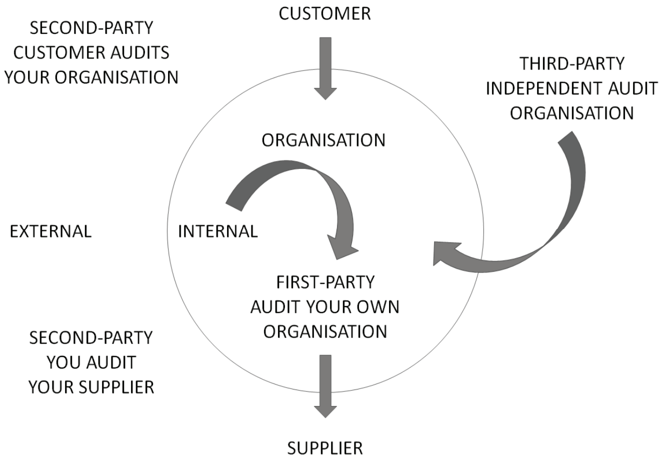


Figure 1: First-, second- and third-party audits (adapted from Russel 2005)

ISO/IEC 17021-2 is a normative standard intended for use by accreditation bodies when assessing management systems, while ISO 19011 provides guidelines for first-, second- and third-party auditors when auditing management systems. The third-party certification industry will use ISO 17021-2 to define requirements for audits and audit arrangements and accreditation bodies will determine whether a certification body's auditing arrangements and activities comply with those requirements. ISO 19011 identifies best practice and provides information on what should be done when carrying out an audit without specifying how it must be done. ISO 19011:2011 edition includes an extension of the standard's earlier scope of application from quality and environmental management systems to all types of management systems auditing. Continuing development of management systems standards for information security, for example, means that ISO 19011 must be able to accommodate differing requirements while still providing useful guidance (ISO 19011 vs ISO/IEC 17021-2 - IRCA – Home).

The three things that make a management system audit different from other types of assessments are that the audit must be 1) systematic, 2) independent and 3) documented. In order to conduct systematic management system audits, there is a need for both audit procedures and an audit programme. From an independence point of view, auditors cannot audit their own work or that of their colleagues', as there would be a conflict of interest. Audits need to be structured, to ensure they are free from bias and conflicts of interest. Audits must be documented, because they are all about making decisions and taking action (ISO 19011 Expert).

1.3 Competence and evaluation of auditors

ISO 19011:2011 includes a section that deals with auditor competence. The section covers determining auditor competence to fulfil the needs of the audit programme, personal behavioural aspects, discipline or sector-specific competence, as well as evaluation and maintenance of competence. In relation to behavioural aspects, auditors should be, for example, open-minded, perceptive, tolerant of pressure, versatile, culturally sensitive and collaborative.

Management system auditors should have generic knowledge and skills needed to audit multiple discipline management systems and implement other parts of ISO 19011:2011. For example, auditors should understand the types of risk associated with auditing. They need knowledge of organisational types, general business and management concepts, processes and related terminology, including budgeting and management of personnel. Auditors should be able to position discipline and sector requirements and audit findings in the wider context of the organisation's business activities, governing agencies, business environment, legal and contractual requirements and management's policies and intentions for the organisation.

Annex A.7 of ISO 19011:2011 describes the knowledge and skills that information security management auditors should have. Auditors who intend to examine information security management systems need to have information security management knowledge and skills. They should be able to apply information security management methods, techniques, processes and practices. They must have the knowledge and skills needed to examine information security management systems and to generate appropriate audit findings and reach valid conclusions.

According to ISO 19011:2011, an audit team leader must have the knowledge and skills 1) to balance the strengths and weaknesses of the individual audit team members, 2) to develop a harmonious working relationship among the audit team members and 3) to manage the uncertainties involved in achieving audit objectives.

1.4 KATAKRI audit team leader training

Due to increasing application of KATAKRI, there is a definite need to teach both the content of KATAKRI and the process of security auditing. Today, many organisations are arranging different kinds of KATAKRI training courses, but the academic level, content and requirements of security auditing training have not yet been defined (Rajamäki 2011). At the initiation of the Ministry of the Interior, Laurea University of Applied Sciences (UAS) organises KATAKRI leading auditor training courses. The first course started in February 2012 and ended in December 2012, and the second started in February 2013. The basic assumption behind any security auditing training course is that the authorities can trust the quality of training and the expertise of those people who have undertaken it.

1.5 Structure of the paper

Section 2 of this paper presents the research targets and methods applied in this study, as well as how the research process has proceeded. In Section 3, the research findings are presented and evaluated against the theories presented in Section 1. Section 4 sets out the conclusions of the study and answers to the research questions. The final section also includes an assessment of the study and suggestions for further research.

2. Research method and process

The purpose of this study is to find out what is expected from the National Security Auditing Criteria, the security auditing process and the audit team leader. We tried to discover what kind of competence the auditor should have and compared these to the suggested competencies of ISO 19011. We analysed KATAKRI's different targets. We also give suggestions regarding how security auditing training and qualification should be developed to correspond with the needs of the security field. This study has been carried out according to the case study method of research represented by Yin (2009). The empirical research was conducted in the form of interviews, questionnaires and observations. Nine highly experienced experts in the fields of security and safety were interviewed. They were selected according to their experience and organisations: four of them represented authorities, three represented private companies, one was a researcher and one was a consultant. The interviews lasted 1 to 2.5 hours each and were recorded, transcribed and analysed with the ATLAS.ti computer program.

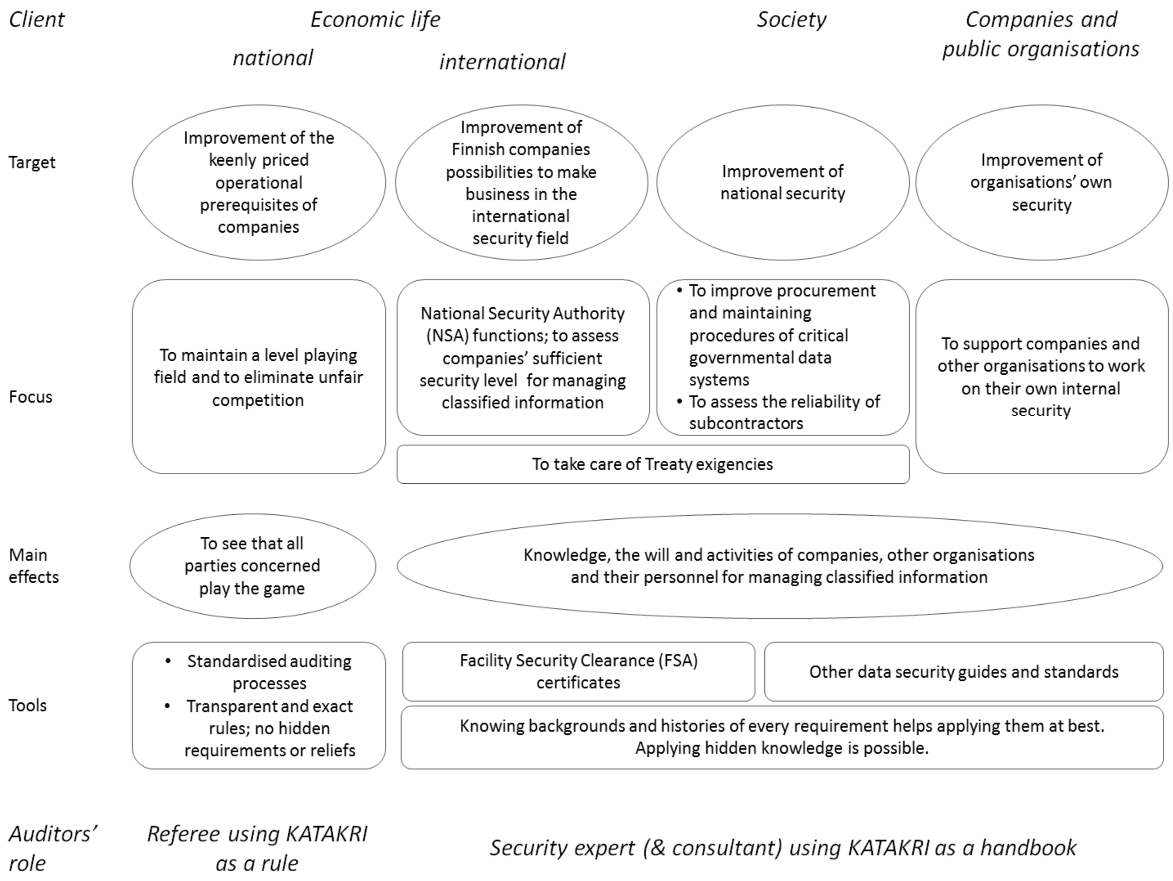
Two different Webropol questionnaires (N=31, N=14) were circulated to graduate security management and ICT students at Laurea UAS. The aim was to find out whether students would be interested in security auditing studies and their opinions on the content of such studies.

The first KATAKRI leading auditor training course was executed between 2/2/2012 – 12/12/2012. One of this paper's authors developed the course; another was one of its seventeen participants. This paper provides research results and lessons learnt from the course.

3. Findings and discussion

3.1 Multiple targets of KATAKRI and security audits

As shown in Figure 2, KATAKRI and the security auditing process serve three main clients to: 1) the economic life (companies that develop and sell security products and services), 2) society and 3) companies and other organisations seeking to improve their own internal security.



Client — *Economic life* (*national*, *international*) — *Society* — *Companies and public organisations*

Target

Improvement of the keenly priced operational prerequisites of companies

Improvement of Finnish companies possibilities to make business in the international security field

Improvement of national security

Improvement of organisations' own security

Focus

To maintain a level playing field and to eliminate unfair competition

National Security Authority (NSA) functions; to assess companies' sufficient security level for managing classified information

- To improve procurement and maintaining procedures of critical governmental data systems
- To assess the reliability of subcontractors

To support companies and other organisations to work on their own internal security

To take care of Treaty exigencies

Main effects

To see that all parties concerned play the game

Knowledge, the will and activities of companies, other organisations and their personnel for managing classified information

Tools

- Standardised auditing processes
- Transparent and exact rules; no hidden requirements or reliefs

Facility Security Clearance (FSA) certificates

Other data security guides and standards

Knowing backgrounds and histories of every requirement helps applying them at best. Applying hidden knowledge is possible.

Auditors' role — *Referee using KATAKRI as a rule* — *Security expert (& consultant) using KATAKRI as a handbook*

Figure 2: Multiple tasks of KATAKRI audits

From economic life's point of view, the task of security audits could be divided into two parts: a national and an international viewpoint. The national viewpoint is taken into consideration when companies are acting as service providers and subcontractors for Finnish Defence Forces or other national (security) authorities. A normal procedure in these situations is that the new services are put out to tender, and fulfilling KATAKRI's requirements is a mandatory precondition for companies. From this perspective, the main aim of security audits is to eliminate unfair competition and maintain an equal-opportunity market for all companies. To achieve these results, security auditors have to inspect the workings of 'the system' so as to determine that all parties concerned observe their responsibilities. And here, by the 'system', we mean that no organisation will benefit from breaking KATAKRI's requirements on purpose. If an attempt to do this is made, it will be detected, resulting in the organisation fouling its own nest. The role of security auditors is to act as a referee between companies and use KATAKRI as a rule. With regard to this function, auditing processes should be firmly standardised, having transparent and exact rules without any hidden requirements or reliefs.

From international economic life's point of view, the target of security audits is to improve Finnish companies' business opportunities within the international security field. Facility security clearance (FSC) certificates provided by NSA under the terms of bilateral treaties between countries enable Finnish companies to take part of calls for offers with regard to international security critical business. From this perspective, the main aim of security audits is to assess that the company concerned has sufficient security procedures and facilities in place for managing classified information.

From society's point of view, KATAKRI's target is to improve national security. KATAKRI is applied when Finnish Defence Forces or other authorities purchase or subcontract security products or services. From this perspective, KATAKRI's aims are to improve the procurement and maintenance procedures for critical governmental data systems, as well as to assess the reliability of subcontractors. The way to administrate for national security is to attempt to contribute to the knowledge, will and activities of all individuals, companies and other organisations so that security becomes one of the true values which conducts the thinking behaviour behind their activities. To achieve these results, KATAKRI should be used as a handbook. The role of a security auditor is to act as a security expert and consultant. Knowing the background and target of every requirement helps to best apply these requirements at best. Also, utilising hidden knowledge could improve security.

KATAKRI is also intended to help companies and public organisations developing their own security mechanisms on a voluntary basis. This is why KATAKRI contains recommendations that are separate and outside the official requirements. From this perspective, KATAKRI could be applied as a handbook and security auditors could act as security experts and consultants.

3.2 Multiple roles of KATAKRI auditors

As stated earlier, security auditing involves multiple tasks. This means that security auditors have multiple roles. The two main roles are 1) to referee on the playing field between companies and 2) to act as a security expert and consultant. In most cases, the role of a referee conflicts with the role of a consultant. When auditors are seeking to maintain a level playing field, in principle, they are not able to consult.

In the worst situations, maintaining a level playing field for companies does not improve their security level at all. This is the case if a certain criterion does not actually measure the facility it is meant to. According to Finnish law, security auditors are acting civil servants. So, as shown in Figure 3, with this 'hat' on, the security auditor should also ensure that the governance system is a functional one – here, by the governance system, we mean national legislation, KATAKRI and standardised auditing processes. As a result of this, a very important role for security auditors is continuous monitoring of the security auditing criteria. When needed, security auditors must react and participate in requirement renewals. This includes KATAKRI renewals as well as development of standardised auditing processes. However, because audit findings are confidential information, enforcement of this role is not an undemanding task.

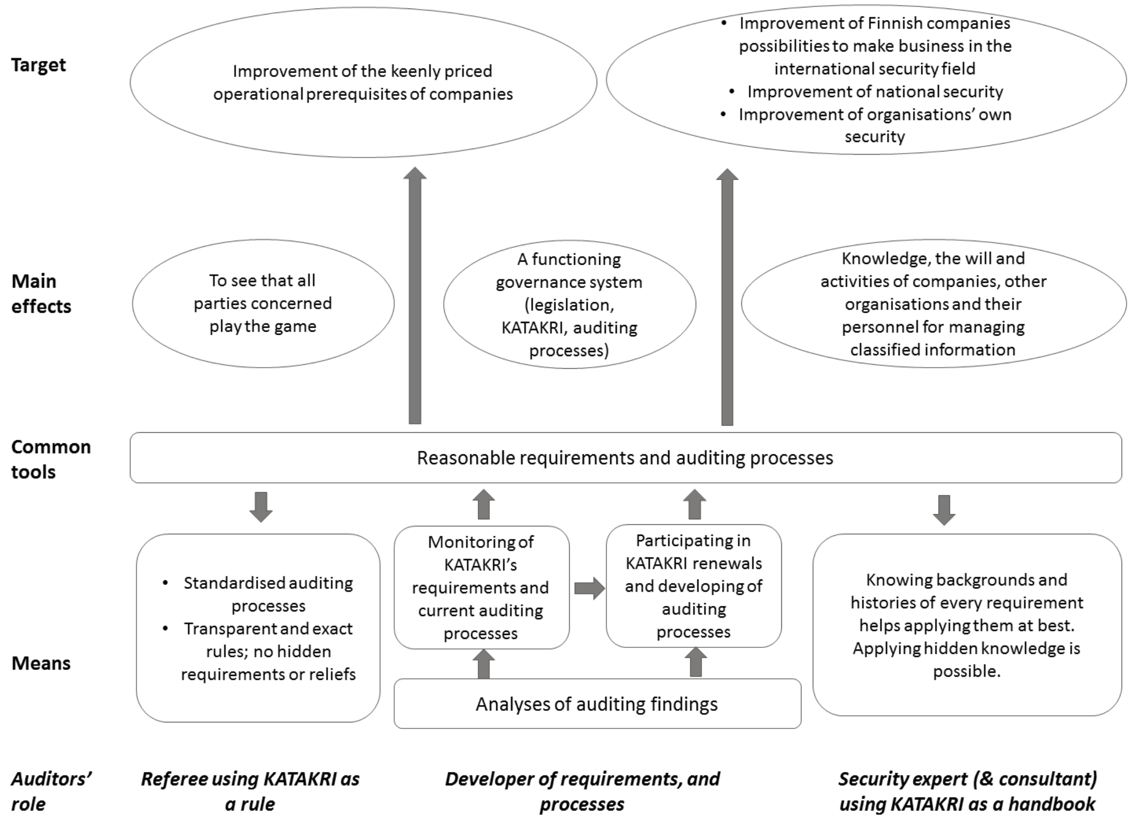

Target
Improvement of the keenly priced operational prerequisites of companies
• Improvement of Finnish companies possibilities to make business in the international security field
• Improvement of national security
• Improvement of organisations' own security
Main effects
To see that all parties concerned play the game
A functioning governance system (legislation, KATAKRI, auditing processes)
Knowledge, the will and activities of companies, other organisations and their personnel for managing classified information
Common tools
Reasonable requirements and auditing processes
Means
• Standardised auditing processes
• Transparent and exact rules; no hidden requirements or reliefs
Monitoring of KATAKRI's requirements and current auditing processes
Participating in KATAKRI renewals and developing of auditing processes
Analyses of auditing findings
Knowing backgrounds and histories of every requirement helps applying them at best. Applying hidden knowledge is possible.
Auditors' role
Referee using KATAKRI as a rule
Developer of requirements, and processes
Security expert (& consultant) using KATAKRI as a handbook

Figure 3: Security auditors' roles, means, tools, effects and targets

3.3 Auditors competences and auditing training

The combined results of the interviews, the questionnaires, and examination of the content of the existing training modules showed that deep knowledge of the security field and the competence to manage the overall security picture is required from security auditors. Furthermore, it was concluded that qualifications for security auditors should be created in accordance with ISO 19011:2011, because it provides a very strong competence model.

In light of the above, it is recommended that the academic level, content and requirements of future audit and security auditing training should be clearly defined, and the quality of the training should be standardised and certified. It would then be possible to plan and implement a different kind of security auditor course for different purposes. For example, the lead auditor course module is a natural module for Laurea UAS to offer to experts from different security branches who want to deepen their know-how regarding leading security audits.

The results also indicate that KATAKRI version II still has defects due to its inconsistency. One task of auditing processes should be collecting information about KATAKRI's shortcomings, and they should be systematically analysed. From the experience of the first leading auditor courses, most participants (students, lecturers) are real security experts with experience of taking part in KATAKRI audits as team members. Future leading auditor courses would be suitable scenes to analyse shortcomings and propose improvements to KATAKRI. KATAKRI should be revised every second or third year.

4. Conclusions

This section evaluates the research process and the findings of this study from the viewpoint of the study's research questions. Finally, suggestions for future research avenues are made.

4.1 Answers to research questions

The main objective of this study was to find out what is expected from KATAKRI, the security auditing process and the leading auditor. As Figure 2 shows, KATAKRI audits have different objectives depending upon the reason for the auditing process being executed. The audit team leader must be aware of these objectives and act according to them. However, the most important tool for auditors to carry out their work is a functioning governance system. This means that auditors should invest in improving criteria so that they are reasonable, topical and functional. In practice, this means that auditors should analyse audit findings as well as monitor KATAKRI's requirements and auditing processes. When needed, they should participate in KATAKRI renewals and develop auditing processes.

The new version of ISO 19011 defines quite well the kinds of competence that auditors and the audit team leader should have. It identifies the necessary auditor competence, including generic knowledge and skills of management systems, discipline and sector (e.g. aerospace) knowledge and skills. Informative Annex A gives examples of auditors' discipline-specific knowledge and skills, including e.g. information security. However, no guidance is given regarding auditors' sector-specific knowledge and skills.

Leading auditor courses are forums to disseminate expertise from earlier audits in which the participants have been audit team members. As stated previously, auditors should also monitor the criteria concerned. To put this task into action, every leading auditor course should include a period in which the participants analyse the possible shortcomings and incoherence in KATAKRI. Unfortunately from this perspective, audit findings are confidential information, which limits the possibilities to discuss these issues. Anyway, a suitable academic level for these studies is a graduate school, because the students should have abilities of analysing of findings. These courses could be part of a master's degree in security management or information systems.

4.2 Future research proposal

The security auditing training and qualification should have different steps to correspond with the needs of the security field. Figure 4 offers a rough sketch of 'the ladder' of an individual security expert moving towards becoming the leader of a security audit team. It also outlines the roles of academies (left) and authorities (right), as well as how the organisation concerned could learn within this development process (middle). However, future research is needed to define and design the right steps, their number and their levels. The concrete roles of academies and authorities in achieving each step should also be clarified.

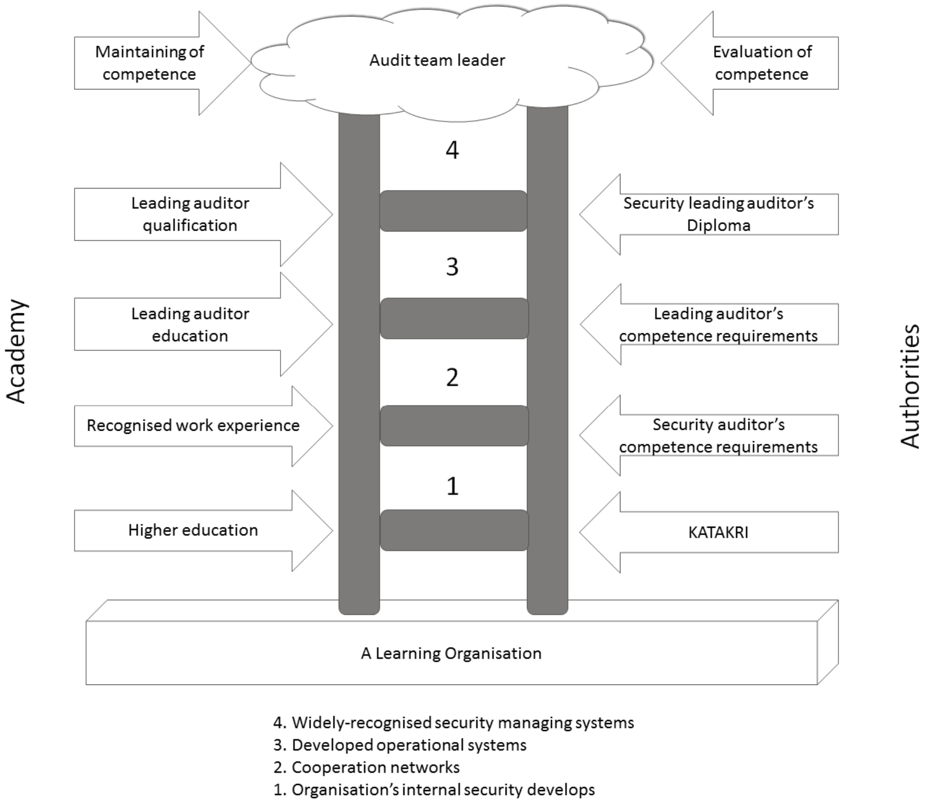

Maintaining of competence
Audit team leader
Evaluation of competence
4
Leading auditor qualification
Security leading auditor's Diploma
3
Leading auditor education
Leading auditor's competence requirements
Academy
Authorities
2
Recognised work experience
Security auditor's competence requirements
1
Higher education
KATAKRI
A Learning Organisation
4. Widely-recognised security managing systems
3. Developed operational systems
2. Cooperation networks
1. Organisation's internal security develops

Figure 4: Authorities', organisations', academies' and individuals' contributions and throughputs when applying the competence model of a security auditor

References

ISO 19011 (2011) 'Guidelines for auditing management systems', Geneva: ISO.

ISO 19011 Expert, http://www.iso19011expert.com/ (accessed February 28, 2013).

ISO 19011 vs ISO/IEC 17021-2 - IRCA - Home | International .., http://www.irca.org/en-gb/resources/INform/archive/issue27/Features/Building-on-safety21/ (accessed February 28, 2013).

ISO/IEC TS 17021-2 (2012) 'Conformity assessment -- Requirements for bodies providing audit and certification of management systems -- Part 2: Competence requirements for auditing and certification of environmental management systems', Geneva: ISO.

Ministry of Defence (2011) *National Security Auditing Criteria (KATAKRI),* version II, 2011. http://www.defmin.fi/files/1871/KATAKRI_eng_version.pdf

Ministry of Defence of Finland – National Security Auditing criteria (KATAKRI). [Online], http://www.defmin.fi/en/administrative_branch/defence_security/national_security_auditing_criteria_(katakri) (accessed February 26, 2013).

Rajamäki, M. (2011) *Pätevyysmalli turva-auditoijan tutkintokoulutukselle – tapaustutkimus Laurean Auditoinnin johtaminen –opintojaksosta (Developing a competence model for security auditor specialization studies – Case study: Laurea's Management of Auditing study module).* Master's thesis. Theseus. Espoo: Laurea (in Finnish).

Russel, J. (ed.) (2005) *The ASQ Auditing Handbook*, Milwaukee: ASQ, Quality Press.

Yin, R. K. (2009) *Case Study Research: Design and Methods*, 4th ed., California: SAGE Publications.

Russian State Leaders’ Contradicting Narratives on Social Media

Jari Rantapelkonen and Margarita Jaitner
Finnish National Defense University, Helsinki, Finland
jari.rantapelkonen@mil.fi
m.a.jaitner@gmail.com

Abstract: Russian top state leaders have a long tradition of using mass media to promote national security goals focusing on one-way communication. President Vladimir Putin is known for maintaining an image of a fearless hero in the mass media. With the rapid spread of social media throughout the world, it has become also very popular amongst Russians in a short period of time. After a long time of state-led one-way communication, Russian politicians are looking for ways to master social media as a means of quick and potentially two-way communication, enabling it to be a tool for themselves and for promoting national security goals. The intention of this article is to bring insights into Russian views on social media at top strategic levels, particularly those of the state leadership. This article explores the narratives of social media’s “place” in Russian politics at the strategic level in the form of speeches and actions. The article reveals that Russian strategic communicators do recognize the importance of social media, but also that the assessment of its potential, as well as methods of use must be concluded via the Russian cultural perspective. The narratives are consistent but ambiguous, like politics itself.

Keywords: Russia, strategic communication, Putin, Medvedev, information policy, social media

1. Introduction: Russian struggle of values

In the recent history the world has seen indicators of how the Internet in general and the social media in particular is viewed as potentially threatening element to the existing order, as well as how it challenges legal frameworks and the judicial systems. According to Jaitner (2013), "The nature of social media challenges the established, state-centric, viewpoint on exercising power". Social media also challenges the traditional means of communication and requires the top state leaders to adapt their techniques of conveying their message to the population.

One of the biggest information wars on the roles and responsibilities of mass media and the use of social media is waged in today’s Russia. After a long history of relying on elaborate and sometimes blunt methods of mass one-way communication, persuasion, and propaganda, today the state leadership finds its communication skills tried by a force of bloggers and twitters. The emergence of cybercrime that came alongside with the rise of Internet use appears to be another concern for the state leadership. The attempts to tackle the newly arisen challenges lead to questions within the international community: In a recent pursuit to protect the younger citizens from illicit, potentially harmful information on the internet, the authorities temporary blocked YouTube and Google. (Securitylab 2012; Blagoveshensky 2012).

Understanding the narratives and actions of Russian key actors brought forward in this article requires putting them into the context of the Russian reality. This reality is multidimensional where values are deeply intertwined with the country’s history. It is necessary to take in account “fundamental values as love for Russia, public unity, the family, individual freedom, democracy, equality of rights, selflessness in Russia's defense, territorial integrity, collectivism, perseverance, conscientious labor, social justice, a multinational culture, and spirituality” (Manilov n. d.). Furthermore, methods and ways Russians have developed in order to cope with the historical, social, and political reality need to be considered. Simply put, one needs to understand the common Russian definition of freedom in order to be able to assess the level of freedom in the country.

Manilov (n. d.) argued that “The systemic crisis that seized the USSR was above all a crisis of values: the loss of common goals, and the growth of pessimism, bitterness, and other negative feelings among the population. Today, a dramatic process of reappraisal of many seemingly inviolable values is occurring. A kind of spiritual vacuum has emerged, in which the nation has become dangerously indifferent towards the absence of common public ideas, of clear notions and traditions that meet peoples’ ''deep feelings””.

Indeed, the rhetoric of the common values that are needed to recreate a strong, independent, successful Russia has been repeatedly included in Mr. Vladimir Putin’s speeches, in 2007 he stated that: “We have an old Russian game - search for the national idea, a search for the meaning of life of sorts. (While) generating novelty, we must at the same time rely on the basic values our people have developed through our more than a thousand year old history. Only then will we achieve success" (Novye Izvestiya 2007). Promotion of common Russian values is a recurring element of Putin's (2000; 2012) speeches and articles through the years of his position at the top of the Russian political machine. Common values are a necessity to recreate the "sacred power" and the "mighty will", and to regain the "great glory" - Russia as it is presented in its national anthem.

2. Social media landscape in Russia

The Russian social media landscape differs significantly from its "western" counterpart. The "RuNet", as Russians themselves call it, is divided from the global Internet by a language barrier and it's historical, political, and social context (Lonkila 2012). This results in different patterns of use of the Internet as a whole, and in the popularity of different platforms. Instead of the worldwide leader Google, the majority of Russians turn to the domestic search engine portal Yandex while Facebook lags behind social media platforms Odnoklassniki and VKontakte. (Daveluy 2012). The American-founded, now Russian-owned blogging platform LiveJournal had upon its introduction to the market quickly gained popularity amongst Russians in general and amongst those opposing the powers in Kremlin in particular.

Russians have embraced social media ever since Internet access became available. Out of approximately 70 million Internet users 83% are active within social media spending about 10.4 hours per month on average surfing the sites of LiveJournal, VKontakte and co. (ComScore 2011). A likely explanation for the intense use of social media in Russia is the comparably young audience: the absolute majority of users are between 25 and 40 years old (Butenko, Hraybe 2012). The prevailing mistrust in official mass-media outlets seems to be another plausible factor for the popularity of the self-selected and self-created online content.

The idea of self-created content is not new to Russia. During the time before the October Revolution self-created content was produced and disseminated in the underground by activists. A culture of the so-called samizdat, literally self-publishing, developed in the Soviet Union and became a backbone of the dissident activity. In this way the avid attraction to social networks can be seen as a continuation of a discourse aside from the state-friendly or potentially state-controlled mass media, in a domain that promises a certain level of anonymity.

This offers an explanation for the politicization of the Russian social media. The protests of 2011/12 have shown a widespread strategic use of social media for the political narrative, and for the organization of off-line sociopolitical action. In the last century critics of the Russian government would spread hand-typed and copied anecdotes in the underground, now they do so in VKontakte groups. Social media constitutes an alternative platform for exchange of ideas by active news consumer as a contrast to passive consumption of state-controlled mass media. According to Liudmila Novichenkova this "enabled ordinary citizens to engage in political and social activism" (Daveluy 2012).

3. Putin: "Internet is like a knife in the hand of a criminal or a doctor. In one case it kills, in the other it heals"

Technological advancement is crucial for meeting the Russian economic and societal needs. Already in 2000 Putin recognized the importance of information technology: "Our country is involved in all international processes including economic globalization. We also have no right to "sleep through" the information revolution that is unfolding in the world" (Putin 2000). Speaking at a meeting with the Supervisory Board of the Agency for Strategic Initiatives in 2012, the President suggested the feasibility of a special fund "through which Internet initiatives will be selected and funded, that have a high social value, to address public interest issues" (RIANovosti 2012a).

In early 2012 the public protests that were fueled by zealous actions online, made denying the relevance of the Internet, and social media in particular, for the political discourse impossible. In February he told RIANovosti (2012b) that "social media is a serious means of modern communication". He expressed little concern for the "false material" about himself that was spread online by the opposition, instead, he urged his supporters to adapt to the new media and to voice their opinions in a more effective and talented way than the opposition, using the same platform: "(Our) Response has to be on the same platform. (We) need to respond on the same platform. (We) shouldn't prohibit and expel, act upon the principle "grab and don't let go"". "So that the people [...] can get a different point of view, formal or informal, but one that appeals to them and is based on the realities of life." he continued. He also stated that censorship is impossible and prohibition would not be an adequate response to opposing forces: "Is it possible to control the Internet? It can only be banned... It is the worst that can be done".

Putin's disapproval of online censorship, however, does not include malicious activity, that has to be met with strength and determination: "Nothing should be prohibited. Criminals on the Internet are the only actors the state has to keep at bay. I think everyone sitting here would agree that when... Internet, let's say, is used by paedophiles, (or) by other criminals, that the society must find some ways to protect itself" (RIANovosti 2012b). Acknowledging that the Internet is not solely a platform for benign political discourse and repeatedly calls for virtue in its use: "Internet is like a knife in the hand of a criminal or a doctor. In one case it kills, in the other it heals. Let us not forbid anything. Let us simply work effectively using this tool in a more talented and efficient way than those people who use it for vile purposes" (RIANovosti 2012b).

But Russia's interests do not end at its national borders. "Russia is a part of the larger world whether we are talking about the economy, or information dissemination, or culture." Putin (2012b) wrote in an article for Moskovskye Novosti. "We cannot and we do not want to isolate ourselves." However, Russia will act upon it's own "interests and goals, rather than based on decisions dictated by others." he stated, recognizing that "the internet, social networks, mobile phones and the like, along with the TV have become an effective instrument of both domestic and international politics. [...] The concept of "soft power" is also gaining popularity - a set of tools and methods to achieve foreign policy goals without the use of arms, but through informational and other levers of influence." " Regrettably," Putin added, "such methods are all too often being used to develop and provoke extremist, separatist and nationalistic attitudes, to manipulate the public and exercise direct interference over the domestic policies of sovereign states."

The president has been very consistent in stating that policymaking is nothing that can or should be copied, and that Russia should strive for its own political order based on its own values, meeting its own needs. In the yearly State of the Nation on 12th December 2012 Putin told the Russian Duma and the Federal Council that there is no other choice for Russia but to be a democratic country, on it's own terms. "Russian democracy is the rule of the Russian people, with their own tradition, and not [...] standards forced upon us from abroad". Russia is a multinational country that has to remain unified by language and culture, he stressed, reminding that Russia has "1,000 years of history, not only world war I or 1917". This patriotism and what it is to be a Russian should give the people "inner strength" the President pleaded, "Today the Russian society clearly experiences a deficit of spiritual ties that at all times in our history have made us stronger". And strength is, according to Putin, a necessity for building and upholding democracy. In an article for Foreign Policy Journal Putin (2012c) wrote that "We will not be able to strengthen our international position or develop our economy or our democratic institutions if we are unable to protect Russia".

In the strive for Russia's independence Putin (2012a) expresses concern about foreign influence: "Any direct or indirect outside interference in our internal political process is unacceptable," he said, particularly concerned with political activity that is financed by foreign actors. This concern had been addressed in practice earlier in 2012 with a law requiring NGOs that are financed from abroad to register as "foreign agents" (RIANovosti 2012c). "People who receive money from abroad for their political activities – most likely serving foreign national interests – cannot be politicians in the Russian Federation" declared Putin.

The task of preventing such interferences online partly falls under the responsibilities of the FSB. Subsequently the President urged the organization to "continue to act systematically and aggressively. Including areas such as counter-intelligence, protection of strategic infrastructure, the fight against crimes in economy and cyberspace" in late 2012. "Protection of the rights and freedoms of citizens, countering terrorism and extremism, crime, and corruption" are the first and foremost priorities of law enforcement and intelligence agencies, he said according to RIANovosti (2012d).

Although Putin encourages the administration and his supporters to master the Internet and the social media, he shows reluctance in using these tools himself. According to Howard (2012) "Putin is media savvy, but his skills are in broadcast media. The Kremlin knows how to manage broadcast media". Ever since his first presidency, Putin has been able to keep big PR catastrophes at a minimum and to retain a heroic image of himself. He has played with tigers, trained martial arts, flown military jets, and participated in firefighting efforts. However, the latest polls show a growing dissatisfaction with Putin's public relations campaign reports

Osipov (2012). Suddenly president Putin is “waking up to the fact that Russia’s media landscape is not the one he inherited in 2000” (Weaver 2012). Social media cannot be managed applying the principles that have proven themselves successful in traditional media and thus poses a challenge for Putin and his PR team.

With so much recognition for the importance of social media it is almost surprising that Putin isn’t a social media user himself: “And I won’t hide: I don’t use it. Honestly. I don’t have the time to sit and poke in there, read, reply, write. It is pointless to entrust someone else with the task. The result would be formalized and not very interesting, you’d get it soon. And then they’ll say: see, it’s not he himself who’s posting”, as quoted by RIANovosti, (2012e). Nonetheless, he keeps the option open to take the step into the world of social media: "However, I will assume that this is your kind comment, suggestion, which would improve the situation to some extent, I will consider it", he said replying to whether he is going to register a personal account in social media. At the time of writing there are two pairs of Twitter accounts associated with the President - accounts in Russian are accompanied by mirror accounts in English. The content consists for the most part of links to news posted on www.kremlin.ru. However, there is also room for Putin to communicate himself: the tag #ВП (#VP), as it says in the Twitter profile of PutinRF (PutinRF_Eng), would identify his personal Tweets. To date the hash tag has not been used.

4. Medvedev: “If a politician can’t master these tools, he has no future”

Prime Minister Dmitry Medvedev has a different relationship with social media than Putin. Medvedev became known as the “blogging president”, who inspired Russians to use Internet and social media. Medvedev is active on the most popular social media platforms including Vkontakte, Facebook, and Twitter. It is disputable whether Medvedev was the first Russian politician use social media, however he is the only Russian top-level politician with a well-organized personal online presence. During his presidency Medvedev maintained two presidential and two personal Twitter accounts. The presidential accounts were handed over to Putin as he was inaugurated for his third term as a president (dp.ru 2012). The personal accounts, however, are still in use. Medvedev also has accounts in Facebook and Vkontakte. He disclosed having an account on Odnoklassniki.ru, noting however that it is difficult to find him there because of the about 600 other accounts using the same name, a part of those people “seem very similar to myself, almost like twins”, Medvedev (2008) said.

Medvedev also maintains several regular and video blogs and a well-used account on Instagram. Here he regularly posts pictures taken during his travels. Since the new posts on his social media presences seem to be very coordinated, redundant and do not necessarily have a “personal feel” to them, the use of privately taken pictures seems to add “personality”. Furthermore the prime minister sometimes loosely engages in discussions that arise in relation to his posts. All in all he is a versed social media user and he is proud of it: “I have some 700,000 likes on Facebook or close to it” he said during a meeting with his Finnish counterpart Jyrki Katainen. “I am always ready to share with friends, “ he added after the comparison to Katainen’s moderate popularity in social media (Medvedev 2012a).

When the government office analyzed citizens’ complaints that were submitted to the White House, they found that a lot of citizens’ complaints, many of them regarding municipal problems, were directed at Medvedev personally (MKRU 2013). Of course the perception of Medvedev as a politician who is close to people and genuinely concerned with their grief cannot be totally contributed to his activity in social media. However, it is certainly a contributing factor, a picture Medvedev endorses in interviews: “I personally read such requests [...] via Facebook, Twitter and other social networks or through my website - at least 50 a day. And on substantial matters I issue orders directly, and sometimes I, before leaving to work, go online myself, see something utterly important, something extremely difficult for the country, I push the printer button, print the relevant document and give orders right on it [same document]”.

Medvedev (2011) also urges his peers to do the same: “I believe that the government, of course, without experiencing a pressure as a whole, has to respond to what is happening in this sphere, has to be up to date and take in account the opinions that are expressed, including (those) on the Internet. It is an obligatory requirement for any politician at any level in today’s life”. “If a politician can’t master these tools, he has no future”, he told the British newspaper Times (2012b).

Picking up on Russian’s avid political discourse in the social media Medvedev passionately advocates the Open Government (formerly Big Government) initiative (RIANovosti 2011). The initiative is not only a web portal for state and government-related information but is also meant to add a crowdsourcing element to legislative work (Adomanis 2012).

“The new information environment” is in Medvedev’s (2012) opinion “is the best guarantee, the best inoculation against totalitarianism and the return to the [our] sad past”. Government critics are nothing that state leadership should be overly concerned about, he said: “What can I say? Let them criticize both President Putin and Prime Minister Medvedev. I think this is what democracy is all about. This is absolutely normal and will continue to be in social media. I believe my colleague and other officials are also being criticized in social media in Finland. This is nothing special. This is normal” Medvedev (2012b) said during the earlier mentioned meeting with Katainen.

However, Medvedev differs between criticizing the government and engaging in dissemination of humiliating false information, which he equates to spreading child pornography and advocating terrorism: "I think that today, no one has any doubt that the online publication of false information that discredits and humiliates personal dignity or discredits professional reputation, dissemination of child pornography, promoting terrorism, ethnic or religious hatred - must be severely punished”. Nevertheless, “This is not, and never will be about any kind of Internet censorship”, he said, quoted by BBC (2012). “It is impossible, I have talked about this many times. It is simply senseless.” Creating meaningful regulations on the Internet is not a simple process, according to Medvedev (2012c) “the Internet has to be managed with a set of rules that the humanity yet has to develop. This is the most difficult process because everything must not be regulated, on the other hand everything can’t be left outside the legal field”.

5. Conclusions

There is a share of ambiguity in Putin’s personal relationship with the Internet and social media: Although urging others to use and master these tools he himself remains offline. As a reporter confronted him with rumors about the president’s poor health during a large press conference in December 2012, “because there is a lot of information about it on the internet”, Putin jokingly replied: “Don’t look at it (the internet) too much, they will teach you bad things”. Even though this was a joke, President Putin’s narratives make clear that he, personally, is not a social media man.

Medvedev on the other hand is almost omnipresent in the social media, creating an image of himself as open, modern, working to create a strong relationship with the population. For the now-prime minister, the Internet is a tool for politicians to improve the situation in the country, and for the people to participate in politics. Although Medvedev leaves room for online regulation, in his words it is rather about game rules than prohibition. The belief that Internet should be largely free prevails in his statements.

Medvedev’s positivity towards the Internet and social media seem visionary if not naïve in the contrast to Putin’s verbal hardness towards the dangers to the society: Any foreign influence is regarded to be an intrusion into Russia’s business. The traditional concept of sovereignty is vital in the set of cultural values that Putin regards to be the glue of the Russian society, the foundation for progress. Ultimately the Internet poses a threat to sovereignty and in the light of this the question is whether Putin may be just as internet-savvy as Medvedev - by largely staying out of it.

“All progress will happen only with power” - Leo Tolstoy wrote in his masterpiece Anna Karenina. A concept that Russian leaders seem to comply with: Strengthening the population and creating opportunities for the “honest man” through aggressive action against those who wish to do harm is what the state leadership seems to be communicating in regard to social media. A contradiction in itself, a message that can be perceived in a number of contradictory ways.

Hidden in the overall context, there are messages that are brought forward by the Russian state leadership, seemingly aimed at the internal audience, that are also meant for the international arena. Particularly the use of English-language mirror accounts is evidence for this. Moreover messages that are directed at the internal audience reach the international community, where they can be taken out of context and thus be perceived for what they are not. This poses a challenge for Russia’s leadership to design their messages to fit both the internal and external audience, demanding a new level of cultural awareness.

Russians have lived trapped between the soviet era's conformity and disbelief in the political system for a long time. After the liberation from a restrictive system the bare mention of regulating the allegory of western-like freedom, the Internet, can be a reason for conflict. However, thrown from a rigid system into the near-lawlessness of the 90's the nation also longs for a rule of law that would protect the honest citizens from criminals, off- and online, domestic and foreign. Thus, the seeds of Internet regulation fall onto a fertile ground of the Russian reality. Furthermore, it's not new for Russia to seek prosperity in strength. This ambiguity of popular views might well reflect the need for both leaders' stances to be present - the progressiveness of Medvedev and traditionalism of Putin.

"Russia is a museum of contradictory truths," wrote Remy de Gourmont. Contradiction is part of the life and narratives in Russia that is fighting to recreate itself in a new, connected world.

References

Adomanis, M. (2012) "Open Government a la Russe: How the Russian Government is Trying to Modernize", [online], *Forbes*, 12. November, accessed 1 February 2013, http://www.forbes.com/sites/markadomanis/2012/11/12/open-government-a-la-russe-how-the-russian-government-is-trying-to-modernize/.

BBC (2012) Медведев: цензура в интернете нереальна и бессмысленна, Medvedev: censorship on the internet is unrealistic and meaningless, [online], *BBC Russkaya Sluzhba*, 18 April, accessed 1 February 2013, http://www.bbc.co.uk/russian/mobile/russia/2012/04/120418_internet_cenzura_medvedev.shtml.

Blagoveshensky, A. (2012) Google заблокировали по ошибке, Google blocked by mistake, [online], *Rossiyskaya Gazeta*, 26 November, accessed 1 February 2013, http://www.rg.ru/2012/11/26/google-site.html.

Butenko ,V. and Hraybe, F. (2012) Рунет вырос, но не повзрослел, RuNet has grown, but not matured, [online], *Forbes*. 18. April, accessed 1 February 2013, http://www.forbes.ru/sobytiya-column/rynki/81236-runet-vyros-no-ne-povzroslel.

ComScore (2011) Social Networking Leads as Top Online Activity Globally, [online], *ComScore*, accessed 1 February 2013,http://www.comscore.com/Insights/Press_Releases/2011/12/Social_Networking_Leads_as_Top_Online_Activity_Globally.

Daveluy, A (2012) The landscape of digital technologies in Russia, [online], *Digital Tech, The review of creative industries and media*, 24 October, accessed 1 February 2013, http://www.inaglobal.fr/en/digital-tech/article/landscape-digital-technologies-russia

dp.ru (2012) Дмитрий Медведев поменяет адрес своего сайта, но аккаунты в соцсетях сохранит, Dmitriy Medvedev to change the address to his website, but to keep the social media accounts, [online], *dp.ru*, 5. May, accessed 1 February 2013, http://www.dp.ru/a/2012/05/05/Dmitrij_Medvedev_pomenjaet/.

Howard (2012) Social media and the new Cold War. [online] Reuters, 1. August, accessed 1 February 2013, http://blogs.reuters.com/great-debate/2012/08/01/social-media-and-the-new-cold-war/.

Jaitner (Forthcoming, 2013) "The Power of Social Media", in Rantapelkonen, J. and Salminen M. (eds), The Fog of Cyber Defence, Series 2: Article Collection N:o 10, National Defence University, Helsinki, Finland.

Lonkila, M. (2012) *Russian Protest On- and Offline: The role of social media in the Moscow opposition demonstrations in December 2011*, 16 February, The Finnish Institute of International Affairs, Helsinki, Finland.

Manilov, V. (n.d.) *National Security of Russia*, Occasional Paper, Strengthening Democratic Institutions Project, Belfer Center for Science and International Affairs, Harvard Kennedy School, Cambridge, Mass..

Medvedev, D. (2008) Медведев про Одноклассников, Medvedev on Odnoklassniki, [online video], 3 April, accessed 1 February 2013, http://www.youtube.com/watch?v=hUm905_8D0E.

Medvedev, D. (2011) in an interview with tv-stations "Pervy", "Rossiya" and NTV, [online] 30 September, accessed 1 February 2013, http://kremlin.ru/transcripts/12880.

Medvedev, D. (2012a) in a joint conference with Katainen, as recorded on government.ru, [online], 14 November, accessed 1 February 2013, http://government.ru/eng/docs/21472/.

Medvedev, D. (2012b) in an interview with The Times, [online] 30 July, accessed 1 February 2013, http://government.ru/docs/19842/.

Medvedev, D. (2012c) Сеть должна быть свободной, но в ней должны соблюдаться элементарные права людей, The network has to be free, but people's basic rights have to be observed, [online video], 12 July, accessed 1 February 2013, http://blog-medvedev.livejournal.com/89944.html.

MRKU (2013) What grief Russians shared with Medvedev, О чем россияне плакались Медведеву, [online], *MKRU*, 7 January, accessed 1 February 2013, http://www.mk.ru/economics/article/2013/01/07/795058-o-chem-rossiyane-plakalis-medvedevu.html.

Novye Izvestiya (2007) Путин: Поиск национальной идеи – старинная русская забава. Putin: The search for the national idea - an old Russian pastime [online] *Novye Izvestiya*, 26 April, accessed 1 February 2013, http://www.newizv.ru/lenta/2007-04-26/68681-putin-poisk-nacionalnoj-idei-starinnaja-russkaja-zabava.html.

Osipov, I. (2012) Россияне устали от пиара Путина, Russians tired of Putin's PR, [online], *Forbes*, 24 October, accessed 1 February 2013, http://www.forbes.ru/sobytiya/vlast/178751-rossiyane-ustali-ot-piara-putina.

Putin V. (2000) *Послание Федеральному Собранию Российской Федерации, State of nation*, [online], 8 July, accessed 1 February 2013, http://archive.kremlin.ru/text/appears/2000/07/28782.shtml.

Putin, V. (2012a) *Послание Президента Федеральному Собранию, State of Nation*, [online audio], 12 December, accessed 1 February 2013, http://kremlin.ru/audio/792.

Putin, V. (2012b) Россия и меняющийся мир, Russia and the changing world, [online], Moskovskye *Novosti*, 27 February, accessed 1 February 2013, http://www.mn.ru/politics/20120227/312306749.html.

Putin, V. (2012c) Being Strong, [online], *Foreign Policy*, 21 February, accessed 1 February 2013, http://www.foreignpolicy.com/articles/2012/02/21/being_strong.

RIANovosti (2011) Медведев пообещал "приглядывать" за сайтом большого правительства, Medvedev promised to “keep an eye” on the big government webpage, [online], *RIANovosti*, 9 November, http://ria.ru/society/20111109/484771261.html

RIANovosti (2012a) Путин предложил учредить фонд для финансирования интернет-проектов, Putin suggested a fund for financing of Internet-projects, [online] *RIANovosti*, 22 November, accessed 1 February 2013, http://ria.ru/economy/20121122/911755293.html

RIANovosti (2012b) Путин призывает своих сторонников быть более креативными в интернете, Putin urges his supporters to be more creative on Internet [online], *RIANovosti*, 1 February, accessed 1 February 2013, http://ria.ru/politics/20120201/553931560.html.

RIANovosti (2012c) NGO 'Foreign Agents' Law Comes into Force in Russia, [online], *RIANovosti*, 20 November, accessed 1 February 2013, http://en.rian.ru/russia/20121120/177597400.html.

RIANovosti (2012d) Путин призвал ФСБ активнее действовать в киберпространстве, Putin urges FSB to be more active in cyberspace, [online], *RIANovosti*, 28 December, accessed 1 February 2013, http://ria.ru/politics/20121228/916628274.html.

RIANovosti (2012e) Владимир Путин против социальных сетей, Vladimir Putin against social networks, [online], *RIANovosti*, 2 February, accessed 1 February 2013, http://ria.ru/society/20120201/553968947.html

SecurityLab (2012) Роскомнадзор заблокировал и разблокировал доступ к YouTube, Roskomnadzor blocked and unblocked access to YouTube, [online], *SecurityLab*, 23 November, accessed 1 February 2013, http://www.securitylab.ru/news/432738.php

Weaver, C. (2012) Social networks pose challenge to Putin, [online] *The Financial Times*, Feb 27, 2012, accessed 1 February 2013, http://www.ft.com/cms/s/0/f21f59d0-5d51-11e1-889d-00144feabdc0.html#axzz2IazQQMxi

Can Keys be Hidden Inside the CPU on Modern Windows Host

Amit Resh and Nezer Zaidenberg
University of Jyväskylä, Jyväskylä, Finland
amit@trulyprotect.com
nezer@trulyprotect.com

Abstract: The “Truly-Protect” trusted computing environment by Averbuch et al (2011) relies on encryption keys being hidden from external software and crackers. “Truly-Protect” saves the keys in internal registers inside the CPU. Such external keys should not be accessible by any software that runs on the machine prior to “Truly-Protect” validation or even after “Truly-Protect” validation. The assumption is that the hackers cannot reverse engineer the CPU and discover the content of these registers. But is it really possible to hide keys in such places? Internal CPU memory is indeed not available for user processes. However, the CPU memory and registers are accessible from the running operating system kernel. Truly protect uses a validation protocol that also verifies the Operating system kernel does not include malicious additions. These tests should ensure a cracker has not modified the OS. But Modern Windows operating system support loading new kernel code segments (drivers) even during the operating system runtime. Can we prevent modifying the kernel (loading drivers) after “Truly-protect” has verified the kernel? In this work we examine modern Intel CPUs available on desktop PCs and the latest releases of Microsoft Windows (windows 7,8) for existence of good hiding places for the encryption keys.

Keywords: virtualization, copy-protection, CPU architecture

1. Introduction

Contemporary digital rights security systems rely mainly on methods of obfuscation or use of plug-in HW devices, such as dongles. Use of HW dongles has been critiqued heavily by users, as being cumbersome and generally inconvenient. Obfuscation methods, by which software protection is realized by introducing code-clutter to conceal the protection mechanism is largely losing the battle to crackers, who on average can break these protection schemes within weeks. A new approach, described by Averbuch et al(2011) suggests a software-only solution, named Truly-Protect, based on encryption and just-in-time decryption of protected software. According to this approach, the protected software shall be stored in computer memory exclusively in its encrypted form. Decryption shall occur "inside the CPU", on-the-fly, as it is being consumed. The decrypted form shall not be stored back into memory. In fact, it shall never leave the confines of the CPU domain. See shaded area in Figure 1.

Software protection based mainly on obfuscation still allows crackers to trace and reverse-engineer the protected software, thereby opening the door to obtaining an unprotected copy. However, by keeping the decryption process and its keys, as well as the decrypted results inside the CPU domain assures that the software remains protected -- unbreakable by any of the currently know cracking techniques.

According to Truly-Protect (Averbuch et al(2011)) the following procedure is used in order to successfully execute protected software on a target computer:

- The target computer communicates to a remote authentication server and transmits proof of eligibility to execute protected software.
- The remote server authenticates the target computer by employing a modified Kennel & Jamieson (2003) procedure. The purpose of this step is to validate the target computer as a real (non-Virtual) machine running a recognized O/S. A side-effect of the validation procedure is exchange of key material.
- The server protects the software by encrypting it using the key material exchanged with the target during the validation procedure.
- The protected (encrypted) software is downloaded to the target's memory and spawned for execution.
- Protected software executes on the target computer using JIT, on-the-fly, decryption: Encrypted instruction code is loaded from memory into the CPU, where it is decoded, executed and then disposed. Decryption keys or decrypted instruction codes never leave the CPU domain.

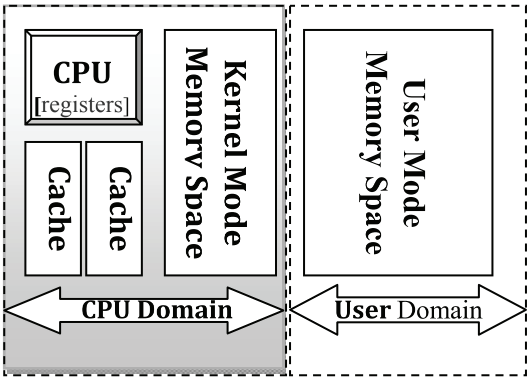


Figure 1: CPU and memory structure

2. Problem definition

A full description of the validation and key-material exchange procedures, detailed in the procedure above, is beyond the scope of this discussion. We will proceed with the assumption that the validation procedure establishes the following:

- The target computer is a real (non-VM) system
- The target is running a recognized O/S that does not include potentially malicious components
- The key material, required for decrypting the protected software, is generated by the validation procedure and stored in the CPU domain

The Truly-Protect scheme is based on maintaining keys and carrying out the protected software decryption exclusively in the CPU domain. This implies that the decryption code runs in Kernel-mode (privilege level 0) on a protected O/S, such as Windows or Linux. This further implies that decryption must either be an integral part of the O/S or a Driver that is loaded into the O/S and operates in Kernel-mode. While this restriction in itself is a complication, it is a blessing in terms of software protection, since it establishes a basis upon which the Truly-Protect goals can be realized.

The protected software is assumed to execute in user-mode. However, according to Truly-Protect, decrypted code cannot exist outside of the CPU domain. This restriction implies that decrypted code must either:

- remain in the CPU register file for the duration of its execution, or
- be latched in cache while the cache method for that space is set to Write-Back

In the former case individual instructions must be decrypted and executed by a VM, while in the latter case, large blocks of code (for example, entire functions) may be decrypted and executed natively, directly from cache. These ideas have been described in detail in Averbuch et al (2011).

The most crucial aspect of Truly-Protect and its "soft-belly" is the decryption-key location. Once generated by the validation procedure, it must be locked in the CPU domain, such that it cannot be accessed under any condition, except, of course, to carry out the JIT decryption. As mentioned above, locating information in the CPU domain restricts its access to the O/S Kernel or driver modules executing in Kernel-mode. Therefore, storing the keys anywhere in the CPU domain will keep it safe from User-mode applications. We assume that during the validation procedure, when the key is initially generated, the CPU domain is clean of malicious code. However, how can Truly-Protect guarantee that malicious Kernel-mode drivers are not loaded thereafter, gain access to the CPU domain and get hold of the key -- thereby using it to completely decrypt the protected software?

Several approaches may be employed to accommodate:

- Lock the key in a memory region that can be accessed only by the decryption engine
- Prevent driver plug-ins to the O/S after the validation procedure has successfully completed.
- Allow Kernel related changes or driver additions – but in the event that these occur – the key must be obliterated

3. Discussion of alternative solutions

Cache

Cache memory is one good storage place, in which to conceal key information deep within the confines of the CPU domain. Most modern age computer systems contain one or more cache units for Instruction, Data or Unified caching. For example, the Intel Pentium processors contain 3 levels of cache units: L1 (Instruction and Data), L2 (Unified) and L3 (Unified).

Cache memory cannot be read or written directly by software (User or Kernel mode) as internal cache mechanisms maintain correspondence between cache and physical memory contents. Therefore, cache contents are read/written only by accessing the memory locations shadowed by cache. However, since delays between introducing new data (writing) to a cached location and when that data is actually committed to physical memory can be taken advantage of to store data in cache while keeping it out of physical memory. A procedure for achieving this is:

- Configure memory location as type WB (Write-Back)
- Read memory location (cache lines are filled)
- Write critical information to memory location (only cache is written)

Following this, Reads from the memory location will return the critical contents from cache. When done, the cache can be overwritten and invalidated. Using this technique, the critical information is never written out to physical memory. This has the distinct advantage of not compromising the critical information to a bus-analyzer, as well as not providing a possibility for physical memory to be polled or extracted from the main board for analysis.

Keys or decrypted data may be manipulated in cache memory using the above technique. For keys, either data-cache or instruction-cache may be used: by storing keys directly in the former case or setting up an instruction sequence that generates a key in the latter.

However, there are several limitations worth mentioning. Storage of critical information in cache, in the interim where it does not get written through to memory, can only be maintained temporarily, since most cache invalidation procedures that occur internally will cause cached data to be written out to physical memory.

Furthermore, cached locations may be read by any process that has access to the address space being cached. Therefore, other processes that gain CPU control while the cache contains critical data may, in theory, obtain access to this data.

Registers

Registers are an appropriate storage location for keys, since they are located deep in the CPU domain and are never implicitly written out to physical memory. Not all registers are suitable for storage of decryption keys. Most contain values that have significant implication on execution flow, such as general purpose registers, registers that point to significant memory locations or registers that contain operational flags.
The Intel architecture includes registers under two major categories:

- Basic Program Execution Registers
- System-Level Registers

Truly Protect focuses on the latter, since most system-level registers are protected from user-applications and may only be accessed from Kernel-mode (privilege level 0). This provides better control over the possibilities for keeping keys locked in CPU and out of reach of malicious code. As mentioned above, system-registers that are suitable for storing arbitrary data without affecting execution flow are the best potential candidates for key storage.

Debug Registers

The Intel architecture contains 8 debug registers (DR0-DR7). DR6 and DR7 are used to report and configure breakpoint conditions; DR4-DR5 are reserved and DR0-DR3 are used to store required breakpoint addresses. Since it can safely be assumed that debugging breakpoints will not be used (and will actually be prohibited) while Truly-protect actively protects a system, the 4 breakpoint address registers, DR0-DR3, can be used to store a key. In a 32bit system, each of DR0-DR3 is 32 bits wide. Therefore, this totals 128 bits of key information.

To ensure that a breakpoint does not occur at some arbitrary address, which happens to be part of the truly-protect key, the DR7 register is configured to disable the 4 DR0-DR3 breakpoints. An extremely useful facility is the DR7.GD[bit 13]. If this bit is set a #DB exception is generated if any of the debug registers (including DR7) are accessed. While not enough to guarantee that no other Kernel-mode program maliciously gains access to DR0-DR3, this facility may be used to control such access as part of a larger key-protection scheme.

Model Specific Registers (MSRs)

The MSRs are a group of system-registers used to report or configure a variety of system-related attributes. They may be used, amongst others, to control debug extensions, performance-monitoring, machine-check and memory type range definition (MTRRs). The majority of these registers cannot be used to store arbitrary values, however we will seek those that can.

Different Intel processor families have slightly different MSRs, so that MSR usage needs to rely on their availability in the current system. This can be verified programmatically with the CPUID instruction.

Performance counters are the most readily available MSR registers for storage. The most basic Intel architecture contains two 32 bit counters. Truly protect takes advantage of the performance counters during the validation process. However once that is complete, the counter registers can freely be used to store decryption keys. Counter register load commands of arbitrary values are supported in Kernel-mode and their corresponding control-registers can be configured to disable counting, thus ensuring that the preloaded values do not change. Both counters total 64 bits of key information.

Dynamic Keys

Dynamic keys are keys, or key modifiers, that are computed temporarily at run-time. They are computed in close propinquity to where they are needed for decryption and then immediately disposed of. Therefore, in a sense they are not stored anywhere, beyond the short period of time when used for decryption. Consequently, they may be stored in any of the CPUs general-purpose registers, provided that no other task can gain access to the CPU during that time.

Dynamic-keys have the distinct advantage of not needing any prolonged storage, therefore no need to find a hiding place somewhere inside the CPU, unreachable to malicious programs. However, while that being true, the code required to compute the key does need to be stored somewhere and because code is relatively large (compared to a key) it must be stored in memory rather than in some internal register in the CPU. While this may seem like a tombstone for that idea – not all is lost. It may still be possible to write a dynamic-key calculation routine, whose instructions are not secret – rather its execution is controlled such that the calculation will be correct only if invoked by a legitimate source.

Two useful tools may be recruited for this task. The first is to make use of the performance counters to count HW side-effects in the process of generating a dynamic-key. The advantage of this is clear: Dynamic-key values cannot be calculated by reverse-engineering the calculation routine. The routine must actually be executed on the target machine in order to achieve the correct results. The second is the validation process, which runs just before any keys are introduced into the system. The validation process guarantees a clean (of malicious code) system. This gives us an opportunity to setup a software "mouse trap" around our dynamic-key calculation routine. Taking this simile a step forward: the rational is that if the mouse (malicious program) goes for the cheese (calculation routine) the trap (exception) is triggered. Dissimilar to the real mouse-trap case, the software incarnation can either catch and be rid of the mouse or it can annihilate the cheese, and so to speak, leave the mouse hungry. To sum, dynamic-key generators may be used to render decryption key material during run-time, alleviating the need for protected key storage – provided that the generator code cannot be reverse-engineered and its execution is controlled.

Avoiding Kernel-mode Plug-ins

A common problem associated with all the above proposed key storage locations is their potential vulnerability to malicious kernel mode drivers. The Truly-Protect system suggests that a validity check shall be carried out as part of the encryption and key setup procedure. The validation verifies that the target system is real (non-virtual machine), running a recognized O/S version and does not contain malicious Kernel-mode drivers. In other words, it is safe to install the encrypted version of the software in the target's memory and store the decryption keys in the CPU domain. From this point on the encrypted software executes while simultaneously being decrypted by the Truly-Protect JIT decrypt engine.

If at any point a malicious Kernel-mode driver is plugged in to the system while the protected software is executing, that driver may access the key storage locations, acquire the keys and use them to decrypt and obtain the protected software. To successfully protect the keys, the Truly-Protect system must either completely prevent plugging in Kernel-mode drivers while the protected software is executing or obliterate the keys if such a plug-in occurs. The Windows O/S, for example, supports driver plug-in as a standard procedure, therefore it is assumed to be difficult to enforce complete driver-load prevention. Consequently, the authors believe that the latter alternative, calling for key obliteration in the event that a Kernel-mode driver is loaded while the Truly-Protect system is active – is a more realistic approach. This warrants that the protection system be aware of any attempt to add a new Kernel-mode driver to the system and be alerted in time to obliterate the key. The success of this approach also heavily relies on the quality of the validation process, which must substantiate a "clean system", in the sense that no malicious Kernel-mode drivers exist at the time the key material is generated.

4. Conclusions and future work

The Truly-protect software-only protection system is based on executing encrypted software by decrypting it just-in-time during execution. Every execution unit (instruction or routine) is decrypted at the moment it is needed and the decrypted incarnation is purged immediately upon its completion. To achieve this, decryption keys must be present during runtime and the decryption keys must be hermetically guarded from malicious programs. The internals of the CPU are considered the safest place to store and guard the keys. Therefore, Truly-protect is designed to use the keys for decryption without the keys ever leaving the internal confines of the CPU. This means they do not exist in memory and are never present on any of the external system buses.
Several storage places, inside the CPU, were considered:

- Cache – has an appropriate storage state which may contain values that are different from the memory it shadows. However, since this state is highly instable it can only be utilized for short periods.
- System Registers – are an appropriate storage place that is protected from all application level programs. However, is susceptible to prying by malicious kernel-mode drivers.
- Dynamic-Keys – these do not need prolonged storage (beyond the decryption process). Nevertheless, the code required to generate the dynamic-keys must be protected against malicious invoking.

Storage of key material inside the CPU domain is safe from User-mode programs but not from Kernel-mode drivers. Means must be provided to safe-guard keys from malicious drivers that may already exist in the system or are loaded while protected software is executing.

No single solution amply solves all aspects of protecting keys on a target system, such that they cannot be confiscated by malicious code. Our future and on-going efforts are focused on combining several such solutions in order to provide fully-protected, software only, DRM solutions.

References

Averbuch A., Kiperberg, M. Zaidenberg N.. An Efficient VM-Based Software Protection. In NSS 2011.

Kennell Rick and Jamieson Leah H.. Establishing the genuinity of remote computer systems. In Proceedings of the 12th USENIX Security Symposium, 2003.

Scan Detection System Using Artificial Neural Networks

Francisco Ribeiro and Henrique Santos
Minho University, Braga, Portugal
fr.fmpt@gmail.com
hsantos@dsi.uminho.pt

Abstract: With the growth and expansion of the Internet, the world has become smaller. Nowadays, when it comes to communication, we don't usually think about borders or distances, since we can easily communicate with anyone anywhere in the world, using a cheap resource like an Internet connection. Furthermore, with the growing complexity of information technology infrastructures, like clouds, even when external deterrence is granted, no one can completely assure that information is safe and controlled. In parallel with the enormous flexibility and capacity of the internet, it is necessary to recall that the presence of sensitive data roaming, the Internet or traversing obscure technological layers may lead to attacks targeting critical data or vulnerable network devices, which is the essence of information warfare activities.
At the moment, with the dissemination of Social Networks, streaming and other popular internet contents, billions of Terabytes of information are transitioned over the internet. This fact allows that malicious activity roams over the internet, hidden in legitimate look alike traffic. Those activities can only be identified by sophisticated Intrusion Detection Systems (IDS). However, and despite their evolution, due to the huge quantity of events they need to look for, IDSs still produce a great number of false positives, leading to a huge efficiency reduction. The main goal of this work is to demonstrate how a modified IDS (Artificial Neural Networks plus Snort) can be used to reduce false positives generation. In order to achieve our goal, it has been developed a Java application, capable of capture network data, which is processed using artificial neural networks and self-learning methods. Those self-learning methods allow the improvement of the neural network false positive generation rate. We set up this prototype, and monitored, for more than 30 days, a general company network serving several employees. During this time, all anomalies were recorded in a MySQL database for posterior analysis. Our detection results were compared with the ones obtained with a default configured Snort. Throughout this paper we present the reasons why we chose not only this subject but also the Neural Networks technology to implement the solution. We also describe the results obtained and how we made it possible to improve the detection of false positives.

Keywords: security, intrusion, detection, internet, false, positives

1. Introduction

The computer network evolved over the years, allowing millions of devices to be connected between them, sharing and distributing information. This feature allows even a small company to have a computer network, with a server containing all data. In circumstance like this, it's really important to guarantee the security of companies information.

Despite all precautions, there are weaknesses on every system or device that is exploited by hackers, leaving sensitive information vulnerable very often. Intrusion Detections Systems are designed to detect attacks to the system where they're installed. [3]

In our study area, computer security, we understand that there are several ways to protect a system. Starting with single machine security, using viruses, spyware, malware detectors, firewalls and others, till company network security using an Intrusion Detection Systems.

It is also possible to achieve this, not only protecting every machine, but also protecting the entire network, allowing that only legitimate traffic and devices are inside the network.[1]

An IDS has an important role, when it comes to guarantee the safety of a network, since it (if well configured) detects intrusions, alerting system administrators or other security applications of that fact. For this reason around the world people are studying ways to deliver more reliable and efficient Intrusion Detection Systems.

At this point we can state that there are two types of IDSs, misuse based IDS and anomaly detection based IDS. A Misuse IDS is a system where intrusions are detected by looking for signatures known to correspond intrusions or vulnerabilities. On the other hand, an anomaly detection IDS is a system where intrusions are detected by looking for abnormal traffic in the network, like wrong attempts to connect with other devices[1]. Previous work proved that is possible to detect network scans counting traffic flags [2] during a specific amount of time. This strategy will be demonstrated later on this paper.

In the work here presented, we propose a solution based on Artificial Neural Networks and Snort IDS which, according to the obtained results, effectively reduces the Snort IDS false positive generation.

Using a multi-layer feedforward artificial neural network with backpropagation algorithm along with snort data, we are able to obtain more accurate results, not only improving the time needed to detect the intrusion, but also lowering the false positive rating.

2. Methology

The application (IDSNN1) was developed using Java language and MySql to store data, and it is divided in two modules. IDSNN1 is the combination of two modules, IDSNN1_ScanDetection and IDSNN1_SnortDetection. Both of them are constituted by an analysis method based on artificial neural networks with feedforward backpropagation algorithm, using data stored in databases. The module IDSNN1_ScanDetection obtains its data by capturing traffic using TCPDump or WinDump,

These data are then stored in the database, creating a profile for each pair (source IP address, destination IP address). In parallel, these data are being prepared and consumed by neural network modules.

IDSNN1_ScanDetection Artificial Neural Network consists of five layers. The first four contains 13 nodes and the last only one. Being the result a double between zero and one. If the value is greater than 0.5 it is considered malicious traffic.

Each node receives a specific result obtained using the following functions:

- ABS(PDa-PDb): positive result of the subtraction between the Number of Host A Destination Ports with Number of Host B Destination Ports;
- SYNa: Quantity of Host A SYN flag;
- SYNb: Quantity of Host B SYN flag;
- SYN ACKa: Quantity of Host A SYN ACK flag;
- SYN ACKb: Quantity of Host B SYN ACK flag;
- RSTa: Quantity of Host A RST flag;
- FINa: Quantity of Host A FIN flag;
- FINb: Quantity of Host B FIN flag;
- SYNa/(RSTa+RSTb): division result between the number of Host A SYN Flags with the sum of Number of Host A RST Flag with number of Host B RST Flag.
- RSTa/SYNa: division result between the number of Host A RST Flag with SYN Flag
- RSTb/SYNa: division result between the number of Host B RST Flag with Host A SYN Flag;
- Max(ICMP): the maximum number of ICMP packages;
- RSTb: the number of Host B RST Flag;

All quantity values are stored on the database, the functions results are calculated live, since variable quantities may change during time.

The ANN's IDSNN1_SnortDetection consists of five layers, with the first 4 containing 8 nodes and the last one, generating a value between one and zero. If the value is greater than 0.5, it is considered malicious traffic.

In both modules the selected data to the neural network is chosen randomly. Thus, it is not possible to predict when a particular profile will be tested.

Since IDSNN1 have the opportunity to work with several threads, the user can test different profiles at the same time, increasing the detection efficiency. Unfortunatly, this process is computationally heavy, requiring a good memory management and machine. For this reason, only strictly necessary data is stored, optimizing the use of available resources.

In order to automatically improve the results of this system, all values greater than 0.95 and less than 0.5 are added to the training dataset, which is used during IDSNN1 execution.

In cooperative mode, IDSNN1_SnortDetection runs firstly, which results are used to trigger the activation of IDSNN1_ScanDetection. In the case that Snort generates an alert IDSNN1_ScanDetection, it tests the specific profile. As a consequence of alert generation by the application IDSNN1, it is also generated a rule using the IPTables that will cause that any traffic originated in the specified IP address is ignored.

3. DataSet

In order to use the system, two separated test datasets were created: one for training the ANN and another for verifying the feasibility of the system. The dataSet training and testing were generated from two distinct sources.

To obtain packages that do not contain any type of scan, we used the DARPA Intrusion Detection Evaluation Dataset, 1999, provided by MIT and commonly accepted as a test dataset. To obtain packets regarding scans, traffic was generated internally using nmap, captured with tcpdump. In order to accomplish this task, we generated different scans, with 6 different machines (described below). It is important to consider that we used 15 different IP addresses in order to get 15 different configurations:

- Acer Aspire One ZG5, Ubuntu 114, 2Gb Ram;
- MacBook Pro late 2009, Mac os X lion, 5 Gb Ram;
- Asus EEEPC 1201NL, Windows XP SP3, 2 Gb Ram;
- Desktop, Intel P4 2.4gh, Windows XP SP1, 768 Mg Ram;
- Acer Aspire 5601, Windows 8Developers Edition, 2 Gb Ram;
- MacBook Pro late 2009, BackTrack 4 r1, 2Gb Ram, Pen Wireless;

In relation to The MacBook Pro, the same machine was used with different configurations of RAM and wireless card. In the remaining machines, the network connectivity was alternated between cable and wireless.

The IDSNN1 was tested in Windows 7, Windows 8 Developers Edition, Ubuntu 114, BackTrack 4 R1, Mac OS X Snow Leopard and Mac OS X Lion.

Some active services:

- Xampp / lampp ports 80, 8080, 443;
- Vuze / utorrent / transmission ports 52531, 56342; • RDP port 3389;
- Stocks IOW'10 port 60000;
- IOW'10 door Order 60001;
- Report IOW'10 port 60002;
- MySQL port 3306;
- sec-kerberos port 88;
- Mobile Mouse port 51101;

The scan test dataset was generated with scans made on two different days. On the second day, scans were generated in two different periods of the day, one in the afternoon and another late at night.

During the test procedure, the machines had no connectivity to the Internet and were connected between them through a Draytek 2820 router with no active firewall.

After this, it was possible to obtain Figure 1 and Figure 2 which represents the ANN training error versus the number of training iterations.

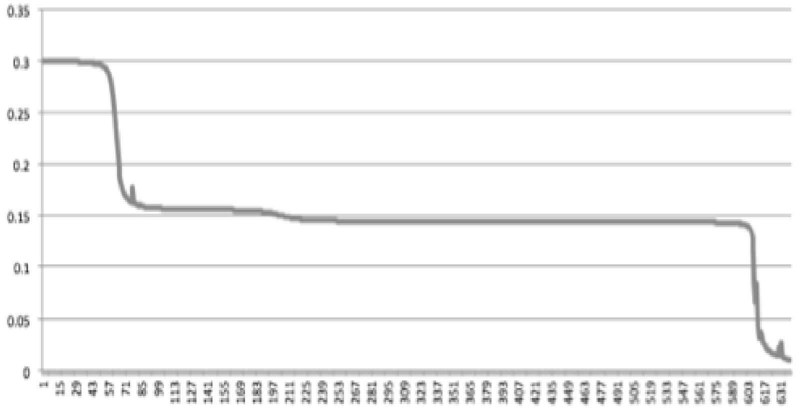


Figure 1: IDSNN1_Scandetection - error vs training iterations

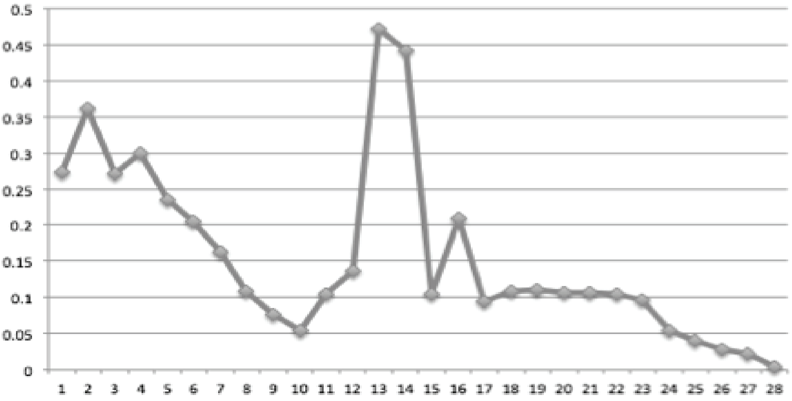


Figure 2: IDSNN1_SnortDetection - error vs training iterations

4. Results

- IDSNN1_ScanDetection

The results of running this module are presented in Table 1 and 2, representing, the first one, an extract of the first packets analyzed, and the second a summary of the overall performance.

Table 1: ISDNN1_ScanDetection DataBase test values

#	Abs(PDa-PDb)	Syna	Syn b	Syn Ack	Syn Ack	RSTa	FINa	FIN b	Syna-(RSTa	RSTa /Syn	RSTb /	Max (ICM	RST b	Resultado
1	981	2788	0	0	189	189	1295	0	2769	0067	0	0	189	0.4698
2	991	1379	0	0	35	35	0	0	1369	0025	0045	0	35	0.4698
3	267	1310	0	0	112	119	0	0	2408	0090	0.807	0	119	0.9172
4	989	1392	0	0	329	329	0	0	1355	0236	0030	35	42	0.4698
5	104	105	0	105	0	0	105	105	105	0	0	0	0	-
6	28	29	0	29	0	0	29	29	29	0	0	0	0	-0.9980
7	0	1	0	1	0	0	1	1	1	0	0	0	0	-0.9866
8	59	60	0	60	0	0	60	60	60	0	0	0	0	-0.9980
9	189	190	0	190	0	0	190	190	190	0	0	0	0	-0.998
10	109	110	0	110	0	0	110	110	110	0	0	0	0	-0.9980
11	3	3	6	3	6	0	9	9	3	0	0	0	0	-0.9955
12	27	28	0	28	0	0	28	28	28	0	0	0	0	-0.9980
13	28	29	0	29	0	0	29	29	29	0	0	0	0	-0.9980
14	75	76	0	76	0	0	76	76	76	0	0	0	0	-0.9980

Analyzing Table 1, it is possible to verify that lines 1,2,3,4 are representative of an alert, while the lines 5,6,7,8,9,10,11,12,13,14 represent legit traffic.

Through this analysis, we could observe how the scans interfere with normal network behavior. The first difference that exists between a scan and legitimate traffic is the generation of a large number of RST's, since the communication is stopped before finishing. The difference between the listened ports and ports that have responded positively is also high, since a higher number of ports are closed. In consequence there is a short generation of SynAck packets by the victim's machine.

On the other hand, legitimate traffic creates a low number RST packets, generating a number of FIN packets very similar to the number of connection initialized/set. So, we could check that, in normal situations, the result of 11.12 columns will be zero, or very close to it, while in cases of scan these values should be higher.

The particular case of line 1 represents the generation of FIN packets, so that the hacker can verify what services are available and what doors are also under surveillance by a firewall.

Table 2 represents the results obtained executing IDSNN1_ScanDetection module. Analyzing it we can figure out that this module contains a low false positive generation ratio, detecting all Scans. This test was performed using DARPA Dataset for legitimate traffic and a custom scan dataset.

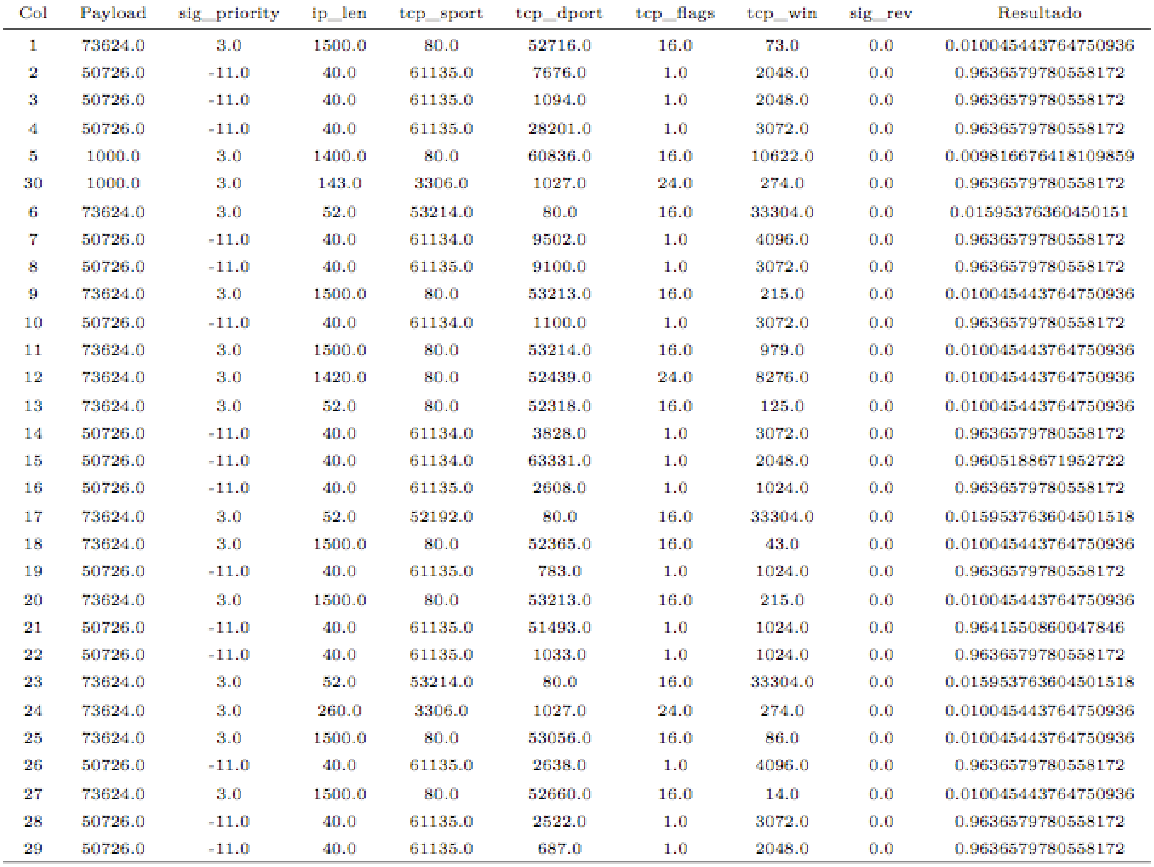
Table 2: IDSNN1_ScanDetection test results

Result Type	Legit Traffic	Scan	False Positives	%
Theorical	5485	15	X	100%
Obtained	5436	15	49	99.10%

- IDSNN1_SnortDetection

Table 3 shows the results obtained by IDSNN1_Snort for a given data set randomly taken from a database.

Table 3: IDSNN1_SnortDetection database data test



Col	Payload	sig_priority	ip_len	tcp_sport	tcp_dport	tcp_flags	tcp_win	sig_rev	Resultado
1	73624.0	3.0	1500.0	80.0	52716.0	16.0	73.0	0.0	0.010045443764750936
2	50726.0	-11.0	40.0	61135.0	7676.0	1.0	2048.0	0.0	0.9636579780558172
3	50726.0	-11.0	40.0	61135.0	1094.0	1.0	2048.0	0.0	0.9636579780558172
4	50726.0	-11.0	40.0	61135.0	28201.0	1.0	3072.0	0.0	0.9636579780558172
5	1000.0	3.0	1400.0	80.0	60836.0	16.0	10622.0	0.0	0.009816676418109859
30	1000.0	3.0	143.0	3306.0	1027.0	24.0	274.0	0.0	0.9636579780558172
6	73624.0	3.0	52.0	53214.0	80.0	16.0	33304.0	0.0	0.01595376360450151
7	50726.0	-11.0	40.0	61134.0	9502.0	1.0	4096.0	0.0	0.9636579780558172
8	50726.0	-11.0	40.0	61135.0	9100.0	1.0	3072.0	0.0	0.9636579780558172
9	73624.0	3.0	1500.0	80.0	53213.0	16.0	215.0	0.0	0.010045443764750936
10	50726.0	-11.0	40.0	61134.0	1100.0	1.0	3072.0	0.0	0.9636579780558172
11	73624.0	3.0	1500.0	80.0	53214.0	16.0	979.0	0.0	0.010045443764750936
12	73624.0	3.0	1420.0	80.0	52439.0	24.0	8276.0	0.0	0.010045443764750936
13	73624.0	3.0	52.0	80.0	52318.0	16.0	125.0	0.0	0.010045443764750936
14	50726.0	-11.0	40.0	61134.0	3828.0	1.0	3072.0	0.0	0.9636579780558172
15	50726.0	-11.0	40.0	61134.0	63331.0	1.0	2048.0	0.0	0.9605188671952722
16	50726.0	-11.0	40.0	61135.0	2608.0	1.0	1024.0	0.0	0.9636579780558172
17	73624.0	3.0	52.0	52192.0	80.0	16.0	33304.0	0.0	0.015953763604501518
18	73624.0	3.0	1500.0	80.0	52365.0	16.0	43.0	0.0	0.010045443764750936
19	50726.0	-11.0	40.0	61135.0	783.0	1.0	1024.0	0.0	0.9636579780558172
20	73624.0	3.0	1500.0	80.0	53213.0	16.0	215.0	0.0	0.010045443764750936
21	50726.0	-11.0	40.0	61135.0	51493.0	1.0	1024.0	0.0	0.9641550860047846
22	50726.0	-11.0	40.0	61135.0	1033.0	1.0	1024.0	0.0	0.9636579780558172
23	73624.0	3.0	52.0	53214.0	80.0	16.0	33304.0	0.0	0.015953763604501518
24	73624.0	3.0	260.0	3306.0	1027.0	24.0	274.0	0.0	0.010045443764750936
25	73624.0	3.0	1500.0	80.0	53056.0	16.0	86.0	0.0	0.010045443764750936
26	50726.0	-11.0	40.0	61135.0	2638.0	1.0	4096.0	0.0	0.9636579780558172
27	73624.0	3.0	1500.0	80.0	52660.0	16.0	14.0	0.0	0.010045443764750936
28	50726.0	-11.0	40.0	61135.0	2522.0	1.0	3072.0	0.0	0.9636579780558172
29	50726.0	-11.0	40.0	61135.0	687.0	1.0	2048.0	0.0	0.9636579780558172

The database contains results, both with and without scans performed, and some records are, consequently, false positives.

Table 3 contains 30 records. Lines 1, 5, 6, 9, 11, 12, 13, 17, 18, 20, 23, 24, 25, 27 and 30 are false positives generated by snort, while lines 2, 3, 4, 7, 8, 10, 14, 15, 16, 19, 21, 22, 26, 28 and 29 represent the real network scans.

It is noticeable that the values change causes an immediate change in the final result generated by ANN. In the case presented, it is apparent that the rows 5 and 30, despite the same amount of payload, have very different results. This variation is caused by the other environment variables. If we compare line 24 with 25, it is apparent that, although the value of Payload is the same, the final result of the processing of the two samples is different. These two examples demonstrate that the end result of the whole ANN is totally dependent on the variables.

Table 4: IDSNN1_SnortDetection + Snort IDS

# Alerts	Snort	Legit	Scan	%
10260	False Positive	10200	121	98,81%
331	Scan	0	331	100%

Table 4 represents the combination of Snort IDS with IDSNN1_SnortDetection.

Using DARPA Dataset, snort generated 10260 false positives. Using scan dataset it detected all 331 scans. Combining Snort IDS with IDSNN1_SnortDetection we we're able to only generate 121 false positives. Resulting in an accuracy of 98,81%, as shown in table 4.

Table 5: IDSNN1_ScanDetection + IDSNN1_SnortDetection result

# Alerts	Legit	Scan	%
121	120	1	99,17%

Table 5 presents results from the combination of the two members of IDSNN1 modules. For its analysis we can see that the 121 results considered by IDSNN1_Snort as scans, 120 were considered legitimate traffic and one was considered as a scan. Thus, one false positive was generated.

Considering this, it is acceptable to say that the IDSNN1 produces a substantial reduction in the level of false positives. This fact leads to an efficiency improvement, to values of 99.17%.

5. Conclusions and future work

The application here presented emerged as a result a year of studying several methodologies, which gave rise to a master's thesis entitled: "Intrusion Detection System in Computer Networks using artificial neural networks".

An innovative approach for a neural network based intrusion detection system along with Snort, has been presented in this paper. It is important to realize that combining Snort with Artificial Neural Network it becomes possible to create a self-learning intrusion detection system with high reliable results. Despite of being more hardware demanding, the results justify the cost, improving significantly the results obtained with an ordinary IDS.

It is also important to note that the application has successfully processing the information supplied by a Dataset internationally accepted for testing IDS and in which Snort failed.

Thus, it was possible to define a hybrid approach that treats the problem of false positives generated by IDS. This methodology is still being tested in order to confirm the viability to automatically detect abnormal behavior by systems which analyze specific deviant events and, therefore, do not need to be constantly monitored by the System Administrator. This is only possible thanks to the implementation of structures behavior analysis with learning ability.

As demonstrated in this prototype, artificial neural networks emerged as a plausible solution to this issue, since it meets requirements as parallelism, speed, efficiency, as well as learning.

Despite the excellent results obtained, we should keep in mind that the application consumes lots of memory and that it was tested in a small network. However, we believe that this technology has great potential and, in the near future, we will put some additional effort on the development of a more efficient and reliable solution.

Considering the time and environment used to test the methodology here presented, the results obtained are quite satisfying. However, the use of an updated and bigger dataset is recommended to improve and confirm the potentialities of this methodology.

As a possible future work, we should create new attack scenarios and test a few other approaches, improving the connection between Snort IDS and IDSNN1 default configurations. The creation of a prototype in a more resource efficient language is also a priority.

References

Lippmann Richard, Joshua Haines, David Fried, Jonathan Korba, and Kumar Das (2000) Analysis and results of the 1999 darpa o-line intrusion detection evaluation. In Hervé Debar, Ludovic Mé, and S. Wu, editors, *Recent Advances in Intrusion Detection*, volume 1907 of *Lecture Notes in Computer Science*, pages 162–182. Springer Berlin / Heidelberg,. 10.1007/3-540-39945-3_11.

Ribeiro, Francisco (2011) Intrusion Detection System in Computer Networks using artificial neural networks master thesis

Zulkernine Mohammad MORADI Mehdi. A neural network based system for intrusion detection and classification of attacks

Mobile Cyberwarfare Threats and Mitigations: An Overview

Libor Sarga and Roman Jašek
Tomas Bata University in Zlín, Zlín, Czech Republic
sarga@fame.utb.cz
jasek@fai.utb.cz

Abstract: Mobile technologies have transformed rapidly with their rate of adoption increasing for several years. Smartphones, tablets, and other small form-factor devices are integrated in educational institutions, medical and commercial facilities with further military, governmental as well as industrial deployment expected in future. However, the complexity from interconnected hardware and software layers opens up multiple attack vectors for adversaries, allowing personally identifiable data exfiltration, malicious modifications of the device's intended functionality, pushing unauthorized code without user consent, or incorporating it into a botnet. Mobile threat landscape has become the next stage of cyberwarfare. Here, users unable or unwilling to adequately protect themselves make decisions based on information originating from untrusted third parties with potentially harmful intents. Recognizing the situation, a comprehensive array of tools and concepts such as ASLR, DEP, closing the source code, sandboxing, and code validation has been implemented. In this asymmetric security model, developers invalidate novel attack vectors while adversaries employ sophisticated techniques to thwart detection for large-scale penetration. The former are further penalized by heterogeneous base of software versions, some entirely defenseless against recent exploits. The paper presents an overview of techniques in current mobile operating systems and best practices the vendors incorporated to minimize unauthorized third-party modifications. It also aims to provide high-level description of exploits malware creators use to target users who, as we further postulate, underestimate capabilities of their devices. Best practices for safer use are briefly outlined, too.

Keywords: mobile, cyberwarfare, exploit, malware, security

1. Introduction

In recent years, mobile phones have undergone rapid transformation. From performing basic operations, they evolved to incorporate functionality on par with desktop stations and laptops while retaining high portability due to their small form factors.

As the user base grows, so do security concerns. Personally identifiable data, GPS (Global Positioning System) coordinates, credit card information, data transfers and others may be correlated to reconstruct provable history of physical location, financial transactions, wireless network trails, and per-user electronic behavior profiling. Moreover, malware makes it possible for perpetrators to exfiltrate such data without user's consent and modify the device without any input required from the victim.

Developers incorporated safeguards and protective measures to mitigate or neutralize existing and novel attack vectors. Ranging from cryptographic engines to hardware-imposed locks, they keep the mobile ecosystem as secure as possible without incurring unnecessary user experience penalties. However, privacy issues have also become a matter of law and research scrutiny.

In this paper, we will provide comprehensive overview of measures implemented into mobile operating systems (OS) and background on existing malware. It is structured as follows: Section 2 includes background on mobile software stack, iOS and Android, their current versions and security additions. Section 3 focuses on mobile exploits on these platforms. Section 4 provides tips on safe smartphone use and concluding remarks.

The terms smartphone and device will be used interchangeably throughout the article.

2. Mobile software stack

As any programmable device, smartphones consist of hardware stack with modules for either general purpose or specialized computations, and software stack housing an OS.

2.1 Background

Mobile OS is an extension of either Unix or proprietary kernel with support for hardware and software specificities. It is "a program that acts as an intermediary between a user of a computer and the computer hardware" and whose goals are to execute user programs and make solving user problems easier, make the computer system convenient to use, and use the computer hardware in an efficient manner. (Silberschatz, Galvin 1994)

The most popular mobile OSs include Android, iOS, BlackBerry OS, and Bada. Due to market and enterprise popularity of Android and iOS, the paper will further consider only the two for security analysis. Respective developers have chosen different approaches to distribute their OSs via OTA (Over-the-Air) updates. However, both allowed third party developers access to API (Application Programming Interface), SDK (Software Development Kit) and documentation, opening the platforms to third party code execution.

2.1.1 Android

Android is dated to 2003 when it was developed as an alternative to existing mobile OSs. In 2005, Android, Inc. was acquired by Google. Envisioned as an open-source product, a Linux-based kernel was chosen due to it allowing any derivative work to be marketed commercially provided an attribution to the original resource is retained.

First version was based on Linux kernel 2.6, as of version 4.0, Linux kernel 3.x was adopted. Architectural diagram is demonstrated on Figure 1. It is divided into four layers: Linux kernel, Libraries, Application Framework, and Applications. As of this writing, the latest available version is 4.2.1, an incremental update over 4.2.

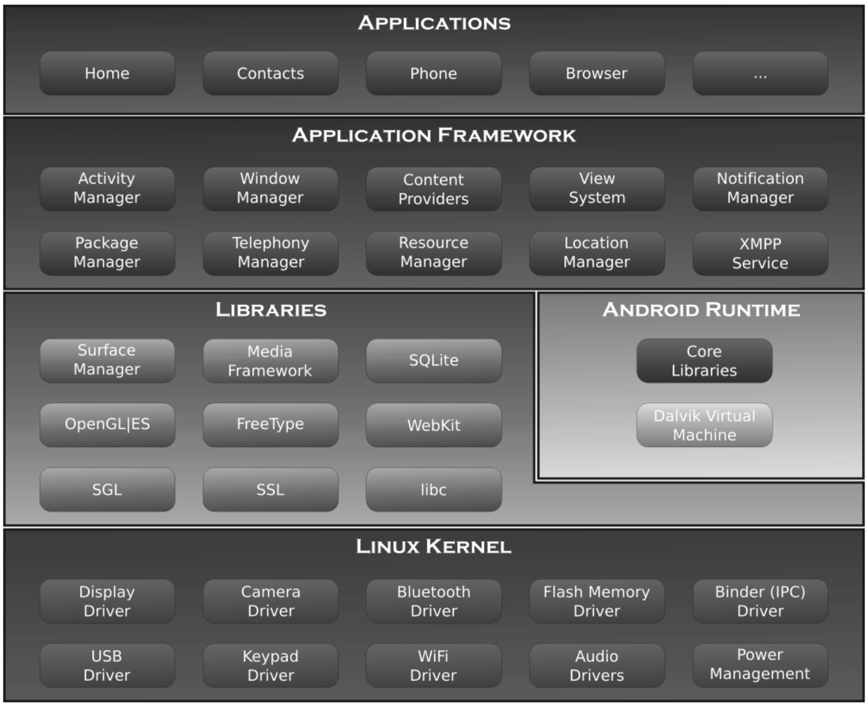


Figure 1: Android architectural diagram (Wang, Stavrou 2012, modified)

Apart from the official sources, a process known as sideloading allows users to install third party applications irrespective of origin. Software and middleware are distributed as self-contained .apk (Application Package File) bundles. They are not tied to any particular device, giving rise to both security and copyright concerns.

Android Security Program includes four elements: design review; penetration testing and code review; open source and community review; and incident response. (AOSP 2012) During the entire lifecycle, the system undergoes security reviews, both from dedicated in-house and external parties. Penetration testing in particular may provide clues as to potential vulnerabilities Android may face upon release by utilizing techniques and mindset of the attacker.

Linux core provides Android with several key security features: a user-based permission model, process isolation, extensible mechanism for secure IPC (Inter-Process Communication), and the ability to remove unnecessary and potentially insecure parts of the kernel. Additionally, newer versions fully support application sandboxing, "… a technique for creating confined execution environments to protect sensitive resources from illegal access. A sandbox, as a container, limits or reduces the level of access its applications have." (Li et al 2009)

The feature is tied to the user-based permission model in which each application mandatorily requests permissions before first execution. Subsequently, a per-application sandbox with the respective privilege set is initiated if the request is granted. User can't select permissions selectively but must either accept the requested ones or deny, prompting the application to abort. As of Android 4.2.1, there are 130 separate permissions.

Each process running on the Android-supported device is assigned a handle and can't communicate with other processes directly. Furthermore, it has only limited privileges when interacting with the OS which stores system libraries, application runtime, framework and applications themselves on a partition designated as read-only to avoid tampering.

Starting with version 3.0, Android supports full system encryption using AES-128 (Advanced Encryption System), a 10-round symmetric-key algorithm. AES has been extensively researched and all current attacks apart from highly specialized scenarios are currently computationally infeasible.

Several additions have been also added to reduce the risk of executing code on stack and heap, two abstract structures commonly used to hold variables and other data during the applications' run. The first is hardware NX (No eXecute), introduced in Android 2.3, which marks the memory space non-executable and enforces the policy for all applications, mitigating risk of unsanctioned code launch. The second is ASLR (Address Space Layout Randomization), introduced in version 4.0, which randomly shuffles locations of key resources in memory using system entropy pool to prevent hard-coding memory addresses into malware. The shuffling also affects stack and heap addresses. The third is PIE (Position-Independent Executable), introduced in Android 4.1, which mandates all application to run correctly regardless of absolute memory location, countering exploitation of predictable offsets. Unlike ASLR, PIE can't be forced entirely by the OS but must be implemented by application developers instead.

To preempt malicious third parties from uploading malware to the official Google Play store, Google introduced an automated verification tool, Bouncer. Details about its inner workings are not forthcoming, presumably to forego reverse-engineering and taking advantage of the vulnerabilities discovered. Nevertheless, it is known Bouncer employs custom-designed virtual machines performing static and dynamic analyses. Virtualizing a system or a component "… maps its interface and visible resources onto the interface of an underlying, possibly different, real system." (Smith 2005) Unofficial application stores may not have such safeguards in place, posing security risk for users.

Android 4.2 introduced a feature enabling users to perform similar analyses on the applications installed from outside Google Play. Based on fingerprinting the .apk package and comparing it to signatures stored on Google servers, user is warned when a positive match is made. Also tuned were prompts for permissions, which show detailed description of each privilege the application requests. (Raphael 2012)

The ecosystem, however, suffers from severely delayed adoption of new OS versions (Android Developers 2013) due to heterogeneous hardware stacks each smartphone class presents, necessitating system vendors to customize them before initiating OTA updates. Tendency to run older versions of Android may be ascribed to either hardware incompatibilities with newer versions, unwillingness to update, or non-existent release from the vendor.

2.1.2 iOS

In comparison with Google who promotes open standards and free innovation, Apple made the decision to tightly control both hardware and software stack using closed-source, proprietary licensing model.

First announced and released in 2007, iOS didn't allow third party developers to create applications at first, relying solely on in-house software modules which were easier to inspect, integrate, update as well as revoke.

Rescinded in 2008, native SDK along with extensive documentation was supplied to Apple Developer Program's participants.

Every user has to accept proprietary EULA (End-User License Agreement) on system's first run. Apple's OS doesn't enable access to the source code, employing the "security through obscurity" model considered unsuitable for large-scale applications.

Despite its restrictive nature, some official information has been released regarding internal iOS structuring, demonstrated on Figure 2.

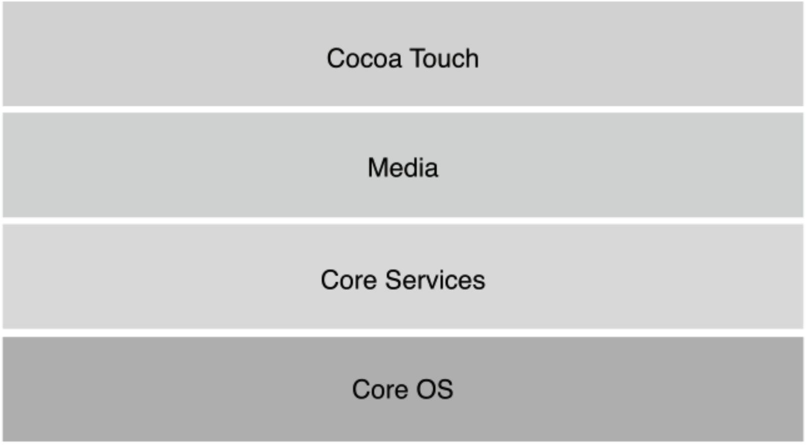


Figure 2: iOS' software stack architecture. (Apple 2012a, modified)

At its core lies a hybrid XNU (acronym for X is not Unix) kernel incorporating several open-source technologies such as BSD (Berkeley Software Distribution) Unix resources. Permissive nature of the accompanying BSD license allows developers to integrate them royalty-free even if when the resulting source code is closed. As of this writing, the latest iOS version available is 6.1. Updates are distributed either wirelessly via OTA, or the iTunes platform when a supported device is connected and network connectivity available. The exclusive source for third party applications on devices not purposefully modified (see Section 3) is the App Store platform as part of iTunes.

A single official document pertaining to iOS security was released in May 2012. Of particular interest is the Secure Boot Chain, a process ensuring no part of the booting routine has been tampered with to gain leverage into the system. The chain of trust starts with a ROM-supplied (Read-Only Memory) public key used to verify each successive module, halting in case an inconsistency is discovered. (Apple 2012b)

Identically to Android, iOS fully supports ASLR as of version 4.3, which was made available only to a subset of devices and opening the rest to known vulnerabilities. Further changes related to positioning of dynamic libraries in memory were introduced in iOS 5 with full kernel space ASLR following in iOS 6. (TiW 2012a) While offering increased protection from exploiting known fixed addresses, Apple again decided to exclude some device classes from the update cycle due to hardware or moral obsolescence.

iOS incorporates XN (eXecute Never) flag used in the ARM (Advanced RISC Machines) hardware architecture. It marks memory pages non-executable akin to DEP (Data Execution Prevention) and NX on Android. The system exempts some applications by assigning them both writable and executable memory region, these are, however, tightly controlled.

Another safeguard in both OSs protects against offline exhaustive password searches. The necessity to choose strong passwords capable of withstanding parallelized offline computational attacks has been stressed previously (Sarga, Jašek 2012) and such recommendations apply to passcodes on smartphones, too. Apple implemented PBKDF2 (Password-Based Key Derivation Function 2), which decreases maximum number of attempts per second by introducing fixed iteration count the search algorithm has to compute locally on the device in order to produce the result. The approach invalidates computing power available to the attacker and instead bounds them to use PBKDF2, itself based on AES. Exhaustive searches on iOS-operated hardware pose a significant time-factor challenge.

All executable code running on an unmodified device is required to be signed by an exclusively issued Apple certificate, and must support sandboxing as all running processes are partitioned. A developer wishing to distribute their applications via the App Store has to be a certified member of the iOS Developer Program with her real-world identity verified. Using her own unique certificate, the application is then signed before being sent for review with all permissions required for its correct functionality also listed.

Description and extent of analyses performed when candidates are submitted to the App Store are vague. They are "… reviewed… to ensure they operate as described and don't contain obvious bugs or other problems" (Apple 2012b) and is assumed they undergo automated static and dynamic analyses on a virtualized machine, identically to Bouncer.

3. Smartphone exploits

This section contains information on existing smartphone reverse engineering techniques, granting attacker access to undisclosed functions, enabling personally identifiable data exfiltration without user consent, or performing operations outside the developer-approved ones.

3.1 Android

Custom-built ROM modules granting access to restricted functions, changing the default UI (User Interface), and optimizing responsiveness and performance by replacing stock Linux kernels with streamlined or overclocked substitutes, have been developed. Changing the vendor-approved version requires account with elevated privileges, namely root. Moreover, with kernel based on Linux, Android inherits vulnerabilities present in the underlying core components.

Rooting necessitates employing reverse engineering techniques such as disassembling, decompiling, analyzing network activity or utilizing vendor-supplied debugging tools. Security implications of such actions are obvious: if an unauthorized third party gains access to these functions, the boot sequence may be unrestrictedly manipulated to include hidden keylogging services or packet analyzers, create hidden partitions, establish network tunnels to compromised servers to download malware, incorporate the device into coordinated initiatives such as botnets etc. In 2011, researchers presented a new kernel hooking rootkit which allows attacker to "… control the [smartphone] remotely via SMS… usurp bank accounts, pin numbers or even replace certain information [with] another… hide or manipulate network status or malware so that a user may perform an attack without knowing it." (You, Noh 2011)

Even without resorting to rooting, users may find their smartphones in an unsecured state when installing applications from official and unofficial sources. Android has a wide range of permissions third party code may request. A research conducted in 2011 focused on evaluating 940 applications as to the breadth of privileges with 35.8 % being classified as overprivileged and 94 % requesting 4 or fewer extra privileges. (Porter Felt at al 2011)

In 2012, fingerprinting techniques were introduced which allow to discern whether the code is being executed in a VM and modify its behavioral patterns accordingly. (Oberheide, Miller 2012) If the attacker is able to perform similar evaluation, seemingly benign application may be submitted to the Google Play store which wouldn't violate any rules as long as it's contained within the VM. However, when run on physical hardware, it could perform arbitrary malicious actions. Coupled with pre-ASLR vulnerabilities, the threat increases substantially.

Further demonstrated were exploits in the NFC (Near Field Communication) standard, present in newer Android devices. A proximity-based data transfer technology, NFC operates by detecting a compatible unit in a 3—10 centimeter vicinity, and initiating a wireless transfer if found. This allows for contactless financial transactions or authentication. By sending maliciously crafted data, the attacker can generate an exception and force the system to perform certain actions, such as visiting a malware-injected web page. (Miller 2012) Bluetooth connections can be also established and data exfiltrated, a process made easier when the smartphone is rooted. Some of the vulnerabilities were patched in Android 4.0.1.

However, the most prevalent class of malware requests a privilege allowing it to send SMS (Short Message Service) from the device, a so-called toll fraud. (Lookout 2012) When accepted, text messages are sent to premium-rate numbers for which the sender is billed substantially higher compared to standard service charges. Another type of intrusive mobile software redirects web page requests or injects them with inline frame (iframe) elements. By correlating leaked user logs, geographical locations, personal preferences and behavioral patterns, highly accurate profiles can be constructed sought by spam campaign operators.

3.2 iOS

Closed-source nature of iOS and a single exclusive source of applications seem to reduce the number of malware samples identified as active.

Nevertheless, severe restrictions on which code can be run on iOS smartphones prompted enthusiasts to seek ways of accessing root account in order to remove the limitations and customize the system by community-developed utilities. Like Android rooting, jailbreaking has gained prominence and a large following. It is defined as "… the process by which full execute and write access is obtained on all partitions…" (TiW 2012b) Jailbreak is a reverse engineering effort to escalate user's privileges to superuser by utilizing existing flaws.

Jailbreaking relies on discovering previously unpatched vulnerabilities, which may be replicated without the need to physically interact or tamper with the hardware stack. As the iOS' complexity has been increasing, a cascade of several exploits is currently needed to successfully perform a jailbreak, with each link creating conditions for the following exploit to work correctly. For example, iOS 5.0.1 efforts linked three separate vulnerabilities. Permanent exploits not easily patchable by OTA updates are preferred, e.g., manipulating the boot sequence which necessitates a hardware revision.

Concerns were raised regarding decreased security of jailbroken devices, though. By sidelining review routines of the Apple software distribution platform, the attack surface has been expanded. This was first demonstrated in 2009 when the ikee mobile worm was released targeting known omission in an SSH (Secure Shell) module on many jailbroken smartphones. The same vulnerability was later used maliciously in the Netherlands to eavesdrop on data pertaining to banking transactions. (Hypponen 2009) Since then, no new malware has been detected.

However, in 2011 a flaw in the code signing procedures was discovered, permitting an application to download unsigned code after it had been vetted suitable for publication. (Miller 2011) As a proof-of-concept, third party code was submitted and passed the review process in spite of it sending a background request to arbitrary untrusted server. There, it checked for unsigned, possibly malicious code to download and install while at the same providing its declared functionality without notifying the user. It further opened a backdoor remote shell with the ability to list directories, running processes, start vibration module or exfiltrate sensitive files such as address book.

4. Discussion, conclusion

Here, an unsorted list of recommendations for safer smartphone use is presented. Many are applicable to any unit supporting wi-fi access, application execution, GPS, and other technologies commonly found in smartphones.

- Users should strongly prefer applications coming from official sources (App Store, Amazon Appstore, Google Play) as they have been vetted and confirmed safe for use.
- Limiting the number of applications kept is advised. Every addition to the code pool expands threat surface by incorporating respective software vulnerabilities. Furthermore, as some applications are frequently used in conjunction, malware may link several exploits to increase probability of successful system penetration.
- Refraining from installing applications coming in email attachments or through any unsolicited channel is recommended, especially on Android as the third party code ecosystem is very easy to exploit for malicious purposes.
- Researching an application the user is about to install should be a priority. Numerous review sites, user experience forums, and developer web pages provide plenty of information. Factors such as update cycle length, UI appearance, standard of English language, reviews as well as user opinions can help in making an informed decisions.

- Application updates should be applied as soon as they appear on release channels. Since they have to pass the same security review, they are guaranteed safe by the platform operator. Adding new functions, patching vulnerabilities, and ensuring compatibility with more devices and operating system versions are three common reasons for updating.
- If an application provides option to create username and password, users should consider using it anonymously or set a reasonably complex password. Generating and storing sensitive data when not required may lead to poor password management, which in turn increases risk of compromise.
- Opting in for sending user experience feedback should be considered only if the data are explicitly stated to be anonymous. Otherwise, an attacker may intercept and correlate it with per-device identifiers and pass them to untrusted third parties. If possible, users should send only encrypted data to any third party.
- Limiting GPS and location services to a strict subset of utilities is encouraged. Collating the information over longer time periods allows the attacker to determine movement patterns and presence in particular locations. Leaking such data also increases threat space by revealing digital trail of user movements. Exceptions can be made regarding services for locating a lost device.
- If rooting or jailbreaking is attempted, users should back up content of their smartphones before starting, a recommendation even the tools' authors frequently mention as important. Backup serves as a baseline to which user can reset their device at any time. An encrypted backup is a viable option, as well.
- Equally important is to consider security implications of rooting/jailbreaking. OS updates address many exploitable vulnerabilities and offer increased stability. Window of opportunity for the adversaries to take advantage of outdated systems may be stretched to several months as many users wait for a successful rooting/jailbreak utility and hold back updating until it has been released. A choice between convenience and increased security should be made prioritizing the latter.
- In case the user is determined to root/jailbreak, reputable sources should be used. Many well-known groups and communities of enthusiasts and power users exist in this area. Attackers frequently exploit such decentralized, authority-less model by providing malicious links and information, especially after new OS updates when root/jailbreak tools are updated internally without public availability.
- Regular device backups are strongly advised, preferably in conjunction with strong passwords set to protect them. iTunes offers offsite cloud backup solution. While security implications of the cloud should be weighted, complex passwords will deter any attempts to access the contents. Periodic backups to external HDD or USB media is recommended for users wishing to exert control over physical location of their data.
- Users should set passcodes or passwords to access their smartphones from the Lock Screen. Setting long, easily memorable passwords ensure the attacker in possession of the device must resort to other means of obtaining data than brute-force attacks due to the time factor involved. Reasonable password changing policy should be set, too.
- Advanced security features include option to wipe the device when password is entered incorrectly in arbitrary number of attempts. Nonjailbroken Apple units have the limit set to 10 while third party Android applications allow to set custom number of attempts. A tradeoff between reasonable security level and user comfort should be made on individual basis.
- Users are advised never to leave their devices unattended, particularly when modules such as NFC or wi-fi are active and no passcode has been set. If misplaced, options exist to locate it using data collected from public wireless networks and signal towers the smartphone is encountering at its current location. In populated areas, the margin of error is usually several meters.
- When selling or passing the device to another user, the owner is strongly encouraged to wipe it and reset settings to factory defaults. If omitted, new owner is able to connect to known networks automatically, access email accounts, extract positioning data from applications, search address book etc. Nevertheless, tools exist claiming to be able to partially recover previously deleted data from flash memory chips. If secure alternatives to wiping are available, such as multiple overwrites with randomly generated data patterns, they should be prioritized in spite of the increased time factor involved.
- One-time passwords are preferable over static ones. Many services support the feature, generating pseudorandom strings instead of original passwords which are automatically discarded after first use so that they can't be entered multiple times.

- Users are advised to connect to secured wireless networks. The suggestion doesn't always hold, though, in which case a VPN (Virtual Private Network) connection with SSL/TLS (Secure Sockets Layer/Transport Layer Security) authentication should be set when performing sensitive actions. All unprotected public networks must be assumed to be under control of an adversary capable of intercepting and modifying network traffic.
- Disabling automatic reconnection to wireless networks and turning off the wireless module when not needed is a security baseline. Especially in cases of public unencrypted networks, users should assume the attacker is capable of setting up rogue APs (Access Points). When encountered, wi-fi logic will automatically prefer it over others with lower relative signal strength and initiate connection, exposing the victim to network packet interception.
- Turning off EDGE (Enhanced Data rated for GSM Evolution)/3G/4G functionality when not needed prevents unsanctioned data transfers, malicious code thus can't establish outbound connections to remote servers to receive or exfiltrate data, assuming the malware doesn't have root access.
- Use of browsers on smartphones should be limited to known and trusted sites, preferably those supporting encrypted transfers. Potentially unsafe websites should be opened only on machines protected with additional measures (antivirus, VM). For known hostile sites, a VM must always be set up in which all actions can be evaluated in an isolated, sandboxed environment.
- Similarly, mobile browsers should not under any circumstances be used for financial operations unless the user is technically proficient to check each and every detail of the connection chain (network, browser, certificate) herself. Credit card data, banking accounts and other financial information must be considered extremely sensitive and highly valuable for any attacker, and protected accordingly.
- Any nonpublic information should not be stored in an unencrypted form such as notes and reminders. Security-conscious users may assume the attacker has the ability to access and exfiltrate them at any time. Both platforms offer third party utilities where sensitive information can be kept securely.

Android developers have chosen openness to provide users with easy way to check and customize their systems or start developing applications. A 2012 study enumerated 1260 Android malware samples and concluded the most common ways of distributing harmful code is repackaging, drive-by downloads, and others. (Zhou, Jiang 2012) Payload functionalities include privilege escalation, remote control, financial charge, and information collection.

While no information about industrial espionage or APT (Advanced Persistent Threat) campaigns have so far been released, as long as mobile cyberspace keeps becoming more ubiquitous and security awareness doesn't outpace its growth rate, it's safe to assume such initiatives will appear and proliferate.

Acknowledgments

The article was supported by the Centre for Security, Information and Advanced Technologies (CEBIA-Tech) project, registration number CZ.1.05/2.1.00/03.0089.

References

Android Developers. (2013) "Dashboards," [online], Android Developer Program, https://developer.android.com/about/dashboards/index.html

AOSP. (2012) "Android Security Overview," [online], Android Open Source Project, http://source.android.com/tech/security/

Apple. (2012a) "iOS Technology Overview," Apple Developer, http://developer.apple.com/library/ios/documentation/Miscellaneous/Conceptual/iPhoneOSTechOverview/iPhoneOSTechOverview.pdf

Apple. (2012b) "iOS Security," [online], Apple Developer Program, http://images.apple.com/ipad/business/docs/iOS_Security_May12.pdf

Hypponen, M. (2009) "Malicious iPhone Worm," [online] F-Secure, http://www.f-secure.com/weblog/archives/00001822.html

Li, Z., Tian, J.–F., and Wang, F.-X. (2009) "Sandbox System Based on Role and Virtualization," *The First International Symposium on Information Engineering and Electronic Commerce, 2009 (IEEC '09),* May 16—17, 2009, Ternopil, Ukraine, pp. 342—346

Lookout. (2012) "State of Mobile Security 2012," [online], Lookout Mobile Security, https://www.lookout.com/resources/reports/state-of-mobile-security-2012

Miller, C. (2011) "Inside iOS Code Signing," *SyScan'11 TPE*, November 17—18, 2011, Taipei, Taiwan.

Miller, C. (2012) “Exploring the NFC Attack Surface,” [online], Black Hat, https://media.blackhat.com/bh-us-12/Briefings/C_Miller/BH_US_12_Miller_NFC_attack_surface_WP.pdf

Oberheide, J., and Miller, C. (2012) “Dissecting the Android Bouncer,” [online], Duo Security, http://jon.oberheide.org/files/summercon12-bouncer.pdf

Porter Felt, A., Chin, E., Hanna, S., Song, D., and Wagner, D. (2011) “Android Permissions Demystified,” *Proceedings of the 18th ACM conference on Computer and communications security (CCS ’11)*, October 17—21, 2011, Chicago, Illinois, pp. 627—638

Raphael, J. R. (2012) “Exclusive: Inside Android 4.2’s powerful new security system,” [online], Computerworld, http://blogs.computerworld.com/android/21259/android-42-security

Sarga, L., and Jašek, R. (2012) “User-Side Password Authentication,” *Proceedings of 11th International Conference on Information Warfare and Security (ECIW-2012)*, July 5—6, 2012, Laval, France, pp. 237—243

Smith, J. E. (2005) “The Architecture of Virtual Machine,” *Computer*, Vol 38, No. 5, pp. 32—38

Silberschatz, A., and Galvin, P. (1994) *Operating System Concepts*, Addison-Wesley, Boston

TiW. (2012a) “ASLR,” [online], The iPhone Wiki, http://theiphonewiki.com/wiki/ASLR

TiW. (2012b) “Jailbreak,” [online], The iPhone Wiki, http://theiphonewiki.com/wiki/Jailbreak

Wang, Z, and Stavrou, A. (2012) “Google Android Platform: Introduction to the Android API, HAL and SDK,” [online], George Mason University Department of Computer Science, http://cs.gmu.edu/~astavrou/courses/ISA_673_S12/Android_Platform_Extended.pdf

You, D.–H., and Noh, B.-N. (2011) “Android platform based linux rootkit,” *2011 6th International Conference on Malicious and Unwanted Software (MALWARE)*, October 18—19, 2011, Fajardo, Puerto Rico, pp. 79—87

Zhou, Y, and Jiang, X. (2012) “Dissecting Android Malware: Characterization and Evolution,” *2012 IEEE Symposium on Security and Privacy (SP)*, May 20—23, 2012, San Francisco, California

Analysis of Services in Tor Network: Finnish Segment

Alexander Semenov
Dept. of Computer Science and Information Systems, The University of Jyväskylä, Jyväskylä, Finland
alexander.v.semenov@jyu.fi

Abstract: Over the last 30 years, we have seen the fast growth of the Internet. Among other things, the Internet supports communication between the people. There are a number of ways to communicate on the Internet, e.g., IRC or WWW chats and the recently emerged social media. Computer-mediated communication allows people to act under virtual identities and conceal their real-life identities. Still, identifying someone's real identity is feasible due to IP addresses left in logs or other traces. Technology allowing for stronger anonymity, Tor, was in 2002 built by the U.S. Navy and initially targeted at military users for concealing their identities and preventing eavesdropping. However, it became popular among other users. The Tor network also contains a number of websites, called "hidden services". Sometimes, they provide services of questionable legality in some jurisdictions, such as sales of controlled substances and illegal drugs. However, due to the anonymous nature of Tor, the IP addresses of Web servers with these sites also cannot be tracked easily. The present paper contains exploratory research that reveals what can be crawled from Tor and what kind of contents there exists. The paper presents a description of the hidden services in the Finnish segment of Tor, their crawling, and an analysis. Also, the paper discusses possibilities of tracking users of Tor's hidden services.

Keywords: crawling; tor; deep Web; hidden web; social media sites; social media analysis

1. Introduction

Over the last 30 years, we have seen the fast growth of a global computer network, known as the Internet. Among many usage scenarios, the Internet acts as a communication medium for the people. There are a number of ways to exchange messages with other people or groups of people, e.g., instant messengers, WWW chats and forums, and social media sites. Popular social media sites such as Facebook ("Facebook Newsroom," 2012), which has more than one billion users, and Twitter, with around 500M users, contain a lot of user-generated content. Such sites allow people to create user profiles, post data about themselves, and connect with other users of the sites. Also, the Internet has allowed criminals or like-minded people to express their ideas or study the experience of other criminals with the goal of copying crimes or avoiding mistakes of earlier perpetrators. Examples of such cases are school shooters. Semenov et al. (2011) show that seven of eleven perpetrators expressed their ideas on the WWW. Another example is Anders Breivik, who detonated a car bomb that killed 8 people in Oslo, and then shot and killed 69 people on the island of Utøya (Norway) in July 2011. Shortly before the attacks, he distributed his 1500-page manifesto, which described his intentions, and uploaded a short YouTube video.

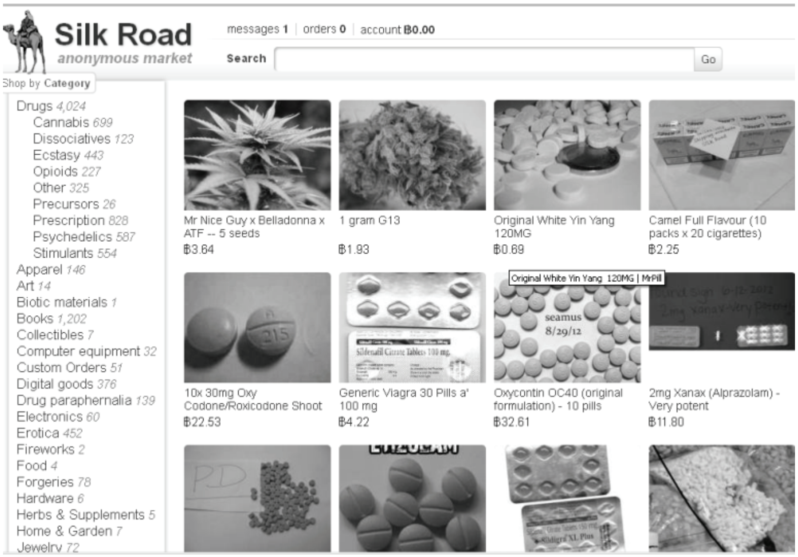
An important aspect of communication through the Internet is the possibility of anonymous or pseudonymous communication; in other words, people communicating with others can use virtual identities. Information describing such identity does not necessarily accurately correspond to information of the real person. However, identification of the real person behind the virtual identity may be possible through the extraction and analysis of various traces left by this person. It might require, e.g., the extraction of IP address used by a virtual identity and a subsequent request for the user's details from a service provider. However, such data may be available only to the service provider, and may be released to police or other authorities only. The retention of IP addresses of connected parties for a certain time is required by the EU Data Retention Directive ("DIRECTIVE 2006/24/EC," 2013). There are many methods to conceal IP addresses from the service provider, such as connection to it through anonymous proxies, SOCKS, or VPN. Still, it may be possible to get the real IP of the person in case of its retention, e.g., when a proxy server stores the logs, which can be extracted from there.

A patent ("United States Patent: 6266704," 2012) issued in 1998 describes the onion routing (OR): technology, which is aimed at protecting the privacy of its users. The idea behind OR is that messages sent there travel from source to destination through a circuit of random number of nodes in the OR network, and each node routes the message further. Circuit is built before the sending of the message, and each node in the circuit knows only its predecessor and successor. Messages are re-encrypted at each node (router), and this prevents the OR network from eavesdropping. Thus, unless the attacker has access to all the nodes in the network, compromising becomes very difficult. The most popular platform that makes use of the OR is Tor ("Tor Project:

Anonymity Online," 2012), developed by the U.S. Navy: "*free software and an open network that helps you defend against a form of network surveillance that threatens personal freedom and privacy*". Tor can be used by individuals to keep websites from tracking them, journalists can safely communicate with whistleblowers, and the military can use it for open-source intelligence gathering. Syverson (2011) underlines that Tor was initially built for military purposes and the presence of a multitude of other users can successfully conceal the military among them. Tor functions as overlay network over TCP/IP protocol.

Tor nodes can be run as clients only, or they can be set up as a relays, in which case other Tor nodes would be able connect to them. The nodes can also be configured to be "exit relays". In this case they provide access to public web sites for other Tor nodes. Tor relays are listed in public directory and are used as introduction points Also, Tor nodes can be configured as bridge nodes. Tor bridges are Tor relays not listed in the public directory ("Tor Project: Bridges," 2013), which may be used to connect to Tor if the public relays are banned by an ISP. Tor client can be manually configured to connect through the bridge. The addresses of the bridges (IP address and port) may be distributed by e-mail, found in social networks, etc.

Tracking Tor users is difficult for the external observers since only the IP address of the latest node in the circuit (exit relay) can be observed by an external web server. Also, Tor tunnels DNS requests so that they are not sent directly from the requesting computer. Despite anonymous access to the WWW, the Tor network provides the possibility of maintaining anonymous *hidden services*, which are servers accessible only from Tor. The addresses of the services are 16-character alphanumeric identities in the .onion pseudo top-level domain zone. Access is provided via connecting to the service. Addresses of the services are stored in a distributed database, and are put there when the server with service is set up. Hidden services may involve Web servers, instant messaging servers, and so on. There is no common directory of all hidden services, however there are number of hidden services which list other services. "Hidden Wiki" project is a directory of hidden services that can act as the entry point. A prominent service is Silk Road, a site that operates an anonymous black market. An article (Gawker, 2012) claims that the yearly turnaround of Silk Road is $22 million. The address of Silk Road (Fig. 1) is silkroadvb5piz3r.onion. It is possible to buy there a number of different things, including drugs (Gawker, 2012). The only currency accepted there is the BitCoin, anonymous digital currency without a central issuer.



Drugs 4,024
Cannabis 699
Dissociatives 123
Ecstasy 443
Opioids 227
Other 325
Precursors 26
Prescription 828
Psychedelics 587
Stimulants 554
Apparel 146
Art 14
Biotic materials 1
Books 1,202
Collectibles 7
Computer equipment 32
Custom Orders 51
Digital goods 376
Drug paraphernalia 139
Electronics 60
Erotica 452
Fireworks 2
Food 4
Forgeries 78
Hardware 6
Herbs & Supplements 5
Home & Garden 7
Jewelry 72

Mr Nice Guy x Belladonna x ATF -- 5 seeds
Ƀ3.64

1 gram G13
Ƀ1.93

Original White Yin Yang 120MG
Ƀ0.69

Camel Full Flavour (10 packs x 20 cigarettes)
Ƀ2.25

10x 30mg Oxy Codone/Roxicodone Shoot
Ƀ22.53

Generic Viagra 30 Pills a' 100 mg
Ƀ4.22

Oxycontin OC40 (original formulation) - 10 pills
Ƀ32.61

2mg Xanax (Alprazolam) - Very potent
Ƀ11.80

Figure 1: The Silk Road screenshot

Currently, considerable attention is being paid to Tor in the popular press: (SC Magazine, 2011) describes how Tor was used in Egypt during the protests, (CSO, 2012) describes online black markets, and (Gawker, 2011) provides a detailed report on buying drugs at Silk Road. ("Iltalehti," 2011) describes the Finnish hidden service Thorlauta and provides screenshots that depict various controlled substances and drugs.

Nowadays, a lot of interest is being expressed in monitoring the World Wide Web (WWW); e.g., DARPA has expressed interest in it ("Federal Business Opportunities," 2012); and large projects propose that soon there will be crisis-monitoring centres that would monitor data from social media and forecast the sentiment of the people (Aliprandi and Marchetti, 2011). Many modern social media sites provide API for data querying, making this activity technically simple. However, information from the sites within the Tor network is inaccessible for traditional Web monitoring tools (e.g., various search engines).

In the present paper, we describe the application of the generic social media monitoring system, described in Semenov et al. (2011) for gathering and analysing data from Tor hidden services, maintained in the Finnish language.

2. Description of Tor

Connection to the Tor network is made through special software (the Tor client), which starts the SOCKS5 server, which tunnels incoming requests to the Tor network. There is software for various operating systems ("Tor Project: Anonymity Online," 2012). For Windows, there is a browser bundle for making connections through Tor. Also, there are plugins for browsers allowing access to Tor. Applications that support SOCKS5 can use Tor through an installed SOCKS5 server. If SOCKS5 is not supported, there are various tools that allow one to "socksify" the applications. For Linux, the "torsocks" software wraps the requests made by applications and sends them to Tor; it tunnels DNS requests as well ("Torsocks," 2012).

Names of the hidden services in Tor are generated automatically and placed in a distributed hash table ("Tor: Hidden Service Protocol," 2012). The hash table contains the names of the hidden services and their introduction points that send the requests further until they reach the destination server, which processes the requests. In the present paper, URLs of the hidden services, maintained in the Finnish language, were extracted from the Hidden Wiki. According to it, there are seven hidden services: three discussion forums, one of which is no longer working (as mentioned on the page), one devoted to "White Power" (Suojeluskunta), and another one without a particular topic; three imageboard sites (anonymous discussion boards); and one site containing advices about the use of Tor and hidden services.

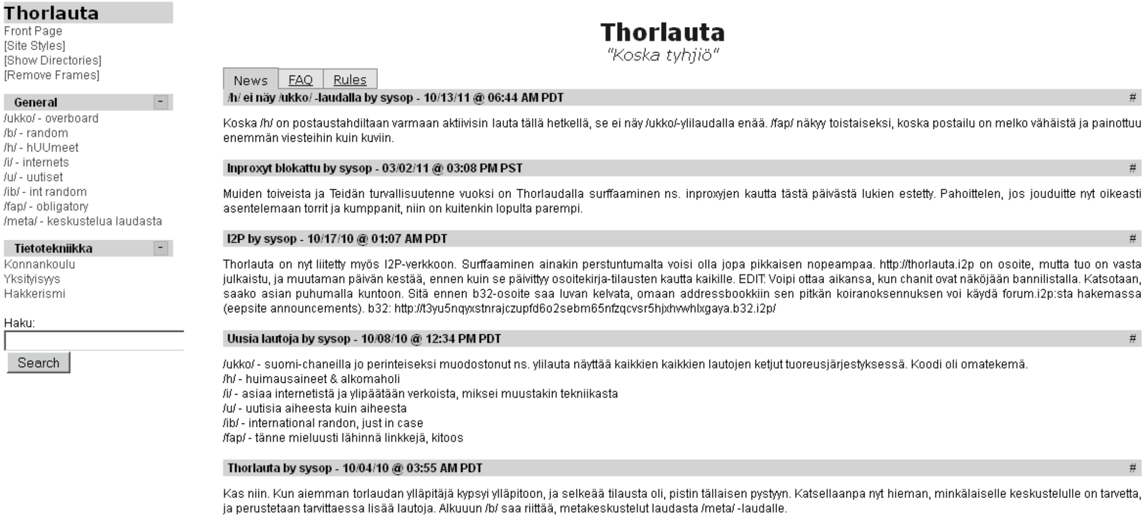
3. Crawling the hidden services

We crawled two hidden services: the imageboard "Thorlauta" and the forum "Suojeluskunta". These are the most active forums found in the Finnish segment of Tor network. Crawling of the Tor hidden services was carried out using the system described in Semenov et al. (2011) and Semenov and Veijalainen (2012a). The developed software system implements a three-layer model, where immediate social reality is modelled by the first level – social media sites, which are modelled by multirelational graph, which is modelled by a third-level model, a database implementation of the graph (Semenov and Veijalainen, 2012b). Currently, the developed monitoring software does not support SOCKS5 proxies, so the crawler was run through the "Torsocks" wrapper, which allowed connection to Tor.

3.1 Thorlauta

Thorlauta is an imageboard site. It is not possible to register or make a profile. Communication there is supposed to be completely anonymous, however, users may leave name and e-mail, if they desire. The Tor platform makes this anonymity stronger. Basically, it is a discussion board that is divided into subforums, "directories" that contain threads with text messages with images. When a new thread appears on the subforum, an old one is deleted in FIFO fashion. Length of the queue is set by the administrator of the site. The site currently contains the following subforums: ukko, which contains links to recent topics from other directories; b, which contains random chat about everything; h, "huumet", topics about drugs; I, "Internets"; u, discussions of news; ib, for international people who do not understand Finnish; fap, mostly containing erotica/pornography; meta, chat about the imageboard; ko, "school for bad guys"; fol, about anonymity; and hak, a board about hacking. Figure 2 contains a screenshot of the imageboard.

The site ontology (Semenov and Veijalainen, 2012a) used during the monitoring is described in Fig. 3 The node "dir", which is shorthand for "directory", has the attribute "directory_name". This node is connected to the node "thread" with the edge "has_thread", which is connected with the node "message", with attributes "message_id" and "message_text"; "message" is connected to the node "author", which has the author's name and e-mail. This multirelational graph describes the structure of the imageboard: there are several directories, which have many threads that contain messages written by authors. Messages also contain information about the date attached to them. Authors can be identified by name and e-mail address.



Thorlauta

"Koska tyhjiö"

/h/ ei näy /ukko/ -laudalla by sysop - 10/13/11 @ 06:44 AM PDT #

Koska /h/ on postaustahdiltaan varmaan aktiivisin lauta tällä hetkellä, se ei näy /ukko/-ylilaudalla enää. /fap/ näkyy toistaiseksi, koska postailu on melko vähäistä ja painottuu enemmän viesteihin kuin kuviin.

Inproxyt blokattu by sysop - 03/02/11 @ 03:08 PM PST #

Muiden toiveista ja Teidän turvallisuutenne vuoksi on Thorlaudalla surffaaminen ns. inproxyjen kautta tästä päivästä lukien estetty. Pahoittelen, jos jouduitte nyt oikeasti asentelemaan torrit ja kumppanit, niin on kuitenkin lopulta parempi.

I2P by sysop - 10/17/10 @ 01:07 AM PDT #

Thorlauta on nyt liitetty myös I2P-verkkoon. Surffaaminen ainakin perstuntumalta voisi olla jopa pikkaisen nopeampaa. http://thorlauta.i2p on osoite, mutta tuo on vasta julkaistu, ja muutaman päivän kestää, ennen kuin se päivittyy osoitekirja-tilausten kautta kaikille. EDIT: Voipi ottaa aikansa, kun chanit ovat näköjään bannilistalla. Katsotaan, saako asian puhumalla kuntoon. Sitä ennen b32-osoite saa luvan kelvata, omaan addressbookkiin sen pitkän koiranoksennuksen voi käydä forum.i2p:sta hakemassa (eepsite announcements). b32: http://t3yu5nqyxstnrajczupfd6o2sebm65nfzqcvsr5hjxhvwhlxgaya.b32.i2p/

Uusia lautoja by sysop - 10/08/10 @ 12:34 PM PDT #

/ukko/ - suomi-chaneilla jo perinteiseksi muodostonut ns. ylilauta näyttää kaikkien kaikkien lautojen ketjut tuoreusjärjestyksessä. Koodi oli omatekemä.
/h/ - huimausaineet & alkomaholi
/i/ - asiaa internetistä ja ylipäätään verkoista, miksei muustakin tekniikasta
/u/ - uutisia aiheesta kuin aiheesta
/ib/ - international randon, just in case
/fap/ - tänne mieluusti lähinnä linkkejä, kitoos

Thorlauta by sysop - 10/04/10 @ 03:55 AM PDT #

Kas niin. Kun aiemman torlaudan ylläpitäjä kypsyi ylläpitoon, ja selkeää tilausta oli, pistin tällaisen pystyyn. Katsellaanpa nyt hieman, minkälaiselle keskustelulle on tarvetta, ja perustetaan tarvittaessa lisää lautoja. Alkuuun /b/ saa riittää, metakeskustelut laudasta /meta/ -laudalle.

Figure 2: Thorlauta imageboard screenshot

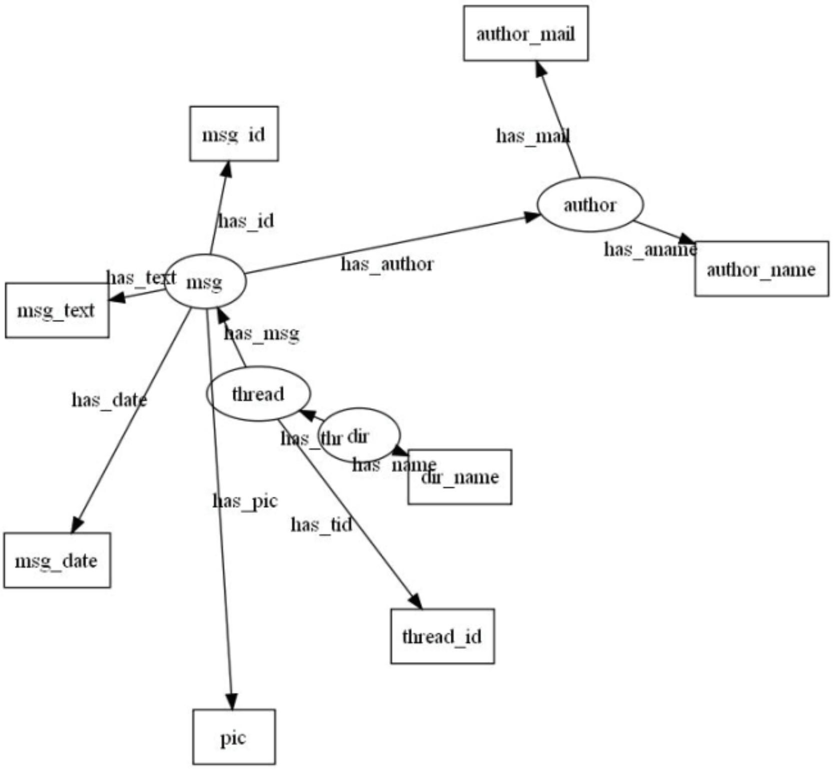


Figure 3: Multirelational directed graph of Thorlauta

Correspondence between entities in the multirelational graph depicted in Fig. 3, and elements of the Web page were established using regular expressions. The Thorlauta imageboard was collected on 23 November 2012, and a static snapshot was saved in the repository, a model of which depicted in Fig. 4. The total size of the downloaded data was 792 MB, and it was downloaded for 5 hours and 15 minutes. Thus, the average speed was 343.28 Kbit/s (entire imageboard, including images). The second time, Thorlauta was collected on 23 January 2013.

3.2 Suojeluskunta

Suojeluskunta translates into English as “White Guard”. The Suojeluskunta organization emerged in 1918 during the Finnish Civil War as a voluntary militia. The organization was abolished in 1944 as a fascist organization (“White Guard (Finland),” 2012). Today, the Suojeluskunta forum contains a number of discussions, many of which some might consider hate speech.

The forum was collected twice. The first collection took place on 8 December 2012. The total size of the downloaded forum was 40 MB (only text information was downloaded, also there were graphical avatars of the users, but this was not collected), and the whole downloading process took 96 minutes. Thus, the average downloading speed was 56.89 Kbit/s, which is almost six times slower than the speed of collection for Thorlauta. The second collection took place on 23 January 2013 and took 80 minutes. Figure 4 shows the site ontology modelling the Suojeluskunta forum.

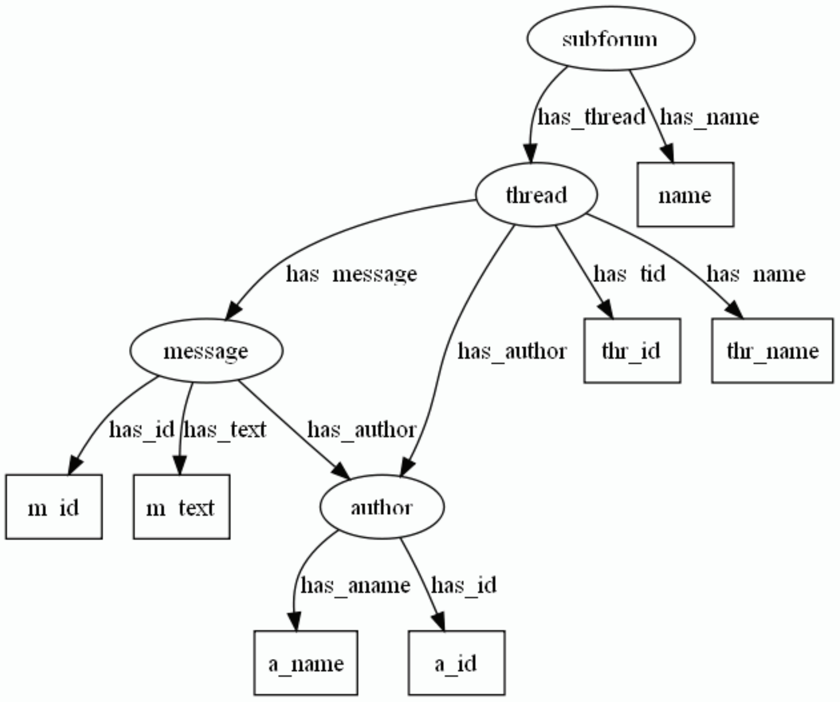


Figure 4: Multirelational directed graph modelling Suojeluskunta

The forum contains several subforums, which have topics containing messages written by authors, who have names and identifiers. If an author is not registered, an identifier is not presented (Fig. 5).

4. Analysis

In the present chapter, we describe the analysis of the collected data from the Thorlauta imageboard and the Suojeluskunta forum.

4.1 Thorlauta

4.1.1 Overall statistics

Table 1: First crawl statistics

Name	N. of msgs	N. of thrs	Minimal date	Maximal date	Oldest last post date
b	2194	190	2011-12-11	2012-11-23	2012-10-04
h	2616	222	2012-09-03	2012-11-23	2012-09-27
i	1588	170	2010-10-14	2012-11-23	2011-07-26
u	753	76	2010-10-05	2012-11-10	2010-10-05
ib	272	36	2010-10-18	2012-11-18	2011-03-24
fap	2529	169	2010-10-31	2012-11-23	2012-05-08
meta	600	66	2010-10-04	2012-11-09	2010-10-05
ko	2143	178	2011-10-15	2012-11-23	2012-01-26
fol	623	77	2011-10-18	2012-11-11	2011-10-24
hak	1370	209	2011-10-14	2012-11-22	2011-10-16

During the first crawl, 14,688 messages were downloaded from the imageboard. Of this, 1005 messages had field “author_mail” filled with mail value, containing ‘@’. From these, 367 e-mail addresses were extracted. 657 messages had e-mail field filled with value not corresponding to e-mail address. Table 1 contains information on the distribution of the messages per directories. “**Name”** is the name of the directory, **“n. of msgs”** is total number of messages, **“n. of thrs”** is the number of the threads, **“minimal date”** is the minimal timestamp of the message**, “maximal date”** is the maximal timestamp of the message, and **“oldest post date”** is the timestamp of the oldest of the last messages in the threads in the directory. Threads on the forum expire depending on the timestamp of the last message. The one with the oldest date expires when new message is posted in the thread. Minimal timestamp of the message through the entire board is maximal in the ‘h’ directory, the directory devoted to drugs (and purchasing them online). Table 2 shows the results of the analysis of the imageboard during the second crawl. The column captions denote the same things, and the added column for “**number of expired thrs”** shows the numbers of the threads that were no longer on the imageboard. Threads in the subforum devoted to selling controlled substances expired much faster than others.

Table 2: Second crawl statistics

Name	N. of msgs	N. of thrs	Minimal date	Maximal date	Oldest last post date	Number of expired thrs
b	1830	167	2011-12-11	2013-01-23	2012-12-10	184
h	1108	147	2012-12-15	2013-01-23	2013-01-13	222
i	1511	166	2010-11-03	2013-01-23	2011-10-21	34
u	912	90	2010-10-05	2013-01-21	2010-10-06	1
ib	301	42	2010-10-18	2013-01-23	2011-03-24	1
fap	2507	162	2010-10-31	2013-01-23	2012-10-20	101
meta	686	70	2010-10-04	2013-01-23	2010-10-10	0
ko	2258	178	2011-10-15	2013-01-23	2012-05-03	41
fol	768	88	2011-10-22	2013-01-23	2011-11-24	7
hak	1521	229	2011-10-15	2013-01-20	2011-11-05	9

4.1.2 Most popular threads

Table 3 contains information on the ten most popular threads during the first crawl. **Dir** is the name of the subforum, and **Number of messages** denotes the number of messages in the topic.

Table 3: Popular topics

Dir	Number of messages
fap	218
h	208
b	201
ukko	156
fap	156
h	141
i	123
ukko	122
h	121
fap	113

In Table 3, we can see that the topic with the largest number of messages is located in the subforum devoted to erotica and pornography (fap), the second is in the one about drugs (h), and the third one is in the random chat forum (b).

Figure 5 presents the dependency of the overall number of messages during the first crawl on the date. The plot shows that the number of messages is growing, with the peak of 403 messages per day. However, since old threads in the forum are deleted when new ones appear and it happens through the directories, it is necessary to analyze similar dependency for individual threads.

4.1.3 Distribution of the messages over time

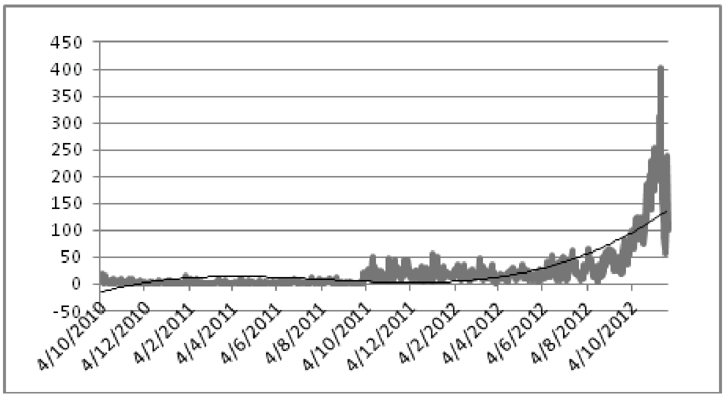


Figure 5: Overall number of messages

4.2 Suojeluskunta

4.2.1 Overall statistics

The first message on the forum was sent on 06 September 2011 and the latest one on 08 December 2012. In total, there were 3,839 messages in 1,204 topics; on average, approximately 8.4 messages were sent per day. Figure 6 shows the distribution of the messages over time. There are several peaks, which can be explained by presence of popular topics (e.g. peak at 13.09.2012 happens due to the discussion about the movie named "Innocence of Muslims").

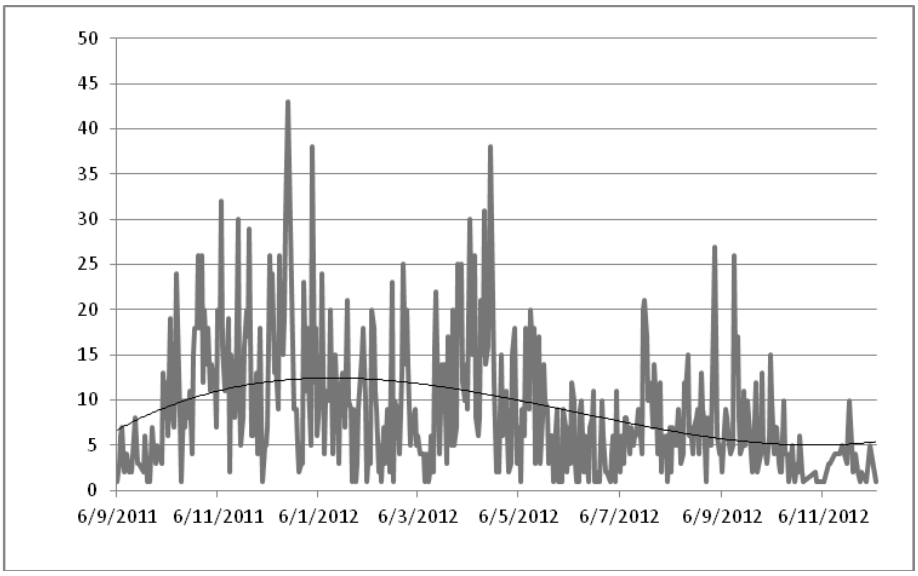


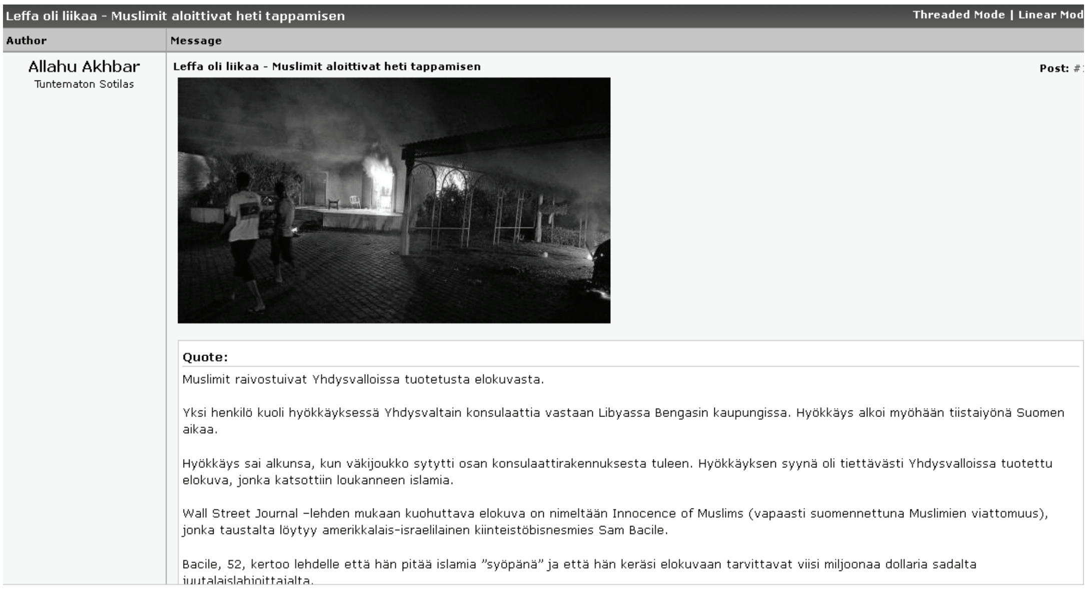
Figure 6: Post distribution over time

765 different user names were found, as well as one 'guest' user, a name that appears when a user does not fill in the "name" field when sending a message to forum. Of these, 85 users were registered at the forum. 2,444 messages were posted by registered users, 1,395 messages were sent by unregistered users, and 449 messages were sent by the guest user. 587 users had sent only one message.

Table 4 contains a description of names in the ten threads having the highest number of posts. The language of the board is Finnish, so the topics were translated. It can be seen that most of the names of the topics are related to hate speech.. Fig. 7 depicts screenshot of the forum message.

Table 4: Popular topics in Suojeluskunta

Finnish name	Translated name	Num. of posts
Leffa oli liikaa – Muslimit aloittivat heti tappamisen	Discussion about the movie *The Innocence of Muslims*	35
Ruutireseptejä	“Recipes of powder”	34
Etnistä puhdistusta Oulussa	“Ethnic cleansing in Oulu”	27
Throlaudan asekaupat	“Purchasing guns at Thorlauta”	26
Ählämeitä kylmäksi Oulussa!	“Muslims iced in Oulu”	23
Neekeristä Mannerheim	“Black Mannerheim”	23
Mainostakaa foorumianne	“Promote your forum”	23
Kaksi neekeriä raiskasi koulutytön Helsingissä	Rape of a school girl performed by two dark-skinned in Helsinki	22
Moi	“Hello”	21
Vallankumous hyvinvointivaltiossa	“A revolution in the welfare state”	21


Leffa oli liikaa - Muslimit aloittivat heti tappamisen

Author: Allahu Akhbar
Tuntematon Sotilas

Message: Leffa oli liikaa - Muslimit aloittivat heti tappamisen

Quote:

Muslimit raivostuivat Yhdysvalloissa tuotetusta elokuvasta.

Yksi henkilö kuoli hyökkäyksessä Yhdysvaltain konsulaattia vastaan Libyassa Bengasin kaupungissa. Hyökkäys alkoi myöhään tiistaiyönä Suomen aikaa.

Hyökkäys sai alkunsa, kun väkijoukko sytytti osan konsulaattirakennuksesta tuleen. Hyökkäyksen syynä oli tiettävästi Yhdysvalloissa tuotettu elokuva, jonka katsottiin loukanneen islamia.

Wall Street Journal –lehden mukaan kuohuttava elokuva on nimeltään Innocence of Muslims (vapaasti suomennettuna Muslimien viattomuus), jonka taustalta löytyy amerikkalais-israelilainen kiinteistöbisnesmies Sam Bacile.

Bacile, 52, kertoo lehdelle että hän pitää islamia ”syöpänä” ja että hän keräsi elokuvaan tarvittavat viisi miljoonaa dollaria sadalta juutalaislahjoittajalta.

Figure 7: Forum screenshot

5. Discussion

From the results presented above, we can see that a generic social media monitoring system is capable of gathering data from the websites and extracting from it various useful characteristics. Some of the characteristics, e.g., the most popular topics, are often shown at the sites explicitly, yet the presence of the data in the repository allows for a more detailed analysis. Also, it is important to consider the time of the data gathering and of the analysis queries. The results show that a board similar to Thorlauta can be completely collected about four times in 24 hours, and the Suojeluskunta forum can be collected completely 15 times per day. Furthermore, collection task may be narrowed down so that only important threads (e.g., those that change faster or contain particular keywords) may be gathered. However, fast crawling of the sites may raise the suspicion of the administrator of the site. Even though Tor does not show the IP of the client connecting to the site, high speed data collection may be observed since it will lead to a burst in the load on the site, which may be traceable.

Analysis algorithm performance should also be taken into consideration. Since, in the present architecture (Semenov and Veijalainen, 2012b; Semenov et al., 2011), the repository implements a multirelational directed graph, many queries result in the graph traversal queries. For instance, queries that return threads for a given subforum or messages for given thread are basically extractions of adjacent nodes, or breadth-first traversal for one level in depth. Extracting adjacent nodes for the presented repository architecture (incidence list) has a time complexity equal to O(|E|), where E is the number of edges. Thus, the extraction of the adjacent nodes would require the same procedure as for the extracted nodes, so a naïve method would take O(|E|^2) time. However, there are indexing techniques and algorithms to optimize this time (Khan et al., 2011). The two-hop neighbourhood extraction (extraction of messages by forum ID, see Fig. 4) presented in the paper takes, on average, 2 ms. However, for larger data sizes, the time would be different. For example, a two-hop neighbourhood query for a dataset of roughly 9M nodes and their connections, the collection of which is described in Semenov and Veijalainen (2012b) takes 2,704 ms, and the extraction of a three-hop neighbourhood takes, on average, 304 sec (the described results were measured in indexed tables, but a non-indexed PostgreSQL DB estimates 3,558,203 ms for the extraction of only a one-hop neighbourhood since it does a sequential scan; estimation result was taken from the EXPLAIN query).

The repository also allows a full text search of keywords from the messages. For instance, the string “suojelus” is found among the messages of Thorlauta four times (query takes 28 ms), and the string “thorlau” is found on Suojeluskunta 11 times, and the query takes, on average, 54 ms. For larger data sets, this result may change, but for small forums, it is acceptable.

6. Related work

(Christin, 2012) describes the crawling of the Silk Road. Daily crawls of the Silk Road were conducted from February 2012 to July 2012. The paper analyses distribution of items per categories, lifetime of availability of the items, lifetime of sellers’ profiles, etc. The paper estimates sales volume as USD 1.9 million/month and states that the main products being sold there are various controlled substances. Brezo and Bringas (2012) analyze the security risks facilitated by cryptocurrencies such as BitCoin: money laundering, illegal trafficking of substances, etc. In Ling et al. (2012b) Tor bridges are analysed. 2,365 Tor bridges were discovered via e-mail, and 2,369 bridges were found using Tor’s middle router. Bauer et al. (2011) provide “ExperimenTor”, a Tor network emulation kit and testbed for conducting research. AlSabah et al. (2011) note the slow performance of Tor and provide methods for its improvement. Wang et al. (2011) describe application-level attacks using a malicious exit relay, which injects code into the browser by modifying traffic transferred from the WWW into Tor. Ling et al. (2012a) describe cell-counting attack, which is able to confirm the communication among Tor nodes. Attack can be carried out having two Tor nodes: malicious exit relay, and relay which could act as entry point in the circuit. Exit relay embeds secret signal into the sequence of the packets. Then, signal is detected at entry relay, what confirms the communication Bernstein et al. (2011) describe an analysis of 4chan.org, a website meant for anonymous communication.

Acknowledgements

The research of the author was supported by the University of Jyväskylä Graduate School in Computing and Mathematical Sciences (COMAS). The author gratefully acknowledges Prof. Dr.-Ing. Jari Veijalainen from the University of Jyväskylä for providing advice and supervision that supported the work.

7. Conclusion

The present paper describes the crawling and analysis of the hidden services existing within the Tor network that are maintained in the Finnish language. The crawling and analysis were carried out by the prototype of a generic social media analysis system (Semenov et al., 2011). The paper provides a description of the crawling of Tor services and an analysis of the size and connection speed of the hidden services. In addition, the paper describes the topics discussed there. Then, the paper describes the possible limits of the monitoring.

Because of the size of the boards and the rate of data transfer, boards may be collected faster than they change, so their development may be followed in detail.

We can see that the most active thread existing on Thorlauta is devoted to selling controlled substances, and Suojeluskunta is wholly about racism and right-wing activism. Right-wing activists in Finland are active not only on Internet forums but in real life. “Keskisuomalainen” (2013) describes a stabbing in the library of the city of Jyväskylä: a man came to the discussion of a book describing right-wing activism in Finland and attacked with a knife the person invited to screen the persons wanting to enter the room.

As follows from the design of Tor, it allows very strong anonymity, and the people posting the messages cannot be easily tracked. However, it may be possible for the police to slip agents into the drug-selling sites so people selling the controlled substances can be detected. The next possibility is a natural language analysis of the content and detection of the authors by authorship analysis methods.

The presence of sites with such topics bemire the legitimate users of Tor, so all people using Tor in countries where the political situation is stable may be under suspicion.

References

Aliprandi, C., Marchetti, A., 2011. Introducing CAPER, a Collaborative Platform for Open and Closed Information Acquisition, Processing and Linking, in: Stephanidis, C. (Ed.), HCI International 2011 – Posters' Extended Abstracts. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 481–485.

AlSabah, M., Bauer, K., Goldberg, I., Grunwald, D., McCoy, D., Savage, S., Voelker, G., 2011. DefenestraTor: Throwing Out Windows in Tor, in: Fischer-Hübner, S., Hopper, N. (Eds.), Privacy Enhancing Technologies, Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 134–154.

Bauer, K., Sherr, M., McCoy, D., Grunwald, D., 2011. ExperimenTor: a testbed for safe and realistic tor experimentation, in: Proceedings of the 4th Conference on Cyber Security Experimentation and Test, CSET'11. USENIX Association, Berkeley, CA, USA, pp. 7–7.

Bernstein, M., Monroy-Hernández, A., Harry, D., André, P., Panovich, K., Vargas, G., 2011. 4chan and /b/: An Analysis of Anonymity and Ephemerality in a Large Online Community, in: ICWSM 2011.

Brezo, F., G. Bringas, P., 2012. Issues and Risks Associated with Cryptocurrencies Such as Bitcoin. Presented at the SOTICS 2012, The Second International Conference on Social Eco-Informatics, pp. 20–26.

Christin, N., 2012. Traveling the Silk Road: A measurement analysis of a large anonymous online marketplace (CMU-CyLab-12-018). CyLab.

CSO, 2012 How online black markets work [WWW Document], CSO. URL http://www.csoonline.com/article/705316/how-online-black-markets-work (accessed 11.23.12).

DIRECTIVE 2006/24/EC [WWW Document], 2013. . URL http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:105:0054:0063:EN:PDF

Facebook Newsroom [WWW Document], 2012. . URL http://newsroom.fb.com/content/default.aspx?NewsAreaId=22 (accessed 3.18.12).

Federal Business Opportunities [WWW Document], 2012. . URL https://www.fbo.gov/index?s=opportunity&mode=form&id=c65777356334dab8685984fa74bfd636&tab=core&_cview=1 (accessed 3.19.12).

Gawker (2011), The Underground Website Where You Can Buy Any Drug Imaginable [WWW Document], 2012. . URL http://gawker.com/5805928/the-underground-website-where-you-can-buy-any-drug-imaginable (accessed 11.20.12).

Gawker (2012), Booming Silk Road Drug Market Boasts $22 Million In Yearly Sales, Fancy Redesign [WWW Document], 2012. . Gawker. URL http://gawker.com/5932924/booming-silk-road-drug-market-boasts-22-million-per-year-in-sales-fancy-redesign (accessed 11.20.12).

Iltalehti, Iltalehti pääsi salaiselle nettisivulle: Huumeita, aseita ja lapsipornoa kaupataan avoimesti | Kotimaan uutiset | Iltalehti.fi [WWW Document], 2011. . URL http://www.iltalehti.fi/uutiset/2011100714534390_uu.shtml (accessed 12.11.12).

Keskisuomalainen [WWW Document], 2013. . URL http://www.ksml.fi/uutiset/kotimaa/useita-henkiloita-kuulusteltu-kirjaston-puukotuksesta/1293536 (accessed 2.7.13).

Khan, A., Li, N., Yan, X., Guan, Z., Chakraborty, S., Tao, S., 2011. Neighborhood based fast graph search in large networks, in: Proceedings of the 2011 ACM SIGMOD International Conference on Management of Data, SIGMOD '11. ACM, New York, NY, USA, pp. 901–912.

Ling, Z., Luo, J., Yu, W., Fu, X., Xuan, D., Jia, W., 2012a. A New Cell-Counting-Based Attack Against Tor. IEEE/ACM Transactions on Networking 20, 1245 –1261.

Ling, Z., Luo, J., Yu, W., Yang, M., Fu, X., 2012b. Extensive analysis and large-scale empirical evaluation of tor bridge discovery, in: 2012 Proceedings IEEE INFOCOM. Presented at the 2012 Proceedings IEEE INFOCOM, pp. 2381 –2389.

SC Magazine ,2011, Egyptians turn to Tor to organise dissent online [WWW Document],. .. URL http://www.scmagazine.com.au/News/246707,egyptians-turn-to-tor-to-organise-dissent-online.aspx (accessed 11.23.12).

Semenov, A., Veijalainen, J., 2012a. Ontology-guided social media analysis System architecture. Presented at the Special Session on Semantic Computing and Ontology Engineering - SCOE 2012, ICEIS.

Semenov, A., Veijalainen, J., 2012b. A modeling framework for social media monitoring. IJWET.

Semenov, A., Veijalainen, J., Boukhanovsky, A., 2011. A Generic Architecture for a Social Network Monitoring and Analysis System. IEEE, pp. 178–185.

Semenov, A., Veijalainen, J., Kyppo, J., 2010. Analysing the presence of school-shooting related communities at social media sites. International Journal of Multimedia Intelligence and Security 1, 232 – 268.

Syverson, P., 2011. A peel of onion, in: Proceedings of the 27th Annual Computer Security Applications Conference, ACSAC '11. ACM, New York, NY, USA, pp. 123–137.

Tor Project: Anonymity Online [WWW Document], 2012. . URL https://www.torproject.org/ (accessed 11.20.12).

Tor Project: Bridges [WWW Document], 2013 . URL https://www.torproject.org/docs/bridges (accessed 11.23.12).

Tor: Hidden Service Protocol [WWW Document], 2012. . URL https://www.torproject.org/docs/hidden-services.html.en (accessed 11.25.12).

torsocks [WWW Document], 2012. . URL http://code.google.com/p/torsocks/ (accessed 11.25.12).

United States Patent: 6266704 [WWW Document], 2012. . URL http://patft1.uspto.gov/netacgi/nph-Parser?Sect1=PTO2&Sect2=HITOFF&p=1&u=%2Fnetahtml%2FPTO%2Fsearch-bool.html&r=1&f=G&l=50&co1=AND&d=PTXT&s1=6266704.PN.&OS=PN/6266704&RS=PN/6266704 (accessed 11.20.12).

Wang, X., Luo, J., Yang, M., Ling, Z., 2011. A potential HTTP-based application-level attack against Tor. Future Generation Computer Systems 27, 67–77.

White Guard (Finland) - Wikipedia, the free encyclopedia [WWW Document], 2012. . URL http://en.wikipedia.org/wiki/White_Guard_(Finland) (accessed 12.11.12).

On the use of Honeypots for Detecting Cyber Attacks on Industrial Control Networks

Paulo Simões, Tiago Cruz, Jorge Gomes and Edmundo Monteiro
DEI–CISUC, University of Coimbra, Coimbra, Portugal
psimoes@dei.uc.pt
tjcruz@dei.uc.pt
jdgomes@dei.uc.pt
edmundo@dei.uc.pt

Abstract: In the last few years, Industrial Control Systems (ICS) have evolved from proprietary systems to open architectures, strongly interconnected with other corporate networks and even with the Internet. Initially, ICS systems were isolated by nature, being limited to the process network. Security was guaranteed by both obscurity and isolation (a bad practice, anyway!). Protocols were proprietary and not publicly documented, creating a false sense of security. Only manufacturers and attackers knew of failures and vulnerabilities, with both parts having no interest in their dissemination. However, the progressive move towards interconnecting the ICS with corporate networks and the internet, together with the use of mainstream ICT technologies and the increasing adoption of open, documented protocols, exposed serious weaknesses in SCADA (Supervisory Control and Data Acquisition) platforms. SCADA systems are becoming increasingly similar to ICT systems. Widely available, low-cost Internet Protocol devices are replacing proprietary solutions, which increases the possibility of cyber security vulnerabilities and incidents. At the same time, ICS are adopting ICT solutions to promote corporate connectivity and remote access capabilities, and are being designed and implemented using industry standard computers, operating systems and network protocols. While this integration introduces new ICT capabilities and tremendous cost optimization opportunities, it also provides significantly less isolation between ICS and the outside world. Many of the protection measures used in standard ICT security frameworks (e.g. firewalls, intrusion detection systems) can be adapted for SCADA environments. However, this also introduces additional security risks, since some of the traditional assumptions regarding ICT networks do not hold for ICS environments, due to their different set of priorities. While ICT systems put more focus on confidentiality and data integrity, ICS systems are built with availability as top priority, often at the cost of confidentiality and data integrity. This calls for domain-specific approaches to cyber attack detection in ICS systems, designed from the ground up to address its specific characteristics. This kind of SCADA-oriented cyber threat awareness constitutes the core topic of CockpitCI, a European project focused on improving the resilience and dependability of Critical Infrastructures (such as energy production and distribution grids). As part of this project, we have been researching novel strategies to develop, deploy and manage low cost honeypots for the SCADA field networks. Such honeypots share the core concepts of traditional honeypots (detect and profile cyber attacks by exposing innocuous and "fake" resources) but require a substantially different approach – from hardware design to operating model. In this paper we examine the role of specialized honeypots in the detection of cyber attacks on SCADA systems, debate how to implement such honeypots – from physical hardware considerations to remote management issues – and provide a complete example of such a honeypot.

Keywords: critical infrastructure protection, SCADA, intrusion detection systems, honeypots

1. Introduction

SCADA (Supervisory Control and Data Acquisition) is a common designation for several technologies, protocols and platforms used in Industrial Control Systems (ICS). SCADA systems are used in several scenarios, such as automation of production lines, control of nuclear or thermoelectric plants, management of distribution grids (electricity, gas, oil, water...) and many other applications.

In the past SCADA systems were restricted to isolated environments, relatively safe from external intrusion. However, with time they became increasingly exposed to more open environments, where they started to show their limits. The progressive move to more open scenarios, together with the use of Information and Communication Technologies (ICT) and the increasing adoption of open, documented protocols, exposed serious security weaknesses.

Moreover, the growing trend towards the interconnection of the ICS network with organizational ICT network infrastructures, and even with outside networks (for instance, for connection with internal company systems or for remote management by external contractors) created a new wave of security problems and incidents. In fact, there is a growing trend in the number of externally initiated attacks on ICS systems, when compared with internal attacks (Kang 2011).

As a result, the old practice of security by obscurity has become unfeasible. Still, the problem of security in SCADA systems has been more or less ignored for several years, and even now serious issues persist. Unsafe protocols such as Modbus (Modbus 2006), for instance, are still widely used in production systems. Moreover, new features such as the auto-configuration capabilities of certain equipment (plug-and-play) only got things worse, since attackers found it to be a valuable resource for attack planning and execution (Clarke 2004).

Nonetheless, the old-school mind-set still persists up to the point that some process managers still think of ICS systems as isolated and implicitly secure (Krutz 2006), disregarding the need for regular security updates or software patching procedures – and thus increasing the probability of successful attacks. While such procedures are trivial matters that are part of the regular maintenance routine in the ICT world, they must be dealt in a different way when it comes to ICS, mainly for two reasons:

- Some ICS components have to work on a continuous basis without interruptions, up to the point of working years without being reinitialized (Fovino 2010) (Zhu 2011).
- Equipment manufacturers must carefully and extensively test any software release before formal acceptance for usage on ICS platforms. Additionally, components still in use on ICS platforms often reach end-of-life support for specific devices or software frameworks.

To counteract threats, ICS have assimilated several security methods, tools and resources from the IT world, such as firewalls, Intrusion Detection Systems (IDS), and honeypots. While this approach provides a cost-effective way to add security to an ICS, it poses additional problems due to context differences (as it is the case with several firewall solutions that operate based on assumptions not applicable to ICS environments). This situation requires the development of domain-specific cyber-security mechanisms for ICS, designed to comply with the operational requirements of such infrastructures while still providing adequate protection.

Such development is one of the main objectives of the CockpitCI project (CockpitCI 2013), a European project focused on improving the resilience and dependability of Critical Infrastructures such as energy production and distribution grids, by automatically detecting cyber-threats and sharing real-time information about attacks among Critical Infrastructure owners. Among the domain-specific ICS cyber-security mechanisms that are being researched in the scope of CockpitCI, field network SCADA honeypots play an important role, providing the means for preventing or detecting attacks that directly target SCADA equipment, like Programmable Logic Controllers (PLC) or Remote Terminal Unit (RTU) devices, used to monitor and control industrial processes. While conceptually similar to the kind of honeypots commonly deployed in the ICT world, these implementations are specifically tailored to fit ICS needs. Their design substantially differs from conventional honeypot implementations, in terms of operation model and hardware components.

In this paper we present the main aspects of the proposed field network honeypot architecture, including the discussion of architectural options and implementation aspects. The rest of the paper is organized as follows: Section 2 addresses the problem of security in ICS, while an analysis of existing honeypot implementations is provided in Section 3. The proposed field network honeypot solution is presented in Section 4, and the issues related with its implementation are discussed in Section 5. The final conclusions we have drawn are presented in Section 6, along with some insights into future work directions.

2. Overview of ICS security issues

When it comes to security, there are several and significant differences between ICT and ICS domains. These differences are deeply rooted in their own specific characteristics. To begin with, ICS systems have a different set of priorities, when compared with ICT infrastructures – and, to a certain extent this inversion is one of the main causes of the security problems related with the SCADA infrastructure. Figure 1 illustrates this situation.

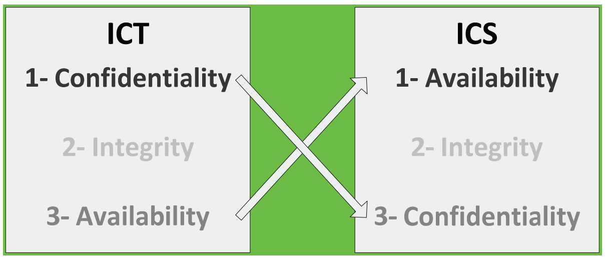


Figure 1: ICT vs. ICS priorities (adapted from (ISA-99.00.01))

On ICT networks, confidentiality and security have maximum priority, followed by communications integrity and, finally, by availability. For SCADA systems and ICS in general, on the other hand, there is an inversion of priorities caused by their critical nature (ISA-99.00.01), and availability comes first, even if at the cost of integrity and confidentiality. This difference of priorities has a real impact when it comes to choosing and implementing security mechanisms. Furthermore, it imposes a significant burden when importing security mechanisms from the ICT world to the ICS domain.

For these reasons, procedures that are trivial in the ICT world, such as frequently patching and updating a system, may become difficult or even impossible in some ICS scenarios. The impossibility or high cost of stopping production or even the explicit prohibition by the system's manufacturer can be pointed as examples. A critical facility cannot install an update on an operating system such as Windows unless the manufacturer of the SCADA software certifies it for the update, resulting in a lag of months or even years between the release of operating system patches and its adoption by critical facilities – even when the old operating system has known vulnerabilities. Another security problem is the lack of control systems knowledge by the IT security staff and vice versa.

Product lifecycle is another matter where the two domains differ. ICT infrastructures have substantially shorter lifecycles, when compared with their ICS counterparts. In ICT infrastructures, equipment and systems are regularly renewed from time to time, something that contrasts with the ICS philosophy of using mature systems, sometimes far beyond their projected lifetime. This limits the possibility of implementing some security mechanisms due to the limited capabilities of existing equipment (Igure 2006).

SCADA communication protocols, which are responsible for the interaction between field devices, such as PLC or RTU components and the stations that control and monitor them, are another example of such differences between ICS and ICT. One of such examples is the Modbus protocol (Modbus 2006), originally developed by Modicon (currently part of the Schneider Electric Group) in 1979 and still one of the most popular protocols for SCADA applications, mainly thanks to its simplicity and ease of use. Still, Modbus suffers from well known security problems: the lack of encryption or any other security measures of Modbus exposes this protocol to different vulnerabilities which have been analysed (Triangle 2007). Despite these known issues, SCADA protocols such as Modbus have a long lifespan and are still being massively deployed and used.

Simply put, when it comes to ICS, technology and platform maturity are valued as an implicit recognition of value and reliability, and even the disclosure of security issues related to them seems to have no effect in discouraging their usage or prompting the adoption of security measures to protect them. This has become the root cause of many ICS security issues that have ultimately been exploited with a variable degree of success, in recent times.

3. Honeypots and ICS

By definition, a Honeypot is a decoy or dummy target set up to attract and detect/observe attacks. By being exposed to probing and attack, its purpose is to lure and track intruders as they advance. Deploying and running a honeypot infrastructure requires a careful approach: it has to be planned in advance so that the infrastructure itself cannot be used to increase the attack surface, while keeping a low profile.

Honeypots can be classified in two groups: research and production. Research honeypots are used to obtain intelligence information about attack methods, while production honeypots are used to implicitly protect and ICT infrastructure by providing advance warning of attacks against the production infrastructure.

Honeypot types can also be distinguished by the ability of the attacker to interact with the application or services (Spitzner 2002):

- High-interaction honeypots can be probed, attacked and compromised. These honeypots let the attacker interact with the system in order to capture the maximum amount of information regarding his intrusion and exploitation techniques. Consequently, these honeypots have no restrictions regarding what the hacker can do, once the system is compromised and, as such, requires a lot of close monitoring and detailed analysis.
- Low-interaction honeypots (Provos 2004) emulate vulnerabilities rather than presenting real ones, therefore restricting the attacker's ability to interact with it. Mainly used as decoys, they are also less

flexible, albeit being more secure since there is little that the attacker can do. Nephentes (Baecher 2006) or Honeyd (Honeyd 2013) are examples of this honeypot type.

Finally, honeypots may also be classified as server honeypots or client honeypots (Riden 2010):

- Server honeypots are designed to passively wait for attacks.
- Client honeypots are able to actively search for malicious servers and behave like victims, an useful feature for detecting client-side browser exploits. Examples of client honeypots are the Shelia (Bos 2009), Honeymonkey (Wang 2006) and CaptureHPC (Hes 2009).

In the context of ICS, a honeypot can be implemented in a different fashion, depending on its operation scope. *In the operations network* a low-interaction honeypot might simulate the operation of a network server (e.g., control station), while *in the field network* a honeypot could be implemented using a system capable of simulating the operation of an RTU (e.g., a SCADA protocol emulator). Finally, *in the process or ICT network*, high-interaction honeypots might be adequate (even in the form of virtual machines, co-located on a same host), as well as low-interaction honeypots simulating minimal services. Moreover, in certain circumstances, some attacks targeting the system can be redirected to the honeypot, therefore providing more information about the attacker and his intentions.

Two previous research initiatives stand out as examples of ICS-specific honeypot implementations:

- The SCADA HoneyNet Project, from Cisco Critical Infrastructure Assurance Group. This project had its beginnings in 2004, and aimed at creating a framework to simulate industrial networks. It was able to simulate several levels of the system:
- *Stack level: simulate the TCP/IP stack of a device.*
- *Protocol level: simulate industrial protocols (for example, Modbus, Ethernet/IP).*
- *Application level: simulate several SCADA applications, such as web services and management consoles.*
- *Hardware: simulate serial ports and modems present in some SCADA devices.*

The project is no longer maintained, but it is still available for download in the project's website (Cisco 2004).

- The research of SCADA honeynets, or honeypot networks, by Digital Bond. Their solution makes use of at least two machines: one monitors the network's activity using a Generation III Honeywall, while the other simulates a PLC with available services for an attacker. The services available are: Modbus/TCP, FTP, Telnet, HTTP and SNMP (DigitalBond 2006).

The solution hereby proposed and described on next chapter distinguishes itself from these initiatives by presenting a low cost, modular and highly configurable solution for a field network honeypot, targeting Modbus environments. This solution is part of a broader platform that encompasses a distributed probe system for cyber attack detection – which is, on its turn, one of the building blocks of the CockpitCI framework.

4. A reference architecture for a SCADA/field network honeypot

SCADA system infrastructures usually comprise two levels: the operations network (where the master stations and databases are) and one or more field networks (where several PLCs and RTUs are deployed, controlling a critical process). Apart from these two levels, which are specific to SCADA operation, there is also the organization's IT network, comprising other devices (workstations, servers) and services (such as e-mail, accounting and stocking). Figure 2 depicts such scenarios.

The Modbus honeypot presented in this paper is designed to operate in a field network of a SCADA/ICS system, coexisting with the existing array of PLCs, RTUs and sensors/actuators that populate the network, binding to the network's unused IP addresses. Its fundamental operating principle is based on the assumption that, by faithfully emulating the behaviour and service footprint of a commercial PLC, the field network honeypot is able to faithfully persuade an attacker that it is a worthwhile target, acting as a decoy which actively reports any suspicious activity, by reporting events to the distributed IDS of the ICS, where they will be processed and correlated.

The honeypot architecture presented in this paper was designed to behave and operate as a PLC. Under normal conditions, the honeypot waits for a connection attempt from someone probing the network or accessing it with the intent of impersonating a master station. In practice, any attempt at contacting the honeypot device may potentially generate a security event since, by definition, any activity in the Modbus honeypot is illegal and unauthorized (or, at least, any activity except management operations).

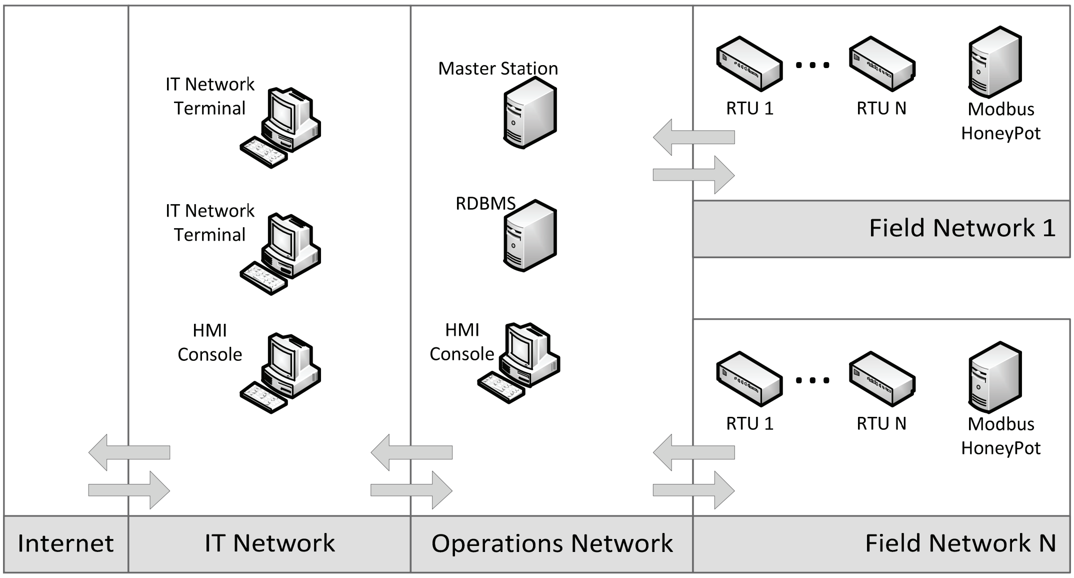


Figure 2: Modbus honeypot placement in an ICS environment

The architecture for the proposed Modbus honeypot for monitoring field networks is presented in Figure 3. This is a hybrid Modbus honeypot architecture, in the sense that it runs both simulated and complete implementations of services commonly available on PLC devices. Its main components and building blocks will be next described into detail.

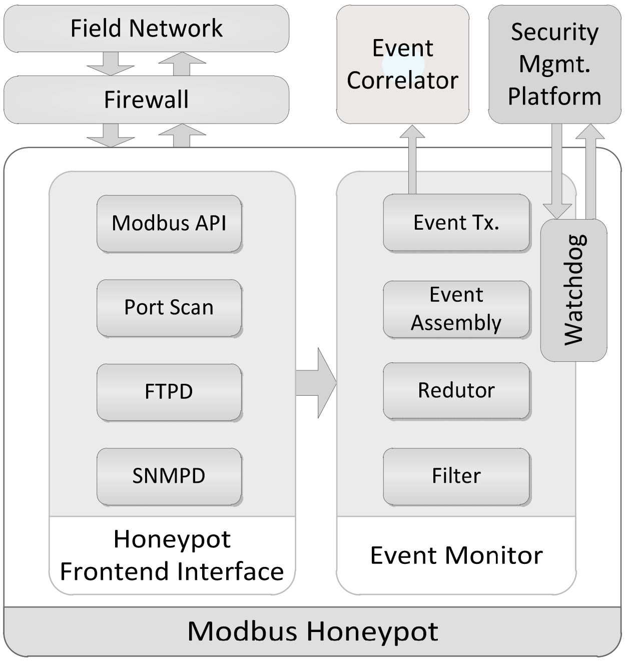


Figure 3: Modbus honeypot software architecture

4.1 Honeypot front-end interface

The connection with the field network is made through the *Honeypot Front-End Interface*. Within this block there are four components:

- The Modbus API simulator (*Modbus API*, in Figure 3), which accepts Modbus commands and behaves like a regular PLC, providing all the minimal protocol functionality (registers, operations, etc.).

- A File Transfer Protocol module (*FTPD*), providing an FTP service like the one commonly found on commercial PLCs.
- A Simple Network Management Protocol module (*SNMPD*), providing the SNMP device management interface and functionalities found on PLCs.
- A *Port Scan* detection module that is able to detect any probing activity in the remaining TCP/IP service ports.

The *Modbus API* module specifically implements the Modbus TCP protocol variant, widely used in ICS systems. This protocol was chosen for the honeypot due to three factors: standardization, popularity and because it is based in an open specification, whose documentation is easily obtainable. The protocol operation is easy to understand: a master station sends commands (mostly read or write operations in the majority of situations) to a RTU/PLC, that responds to them. The master station regularly polls and changes register values of the available RTU/PLC.

Like a real PLC, the Modbus API module implements variables (registers) for storing values, enabling an attacker to interact with it, polling and changing the values, with the *honeypot* responding to the attacker's requests with the corresponding response. The *Modbus API* module has also a Modbus message parser, to separate the various message fields in order to send them to the *Event Monitor* (described below). Additionally to the Modbus message fields (transaction identifier, protocol identifier, length field, unit identifier, function code and data bytes), this module stores additional information about the interaction, such as source IP, and a timestamp.

The *FTPD* and *SNMPD* modules respectively provide file transfer and management services commonly found in various Modbus PLCs. Each module has also a program monitoring the logs produced by these services. The program is aware of any entry in the logs and is capable of reporting it to the *Event Monitor*, for further analysis.

For a better coverage of interactions, a *Port Scan* module is included in the architecture. This module is not directly related to the SCADA technology context, being capable of capturing generic network interactions in order to detect the presence of an attacker. It listens to the remaining ports not covered by the other modules (*Modbus API*, *FTPD* and *SNMP*). Depending on the configuration used in the honeypot, it may do a simple interaction report or, alternatively, it may send detailed information to the *Event Correlator*.

4.2 Event monitor

The information obtained in the *Honeypot Front-End Interface* is analysed by the *Event Monitor*. The module is divided into four sub-modules:

- Filter.
- Event reduction and aggregation (Reductor).
- Event Assembly.
- Event transmission (Event Tx.).

Any event will pass through all the sub-modules in the following sequence: Filter (1); Event reduction and aggregation (2); Event Assembly (3); Event Transmission (4). The *Filter* and *Event reduction and aggregation* modules pre-process security events, optimizing system resources (e.g., processing and network) and contributing to increase the scalability of the solution up to larger ICS network scenarios.

The *Filter* module is used to filter relevant events according to previously defined configurations, which are stored in a file that is read when the module starts or by a *watchdog* module request. The *Filter* module can, for example, discard events of no interest. Relevant events are next sent to the *Event reduction and aggregation* module, which processes events in order to aggregate them by similar characteristics (for instance, grouping related events).

The *Event Assembly* module is responsible for creating the security event messages, using a standardized format. The event message structure is based on the IDMEF (Intrusion Detection Message Exchange Format) (Debar 2007), which is a standard data format designed for Intrusion Detection Systems. Using IDMEF as a standard message format can increase the interoperability from software amongst different vendors. IDMEF messages are based on XML, adopting an object-oriented data model where the top class is the IDMEF-Message Class. This class has two sub-classes: The Alert Class, and the Heartbeat Class. Each of these second level classes encompasses several aggregated classes that contain information about the message (such as sources, classification and detect time). IDMEF is a widely used format, being supported by several types of Host and Network IDS.

In this specific case, IDMEF messages are to be sent using a secure channel between a sensor and a processing node, in this case the honeypot and the *Event Correlator* (which is responsible for event processing and correlation), respectively. The *Event Tx.* module ensures message transmission.

4.3 Honeypot management (watchdog)

The honeypot contains a *watchdog* module for remote management. This module allows security staff to modify the honeypot configurations (for example, configuration of modules such as *Filter* and *Event reduction and aggregation*) from an authorized device. This is the only authorized connection to the honeypot. The watchdog module also allows some actions to be remotely performed, such as restarting a module. The connection to the watchdog module is protected by a secure channel (either using in-band or out-of-band management) and authenticated in both ends, using the Transport Layer Security (TLS) protocol.

4.4 Firewall

Containing measures must be implemented to prevent the attacker from gaining access to the ICS and turn the honeypot into an attack vector. The firewall has an important role in this, as it should allow all incoming connections to the honeypot, but it must deny connections from the honeypot to the remaining system (opposite of a typical firewall configuration). Connections from the honeypot to the attacker are the only outgoing connections that are allowed.

5. Implementation notes

This architecture is designed to run on a SBC (Single Board Computer), thus being a cost-effective solution. In our specific case, the proof-of-concept implementation was done using a Raspberry Pi (RaspberryPi) SBC equipped with a Broadcom BCM2835 System on Chip, 512 megabytes of RAM and Ethernet communication. The Raspberry Pi runs the Linux 3.2.27+ armv6l operating system.

The Modbus API implementation uses two Python libraries to manipulate Modbus data. The python module responsible for the Modbus protocol simulation is based on the modbus-tk library (Modbus-tk 2011). For parsing the Modbus messages a module based on the pymodbus library (pymodbus 2011) was used. The versions of the libraries are, respectively, modbus-tk-0.4.2 and pymodbus-0.9.0. The *SNMPD* and *FTPD* modules are both composed by a complete service implementation and a script to check its logs. The *SNMPD* module runs NET-SNMP version 5.4.3 (Net-SNMP), while the *FTPD* module uses VSFTPd 2.3.5 (VSFTPd). Additionally, each module includes a python-based script that parses service logs, looking for activity. Lastly, the *Port Scan* module is a C program coded using the *libpcap* (Libpcap) library and developed in-house for reduced footprint and computational requirements.

6. Conclusion

This paper presents the architecture for a Hybrid Modbus honeypot. It is able to detect attackers by simulating a Modbus TCP RTU, providing a *Honeypot Frontend Interface* composed by the simulation of the Modbus TCP protocol and the SNMP and FTP services, the last two being real services running on the device. Additionally, a *Port Scan* module is capable of monitoring the remaining network ports in order to detect additional attacks or probes.

One of the strong points of this architecture is the low cost hardware requirements, specifically a SBC (Raspberry Pi) that currently costs, will accessories included, around 50 Euro. Furthermore, its configurability allows security managers to fine-tune the monitoring specifications through the different modules of the honeypot. That enables them to manage the scalability of the solution according to the scenario in which it is deployed. It allows processing the data locally on the deployed nodes (honeypots) or transferring the entire event processing to the centralized *Event Correlator*, thus providing a compromise between network usage and centralized/distributed processing.

In the future, the *Event Correlator* will be further developed, as part of the presented architecture. It will be responsible to process the monitoring data received by the honeypots and will be responsible to report alarms to the system's operator.

Acknowledgements

The authors would like to thank the support of project CockpitCI (FP7-SEC-2011-1 Project 285647) and project iCIS (Intelligent Computing in the Internet of Services – CENTRO-07-0224-FEDER-002003). We would also like to rest of the CockpitCI team for its on-going support and collaboration.

References

Baecher, P., Koetter, M., et al. (2006), *The nepenthes platform: an efficient approach to collect malware,* Recent Advances in Intrusion Detection, vol. 4219 of Lecture Notes in Computer Science, pp. 165–184, Springer, Berlin, Germany, 2006.

Byres, E., Chauvin, B., et al. (2005) *The special needs of SCADA/PCN firewalls: architectures and test results*, 10th IEEE Conf. on Emerging Technologies and Factory Automation (ETFA 2005), 2005.

Cisco Critical Infrastructure Assurance Group (2004), *SCADA HoneyNet Project* [online], http://scadahoneynet.sourceforge.net/

Clarke, G., Reynders, D., and Wrigth, E. (2004) *Practical Modern SCADA Protocols: DNP3, 60870.5 and Related Systems,* Great Britain: IDC Technologies

CockpitCI (2013) *Cybersecurity on SCADA: risk prediction, analysis and reaction tools for Critical Infrastructures*, FP7 SEC-285647 Project Website, [online], http://www.CockpitCI.eu/

Digital Bond (2006) *SCADA Honeynet,* [online], http://www.digitalbond.com/tools/scada-honeynet/

Fovino, I., Coletta, A. and Masera, M., (2010) *Taxonomy of Security Solutions for SCADA Sector*, Project ESCORTS Deliverable 2.2, Version 1.1,

Hes, R., Komisarczuk, P., Steenson, R. and Seifert, C. (2009) *The Capture-HPC client architecture*, Technical report, Victoria University of Wellington.

Honeyd – Virtual Honeypot, [online], http://www.honeyd.org/

Igure, V.M., Laughter, S.A. and Williams, R.D. (2006) *Security issues in SCADA networks*, Computers & Security, Vol. 25, Issue 7, pp. 498-506, ISSN 0167-4048.

ISA-99.00.01 (2007) *Security for Industrial Automation and Control Systems - Part 1: Terminology, Concepts, and Models*, American National Standard.

Kang, D.J., Lee, J.J., Kim, B.H. and Hur, D. (2011) *Proposal strategies of key management for data encryption in SCADA network of electric power systems*, Int. Journal of Electrical Power & Energy Systems, Vol. 33, Issue 9, November 2011, pp. 1521-1526, ISSN 0142-0615.

Libpcap, *TCPDUMP/LIBPCAP public repository*, [online] http://www.tcpdump.org/

Pothamsetty, V. and Franz, M. (2009) *Modbus Firewall* [online] http://modbusfw.sourceforge.net/

Modbus-IDA (2006), Modbus Application Protocol Specification V1.1b

Modbus-tk (2011), [online], http://code.google.com/p/modbus-tk/

Net-SNMP, *Net-SNMP*, [online], http://www.net-snmp.org/

Pauli, S., Krahl, B. and Leuschner, B. (2003) *A guide to specifying, justifying and installing substation monitoring and control systems*, Petroleum and Chemical Industry Conference, pp.71-80, 2003.

Provos, N. (2004) *A Virtual Honeypot Framework*, Proceedings of the 13th USENIX Security Symposium, 2004.

Pymodbus (2011) *A Modbus Protocol Stack in Python*, [online] http://code.google.com/p/pymodbus/

RaspberryPi (2013), [online], http://www.raspberrypi.org/

Riden, J. and Seifert, C. (2010) "A Guide to Different Kinds of Honeypots", Symantec Connect, 2010, [online], http://www.symantec.com/connect/articles/guide-different-kinds-honeypots

Bos, H. (2009) *Shelia*, [online], http://www.cs.vu.nl/~herbertb/misc/shelia/

Spitzner, L. (2002) *Honeypots: Tracking Hackers*, Addison-Wesley Professional.

Triangle MicroWorks, Inc (2002) DNP3 Overview, Raleigh, North Carolina, [online], http://www.trianglemicroworks.com/documents/DNP3_Overview.pdf

VSFTPd, *vsftpd - Secure, fast FTP server for UNIX-like systems*, [online] https://security.appspot.com/vsftpd.html

Wang, Y., Beck, D., Jiang, X. Roussev, R., Verbowski, C., Chen, S. and King, S.T. (2006) *Automated Web Patrol with Strider HoneyMonkeys: Finding Web Sites That Exploit Browser Vulnerabilities*, NDSS 2006.

Zhu, B., Joseph A. and Sastry, S. (2011) *A taxonomy of Cyber Attacks on SCADA Systems*, Proc. of the 2011 Int. Conf. on Internet of Things and 4th Int. Conf. on Cyber, Physical and Social Computing (ITHINGSCPSCOM '11), pp. 380-388, IEEE Computer Society.

Critical National Infrastructure Protection: Evolution of Israeli Policy

Lior Tabansky
Yuval Ne'eman Workshop for Science, Technology and Security, Tel Aviv University, Tel Aviv, Israel
Cyber.ac.il@gmail.com

Abstract: Israel has developed a unique legal and regulatory model for critical national infrastructure protection, and has implemented it since late 2002. Recently, a comprehensive review of cyber security posture has been conducted, and significant policy changes are in progress. The Israeli approach appears to be highly successful, as the nation continues to be a world-class ICT power, and to provide cyber security for its critical infrastructure and beyond, while balancing conflicting interests and fostering cooperation between public, security, academic and private sectors. This article examines the evolution of Critical National Infrastructure Protection policy in Israel and analyses its performance. This study of the evolution of Israeli Critical National Infrastructure Protection policy may assist national policy-making, thus helping the developed countries in facing future challenges.

Keywords: Israel; cyber security; critical infrastructure protection; CIP; policy

1. Introduction

Cyber security issue has reached the national policy-making level in the developed countries. An international comparative study on "cyber defense" of 23 developed countries recently awarded Israel with a top grade, alongside Sweden and Finland. (*Cyber-security : the vexed question of global rules : an independent report on cyber-preparedness around the world* 2012) This article presents and examines the evolution of Critical National Infrastructure Protection policy in Israel and analyses its performance. The CIP policy has been built upon the insights of defense establishment and evolved from a limited involvement with IT branches of government, toward an early adoption of national CIP, and recently moved towards a comprehensive effort aimed at attaining global leadership, to contribute to national security, the economy, and international status. *This study of Israeli Critical National Infrastructure Protection policy may assist national policy-making, thus helping the developed countries in facing future challenges.*

2. Challenges of studying critical infrastructure protection policy in Israel

Public policy is most clearly expressed in its formal aspect. However, the latent aspect of the policy is often highly significant. The informal aspect is not manifested by legal, governmental, or regulatory decisions. It is revealed in an in-depth study which reflects on implementations, out of "reverse engineering" of the historical background. The informal aspects are elusive, but vital for understanding.

Israel's security doctrine is based on its clear numerical inferiority in the region. One major guiding principle is to counterbalance quantitative inferiority with a qualitative edge. This approach has been quite fruitful throughout the Zionist history. Thus scientific R&D in Israel, and the development of computerized applications in particular, should be viewed in context of gaining and maintaining quality advantage to balance quantitative inferiority.

The defense apparatus covered throughout the years the technological and operational aspects of the computerized revolution in wide-ranging fields, in accordance with the guiding perception of superiority that is based on science and technology. The cyber issue, having technical foundations that are closely similar to electronic warfare, was developed by the ongoing involvement of the defense forces with electronics and computers. The defense establishment in general, and the Israeli air force and intelligence services in particular, paid close attention to computers and electronics, including aspects related to information security, encryption and electronic warfare. Naturally, the details are shrouded in secrecy. However, those involved in the field in the relevant agencies have comprehended that a penetration into the enemy's computer systems provides considerable advantages in battle. For some of them it turned out that the exhilarating opportunities also carried new risks. These were the roots of the national CIP policy, finalized in late 2002.

3. Critical infrastructure protection in Israel, 2002 – 2011: regulation and cooperation

Following the accumulated understanding of civilian infrastructures vulnerabilities for cyber-attack, the MoD DR&D (MAFAT) has initiated council work at the National Security Council. Its outcome resulted in Special

Resolution B/ 84 on "The responsibility for protecting computerized systems in the State of Israel", of the ministerial committee on national security of December 11, 2002. After years of localized activity, the decision opened an era of national civilian cyber security policy. *In fact, it might have been the first active national Critical Infrastructure Protection policy in the developed world.*

The definitions stated in the 2002 Resolution are worth examining. First, "cyberspace" was not an independent area of operation, but one interconnected with all physical spaces. Second, "information" system is differentiated from "control" system. An information system *"performs mechanized activities of input reception, processing, storage, processing, and conveyance of information."* On the other hand, a control and supervision system is a *"computer-integrated system that controls and supervises the frequency and regulation of measureable activities, which are carried out by mechanized means within the information system itself."*

The responsibility for protecting computerized systems rests with the users and state regulators. A "user" is a supervised organization, which is in charge over financing all operation, protection, maintenance, upgrading, backup and recovery of its critical IT systems, as it shares information and activities with the regulator.

The regulators are the existing chiefs of security at government ministries, who are professionally responsible for guided bodies (for example, the Ministry of Communication is in charge over the telephone company Bezeq).

Two additional regulators are established: "The top *steering committee* for the protection of computerized systems in the State of Israel," and "The *national unit* for the protection of vital computerized systems."

The steering committee was established into the National Security Council, and comprised of senior government officials, representatives from the Bank of Israel, and the security forces. While the steering committee has a policy perspective, the "national unit" - "National Information Security Authority" (NISA or *Re'em* in Hebrew) - has the professional authority.[1]

The government's decision delegates eight responsibilities for NISA

- To assess the threat landscape – subject to the steering committee approval.
- To suggest classifying systems as critical and suggest oversight to the steering committee
- To develop protective doctrine and methods.
- To integrate intelligence.
- To provide professional instruction to the supervised organization.
- To set standards and operating procedures for the benefit of supervised organization.
- To develop technological expertise and cooperation with partners in Israel and abroad.
- To initiate and support research for developing defensive capabilities, in cooperation with the defense community.

The Israeli law only permits the General Security Service and the police to intervene with civilian matters for security purposes. Designating the responsibility for protecting vital computerized systems in public and privately-owned civilian bodies to the army would create an ethical problem and a legal hurdle. At the time, on top of its common duties, the Israeli police was inundated with criminal and terrorist activity sourced at the Palestinian Authority. In the GSS, NISA was in place within the GSS Protective Security Division long before 2002; it attended to information-security concerns at the governmental departments, Israeli embassies abroad, and state-owned companies. Broadening the authorities was the self-evident track of development, although one resulting in a substantial expansion of the workforce and the unit's budget.(Assaf 2008)

To implement the new arrangement, the "Regulation of Security in Public Bodies act of 1998" was amended, to provide the new bodies – the steering committee and NISA – with authority to supervise public bodies. It should be noted that despite the word "public", private ownership does not diminish the authorities of the law. The law defines "activities for protecting vital computerized systems" as activities required to preserve those vital computerized systems, information stored in them, confidential information related to them, as well as

[1] Israel General Security Service's www.shabak.gov.il/about/units/reem/Pages/default.aspx

preventing damages to those systems or the information in question. The supervised public body will appoint dedicated personnel on its behalf, which will be responsible for implementing the instructions of the authority. The law grants NISA extensive authorities in supervised organization: physical access to the most sensitive installations, the authority to review all of the organization's documents and the authority to overturn appointments of functionaries.

Circa 2006, the steering committee and NISA reached the conclusion that the Tel-Aviv Stock Exchange (TASE) – with operations wholly dependent on computerized systems – should be defined as "critical infrastructure". The legislative process needed to change the addendum of the law enables an organization to state its position on the matter. The CEO of the TASE, Mrs. Ester Levanon, presented two arguments against the decision, both through the official channels and via mass media.

- The head of the TASE personally[2] and the organization as a whole have a sound expertise in information security. Therefore, the organization does not need state oversight.
- TASE is required to administer itself in the global financial market. Any supervision of the stock exchange's computerized infrastructure by a clandestine intelligence service will severely tarnish the Israeli financial market reputation and divest foreign capital. Therefore, the proposed supervision will damage national economy.

After much deliberation, the injunction that defines the stock exchange as a critical infrastructure, thus subject to supervision by NISA - was approved in 2008. In late 2011 the oversight was extended to eight more companies - cellular and internet service providers, totaling over two dozen entities. In the decade that the current arrangement was enacted, there was a single objection.

Despite the costs mandated by the regulatory guidance, the study clearly demonstrates a high level of cooperation between the state authorities and the critical infrastructure owners and operators.

NISA was actively involved in cyber-security policy development and often initiated proposals to the steering committee. Some may consider this as a breach in "separation of authorities": is it legally appropriate for the professional operational body to intervene with the policy work of the steering committee? However, this state of affairs has not raised any legalistic criticism in Israel.

The cyber-risks have indeed intensified rapidly with the accelerated growth of cyberspace. There were growing voices in Israel, stressing the need for defensive updates and revisions. The next chapter is devoted to the review process, which heralded a new era in Israeli cyber-security policy.

4. The national cyber initiative, 2010

Prime Minister Benjamin Netanyahu approached the Israeli National Security Council in 2010 requesting a review on cyber security and Israel's policy. It looks as though the National Security Council did not implement this work. The Prime Minister approached retired brigadier-general professor Isaac Ben-Israel, who was heading the National Council for Research and Development in the Ministry of Science, to take on this mission. He indeed accepted this request, and during 2010, the prime minister's initiative was launched. The National Cyber Initiative has formed the basis for a substantial change in Israel's national cyber policy. The Initiative's activity, organization and structure are presented to properly understand the policy-design process.

The Initiative performed a systematic overview of the challenges and opportunities that the State of Israel faces as the cyber threat develops. The vision that guided the initiative is *"To preserve Israel's standing in the world as a center for information-technology development, to provide it with superpower capabilities in cyberspace, to ensure its financial and national resilience as a democratic, information-based, and open society."*(NCR&D 2011)

The initiative dealt with three key questions:

How can cyber-technology be incentivized and developed in Israel, to ensure Israel's standing as one of the top five leaders in the world by 2015?

[2] Prior to managing TASE, Mrs. Ester Levanon served as the GSS CIO .

Which infrastructures are needed to develop high-performance computing in Israel?

What arrangements (organization, responsibility, policy and regulation) are required in order to best deal with the challenges and threats in cyberspace?

The team composition reflected on the initiative's vision. For six months, eighty experts worked on the project: army representatives, academic experts, research and development institutional directors, and representatives from the ministries of finance and science and technology. The work was divided into seven sub-committees, and a business consultancy contributed an organizational-budgetary analysis. Below is a brief review of the key sub-committees findings and recommendations.[3]

5. The key products of the National Cyber Initiative

The sub-committee on monitoring and supervision acted to produce updated recommendations related to cyber protection in Israel, focusing on the components that should be encouraged to develop the field in Israel, to upgrade the State of Israel posture and ensure its resilience.

An examination of the threat *reemphasized that some specific cyber-attacks may cause widespread national harm,* defined in terms of damage to the uninterrupted functioning of the state. The threats were divided into areas: Governance capacity; State symbols; Provision of services to citizens; Economic damage; number of casualties; Invasion of privacy; Protecting state secrets; Disruption of public order.

In view of the threat, the committee examined the existing response: Israel has implemented policies for the protection of the defense sector and the critical national infrastructures. Yet, the civilian space was not properly addressed and some types of threats are neglected:

- Damage to civic services and services to private homes
- Threats to "concealed" computers, such as navigational devices or controllers in cars
- Degradation of morale by cyber means.

The sub-committee reached the conclusion that the current response is insufficient. The civilian space is more exposed than ever to cyber-attacks, yet the existing protection arrangement does not cover it. No single national agency for comprehensive cyber policy exists in Israel

The main objective proposed is to develop and deploy groundbreaking capacity that will provide Israel with an advantage in cyberspace. Israel's main assets are: A first-rate academic establishment, which contributes to research and innovation, a range of state and security organizations possessing information security expertise, and an extensive high-tech sector, including world leading information security companies. However, various roadblocks for productive collaboration were recognized. The major challenge is to incentivize the industry to invest in innovative cyber research and development, and to motivate higher education to research this field and develop a workforce. Another major challenge is the establishment of a coordinating body, which will overcome organizational obstacles, conflicts of interests, and other problems, in order fashion optimal policy.
The sub-committee formulated *six major recommendations to improve national cyber-security:*

- To raise awareness and education through, beginning with basic best-practice and leading to advanced interdisciplinary R&D.
- To develop knowledge and R&D infrastructure, especially for secure code development; to encourage the academia to launch multidisciplinary programs on cyber defense, and to establish a national cyber-simulator.
- To create a statewide protective shield based on domestic R&D, while addressing privacy concerns.
- To develop national operational capabilities in cyberspace for routine and emergency, while confronting moral, legal, and financial challenges.
- To upgrade the defense by combining technical and legislative measures.
- To deploy unique Israeli technologies, developed cooperatively by scientific and industrial sectors, with the government encouraging local procurement.

[3] One sub-committee dealt with security issues, and its findings are classified.

The sub-committee on cipher and simulation

Cryptography is essential to guard state secrets and intellectual property. The assumption that Israel has leading capacities in the scientific and theoretical areas was reaffirmed. Yet, in terms of industrial implementation, the conclusion was that Israel's scientific potential is not fulfilled. One of the hurdles is export restriction, which makes cipher development unprofitable.

It is recommended to encourage collaboration between the IDF and the academia. Another recommendation is to encourage the public to use available commercial encryption.

The need for simulation for research and training is clearly raised in academia and the defense establishment; it is recommended to develop a large-scale simulation facility that will cater to all consumers.

The sub-committee on super-computing and broadband

HPC is a necessary tool, yet Israel cannot purchase a super computer because of political restraints. The sub-committee's work concluded that Israel has "islands" of super-computing competence in the defense establishment, academia, and industries.

The major recommendation is to establish a national center for super-computing, which will be able to serve all consumers. The report presents suggested alternatives and their costs.

The sub-committee on the financial benefits

"Security" is a public good, and the market cannot fully supply it. But the private sector is fundamental in developing Israeli capacities in cyberspace. Therefore, if the state has to improve cyber security, it must engage the market. Encouraging cyber security industry will not only be a security benefit, but an economic one.

The main recommendations are:

- Increasing relevant defense R&D, while improving the export capacity of the products.
- Increasing transparency and cooperation in government and the Ministry of Defense, defense industrial base, and the civilian industry, while resolving secrecy restrictions.
- Encouraging initial market for instilling innovations.
- Developing cyber-protection indexes for organizations, to help selecting optimal solutions and also to contribute to risks insurance and financial grading.

The sub-committee for examining the academic benefits key recommendation is to establish a research excellence center on cyber-related issues, as part of I-Core: Israeli Centers of Research Excellence plan, adopted in a March 2010 government resolution.[4]

The sub-committee on policy and legislation reviewed and showcased the various relevant Israeli legislation and regulations, assessed the disparities, and recommended solutions to some legal issues. It reviewed national and international activity in ENISA, EU, EU-FP7, EUREKA, OECD, NATO, UN-ITU. The recommendation for Israel is to publish a formal policy document, and to participate in international initiatives, with an emphasis on the Council of Europe Convention on Cybercrime – 2001 (The Budapest Convention) to promote cyber-defense.

All was then integrated in a final report, which was submitted to the government. The fate of the "National Cyber Initiative" report was different than that of many other reviews and reports: the government soon adopted most of the recommendations.

[4] The Higher Education Reform Plan I-CORE. http://www.i-core.org.il/The-Higher-Education-Reform-Plan

6. Critical infrastructure protection: An ingredient of a broad national strategy, 2011

The Government resolution 3611 "Advancing the national capacity in cyberspace" of August 2011, which adopts the recommendations of the "National Cyber Initiative," was intended to *" improve the protection of national infrastructures essential for daily life in Israel, and to strengthen them, as much as possible, against cyber attacks, while promoting Israel's status as a center for ICT development, all through the cooperation of academia, industry, ministries, and the security organizations."*

The key aspect in the resolution is to establish a national cyber bureau (INCB) in the Prime Minister's office, reporting directly to the PM. Similar to the nature of the earlier "steering committee" the INCB is not an operational branch. This new body will coordinate policies, acting to implement the professional recommendations. The resolution designates INCB with these duties:

- To advise to the prime minister, the government and its committees on cyber-related issues and too coordinate the topic (excluding security and foreign relations).
- To council the government on a national cyber policy, to initiate legislation, to advertise the government policy; to follow-up on and inspect its implementation.
- To provide national cyber-threat estimate, combining relevant intelligence from all sources.
- To promote research and development on cyber and super-computing topics by the professional bodies, and to fashion national plans for education and sensible use of cyberspace; to promote cyber industry in Israel.
- To promote public awareness on cyber security and publish information, warnings, and directives.
- To promote domestic and international collaboration on cyber-related issues.

The 2002 critical infrastructure protection arrangement remains in place until further notice.

7. Conclusion and recommendations

A national CIP policy has been evolving in Israel since the mid-90s. The major milestone was the formation of the unique legislative and organizational CIP arrangement in 2002. The latest milestone was National Cyber Initiative in 2010 which looked beyond CIP, and suggested a comprehensive policy to boost overall national security while seizing new economic, industrial and foreign policy cyber- opportunities.

The ambitious vision became a policy objective. The new government resolution on cyber-policy approved in 2011 adopted this report's recommendations, forming the latest phase in Israel's CIP policy. Thus, a central coordinating body was established, reporting directly to the Prime Minister. The INCB will coordinate between activities in the various relevant organizations and handle organizational, legislative, and other roadblocks, in order to achieve the official Israeli policy objective – to be a top five global cyber superpower by 2015.

Reviewing the development of Israel's cyber-security policy raises interesting insights. The understanding that the changing technological reality impacts in a number of ways on trans-social issues, first developed at the professional operational level. The defense sector was the first to adopt cyber technology, due both to the operational challenges, and its ability to fund it. The Israeli case researched here presents another example for the central role that national security issues play in any technological revolution and the information revolution in particular.(Boot 2006, Gat 2006) Further on, the defense leaders were interested and able to transit their understandings to civilian sectors, which understandably were not eager to take on a whole new set of concerns. Eventually, after several years the government began to address information-security issues. Information sharing and knowledge transfer issues occupy researchers of management. This Israeli case demonstrates a dynamic information sharing and organizational flexibility – the same qualities that a government and particularly security organizations perceivably lack. It appears that the State of Israel's characteristics – the small size of the country, the informal culture, military service and the common experience of insecurity, have all contributed to the rather successful cooperation between defense and civilian levels.

The Israeli cyber-security policy appears to be quite effective. The Israeli national CIP commenced quite early (2002), both compared to other leading countries, and compared to the rest of issues that Israeli government has to deal with. Unlike the vast majority of public policy in Israel, it was not driven by a catastrophic event or a crisis. Rather, it should be seen as a *rare proactive government initiative. Despite the costs mandated by the regulatory guidance, the study clearly demonstrates a high-level of cooperation.* The analysis of this success leads to several insights for a model of a national policy coping with a rapid change in the technological environment. One finding is that contrary to the widely-held stereotype, a *government can play a positive role*: it can initiate proactive policy measures, it can be agile and responsive to changing demands, and can deploy a wide array of means to further a comprehensive response to challenges as well as opportunities.

The main recommendations for national CIP policy are to *reduce strictness, allow for innovative initiatives, and nurture the delicate combination of formal frameworks with informal ties*. When designing and implementing a national policy to deal with a novel rapidly evolving issue of technological origin, a state or a company must be able to act quickly and flexibly, to be able to learn in process, accommodate new experience, and adjust the policies accordingly.

The third recommendation is that there is *no one-fits-all CIP blueprint*. As any public policy, CIP has to suite the specific society. Public attitudes, ideology, social structure, economic development, market model, business competition, structure of the political system – are the major factors that influence CIP policy, whether the IT personnel likes it or not. To conclude, Critical Infrastructure Protection is an issue of national importance, and further research should promote the policy-making process in the developed countries.

Acknowledgements

I would like to thank The Neubauer Research Fellowships Program at the Institute for National Security Studies (INSS) and the Yuval Ne’eman Workshop for Science, Technology and Security at the Tel Aviv University

References

Assaf, D. (2008) 'Critical Information Infrastructure Protection Policy in Israel', *The CIP Report , George Mason University School of Law,* 6(12).

Boot, M. (2006) *War made new : technology, warfare, and the course of history, 1500 to today,* New York: Gotham Books.

Cyber-security : the vexed question of global rules : an independent report on cyber-preparedness around the world, (2012) Brussels: Security & Defence Agenda (SDA).

Gat, A. (2006) *War in human civilization,* Oxford; New York: Oxford University Press.

NCR&D (2011) *"The national cyber initiative" – a special report for the Prime Minister* Tel-Aviv:

Comparative Analysis of Open-Source log Management Solutions for Security Monitoring and Network Forensics

Risto Vaarandi and Paweł Niziński
NATO Cooperative Cyber Defence Centre of Excellence, Tallinn, Estonia
risto.vaarandi@ccdcoe.org
pawel.nizinski@ccdcoe.org

Abstract: Nowadays, centralised event log management plays a crucial role in security monitoring and network forensics. While commercial log management solutions are regularly reviewed and compared by independent organisations (e.g. Gartner Magic Quadrant reports), such comparisons are often hard to find for open-source tools, especially for recently created solutions. However, many institutions are using open-source tools for monitoring and forensics, since they allow for implementation of incident detection and analysis frameworks in a cost-efficient way. Furthermore, recently appeared open-source solutions have started a new architectural trend, where the log management system consists of independent and replaceable modules which interact through well-defined interfaces and protocols. In contrast, most commercial systems are essentially monolithic solutions where individual components (such as the GUI or event collector service) cannot be changed for a component from another vendor. In this paper, we will focus on widely used open-source solutions for log management and discuss their recent developments. We will also cover novel technologies and tools which have appeared during the last 2-3 years.

Keywords: log management, network forensics, security monitoring

1. Introduction

Centralised event log management plays a crucial role in security monitoring and network forensics, since it allows for gathering events from thousands of nodes to a few dedicated servers where central analysis is carried out. The analysis can be a real-time process, where security incidents are detected from incoming events through event correlation and other advanced monitoring techniques; it can also be an off-line forensics activity, where past events are investigated in order to study security incidents that have already occurred.

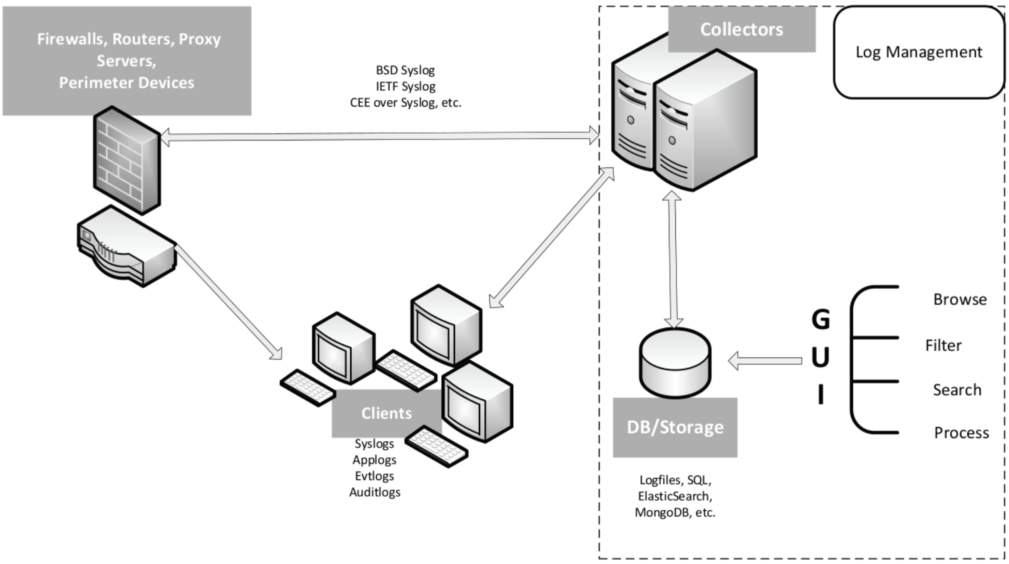


Figure 1: An architectural overview of an event log management framework

Without event log collection to central location(s), monitoring and forensics activities would have to be carried out at individual network nodes, which is time-consuming and prevents the timely resolution of security incidents. Furthermore, the attacker may erase events from the local log in order to remove any traces of his/her malicious activities. For the above reasons, a number of commercial and open-source solutions have been created for log collection and centralised analysis. Figure 1 provides an overview of the essential components of an event log management framework.

As depicted in Figure 1, nodes of the IT system are using protocols like IETF syslog for sending events to the collector service(s) at the central log server(s). Collector services use various techniques for filtering and normalising events and store preprocessed events into some means of storage (e.g. a database or a flat file). Human analysts can access stored data through a GUI for carrying out searches, creating reports, and other analytical tasks.

In this paper, we will focus on most prominent open-source log management solutions. The first contribution of this paper is an analytical comparison of presented tools; the second contribution is a detailed comparative performance evaluation of the tools. For this purpose, we conducted a set of experiments for assessing their resource consumption and event processing speed under a heavy load (all experiments were carried out in December 2012). To the best of our knowledge, such performance evaluations have not been conducted recently for state-of-the-art open-source log management solutions.

The remainder of this paper is organised as follows: section 2 provides an overview of log collection protocols and log storing techniques, section 3 provides a discussion and performance evaluation for leading open-source syslog servers, section 4 presents an overview and performance assessment of graphical log management systems, and section 5 concludes the paper.

2. Log collection and storing

Until the 1980s, event logging was mostly accomplished by writing events to a file on a local disk. In the 1990s, UDP-based BSD syslog protocol became widely used for log collection (Lonvick 2001). The protocol defines 24 message facility (sender type) and 8 severity values which range from 0 to 23 and 0 to 7 respectively. For reasons of convenience, textual acronyms are often used instead of numerals, e.g. *daemon* denotes facility 3 and *warning* denotes severity 4. According to BSD syslog, the payload of the UDP packet carrying the message must have the format *<Priority>Timestamp Hostname MSG*, where *Priority* is defined as a numeral *8*facility_value + severity_value*. For example, the following message represents a warning “ids[1299]: port scan from 192.168.1.102” for *daemon* facility which was issued on November 17 at 12:33:59 by the network node myhost2:

<28>Nov 17 12:33:59 myhost2 ids[1299]: port scan from 192.168.1.102

By convention, the alphanumerals that start the *MSG* field are regarded as the *Tag* subfield which represents the name of the sending program (“ids” in the above example), while the remainder of the *MSG* field is regarded as the *Content* subfield (“[1299]: port scan from 192.168.1.102” in the above example).

Despite its popularity, BSD syslog protocol has a number of drawbacks which are summarised below:

- no support for reliable message transmission over TCP;
- no support for encryption and authentication;
- timestamps are not specific enough, lacking the timezone, year, and fractions of a second;
- apart from *Tag* and *Content* subfields, the *MSG* field has no structure.

In order to make message transmission more reliable, a TCP flavor of BSD syslog protocol was proposed during the previous decade, where a stream of newline-separated messages in BSD syslog format is sent over a TCP connection. In 2009, more advanced IETF syslog protocol was introduced that addresses all drawbacks of BSD syslog (Gerhards 2009; Miao, Ya and Salowey 2009; Okmianski 2009). The IETF syslog supports secure message transmission over TLS, and uses a new message format with more detailed RFC3339 timestamps (Klyne and Newman 2002) and structured data blocks. The following example depicts the previous sample message in a new format:

<28>1 2012-11-17T12:33:59.223+02:00 myhost2 ids 1299 - [timeQuality tzKnown="1" isSynced="1"] port scan from 192.168.1.102

The priority specification *<28>* is followed by the protocol version number (1). In addition, the sender is passing a structured data block *timeQuality* to the receiver, indicating that sender’s clock is synchronised to an external reliable time source.

Another currently ongoing effort to introduce structure to log messages is the Common Event Expression (CEE) initiative. CEE has proposed JSON and XML formats for events, while also suggesting the use of BSD and IETF syslog protocols for transporting JSON-formatted events (CEE 2012). In addition, there are some application-specific protocols for structured logging, e.g. GELF.

When logs are collected to central server(s), they need to be written to a permanent storage. In many cases, incoming log messages are written into flat files on the disk of the central server. While this consumes little CPU time and allows for receiving large volumes of events per second, searching relevant events from flat files can be time consuming and resource intensive. Therefore, log messages are often stored into SQL databases that facilitate fast and efficient searching. However, each database table contains a fixed number of columns, with each column representing a certain message field of a fixed type (e.g. integer or string). As a consequence, the fields of a log message have to comply with the structure of a database table that is holding the log data. In order to address this requirement, fields of all collected messages must be well known in advance, so that appropriate database schema can be defined.

Unfortunately, if logs are received from a wide variety of sources, log messages in previously unseen formats are likely to appear. In order to address this problem, *document-oriented databases* have emerged recently as alternative log storage solutions. Although the implementations of document-oriented databases vary, they can be viewed as a collection of *documents*, where each document is usually a record of fieldname-value pairs. It is important to note that each inserted document can have a unique set of fields which do not need to be known in advance.

During the last 1-2 years, Java-based *elasticsearch* has emerged as one of the most widely used document-oriented database engines for storing log data (Elasticsearch 2013). *Elasticsearch* accepts new documents in JSON format over a simple HTTP interface, inserting the incoming document into a given index and thus making the document searchable for future queries. Support for distribution is built into the core of *elasticsearch* and several instances of *elasticsearch* engines can be easily joined into a single cluster. Furthermore, each database index can be split into so-called *shards,* which can be located at different cluster members. Also, each index can have one or more *replicas* for implementing a fault tolerant cluster.

3. Syslog servers

In this section, we will cover leading open-source syslog servers *rsyslog* (Rsyslog 2013), *syslog-ng* (Syslog-ng 2013) and *nxlog* (Nxlog 2013). The discussion and experiments presented in this section are based on *rsyslog* 7.2.3, *syslog-ng* 3.3.7 and *nxlog* ce-2.0.927.

3.1 Rsyslog, syslog-ng and nxlog

Rsyslog, *syslog-ng* and *nxlog* have been designed to overcome the weaknesses of traditional UNIX *syslogd* server which supports only BSD syslog protocol, and is able to match and process messages by facility and severity. *Rsyslog*, *syslog-ng* and *nxlog* support not only such simple message matching, but advanced message recognition with regular expressions, conversion of messages from one format to another, authenticated and encrypted communications over IETF syslog protocol, etc. *Syslog-ng* and *nxlog* have also a commercial edition with extended functionality. *Rsyslog* and *syslog-ng* run on UNIX platforms, while *nxlog* is also able to work on Windows.

The configuration of all servers is stored in one or more textual configuration files. *Syslog-ng* uses a highly flexible and readable configuration language which is not compatible with UNIX *syslogd*. The message sources, matching conditions and destinations are defined with named blocks, with each definition being reusable in other parts of the configuration. *Syslog-ng* is also well documented, featuring a detailed administrator's manual consisting of hundreds of pages. This makes it easy to create fairly complex configurations, even for inexperienced users.

Rsyslog is an efficient syslog server that was specifically designed for handling heavy message loads (Gerhards 2010). *Rsyslog* uses a quite different configuration language that supports UNIX *syslogd* constructs. This allows for easy migration of old *syslogd* setups to *rsyslog* platform. Also, there are many additional features in the *rsyslog* configuration language. Unfortunately, over time, several different syntaxes have been included in the language which has introduced inconsistencies (Gerhards 2012). Functionality-wise, *rsyslog* supports several highly useful features not present in open-source versions of *syslog-ng* and *nxlog*. Firstly, it is possible to set up temporary message buffering to local disk for log messages which were not sent successfully over the network. The buffering is activated when the connection with a remote peer is disrupted, and when the peer becomes available again, all buffered messages are retransmitted. Secondly, the latest stable release of *rsyslog* has an efficient support for *elasticsearch* database (Rsyslog-ver7 2012).

Nxlog uses the Apache style configuration language. As with *syslog-ng*, message sources, destinations and other entities are defined with named blocks which allows them to be reused easily. Also, *nxlog* has a solid user manual. The advantages of *nxlog* over other syslog servers include native support for Windows platform and Windows Eventlog. Also, *nxlog* is able to accept input events from various sources not directly supported by other servers, including SQL databases and text files in custom formats. Finally, *nxlog* is able to produce output messages in the GELF format, allowing for seamless integration with the *Graylog2* log visualisation solution.

In order to illustrate the differences between the configuration languages of syslog-ng, rsyslog and nxlog, we have provided configuration statements in three languages for the same log processing scenario:

```
##### configuration for syslog-ng

@version:3.3

source netmsg { udp(port(514)); };
filter ntpmsg { program('^ntp') and level(warning..emerg); };
destination ntplog { file("/var/log/ntp-faults.log"); };

log { source(netmsg); filter(ntpmsg); destination(ntplog); };

##### configuration for rsyslog

$ModLoad imudp
$UDPServerRun 514

if re_match($programname, '^ntp') and $syslogseverity <= 4 then {
  action(type="omfile" file="/var/log/ntp-faults.log")
}

##### configuration for nxlog

<Input netmsg>
   Module  im_udp
   Host    0.0.0.0
   Port    514
   Exec    parse_syslog_bsd();
</Input>

<Output ntplog>
   Module  om_file
   File    "/var/log/ntp-faults.log"
   Exec    if $SourceName !~ /^ntp/ or $SyslogSeverityValue > 4 drop();
</Output>

<Route ntpfaults>
   Path    netmsg => ntplog
</Route>
```

First, the above configurations tell syslog servers to accept BSD syslog messages from UDP port 514 (in the case of *syslog-ng* and *nxlog*, the name *netmsg* is assigned to this message source). Then, the message filtering condition is defined for detecting messages with the *Tag* field matching the regular expression *^ntp* (in other words, the name of the sending program begins with the string "ntp"), and with the message severity falling between *warning* (code 4) and *emerg* (code 0). Note that for *nxlog*, the inverse filter is actually used for dropping irrelevant messages. Finally, the file /var/log/ntp-faults.log is used as a destination for storing messages that have passed the filter (in the case of *syslog-ng* and *nxlog*, the name *ntplog* is assigned to this destination).

3.2 Experiments for evaluating the performance of rsyslog, syslog-ng and nxlog

In order to evaluate how well each server is suited for the role of a central syslog server, we conducted a number of experiments for assessing their performance. During the experiments we used three benchmarks for stress-testing the servers, and measured the CPU time consumption and overall execution time of each server during every test run. We call the benchmarks BSD-Throughput, IETF-Throughput and Filter-Throughput, and define them as follows:

- BSD-Throughput – 1 client sends 10,000,000 plain-text BSD syslog messages to the syslog server over TCP. The messages are written to one log file without filtering.
- IETF-Throughput – 1 client sends 10,000,000 encrypted IETF syslog messages to the syslog server over TCP. The messages are written to one log file without filtering.
- Filter-Throughput – there are 5 clients, each sending 2,000,000 plain-text BSD syslog messages to the syslog server over TCP. The messages are identical to messages of the BSD-Throughput benchmark which allows for making performance comparisons between two benchmarks. The server is configured to process incoming log data with 5 filters and to write messages into 5 log files. All filters include regular expression match conditions for the message text (*Content* field) and/or program name (*Tag* field), and some filters also have additional match conditions for message facility and severity.

Table 1: Comparative performance of rsyslog, syslog-ng and nxlog

	rsyslog	syslog-ng	nxlog
BSD-Throughput maximum, minimum and average CPU time consumption (seconds)	17.994 14.601 16.024	97.094 88.942 94.090	86.809 83.929 85.100
BSD-Throughput maximum, minimum and average execution time (seconds)	12.778 10.736 11.853	98.961 90.715 96.079	54.261 52.253 53.281
IETF-Throughput maximum, minimum and average CPU time consumption (seconds)	47.190 41.448 43.883	115.684 106.813 111.823	166.455 161.536 164.605
IETF-Throughput maximum, minimum and average execution time (seconds)	33.268 30.184 31.337	128.055 108.911 115.084	71.605 69.434 70.404
Filter-Throughput maximum, minimum and average CPU time consumption (seconds)	50.265 45.626 47.661	216.093 211.143 213.683	2237.954 2191.502 2216.758
Filter-Throughput maximum, minimum and average execution time (seconds)	44.389 39.941 41.624	60.496 58.886 59.792	715.320 701.210 707.933

Note that *rsyslog* and *nxlog* always run in multi-threading mode, while for *syslog-ng* this mode has to be enabled manually. During the testing we discovered that for BSD-Throughput and IETF-Throughput *syslog-ng* performance decreased in multi-threading mode (according to the *syslog-ng* manual, this mode yields performance benefits in the presence of many clients, filters and message destinations). Therefore, we ran *syslog-ng* in a default single-threaded mode for BSD-Throughput and IETF-Throughput tests. Also, we discovered that the tested *nxlog* version was not able to handle IETF syslog messages as required by RFC5425 – in a stream of incoming messages, only the first syslog frame was properly recognised. Also, the tested version was not able to parse some timezone specifications in RFC3339 timestamps. For these reasons, we modified the IETF-Throughput benchmark for *nxlog*, so that instead of proper RFC5425 frames, a stream of newline-separated IETF messages was sent to *nxlog* over TLS connection (this unofficial data transmission mode is supported by all tested servers as an extension to standard modes). The tests were carried out on a Fedora Linux node with 8GB of memory and an Intel Core i5 650 processor. We repeated each test 10 times, and the results are presented in Table 1.

The results reveal several interesting aspects of server performances. Firstly, the performance of *rsyslog* is superior to other servers, both in terms of raw message throughput from single client and efficiency of message filtering for multiple clients. Also, *rsyslog* is able to share its workload between several CPU cores with multi-threading, and thus the execution times are less than overall consumed CPU times. Multi-threading is used very efficiently by *nxlog*, resulting in execution times being 1.5-3 times lower than used CPU time. Unfortunately, the performance of *nxlog* filters is poor – compared with the BSD-Throughput test, the average CPU time consumption increased about 26 times. In contrast, CPU time consumption for *syslog-ng* increased only 2.27 times, while the average execution time actually decreased by almost a half due to the manually enabled multi-threading mode.

4. Log visualisation and preprocessing applications

While the use of databases for storing log messages facilitates fast searching with flexible query language, it is tedious and time-consuming for the user to write separate queries for each search, report or other analytical task. Furthermore, the output from database queries is textual and the user would have to use a separate tool, or even programming language, for visualising this output. For solving this problem, several open-source log visualisation applications have been developed during the last 2-3 years, which are all using *elasticsearch* as their main database engine. In this section, we will review and conduct performance evaluation experiments for *logstash* (version 1.1.5), *Graylog2* (version 0.9.6) and *Kibana* (version 0.2.0).

4.1 Logstash

Logstash is a Java-based utility where a graphical user interface and embedded *elasticsearch* engine are encapsulated into a standalone jar-file (Logstash 2013). This eases the installation of *logstash* since the user does not have to download and install all product components separately. One advantage of *logstash* is its support for many different input and output types. Currently, there are input plugins for accepting syslog messages over TCP and UDP, but also for many other messaging protocols like AMPQ, RELP, GELF, IRC, XMPP, twitter, etc. Among outputs, other monitoring and visualisation systems are supported, including *Nagios*, *Zabbix*, *Loggly*, *Graphite* and *Graylog2*.

Another advantage of *logstash* is a number of different message filter types which allow for flexible recognition, filtering, parsing and conversion of messages. For instance, it is possible to convert multi-line messages into single line format, filter out messages with regular expressions, add new fields to messages from external queries and accomplish many other advanced message manipulation tasks.

One of the most commonly used *logstash* filter types is *grok*. While most log management tools use regular expression language for message matching and parsing, *grok* filters employ many predefined patterns that represent regular expressions for common matching tasks (GrokPatterns 2013). For instance, the pattern PROG is defined as the regular expression *(?:[\w._/%-]+)* and is designed to match the name of the logging program. Using predefined *grok* patterns, a person who is not familiar with the regular expression language can accomplish event parsing tasks in an easier way.

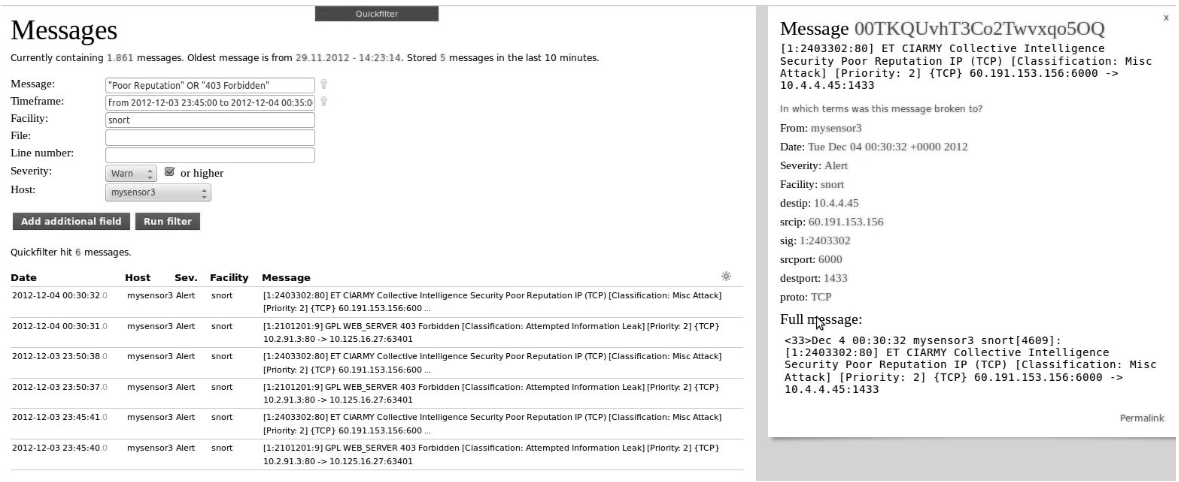
In order to use the GUI of *logstash*, it must be configured to insert events into its embedded *elasticsearch* database. With the GUI it is possible to carry out basic searches from log messages in the embedded database. Unfortunately, compared with other log visualisation tools, the GUI of *logstash* has quite limited functionality. However, since *logstash* has powerful event filtering and conversion capabilities, it is used mostly as an event preprocessor for different systems, including other log visualisation systems.

4.2 Graylog2

Graylog2 (Graylog2 2013) is a log management system which consists of a Java-based server and a web interface written in Ruby-on-Rails. *Graylog2* server can receive BSD syslog messages over UDP and TCP, but it also features its own GELF protocol that facilitates structured logging (GELF 2013). In addition, *Graylog2* can accept syslog and GELF messages over the AMPQ messaging protocol. Unfortunately, *Graylog2* is not able to parse structured IETF syslog messages and recognise already defined fieldname-value pairs. For BSD syslog messages, *Graylog2* can recognise the standard *Priority*, *Timestamp*, *Hostname* and *MSG* fields, but cannot by default parse the unstructured *MSG* field.

The parsing problem for unstructured messages can be overcome in several ways. First, *Graylog2* server supports message parsing and rewriting through Drools Expert rules and regular expressions (Graylog2-Drools 2013). Second, many sites use *logstash* for receiving syslog messages, parsing out relevant message fields with *grok* filters, and finally sending the parsed messages in structured GELF format to *Graylog2*. Finally, since the *nxlog* syslog server supports the GELF protocol, it can be used as a frontend for receiving both encrypted and plain-text syslog messages and converting them into GELF messages.

For storing parsed messages, *Graylog2* uses *elasticsearch* as its main backend. Unfortunately, all log messages are stored into a single index*,* which might cause performance issues as many log messages are inserted into it over time. This issue can be addressed with creating a separate *elasticsearch* index for each week or day (the latter technique is used by *logstash*). Fortunately, the developers of *Graylog2* are aware of this problem and it is supposed to be fixed in the next version.



Messages

Currently containing 1.861 messages. Oldest message is from 29.11.2012 - 14:23:14. Stored 5 messages in the last 10 minutes.

Message: "Poor Reputation" OR "403 Forbidden"
Timeframe: from 2012-12-03 23:45:00 to 2012-12-04 00:35:0
Facility: snort
File:
Line number:
Severity: Warn ☑ or higher
Host: mysensor3

Quickfilter hit 6 messages.

Date	Host	Sev.	Facility	Message
2012-12-04 00:30:32.0	mysensor3	Alert	snort	[1:2403302:80] ET CIARMY Collective Intelligence Security Poor Reputation IP (TCP) [Classification: Misc Attack] [Priority: 2] {TCP} 60.191.153.156:600 ...
2012-12-04 00:30:31.0	mysensor3	Alert	snort	[1:2101201:9] GPL WEB_SERVER 403 Forbidden [Classification: Attempted Information Leak] [Priority: 2] {TCP} 10.2.91.3:80 -> 10.125.16.27:63401
2012-12-03 23:50:38.0	mysensor3	Alert	snort	[1:2403302:80] ET CIARMY Collective Intelligence Security Poor Reputation IP (TCP) [Classification: Misc Attack] [Priority: 2] {TCP} 60.191.153.156:600 ...
2012-12-03 23:50:37.0	mysensor3	Alert	snort	[1:2101201:9] GPL WEB_SERVER 403 Forbidden [Classification: Attempted Information Leak] [Priority: 2] {TCP} 10.2.91.3:80 -> 10.125.16.27:63401
2012-12-03 23:45:41.0	mysensor3	Alert	snort	[1:2403302:80] ET CIARMY Collective Intelligence Security Poor Reputation IP (TCP) [Classification: Misc Attack] [Priority: 2] {TCP} 60.191.153.156:600 ...
2012-12-03 23:45:40.0	mysensor3	Alert	snort	[1:2101201:9] GPL WEB_SERVER 403 Forbidden [Classification: Attempted Information Leak] [Priority: 2] {TCP} 10.2.91.3:80 -> 10.125.16.27:63401

Message 00TKQUvhT3Co2Twvxqo5OQ

```
[1:2403302:80] ET CIARMY Collective Intelligence
Security Poor Reputation IP (TCP) [Classification: Misc
Attack] [Priority: 2] {TCP} 60.191.153.156:6000 ->
10.4.4.45:1433
```

In which terms was this message broken to?

From: mysensor3
Date: Tue Dec 04 00:30:32 +0000 2012
Severity: Alert
Facility: snort
destip: 10.4.4.45
srcip: 60.191.153.156
sig: 1:2403302
srcport: 6000
destport: 1433
proto: TCP

Full message:

```
<33>Dec 4 00:30:32 mysensor3 snort[4609]:
[1:2403302:80] ET CIARMY Collective Intelligence
Security Poor Reputation IP (TCP) [Classification: Misc
Attack] [Priority: 2] {TCP} 60.191.153.156:6000 ->
10.4.4.45:1433
```

Figure 2: Graphical user interface of Graylog2

In order to visualise collected log data, *Graylog2* provides a well-written and comprehensive web interface. For accessing the interface, different password-protected user accounts can be set up, with each user having either full or limited rights (the remainder of the discussion will concern the user interface with full admin rights). The interface is divided into several parts. The message view (see Figure 2) allows for getting an overview of stored log messages, presenting the messages in a browser with the most recent messages coming first. In the browser, the timestamp, host, severity, facility and message text fields are displayed. By clicking on an individual message, a detailed view of the message is provided which contains all fieldnames with values (see the right-hand part of Figure 2). Clicking on each individual value will carry out a search for messages with the same fieldname-value pair, and discovered messages will be displayed in the main message browser. Message search can also be initiated through a separate 'Quickfilter' button in the message view which allows for specifying more than one search condition. For searching a message text field, the Apache Lucene query syntax can be used (Lucene 2012). It supports searches for individual strings, approximate string matching based on Levenshtein distance, proximity searches (e.g. find two strings which have up to 10 words in between), and combining individual search conditions with Boolean operators. Figure 2 depicts an example search for Snort IDS alarms.

Apart from viewing all messages, the user can configure streams that are collections of messages satisfying some message filtering condition. Streams are updated with matching incoming messages in real-time. Messages under each stream can be viewed separately, and also it is possible to configure thresholding and alarming conditions for each stream (e.g. send an alarm to a security administrator if more than 10 messages have appeared under the stream during 1 minute). In order to define a filtering condition for a stream, message field values can be compared with fixed values, but in the case of some fields, also with regular expressions. In addition to streams, *Graylog2* also contains a so called analytics shell which supports a flexible query language for finding individual messages and for creating various reports. Unfortunately, currently all reports are text-based, although in the future releases support for graphical reporting might be added.

During our experiments, we attempted to establish the performance of *Graylog2* in terms of event throughput. We expected the performance of *Graylog2* to be significantly slower than for syslog servers tested in the previous section. First, both *Graylog2* server and *elasticsearch* database engine are written in Java, while *rsyslog*, *syslog-ng* and *nxlog* are coded in C. Second, *Graylog2* server has to insert log messages into *elasticsearch* index, which requires considerably more CPU time than writing messages into flat files. During the tests, we ran *Graylog2* on a Fedora Linux node with 8GB of memory and an Intel Core i5 650 processor. We set up one client for *Graylog2*, which was issuing large amounts of BSD syslog messages over TCP. The combined performance of *Graylog2* server and *elasticsearch* backend was measured in terms of message transmission throughput observed at the client side. During several tests, we were able to reach a throughput of 3,500 messages per second. This illustrates that one *Graylog2* server instance is not scalable to very large environments with many thousands of logging hosts and heavy message loads. Fortunately, the developers are planning to add support into *Graylog2* for multiple server instances, which should substantially increase its overall scalability.

4.3 Kibana

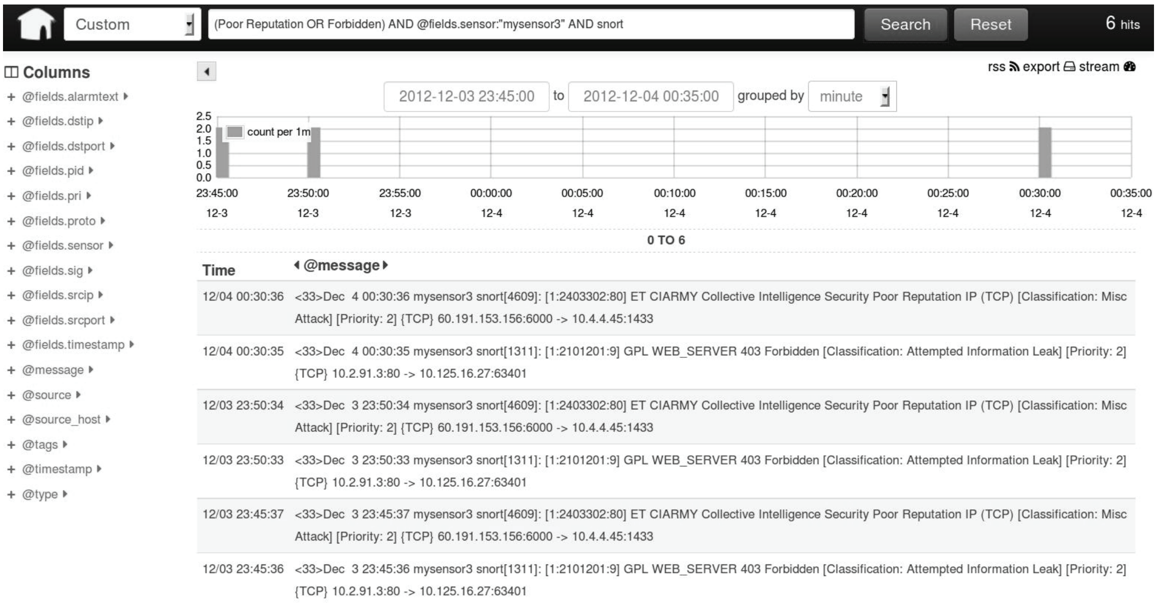
Kibana is another application for visualising collected log data (Kibana 2013). Unlike *Graylog2*, *Kibana* consists only of a Ruby-based web interface which uses *elasticsearch* as a backend, and there is no server to receive log messages over the network and store them into a database. For this reason, *Kibana* cannot run as a standalone system, but has to be used with an application that receives, parses, and stores log messages into *elasticsearch*. Many sites are employing *logstash* for this task, and *Kibana*'s default configuration is *logstash*-compliant. Also, *Kibana* expects that log messages in *elasticsearch* database have some fields that are created by *logstash* (for example, *@timestamp* and *@message*). However, if another application is configured to insert log messages into *elasticsearch* database with these fieldname-value pairs, *Kibana* is able to work on stored log data.

In order to search log messages, *Kibana* supports full Apache Lucene query syntax for all message fields (Figure 3 depicts an example search for Snort IDS alarm data). One advantage of *Kibana* over *Graylog2* is the support for the creation of various graphical reports. Reports can be created based on a selected message field and time frame, either for all messages or for some message matching criteria. *Kibana* supports the creation of pie charts which reflect the distribution of field values, trend analysis reports and count reports for field values. By selecting some value from the report form, the user can go to relevant log messages. Reports can also be created directly from searches – for example, the query *@fields.srcip=10.1.1.1* selects all messages where the field @fields.srcip (source IP address) has the value 10.1.1.1, while the query

@fields.srcip=10.1.1.1 | terms @fields.dstip

creates a pie graph about the distribution of @fields.dstip (destination IP address) values for source IP 10.1.1.1.

Since *rsyslog* has had *elasticsearch* support since 2012, it can be used instead of *logstash* for receiving and preparing log data for *Kibana*. In order to assess the performance of *logstash* and *rsyslog*, we installed *Kibana* with *elasticsearch* on a Fedora Linux node with 8GB of memory and an Intel Core i5 650 processor, and set up both *rsyslog* and *logstash* at this node. Both solutions were configured to insert messages into *elasticsearch* in bulk mode (for *rsyslog*, the message batch size was 16, while for *logstash* a batch size of 100 was used). For performance evaluation, we sent 100,000 BSD syslog messages over TCP to the receiver, and measured the processing time of these messages. At the end of each test, a query was made to *elasticsearch* for verifying that all messages were successfully inserted into the database. We repeated this test 100 times for *rsyslog* and *logstash*, deleting all inserted messages between consecutive test runs. The results of our experiments are presented in Table 2. For *rsyslog*, 100,000 messages were processed in an average of 17.066 seconds, yielding the average processing speed of 5859.6 messages per second. In the case of *logstash*, the average processing speed was 1732.952 messages per second. In other words, *rsyslog* is able to insert messages into *elasticsearch* more than 3 times faster.



```
Time            @message
12/04 00:30:36  <33>Dec  4 00:30:36 mysensor3 snort[4609]: [1:2403302:80] ET CIARMY Collective Intelligence Security Poor Reputation IP (TCP) [Classification: Misc
                Attack] [Priority: 2] {TCP} 60.191.153.156:6000 -> 10.4.4.45:1433
12/04 00:30:35  <33>Dec  4 00:30:35 mysensor3 snort[1311]: [1:2101201:9] GPL WEB_SERVER 403 Forbidden [Classification: Attempted Information Leak] [Priority: 2]
                {TCP} 10.2.91.3:80 -> 10.125.16.27:63401
12/03 23:50:34  <33>Dec  3 23:50:34 mysensor3 snort[4609]: [1:2403302:80] ET CIARMY Collective Intelligence Security Poor Reputation IP (TCP) [Classification: Misc
                Attack] [Priority: 2] {TCP} 60.191.153.156:6000 -> 10.4.4.45:1433
12/03 23:50:33  <33>Dec  3 23:50:33 mysensor3 snort[1311]: [1:2101201:9] GPL WEB_SERVER 403 Forbidden [Classification: Attempted Information Leak] [Priority: 2]
                {TCP} 10.2.91.3:80 -> 10.125.16.27:63401
12/03 23:45:37  <33>Dec  3 23:45:37 mysensor3 snort[4609]: [1:2403302:80] ET CIARMY Collective Intelligence Security Poor Reputation IP (TCP) [Classification: Misc
                Attack] [Priority: 2] {TCP} 60.191.153.156:6000 -> 10.4.4.45:1433
12/03 23:45:36  <33>Dec  3 23:45:36 mysensor3 snort[1311]: [1:2101201:9] GPL WEB_SERVER 403 Forbidden [Classification: Attempted Information Leak] [Priority: 2]
                {TCP} 10.2.91.3:80 -> 10.125.16.27:63401
```

Figure 3: Graphical user interface of Kibana.

Table 2: Comparative performance of rsyslog and logstash for elasticsearch bulk insert operations.

	Minimum processing time (seconds)	Maximum processing time (seconds)	Average processing time (seconds)	Average event processing speed (events per second)
rsyslog	15.998	21.031	17.066	5859.604
logstash	56.084	73.695	57.705	1732.952

5. Summary

In this paper, we have reviewed a number of widely used and efficient open-source solutions for collecting log data from IT systems. We have also described some novel technologies and setups for tackling the logging in large networks. One of the major contributions of this paper is the performance evaluation of described solutions through a series of benchmarks which mimic heavy workload in real-life environments. Through the experiments conducted, we have been able to identify specific advantages and weaknesses of each tool (e.g. efficient multi-threading or event filtering). Although our tests indicate that some tools have superior performance under specific circumstances, each tool offers a unique set of features to the end user. Therefore, the selection of a log management tool depends heavily on the specific nature of the environment. In order to automate the experiments for evaluating log management tools, we have written a simple toolkit consisting of a few Perl and UNIX shell scripts. For future work, we plan to elaborate our testing tools and release them to the public domain.

References

CEE (2012) *Common Event Expression, version 1.0beta1*, http://cee.mitre.org/language/1.0-beta1/

Elasticsearch (2013) http://www.elasticsearch.org

GELF (2013) http://www.graylog2.org/about/gelf/

Gerhards, R. (2009) *The Syslog Protocol*, RFC5424, http://www.ietf.org/rfc/rfc5424.txt

Gerhards, R. (2010) "Rsyslog: going up from 40K messages per second to 250K", Linux Kongress 2010, http://www.gerhards.net/download/LinuxKongress2010rsyslog.pdf

Gerhards, R. (2012) *BSD-Style blocks will go away in rsyslog v7*, http://blog.gerhards.net/2012/09/bsd-style-blocks-will-go-away-in.html

Graylog2 (2013) http://graylog2.org

Graylog2-Drools (2013) https://github.com/Graylog2/graylog2-server/wiki/Message-processing-rewriting

GrokPatterns (2013) https://github.com/logstash/logstash/blob/master/patterns/grok-patterns

Kibana (2013) http://kibana.org

Klyne, G. and Newman, C. (2002) *Date and Time on the Internet: Timestamps*, RFC3339, http://www.ietf.org/rfc/rfc3339.txt

Logstash (2013) http://logstash.net

Lonvick, C. (2001) *The BSD syslog Protocol*, RFC3164, http://www.ietf.org/rfc/rfc3164.txt

Lucene (2012) *Apache Lucene – Query Parser Syntax*, https://lucene.apache.org/core/old_versioned_docs/versions/3_5_0/queryparsersyntax.html

Miao F., Ya, M. and Salowey J. (2009) *Transport Layer Security (TLS) Transport Mapping for Syslog*, RFC5425, http://www.ietf.org/rfc/rfc5425.txt

Nxlog (2013) http://http://nxlog-ce.sourceforge.net/

Okmianski, A. (2009) *Transmission of Syslog Messages over UDP*, RFC5426, http://www.ietf.org/rfc/rfc5426.txt

Rsyslog (2013) http://www.rsyslog.com

Rsyslog-ver7 (2012) *Main Advantages of rsyslog v7 vs v5*, http://www.rsyslog.com/main-advantages-of-rsyslog-v7-vs-v5/

Syslog-ng (2013) http://www.balabit.com/network-security/syslog-ng/

Determining Trust Factors of Social Networking Sites

Namosha Veerasamy and William Aubrey Labuschagne
Council for Scientific and Industrial Research, Pretoria, South Africa
nveerasamy@csir.co.za
walabuschagne@csir.co.za

Abstract: Social networking sites have become part of everyday use with many users posting updates of their status as well as establishing contacts. While many users use social networking sites to connect and maintain contact, attackers may see social networks as a prime target for spreading malware, propaganda or marketing. Many activities can thus be carried out using these platforms- both malicious and beneficial in nature. Social networking sites can be used for social engineering attacks as users may be eager to interact and engage with a new contact made on a social networking site. However, they may not be aware that the profile they are engaging with may be valid. In addition, malware is also being developed to target users of social networking sites. This paper entails the investigation of the reasons that users trust these sites. Users develop trust based on certain factors. Survey data is presented to indicate some of these trust factors. Users will also maintain a certain level of privacy based on the trust is established. Therefore, identifying why users trust social networking sites can be beneficial in understanding why certain information is divulged. Therefore, important questions that need to be answered are why do people trust these platforms and how much do users trust these platforms. These factors are important for security awareness to protect users from being attacked with social engineering techniques. Social engineering makes use of trust as a component to influence users to perform actions detrimental to themselves and others. In addition, this paper users survey data to determine whether users are aware of these potential malicious objectives. Thereafter, the paper looks at the various indicators that could signal to users that they are not communicating with a genuine user but instead a fake profile. Another goal of this paper is to show users the dangers of social networking malware before they infect themselves. Once insight is gained into the trust factors, the study can also show users how social networking sites can be manipulated and thus help users protect themselves against being the target of attacks.

Keywords: trust factors, social networking sites, malware, profile, infection

1. Introduction

Social networks consist of nodes of individuals or groups with similar values, visions, goals, interests or friendships. Online social networks has helped encouraged such interactions due to its ability to rapidly share information and support communications. Online social networks allow users to create profiles that can be shared with contacts. A person's social status is often enhanced by their social profile. People are drawn together based on similar or shared social statuses. Thus, a social networking profile serves as an opportunity to grow a user's connections and thus expand one's contact circle.

Social networks have grown tremendously in popularity due to a number of useful features. This includes:

- Profile page showing details and interests
- Ability to connect to like-minded individuals
- List of contacts
- Photo albums
- Messages
- Status updates
- Comments
- Make new contacts
- Contact lookup
- Integrated applications, gadgets, add-ons

Due to the popularity of these sites, users may be eager to accept friend requests and posted links. Invitations may be easily accepted without any consideration of the consequences.

It is, however becoming vitally important that users be educated on the dangers of malware dispersal and fake personas. Social networking sites and the dangers of implicit trust have become a very important issue. Users need to made aware of the signs that perhaps that are not dealing with a legitimate user and how to evaluate the various components of a social network account. Also when users seek information, it is often overlooked that people may turn to social networks of individuals to make trust decisions (Thomas, Enrico & Marian 2006). By assessing various components of a social networking site they will be better empowered to determine whether they are dealing with an authentic person or an automated account that may have malicious intent. The paper commences with a brief explanation of how trust in social networks is built. It then presents results from a survey polling users on their trust levels of social networks. It concludes with a model that summarises critical indicators that may arouse suspicion that users are dealing with an invalid profile.

2. Trust

Trust grows when one party believes the other party will perform in such a way as to produce a positive outcome (Rahman, Haque & Khan 2011). Trust can also established through a brand to deliver a product or service of a good quality. Trust helps build up an organisation's credibility, track record and thus helps develop a person's/organisation's reputation. Associated with trust is the aspect of privacy. Users also want to believe that social networks will protect their privacy and not divulge personal information. If users have a strong sense of trust they may recommend it to other users and will themselves continue using social networks.

A study was carried out by Nikolaos Volakis at the University of Edinburgh to investigate "Trust in Online Social Networks". A summary of his findings is given (Volakis 2011).

Trust looks the relationships that people have, such that they place their trust in someone to behave in an expected way based on their previous behaviour. In social networks, trust often is determined by:

- Number of existing friends
- Number of message sent
- Number of replies to the sent message
- Degree of influence on other users
- Position within a network of friends

Many people may form groups of interest to discuss a mutual topic. Thus, circles of trust may form in the group based on the topic of interest. A person who is not as knowledgeable on the topic may not be as influential as a hot avid fan or expert.

In an on-line environment, for trust to exist:

- The contacts should share a common background with regard to culture, topic or organisation
- The contacts should be certain of the other's identity.

These two points from Volakis can be argued in that a common culture, topic or organisation is not always possible. In other instances, people only find and form contact in a digital manner and thus there is no way to prove the other's identify. Thus, these points may not always be feasible. Furthermore, trust can be built based on reputation.

Critical to online trust is trustworthiness based on transparency and honesty. This paper entails the investigation of the reasons that users trust social networking sites. Users develop trust based on certain factors. Therefore, the next section looks at identifying signs why users should be suspicious of certain social networking accounts.

3. Identifying fake profiles

Social engineering uses trust as a component to influence users to perform actions detrimental to themselves and others. There are various ways to carry out social engineering and influence users. One of these ways is the way a social network account is portrayed. The number of friends, types of photos and friends are all ways that can help create the impression of popularity. An automated account may also use abstract images. While there is no categorical manner to determine the validity of a social networking account there are various components of a profile that a user could evaluate to help indicate to them they should be suspicious. This sections looks at possible signs that users may be dealing with an invalid social network profile. The initial discussion looks at evaluating profile pictures and pictures/albums on social network accounts.

Fake Facebook profiles often have (IdentifyfakeFacebook.com 2012):

- Have pornographic, attractive or related photos
- Profile Picture
- *Profiles with one profile pictures*
- *Profiles with no profile pictures*
- *Profiles with other profile pictures*

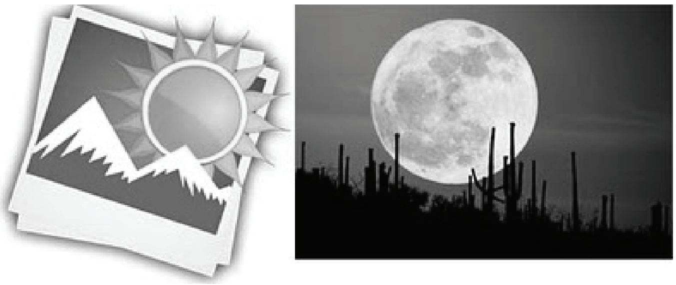


Figure 1: Images as profile pictures (IdentifyFakeFacebook.com 2012)

Figure 2: Provocative profile pictures (IdentifyFakeFacebook.com 2012)

A good indicator of a fake profile is thus the profile picture. Fake profiles may use pornographic, provocative or sexually orientated pictures as their profile pictures in order to attract people.

Another method of identifying a fake profile is the presence of a profile picture, which could be a photo of an attractive boy or girl, photos of actors or actresses.

A profile with no profile picture or random images could also be indicative of a false profile. However, some people may not use a profile picture or use the default pictures as they prefer not to display pictures of themselves. A better way to get a sense of a person is to look at their albums.

If a person uses pictures of professional models as their profile picture, it could be a sign of a fake profile. Other signs of fake profiles include (Hellbound bloggers 2012):

- Many male friends
- Many posts for requests for “Tag me!”. It is unusual if a person is tagged in various cartoon pictures rather than their own pictures
- Many applications requests like “Can you send me a chicken” or “Here a apple martini drink for you”
- In Facebook albums, fake profiles often have all their photos open to everyone. Other signs include:
- *Really tiny photos*
- *Many pictures but no tags indicating other profiles*
- Short profile- fake profile creators do not have time to create long and interesting profiles. However, a friend may have recently joined Facebook or is not very technology literate, and thus might have short profiles as they are still learning how to use the functionality.

- Many fake profiles have descriptions like “Accept my Farmville Request” or “Add me in Mafia Wars”. However, based on a conversation with a colleague about games/applications, one of them might send out an invite. In such cases, the request may be authentic.
- Conservative girls would not send requests to strangers. If a male receives random requests from girls, first verify the details (especially if the male is not particularly good-looking and the requests are from really attractive girls- this might be a way of tricking a person into accepting the request)
- The previous point also applies to females. Be wary of requests from very attractive males to average-looking females.
- Fake profiles may not have very frequent status updates. However, many users of Facebook are not very frequent users. They have created their accounts in order to get updates from friends and family around the world but rarely post anything about themselves. However, if their profiles are examined, there could be a legitimate profile picture, marital status, interests, a few family photos, etc. Their profile would appear legitimate without too many applications or tags.
- Many friends- On average fake profiles have 726 friends. If a person has an unusual profile picture (model, provocative or random image), does not update their status, only has friend’s add-ins as feeds, recently joined, has only cartoon pictures in their albums, has many tags on their photos, has many app requests and has hundreds of friends this is suspicious.

If a person joined recently and has many friends this could be suspicious. If a friend is in regular contact and they have communicated that they have just created a profile and send through an invite, this request could be trusted to be legitimate. However, random friend requests from people who joined recently, have models have profiles pictures and have many friends of one sex is slightly out of the ordinary.

Fake profiles can use lucrative pictures and catchy posts in order to attract more contacts. Individuals with a large and diverse network of contacts are thought to have more social capital than individuals with small, less diverse networks (Valenzuela, Park & Kee 2009). Spammers, marketers and attackers may then use this ploy to attract more users that can be used for their scams, product sales or attacks.

A few statistics based on fake and real profiles is given next in order to help identify fake profiles (Barracuda Labs 2012):

- Gender – 97% of Fake Facebook Profiles identify themselves as females
- 58% of females from fake profiles are interested in males and females
- 40% of real people claimed to have gone to a college versus 68% of fake profiles who make the same claim
- Fake profiles have an average of 136 tags for every four pictures whereas real people with profiles have on average 1 tag for every 4 photos
- 43% never update their status, compared to 15 % of real people that never update their status
- 35% of fake profiles lists entertainment interests

Users need to understand that spam can be sent out on social networking sites to lure them into phishing scams.

- Spear-phishing high ranking targets is called “whaling
- Spies set up a fake profile to lure a NATO official (Brean 2012)

Spam can also be used to market a specific company or product. It may be normal to like a clothing or food brand but be suspicious if a profile has constant support or posts about a particular company. Persuasion can consists of a number of influence processes, like those that increase awareness of a product or its features or change expectations about known features of the product (Aral 2011).

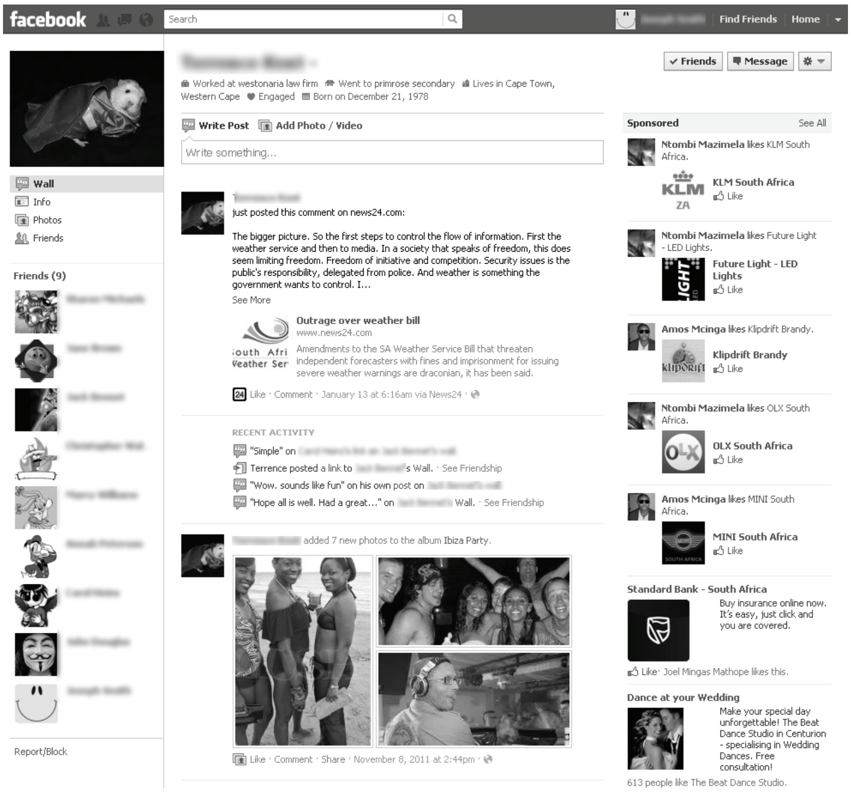
In essence, in order to determine where a contact is legitimate, study the profile. Sometimes a known contact is not very active on Facebook and therefore does not have regular status updates. They might have some photos showing their family and activities. A real profile is unlikely to have hundreds of tag requests to cartoon pictures. Users need to be careful of a profile with a provocative picture, is interested in males and females, has many friends, never updates status, excessively tags every photo, or has many app requests as there is a high probability it is fake

Others questions that may be asked to help identify fake profiles include (Facebook.com 2012):

- Are there pictures available of the person? Some people like to be private, but there is an option to show photos to friends.
- Is there constant blogs or posts about a company? Some people may like a particular brand of cold-drink or sportware but if a person constantly supports XYZ Incorporated, this could be a marketing ploy
- Has anyone ever met the person? If there are mutual friends, someone has had to have met the person. Mcknight et al explain the personal traits, structural assurance and normality of the Web, initial impressions and personal interactions play an important role for the formation of trust (Harrison McKnight, Cummings & Chervany 1998).
- Does the person make the same comments on multiple posts? This could be generic postings of spam messages of supportive comments in order to show interest in a profile. Is the only thing on their wall messages like “Looking good” on 10 different people’s accounts?
- Do they have a web site? The web site should be about them and not XYZ Incorporated.
- Do they make requests to support a prominent public figure but hide their identity? In such cases, the person just might be stirring trouble (especially if they remain hidden but prompt people to make controversial statements against well-known figures)
- Do they constantly take stabs at an opposition company?
- Some people may just pay compliments in order to get money. They will claim anything, including heart-breaking stories or even flirting. Evaluate carefully.
- Do they make meaningful comments or are they just giving a constant marketing or sales pitch?

People may choose to trust someone based on the past experience with the person or his friends, one’s opinion of actions taken by the person, psychological factors that are impacted by a lifetime of history and events, rumour, influence by others opinions and motives to profit amongst others (Golbeck 2005). However, users may choose to openly trust without being aware that they are dealing with an illegitimate profile.

In **Figure** an example of a fake profile is given. It has cartoon images as profile pictures as well as all their friends. There are no real albums and feeds are not meaningful.



[redacted]

Worked at westonaria law firm · Went to primrose secondary · Lives in Cape Town, Western Cape · Engaged · Born on December 21, 1978

[redacted] just posted this comment on news24.com:

The bigger picture. So the first steps to control the flow of information. First the weather service and then to media. In a society that speaks of freedom, this does seem limiting freedom. Freedom of initiative and competition. Security issues is the public's responsibility, delegated from police. And weather is something the government wants to control. I...

See More

Outrage over weather bill
www.news24.com
Amendments to the SA Weather Service Bill that threaten independent forecasters with fines and imprisonment for issuing severe weather warnings are draconian, it has been said.

Like · Comment · January 13 at 6:16am via News24

RECENT ACTIVITY

"Simple" on [redacted]
Terrence posted a link to [redacted]'s Wall. · See Friendship
"Wow. sounds like fun" on his own post on [redacted]
"Hope all is well. Had a great..." on [redacted] Wall. · See Friendship

[redacted] added 7 new photos to the album Ibiza Party.

Like · Comment · Share · November 8, 2011 at 2:44pm

Ntombi Mazimela likes KLM South Africa.
KLM South Africa Like

Ntombi Mazimela likes Future Light - LED Lights.
Future Light - LED Lights Like

Amos Mcinga likes Klipdrift Brandy.
Klipdrift Brandy Like

Ntombi Mazimela likes OLX South Africa.
OLX South Africa Like

Amos Mcinga likes MINI South Africa.
MINI South Africa Like

Standard Bank - South Africa
Buy insurance online now. It's easy, just click and you are covered.
Like · Joel Mingas Mathope likes this.

Dance at your Wedding
Make your special day unforgettable! The Beat Dance Studio in Centurion - specialising in Wedding Dances. Free consultation!
613 people like The Beat Dance Studio.

Figure 3: Fake profile

4. Survey of trust factors

A research collaboration was undertaken with the University of Michigan in order to gain a more global perspective of the users' attitudes and views on security in the Information and Communication Technology (ICT) field. The survey contained questions relating to security in social networks. Some of the critical findings are presented in this section. There were 286 usable questionnaires. The data was collected over a period of six months in 2012.

The survey instrument included demographic questions. Some questions used a 5-point Likert answer scale from strongly disagree to strongly agree. The surveys were posted on-line and users were requested to complete the questions voluntarily. The survey participants included the following nationalities: South African, European, Namibian, Lesotho, Zimbabwean, Kenyan, Nigerian, Slovakian, Ugandan, and American. The majority of the respondents (61%) were male.

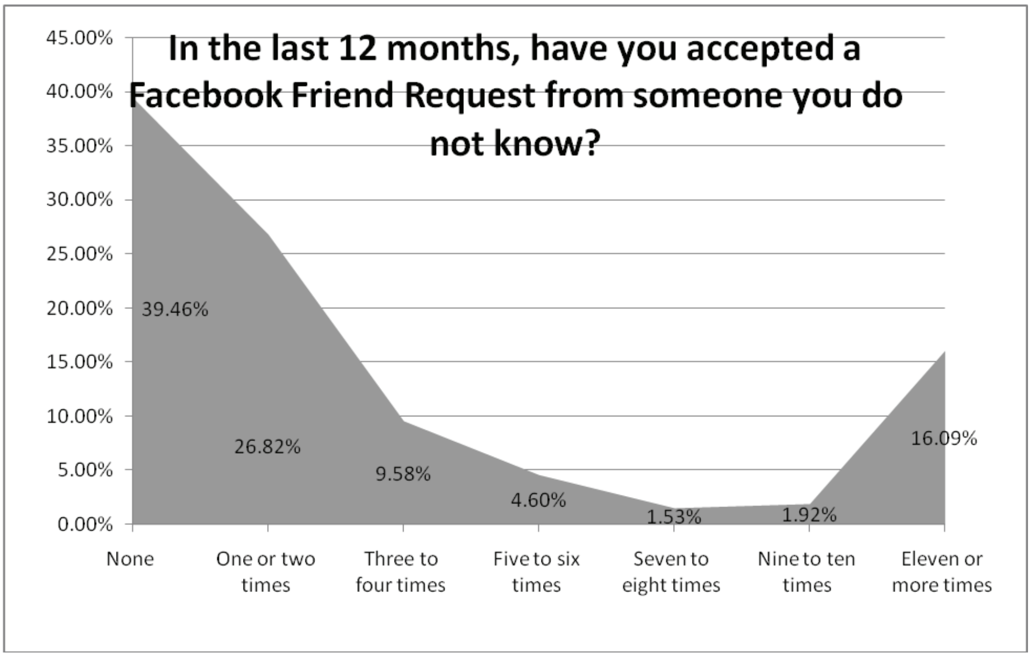


Figure 3: Acceptance of unknown friend requests

While 39.46% of users have never accepted a Friend request from an unknown contact, the rest of the responses are spread out across accepting such requests a few times to more than eleven times. This is alarming as users may not be aware of malware in social networks that send out random friend requests in order to make fake accounts appear more authentic. Koobface, a well-known Facebook malware has zombies that are required to find friends from existing accounts. Zombies would query the Command and Control server for logic credential to Facebook and would receive command like REG to register a new account or ADD, to log into an existing account (Thomas 2010). Fake accounts are then able to spam their list of contacts and spread more malware.

Participants of the survey were posed the question “Do you use privacy controls to protect your personal data on Facebook.” **Figure 4** shows the results. 73.18% of users utilise the privacy controls on Facebook whereas 12.64% of the respondents do not and 14.18% use them selectively. Even though the majority of the respondents did use the privacy settings, it is still an important message that users need to be made aware of the necessity of using privacy controls in order to prevent their sensitive information from viewed by any contact. With regards to fake profiles, if contacts are merely accepted, it will be possible for these contacts to gain a great deal of personal information. It is therefore imperative that users apply privacy controls in order to protect themselves.

Table 1 shows the responses to the question “What personal data on Facebook has the most value for you?” One is the least valuable and six is the most valuable.”

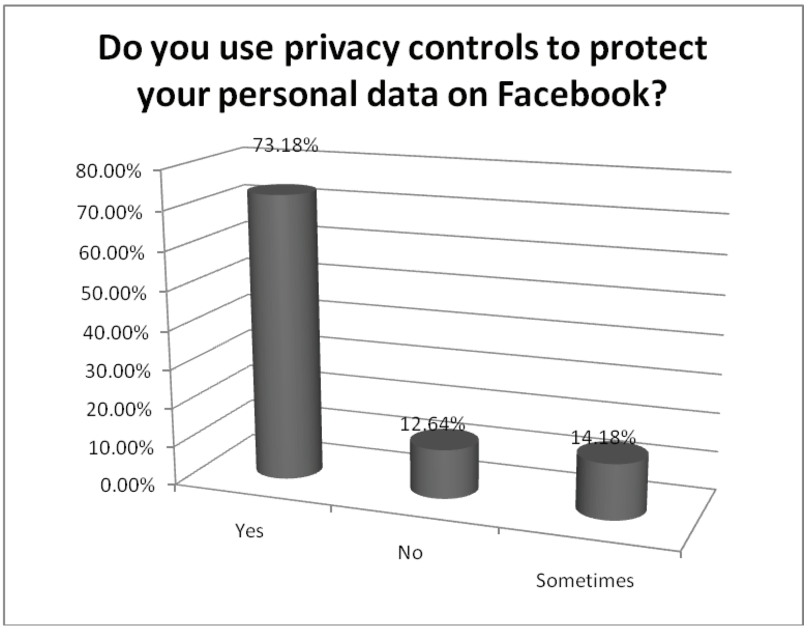


Figure 4: Use of privacy controls

Table 1: Response of question posed on value of personal data on Facebook

	Least valuable	Little Valuable	Not so Valuable	Valuable	Quite valuable	Most Valuable
List of Friends	21.84%	11.88%	18.39%	19.54%	12.26%	16.09%
Photos	6.51%	5.36%	11.11%	13.03%	21.07%	42.91%
Wall Posts	14.56%	9.20%	19.92%	16.09%	21.07%	19.16%
Contact Details	8.05%	7.66%	6.90%	8.43%	29.96%	59.00%
Interests and Activities	21.07%	19.54%	20.69%	14.94%	9.58%	14.18%
Messages	4.98%	3.07%	8.43%	11.49%	18.39%	53.64%

Repondants were asked to rate the value each of the components on a social networking account. The results show that feel that their contact details (eg. email, mobile number, etc) was the most valuable information to them (59% strongly agreed). 53.64% also rated messages as most valuable and 42.91% believed their photos had most value.

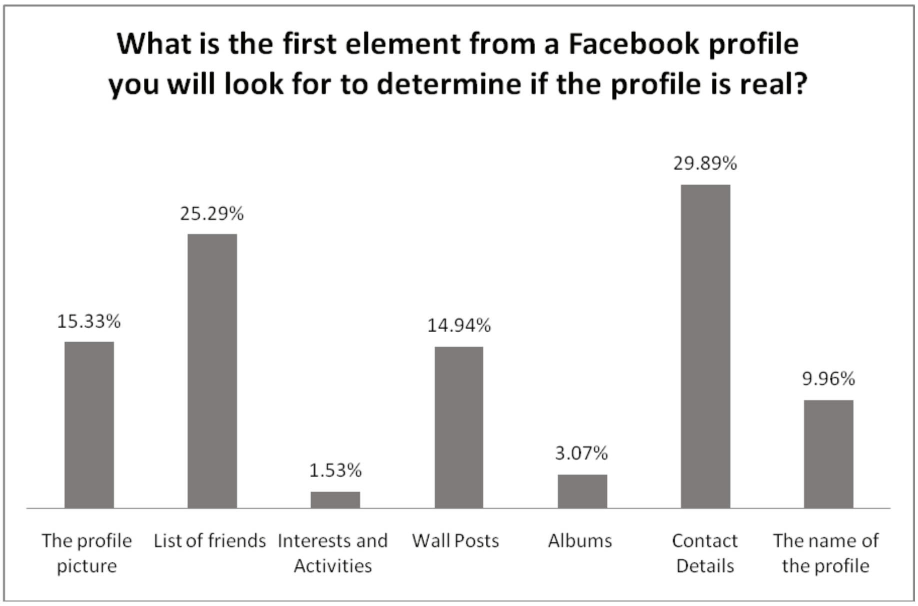


Figure 5: Determination of fake profiles

The response was widespread with regards to determining whether a profile is real (**Figure 5**). The respondents mainly determined whether a Facebook profile was real based on the contact details provided (29.89%) or their list of friends (25.29%). However, 15.333% also looked at the profile picture and 14.94% would initially judge from the wall posts.

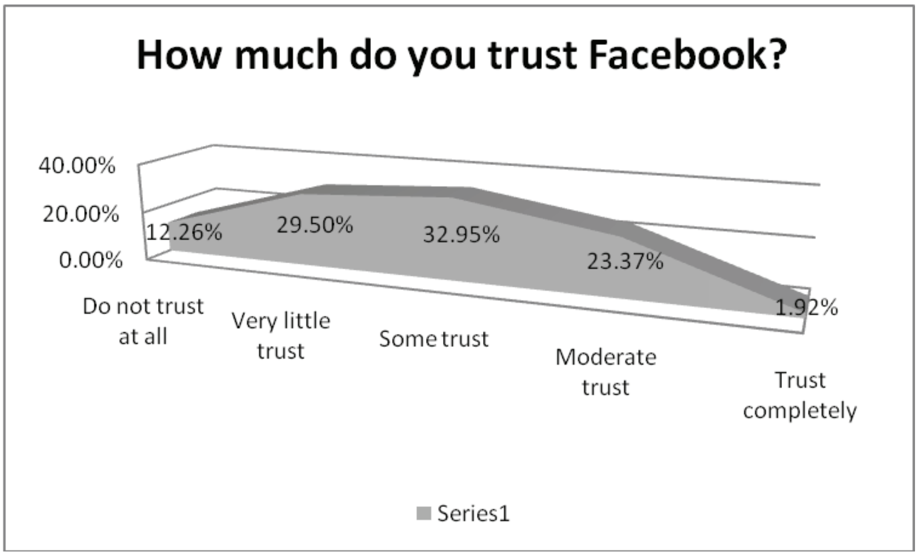


Figure 6: Trust in Facebook

The results concerning whether users trust Facebook was quite dispersed (**Figure 6**). The majority of respondents have some-to-moderate trust in Facebook with a very minor portion (1.92%) completely trusting Facebook.

5. Critical indicators

The previous sections describe some indicators that a social networking profile could not be genuine. Survey data was also presented concerning users' levels of trust concerning social networks. This section presents a summary model of the critical indicators that could alarm users of a suspicious social networking account.

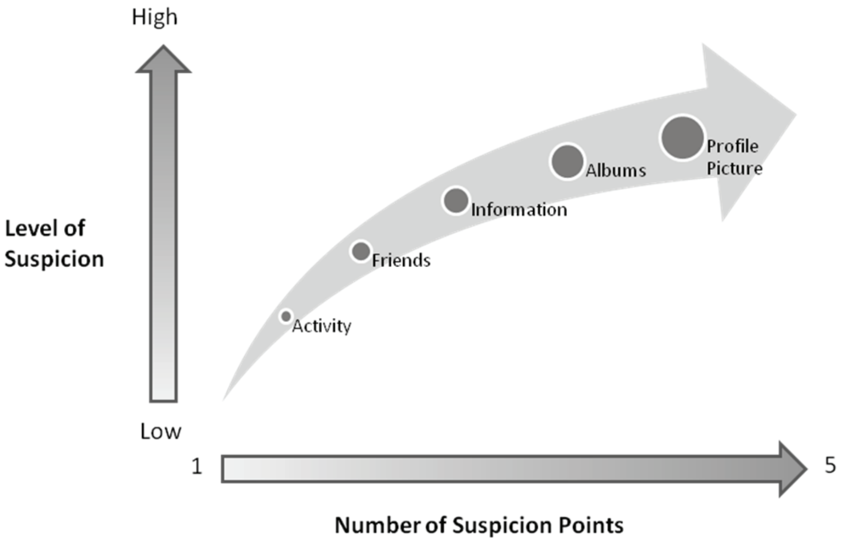


Figure 7: Model of critical components to identify a suspicious profile

While it may not always be possible to determine the validity of social networking profile, the following components of a profile can be evaluated to indicate to a user whether they should be suspicious and exercise care. These include but are not limited to the timeline activity, friend's list, profile information, albums and the profile picture. These will be discussed in order of importance. **Figure 7** shows a summary of the critical components identified that can help a user determine whether a profile is valid.

The first component to evaluate is the activity on the user's timeline. The term timeline is synonymous with the term wall, and will be used interchangeable. This can be used to post information bits, which was visible to the user, as well as the friends. This information bits include but are not limited to interesting links, describing what they are doing, posting of pictures, commenting on other posts in the form of comments and adding personal information to their profile. The timeline could provide valuable information in the attempt to identify a fake profile. A timeline should be inspected for valid activities which indicates active creation of content, for example if a timeline contains a post about the user's plan for the weekend and a few of his friends make human-like comments on the post, then the legitimacy of profile increases. The activity on the wall is has the highest priority since many users who prefer anonymity might not post images or provide personal identifiable information, but still have an active timeline with legitimate content which indicate human interaction.

The friend's list should be examined next. Studies have shown that Facebook users have an average number of 130 friends, whereas many fake profiles have in excess of 600 friends (Barracuda labs, 2012). Not only the number but also the composition of the friends should be analyzed. For example, a fake Facebook profile might have a large number of men if the profile under investigation portraits an attractive female using sensual pictures.

Thereafter, the personal information associated with the profile can be examined. This information discloses gender, interests, activities, current location, education, real name, relationship status or date of birth. These are usually associated with real human beings. Although the absence of this information might raise suspicion, it should be noted that some users might withhold this information to protect their identity. The information provided should be logically analyzed for consistencies, for example if the data of birth indicates that the user of the profile is 18 years of age, then its unlikely that the user could be a doctor, worked in the banking sector for 5 years and have three children.

The albums of the profile should be inspected next. A real profile should have some pictures about the person. The legitimacy of the profile should be questioned, if the pictures contain non-human objects

or have pictures containing only one person. For example only having pictures without any other humans (friends) should be a concern, as well as having unrelated pictures uploaded. If the user has a high number of friends but does not have any images of friend interaction, then caution should be taken when contemplation becoming friends with this user. However, many users might not upload pictures to protect their privacy.

The last component of profiles that should be analyzed is the profile picture. Many fake profiles use provocative images to lure people of the opposite sex to accept friendship requests. Granted attractive people do exist but all the other components of a profile should be taken into consideration. Other users also use non-human images, for example a picture of nature or their favourite cartoon. These users protect their identity through obfuscation. A user should analyze all these components for consistencies before trusting the validity of the profile. Furthermore, social networking sites are used as a platform to promote socialization between users and lack of these components should be suspicious while the disclosure of these promotes trust. However, cognisance should be taken into consideration by looking at the information logically before implicitly trusting an unknown profile.

6. Conclusion

The use of social networking sites has exploded with users making contacts and posting updates about their various daily activities. While many users use social networking sites to connect and maintain contact, attackers may see social networks as a prime target for spreading malware, propaganda or marketing. Many activities can thus be carried out using these platforms- both malicious and beneficial in nature. This paper describes various indicators to arouse suspicion that a social networking account make be fake. It provides survey data and a model to demonstrate to users their level of trust. This paper takes a very practical approach at educating users on how evaluate a profile to determine whether it is potentially fake.

References

Aral, S. 2011, "Commentary—Identifying Social Influence: A Comment on Opinion Leadership and Social Contagion in New Product Diffusion", Marketing Science, Vol. 30, no. 2, pp. 217-223.

Barracuda Labs, Social Networking Analysis, [online], http://www.barracudalabs.com/fbinfographic/, Accessed 20120615.

Brean J, (2012) [online], National Post, http://news.nationalpost.com/2012/03/11/chinese-cyber-spies-set-up-fake-facebook-profile-to-friend-top-nato-officials/.

Facebook, How to Spot a Fake Profile, [online], http://www.facebook.com/notes/my-social-practice/how-to-spot-a-fake-profile/199791076710071, Accessed 20120615.

Golbeck, J.A. (2005)" Computing And Applying Trust In Web-Based Social Networks", Dissertation Submitted to the Faculty of the Graduate School of the University of Maryland, College Park

Harrison McKnight, D., Cummings, L. & Chervany, N. (1998), "Initial Trust Formation in New Organizational Relationships", The Academy of Management Review, Vol. 23, no. 3, pp. 473-490.

Hellbound Bloggers, 10+ Tips To Identify Fake Profiles on Facebook, [online] http://hellboundbloggers.com/2010/05/17/identify-fake-facebook-profiles/, Accessed 20120615.

IdentifyFakeFacebook.blogspot.com, How to Identify Fake Profiles on Face book, [online], http://identifyfakeinfacebook.blogspot.com/, 20101020.

Rahman M, Haque M & Khan M, (2011) "The Influence of Privacy, Trust towards Online Social Networks: An Online Exploratory Study on Bangladeshi Customers Perception", European Journal of Economics, Finance and Administrative Sciences, ISSN 1450-2275, Issue 35.

Thomas, H., Enrico, M. & Marian, P. (2006), "Person to person trust factors in word of mouth recommendation", Conference on Human Factors in Computing Systems (CHI '06). Montreal, Quebec, Canada, 2006.

Valenzuela, S., Park, N. & Kee, K. (2009), "Is There Social Capital in a Social Network Site? Facebook Use and College Students' Life Satisfaction, Trust, and Participation", Journal of Computer-Mediated Communication, Vol. 14, no. 4, pp. 875-901.

Volakis Nikolaos, (2011) "Trust in Online Social Networks", University of Edinburgh, School of Informatics, Master of Science dissertation, [online], http://www.inf.ed.ac.uk/publications/thesis/online/IM110932.pdf.

The Influence of Joining the European Union on Human Trafficking

Ineke Weijer[1], Karin Ehnberg[2], Dijana Grd[3], Boris Kišić[3], Juan Muro Garrido-Lestache[4] and Marc-Johan Tsalatsouzy[5]
[1]University of Applied Sciences, Amsterdam (HvA), The Netherlands
[2]Sweden University - Archive and Information Science, Sweden
[3]University of Zagreb, Croatia
[4]Universidad de Alcala, Madrid, Spain
[5]ESIEA, Laval, France
ineke.weijer@gmail.com
karin.ehnberg@viasilva.se
dijana.grd@gmail.com
bkisic@gmail.com
juanmurogl@yahoo.es
tsalatsouzy@et.esiea-ouest.fr

Abstract: Recently the European Commission (EC) received data that might possibly contain crucial information on topics that are of primary concern to the European Union (EU), one of them being human trafficking. This report describes the research done on the data by an international group of students in IT, Law, Archiving and Media during a two-week Intensive Programme in Laval, France. We analysed the signals in the data to help the European Commission decide whether or not human trafficking should be of higher or highest concern for the EU. Another issue we addressed is if there were any signals on this topic that the US could use against the EU. Over the last couple of years new countries have joined the European Union and became part of the Schengen Treaty. As a result of this the boarder of the European Union (EU) has been moved further east. This has rapidly increased the problem with human trafficking, especially trafficking for prostitution in the EU and it also affected the United States, as a destination country. Roughly a year after Bulgaria and Romania joined the EU in 2007, the increase in number of people being trafficked became obvious. We found evidence that in Bulgaria corruption of the government and maffia are part of the problem. Now routes through those countries and also through Hungary are more attractive for traffickers. Since Hungary joined the EU it has become slightly easier for traffickers to move people in and out of Hungary e.g. using tourist visa. Croatia currently is a country with already a high trafficking number, and when Croatia enters EU, it could become an important transit country for human trafficking. Croatian government is already introducing standards to reduce human trafficking, but the problem is still present in the country. The EU should work with Croatia to implement right measures, to avoid problems as with Bulgaria and Romania. Our recommendation to the European Commission is that it has to think carefully what the implications of new countries, like Croatia, entering the EU would be, given the experience the EC had when Bulgaria and Romania entered the EU. It is in the interest of the US to help Europe fight against human trafficking, since the US is a country of destination for trafficking in people. Fighting trafficking in Europe thus resolves their problem too. Europe and the US should work very closely together on this topic to solve the problem and try to define uniting laws.

Keywords: human trafficking, prostitution, transit service, corruption, traffickers, smuggling

1. Introduction

This paper has been created as a part of the European IP E-Discovery 2013 in Laval, France. An international group of students from branches as IT, law, archiving and media gathered and analysed data that the Europeen Commission (EC) recently received. The data on an USB-image were presented to the EU by a non-profit organisation. The image contained crucial information on topics as Drug trafficking, Human trafficking, Money laundering, Monitoring (datacommunication), Intelligence (capabilities), United Kingdom, Turkey, Ukraine, China (technology) and Syria.

The commission is worried about the possibility that this confidential information will be disclosed and be harmful to individuals and governments.

Therefore the commission had three questions it wanted to be investigated by the groups of the IP:

- do the data contain signals that indicate that the topics should be of higher or highest concern for EU?
- is the EU dealing properly with those topics? and
- what signals can be found that the US could use against the EU?.

In this paper only one of the topics will be discussed: Human Trafficking or trafficking in person (TIP).

Human Trafficking is generally defined as "a global phenomenon that involves obtaining or maintaining the labor or service of another through the use of force, fraud, or coercion in violation of in individual's rights" (International Association of Chiefs of Police). We used this definition as the starting point of our investigation, which we performed following the well described procedures of the Electronic Discovery Reference Model (EDRM). A description of the model, and the way we implemented it in our research is described below.

2. The electronic discovery reference model

The EDRM Model defines stages and procedures for the discovery and recovery of digital data, in such a way that the findings can be used in a possible courtcase. As the EDRM model was established in the U.S. and because of the differences in litigation concerning the data (e.g. differences in privacy) between the U.S. and Europe on civil litigation we choose to follow the structured (Dutch) approach. A Schema of this approach can be found in figure 1, with the different stages divided into five main phases (marked I to V). Note how all stages are interconnected, meaning that during the investigation one can always go back to an earlier stage if the evidence found leads to new insights on the incident. However, going back to an earlier stage still needs to be done in such a way that it does not hamper the forensic soundness of the data, as we will explain in the preservation stage.

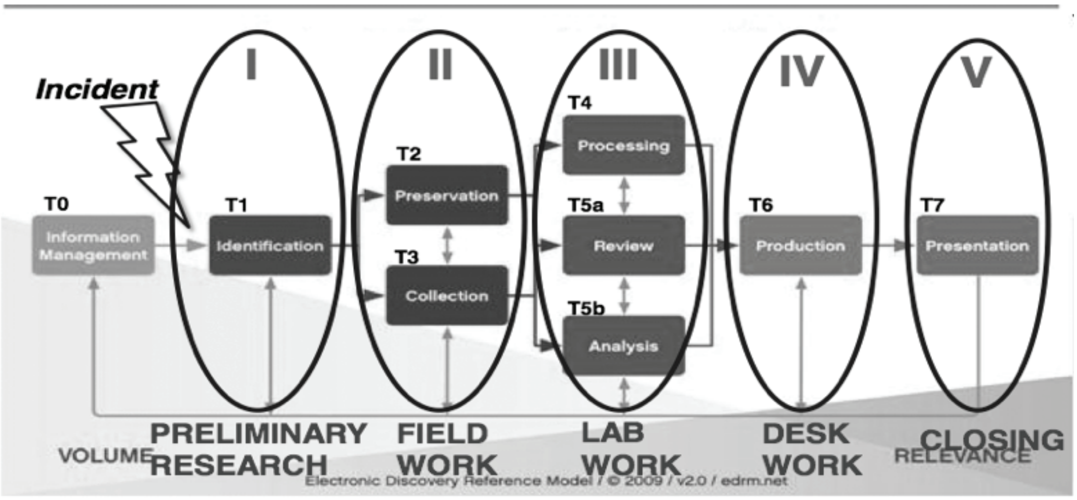


Figure 1: The EDRM model and the 5 stages (EDRM, 2009)

The model consists of the following stages:

Information Management: Description of how to manage all sources of Electronically Stored Information (ESI) in the information lifecycle. It starts with creation, goes from usage to archiving or deletion. Proper management of information should lead to easier future eDiscovery processes. Technically, information management is not part of the investigation. It helps in establishing background information and in the process of Identification.

Identification: After an incident has happened, triggering an investigation, location and definition of potential sources of ESI is necessary. To do so, background information should be checked, using open sources. For our background research we used Google and DuckDuckGo search engines. The essence is to gather as much information as possible on the situation leading to the incident. In other words it is essential to master the subject before answering the problem. In our case we used the publicly available WikiLeaks files, specifically that part of the files commonly known as "Cablegate". We searched for information on what human trafficking entails, which countries are involved, what means of transport are used and possible information on people or organisations involved. We started with the keywords "human trafficking" or "trafficking in persons" + Europe . In the data, one word repeatedly turned up in connection to these keywords: prostitution. There are many possible causes for human trafficking but prostitution seems to be the main reason. Refining our research with this new keyword, we then searched for names of people, organisations and countries in Europe connected to the problem of human trafficking for prostitution.

This idenfication stage leads to the next stage.

Preservation: What ESI should be stored and preserved. Storage and preservation should happen in a forensically sound way. Forensically sound means that the data that are retrieved are not tempered with in any way. This stage was handled by the IT experts of our group, as was the stage of Collection.

The data delivered for this investigation were delivered on a USB device. When obtaining our copy of these data, we first made sure that we had an exact copy of the original material, by comparing the hash-values of the images by using 3 hash algorithms: MD5, SHA1 and SHA256 sum.

The ESI consisted of cable files in CSV format and other spyfiles (documents and e-mails) all contained in a single RAR archive.

Scanning the documents and the spyfiles led to our understanding that for our investigation the important information would be found in the cable files (the *Processing* stage).

We developed a tool to facilitate investigating the data, by using PostgreSQL and C# so that not only our IT people could work with the data, but also the Archive and Media experts could join in these stages of the investigation for the *Review* stage. In this stage, the ESI were evaluated for relevance & privilige. The latter meaning this was also the moment to decide what information out of the data found should and could actually be presented to the EU and in what way. One should be careful not to present information that can either endanger people (names etc) or that cannot be used in court because of other privacy rules. During the *Analysis* stage the final evaluation of the relevant information takes place. The gathered information is reviewed and conclusions are drawn. These are written in the report, the result of the *Production* stage. For a detailed description of our research and conclusions we refer to the report written during the IP in Laval. During the Analysis and Review Stage we were not only looking for signals, but we also tried to corroborate the found evidence with outside sources, to define the actual "truth" of the found evidence. The information was initially judged as "Totally Fact, Partly Fact, Partly or Totally Fiction" and only if we did find outside information that confirmed our findings, we judged the information to be "totally fact". In the next chapter we summarize our main findings.

3. Main findings

The keywords we initially used for searching were ("Human Trafficking" or "trafficking in persons" + Europe). We found 2659 cables in which the combination of these keywords were used After this first step, we noticed that one topic was showing up in relation to the human trafficking all the time, the topic of prostitution. We then looked for signals in the Wikileaks Cablegate cables reflecting the assumption that prostitution is one of the biggest problems, in this way refining our research. We then checked for countries that were most mentioned. These were the Balkan countries, Bulgaria, Romania, Hungary, Croatia, Slovenia. All of this information is also shown on a map we found on the internet, thus an open source corroborated our findings, showing the countries of origin and the countries of destination, see Figure 2.

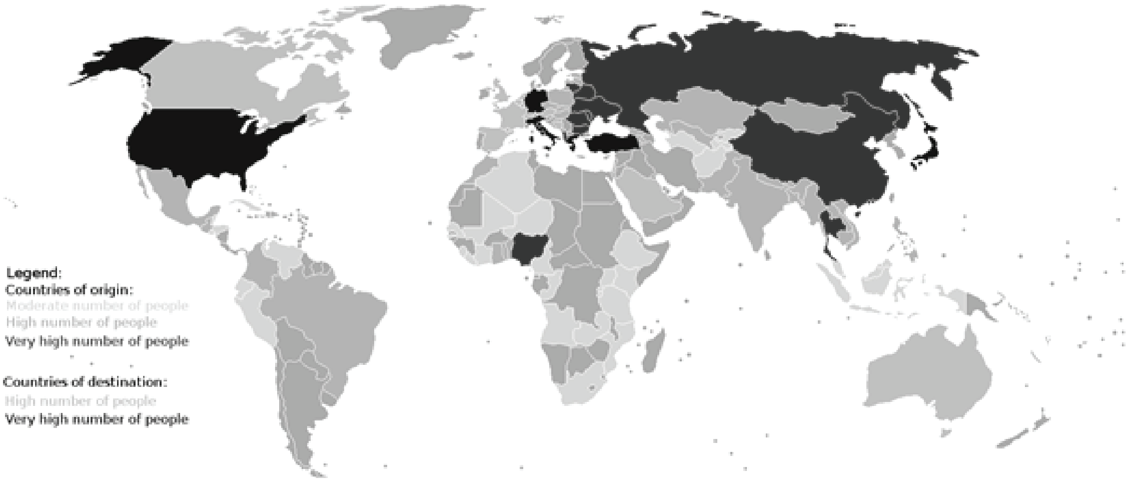

Legend:
Countries of origin:
Moderate number of people
High number of people
Very high number of people
Countries of destination:
High number of people
Very high number of people

Figure 2: Countries involved in human trafficking (AAUW, 2012)

We then searched for organizations and individuals that are involved in this trafficking. The role of traffickers is to invest, to recruit, and to transport people. Figure 3 is the recap of this part of research, showing how we tried to arrive to the answers to the main questions about human trafficking, the starting points of the investigation.

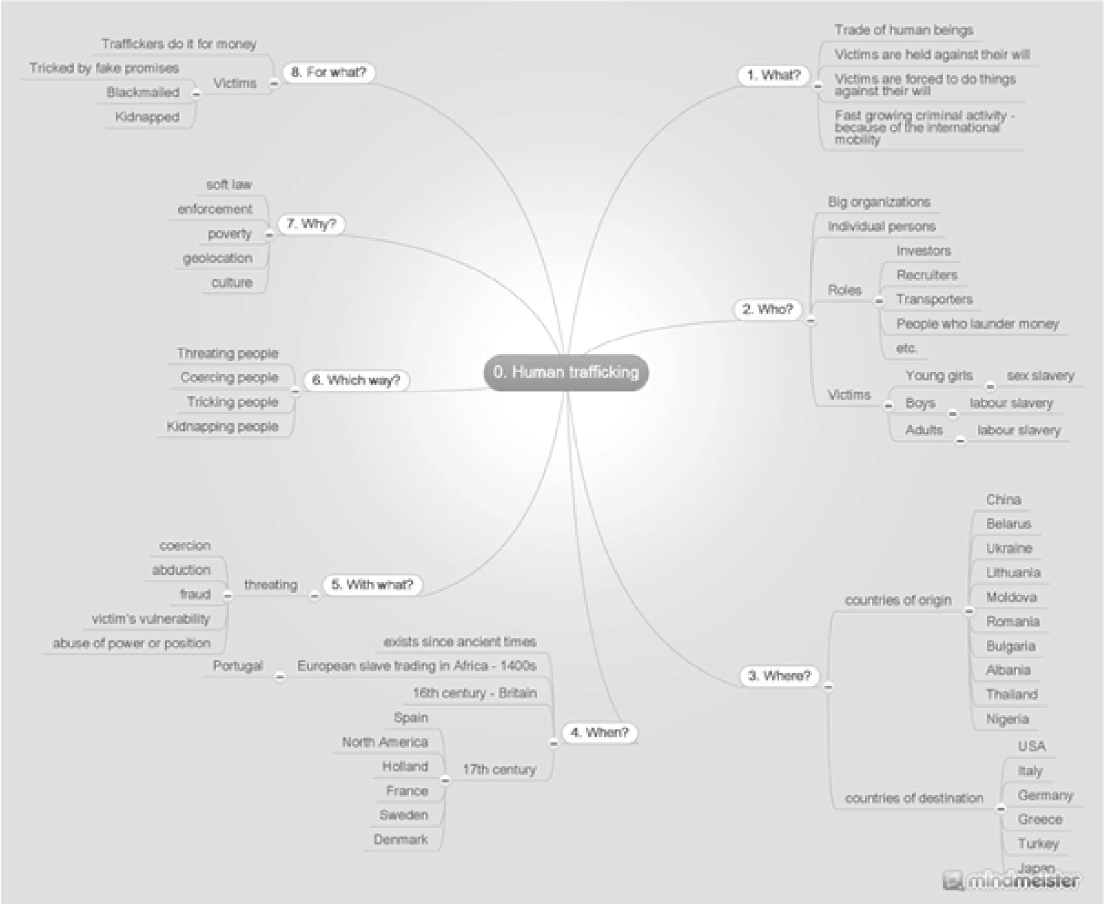


Figure 3: Mind map used in our research on human trafficking

3.1 Importance of human trafficking to the EU

In the Wikileaks cables we searched for signals identifying the ways in which European countries are dealing with human trafficking. Are governments (together with EU) trying to reduce human trafficking and if so what (legal) action has been taken? We found a lot of confidential documents, mostly government reports, about investigations, raising people's awareness and other actions.

Many of these documents concerned Bulgaria, where human trafficking is widely spread. There was an EU intervention at the Embassy in Bulgaria to close duty-free shops and gas stations on external borders, because it's a source of crime, such as human trafficking and forced prostitution. The Bulgarian government arrested Yordan Tonov along with 13 others on charges of organized crime, drug trafficking, human trafficking, and extortion. One document mentioned how the European Commission and Bulgarian public are skeptical about reducing crime because criminals are not punished. There are also some documents showing how the corruption in Bulgaria is one of the gravest problems. The Bulgarians are trying to break this corruption, but the big effort needs to be made yet. Still, the Bulgarians are fighting against the mafia involved in human trafficking and other crimes, despite the connections between the mafia and the government. Related to these actions, France (together with EU partners) is strongly fighting against trafficking in people, and the French are emphasizing the problem with Bulgarian citizens. France is not a country of origin, but it is a destination country (mainly for women trafficked for prostitution). Other countries of destination mentioned are Luxembourg and Norway.

There are also some other countries mentioned in relation to the problem of human trafficking. One of them is Hungary and we found documents that mention how Hungary has made it slightly easier for human traffickers to traffic people in and out of other countries, making Hungary a transit country.

The information found in the image showed that a lot of European countries started seriously fighting human trafficking in 2008 (for example France ratified the EU Convention on Action Against Trafficking in Persons – information found in cables). This is a year after Bulgaria and Romania joined EU. These are countries of origin of people in trafficking. Looking at the problems these countries caused to the EU when they became members and thus part of the Schengen Treaty, we searched for the documents which will show us how is Croatia (which is entering EU this year) dealing with this problem. We saw that The Government of Croatia has demonstrated strong political will to combat Trafficking in Persons and has continued to strengthen its legal framework to both protect victims and criminalize traffickers. Croatia has full compliance with the minimum standards for the elimination of trafficking in persons, but the problem is that Croatia is still considered a country of origin, transit and destination for internationally trafficked, men, women and children. Trafficking does occur within the country's borders. Croatia is no longer an attractive transit route to the EU for Romanians, Hungarians, Ukrainians, Moldovans and other Eastern Europeans due to the accession of Bulgaria and Romania into the EU in 2007. As we said before, because routes through Romania, Bulgaria and Hungary and on to the western part of the EU are now more attractive and easier for traffickers. That could change when Croatia joins EU. Croatia could become a transit country for human trafficking. Croatian government is already introducing standards to reduce human trafficking, but this problem is still present in the country. EU needs to help Croatia to implement right measures, so that there won't be so many problems as with Bulgaria and Romania.

3.2 Cooperation between EU and US

We found evidence that EU cooperates with the US in reducing human trafficking.. One of the documents describes the meeting between the US Ambassador and the Romanian Orthodox Patriarch about the problem with human trafficking and prostitution. There is a problem that Romania is now a transit country, but in the future it can be destination country. This should be a signal to the EC too. More evidence was found on cooperation between EU and US. Council UN expert Paulo Oliviera suggests that US and EU can work together on resolving the problem related to human trafficking given their broad areas of agreement. US, however, has had problems with the Dutch on trafficking. This was based on differences with regard to prostitution laws rather than the involvement of minors. The Dutch delegation has been firm in their affirmation of the need to implement UN protocols, including those relating to the age of consent.

4. Legal aspects

Currently, Human Trafficking is high on the list of important topics on the EC's Agenda. The EU has adopted a five year strategy (2012-2016) and formulated a strategy with five key priorities. It is expected to be fully transposed by April 6, 2013 (EUR-Lex, 2012). These key priorities are:

- Identifying, protecting and assisting victims of trafficking;
- Stepping up the prevention of trafficking in human beings;
- Increased prosecution of traffickers;
- Enhanced coordination and cooperation among key actors and policy coherence;
- Increased knowledge of and effective response to emerging concerns related to all forms of trafficking in human beings

With this in mind, during the IP we found the following EU directives already formulated on Human Trafficking:

2011/36/EU (European Commision, 15.4.2011)

> *"Prevention and combating trafficking in human beings and protecting its victims in the official Journal of the European Union."*

2011/92/EU (European Commision, 17.12.2011)

> *"Directive rules concerning the definition of criminal offenses and sanctions of sexual abuse and exploitation of children, child pornography and solicitation of children for sexual purposes."*

2004/80/EC (European Commision, 2004)

> *"Ensures that the Member States of EU have a national scheme which guarantees compensation to victim of crime, including victims of human trafficking"*

In International Law, however, we found the *Council decision of the United Nations, dated December, 8, 2000* (United Nations, 2000)**.** This Protocol is about human trafficking and is a supplement to the United Nations Convention against Transnational Organized Crime and adopted by the General Assembly of the United Nations on 15 November 2000. The Council of the European Union agreed with this protocol, as did the United States. This protocol could lead to further joint legistation on the matter of Human Trafficking.

5. Conclusion

The signals point towards a strong position by the EU on anti-Human Trafficking. But the data found in the cables indicate however that there are certainly still problems with Human Trafficking in several countries within the EU that should be brought to the attention of the EC. Western countries like Luxembourg, France, Norway, The Netherlands are working with eastern countries like Bulgaria, Romania, and Hungary to fight TIP. The US also gives strong signals to the eastern European countries to fight human trafficking. However the problem is that in some of these countries the maffia is interconnected with the government, which makes it hard to fight against TIP. Also, the Schengen Treaty, makes it more difficult to actually trace people once they are inside the EU. The EU should act in advance to prevent the human trafficking from increasing when new countries like Croatia join. It should also act in order to prevent traffickers to find new routes. The US are trying to help Europe to fight human trafficking because it is also in their interest to resolve this problem. The EU and the US should cooperate to find solutions together and try to define uniting laws.

While governments and the international community have responded to the frequency increasing traffic, much remains to be done to combat this problem.

We can see that many European countries working together, but it still lacks collaboration between U.S. and EU to really stop human trafficking. It is probably foolish to think that the phenomenon will be resolved without this collaboration which is essential.

References

AAUW (2012) „Do You Know How Many Slaves Work for You?", [online], http://www.aauw.org/2012/03/26/how-many-slaves-work-for-you/

Electronic Discovery Reference Model (EDRM), [online], http://www.edrm.net/

European Commision (2004) 2004/80/EC, „Council Directive of 29 April 2004 relating to compensation to crime victims", [online], http://ec.europa.eu/antitrafficking/entity.action?path=Legislation+and+Case+Law%2FEU+Legislation%2FCriminal+Law%2FCouncil+Directive+of+29+April+2004+relating+to+compensation+to+crime+victims

European Commission (15.4.2011), „Directive 2011/36/EU of the European Parliament and of the Council", Official Journal of the European Union, [online], http://ec.europa.eu/antitrafficking/download.action?nodePath=%2FLegislation+and+Case+Law%2FEU+Legislation%2FDirective+THB+L+101+15+april+2011.pdf&fileName=Directive+THB+L+101+15+april+2011.pdf&fileType=pdf

European Commission (17.12.2011), „Directive 2011/92/EU of the European Parliament and of the Council", Official Journal of the European Union, [online], http://ec.europa.eu/antitrafficking/download.action?nodePath=%2FLegislation+and+Case+Law%2FEU+Legislation%2FCriminal+Law%2FDirective+2011_92.pdf&fileName=Directive+2011_92.pdf

EUR-Lex (2012) „Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions the EU Strategy towards the Eradication of Trafficking in Human Beings 2012–2016", [online], http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52012DC0286:EN:NOT

International Association of Chiefs of Police, „The Crime of Human Trafficking: A Law Enforcement Guide to Identification and Investigation", [online], http://www.vaw.umn.edu/documents/completehtguide/completehtguide.pdf

United Nations (2000), „Protocol to prevent, suppress and punish Trafficking in Persons, especially Women and Children, supplementing the United Nations Convention against Transnationalorganized Crime", [online], http://treaties.un.org/doc/source/RecentTexts/18-12-a.E.htm

Wikileaks, Cablegates and Spy Files. [online], http://wikileaks.org/, http://www.cablegatesearch.net/

Covert Channel Based Eavesdropping Malware Analysis and Detection for Android Systems

Lu Yu-Chun, Sung Je-Guang, Yang Chu-Sing and Yang Ya-Yin
Department of Electrical Engineering, National Cheng Kung University, Tainan, Taiwan
lur.ncku@gmail.com

Abstract: Nowadays we have highly developed semiconductor engineering and can see the increasingly popular use of mobile computing devices and smartphones which are not only equipped with high speed CPUs and enormous storage capabilities but also have various built-in auxiliary processors and sensors. This advanced hardware and technology brings great convenience, however users are faced with a growing threat to their personal privacy due to various information security issues. This is especially true for the non-official APP markets which might provide malicious cybercriminals with a breeding ground from which to spread their malware and viruses targeting Android mobile devices. Due to their growing popularity, mobile phones and smartphones and tools for voice communication and information-passing may be endangered by the threats mentioned above once there is malicious eavesdropping malware which targets these smart mobile devices and they start to spread themselves. Regardless of whether users are passing information via a telephone network, their voices over Internet Protocol communication system, or simple text messages and email, malware will inevitably crop up, causing negative consequences which smartphone users must face due to the great threat to their personal privacy and information security. The very existence of Covert Channels on Android systems provides a pathway for stealthy data transfer between different Android APPs. Malicious Android APPs can utilize system resources such as screen brightness, volume and external storage to launch a covert channel communication. If no appropriate countermeasure is deployed, malicious Android Malware will use this approach to lower Android Permissions required to block each malware's entry, secretly transmitting/receiving private data, and jeopardizing smartphone users' privacy and information security. Therefore, we have to pay attention to these kinds of threats. In this paper, we analyze various scenarios and examine the possibility of android smartphones being eavesdropped upon by malicious APPs. For the purpose of experiment and analysis for our anti-eavesdropping framework design, we implemented a test malware which integrates VoIP technology and an Android covert channel. In our conclusion, we propose a malware eavesdropping countermeasure solution composed of a Covert Channel Detection Module and an Eavesdropping Behavior Analysis Module. Based on this solution, we implement an Android APP and prove that our APP can execute malicious eavesdropping behavior analysis using limited Android Permissions and mobile computing resources.

Keywords: android, smartphone, eavesdropping trojan, malware

1. Introduction

If eavesdropping Trojan horses designed to steal voice data in personnel computers have already been discov`ered (NICHOLAS KULISH, 2011), then it can only be a matter of time before smartphones responsible for communications will become targets of this kind of malicious eavesdropping malware. Soundcomber (R. Schlegel et al, 2011) has already proven that not just Personal Computers, but also handheld mobile computing devices are may be compromised by voice eavesdropping attacks.

In recent years, smartphone sales have been flourishing, and apparently will continue to sell briskly even in our present bad economy (IDC, 2012). With this economical trend, malicious cyber criminals will target smartphone users for the enormous illegal benefits. Ever since the mobile phone was developed, its main purpose has been to establish reliable voice communication, and this purpose hasn't disappeared with the development of the smartphone, which has added powerful and portable computation capability to daily life. These new products integrate advanced technology and improved functions, but in spite of all these improvements, smartphone users are more vulnerable to malicious cyber-attacks than ever. For now, Android malware mainly targets mobile users private information (such as contact info, SMS text…etc) and uses this data for financial fraud (Xuxian Jiang, and Yajin Zhou, 2012). Eavesdropping mobile malware might still be rare, but the potential losses caused by this kind of threats will not be less than losses caused by conventional malwares.

Eavesdropping can be achieved by either recording through mobile microphone, which is the method used by the Soundcomber, or by intercepting voice packets during the transmission, VoIP is an example of this. If RTP packets are not well encrypted, then packets can be easily intercepted and reorganized into audio files, and in this manner the purpose of eavesdropping achieved. Fortunately, proper media protection and key management protocols reduce the possibility of such attacks (O. Jung, M. Petraschek, T. Hoeher, I. Gojmerac, 2008). In addition, higher mobile privileges are required to carry out this scheme. Therefore, this paper will focus on analyzing eavesdropping malware utilizing the mobile microphone for each potential scenario. In the end, we design and implement an eavesdropping warning Android APP.

The study of R. Schlegel (R. Schlegel et al, 2011) also depicts the possibility of malware deploying cooperative attack through Covert Channels on the Android platform. This Covert Channel based cooperative malware attack would effectively reduce the number of android permissions required by each malware app, separating each attack component into different APPs and communicating through an irregular approach which by violating android's Inter Process Communication design would avoid being detected by the anti-virus APP. This kind of malware attack would pose a great threat to smartphone users. Therefore in our study we designed a Covert Channel based Eavesdropping malware app as an experimental object to verify our anti-eavesdropping design.

The rest of this paper is structured as follows: section 2 is related to work which introduces the covert channel and its counter measure. Section 3 is our designed defense framework for malicious eavesdropping malware. Section 4 is our experiment & results for implementing this framework, and contains data regarding our experiment using this design. Section 5 is the conclusion.

2. Related work

A covert Channel is an abnormal communication method; it can cause a great threat to system security. Since the covert channel concept was developed by Lampson in 1973 (B. W. Lampson, 1973), this topic has been an important issue for system security.

There are mainly two types of covert channels: Timing Covert Channels and Shared Storage Covert Channels. In Okhravi's research (H. Okhravi, S. Bak, and S.T. King, 2010), more detailed covert channel types are introduced and a hybrid covert channel is proposed.

The study of R. Schlegel (R. Schlegel et al, 2011) introduces Soundcomber, an android malware application which activates and records victim's phone-call content, extracts the necessary information to a text file, and sends this text file to another malicious APP through a covert channel which has network access permission. In this manner the victim's sensitive information is stolen without the user being aware. This attack approach not only reduces the system permissions needed for single APP to carry out such an attack but also the malicious server's receiving network traffic load. Soundcomber presents a well-designed eavesdropping Trojan horse.

Lin (Y. B. Lin, and M. H. Tsai, 2007) provides some mobile phone eavesdropping approaches, namely making a phone call to a modified mobile phone which automatically answers the phone call from a specific phone number without notifying the user. This converts a simple user into an innocent victim who brings an eavesdropping bug with their mobile phone and can post a great threat to the victim's privacy. If this approach is adapted by malware, then it should be handled more carefully.

3. System design

Our study analyzes this Covert Channel based malicious eavesdropping malware first, then presents a solution which can be applied to detect this kind of threat. Furthermore, our designed system will try to jam the communication channel between malwares during Covert Channel Communication, so that malware will not be able to send messages through covert channels successfully.

3.1 Valware

In addition to the above mentioned Soundcomber, which achieves malware cooperative attack through a Covert Channel, we propose one experimental real time eavesdropping malware—Valware. This is malware which combines VoIP and Covert Channel. Valware behaves like most malware applications. It repacks malicious code into a benign APP in order to propagate itself. Once the smartphone user downloads Valware, it tricks users into downloading other malware—SMS_Sniffer.

Valware behaves like most malware applications. It repacks malicious code into a benign APP in order to propagate itself. Once the smartphone user downloads Valware, it tricks users into downloading other malware—SMS_Sniffer.

Covert Channel based malicious eavesdropping malware— Valware works as follows: first SMS_Sniffer acquires the smartphone user's private schedule information by analyzing the user's SMS or Calendar data. It passes this user private information to Valware through a covert channel. After receiving the user's private data from the Covert Channel, Valware acquires the real-time voice information with integrated VoIP technology without the user's knowledge. Finally, the Malicious Remote Eavesdropper will be able to access a mobile microphone and get the smartphone user's real time audio information, thus the mission of Valware is accomplished, as shown in Figure 1.

Due to this ability, we have to take such eavesdropping behavior pattern into consideration when designing an anti-eavesdropping framework and properly design respective countermeasure to remind smartphone users of such threats. In the following section we analyze Covert Channel based eavesdropping malware behavior patterns and introduce the system design of our anti-Eavesdropping solution—Android Malicious Eavesdropping Analysis APP.

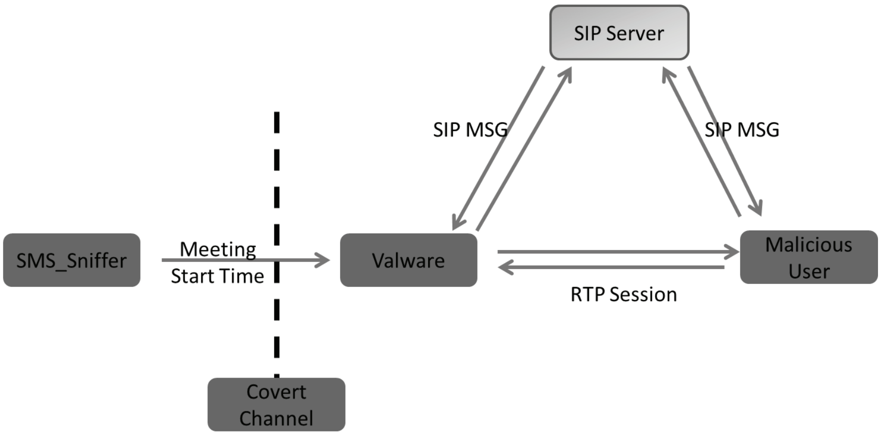


Figure 1: Valware illustration

3.2 Eavesdropping analysis and detection

Here we will explore the behavior pattern of Covert Channel based malicious eavesdropping malware by analyzing the malware apps mentioned in the previous sections. We simplify the workflows of Covert Channel based eavesdropping malware and Valware, as illustrated in figure 2 and figure 3.

From figures 2 and 3, we can determine some characteristics of this kind of malware: 1: The use of covert channels—to avoid any single malware from possessing too many android permissions and thus alerting users to this APP, the designers of malicious malware split different functions into different APPs, thus each malicious piece will require fewer android permissions and so receive less attention from the victim. Spitted malware apps need a stealthy method to communicate with each other to avoid raising attention from the victims or antivirus APPs. Therefore, Covert Channels are used for inter process communication between malwares. In one word, the usage of covert channel should be regarded as malicious behavior if detected in the first place.

Second, a mobile microphone is used by other APPs during important time periods or phone calls—however, this characteristic should not automatically be regarded as malicious behavior, because the mobile phone user might intend to record this phone conversation or use the phone dial out to call someone to discuss important issues during the meeting. This should not be regarded as malicious behavior, but the smartphone user should definitely be alerted to this kind of activity.

Third, VoIP phone calls are made during the important time period or phone call—this is similar to the second characteristic, and should be dealt with the same way, i.e. users should be notified when this behavior is detected.

After taking these three characteristics into consideration, we designed the Anti-Malicious Eavesdropping Solution and implemented it on android mobile phone. Anti-Malicious Eavesdropping Solution have two behavior analysis modules: Covert Channel Detection Module and Eavesdropping Behavior Analysis Module.

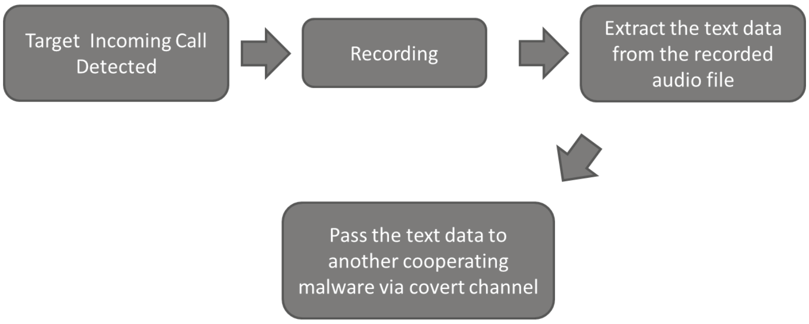

Target Incoming Call Detected
Recording
Extract the text data from the recorded audio file
Pass the text data to another cooperating malware via covert channel

Figure 2: Simplified local eavesdropping workflow

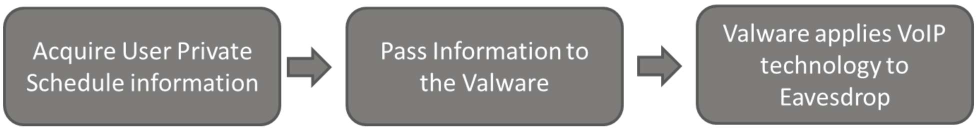

Acquire User Private Schedule information
Pass Information to the Valware
Valware applies VoIP technology to Eavesdrop

Figure 3: Simplified workflow for valware

3.3 Covert channel detection module

The task of the Covert Channel Detection Module is to detect the deployment of Covert Channel communication between malwares. Detecting Covert Channel at the Android Application level is more difficult than in the Android framework level because Android APP has limits in acquiring other APP's information due to Android's system design. For instance, we cannot design an APP explicitly to get the information to understand which android APP modified the system volume. We can't even directly learn of possible suspect APPs which have the ability to modify the system settings if this modification doesn't require any android permission (such as modifying system volume).

Despite all these limitations, we were able to get the possible Covert Channel message sending APP and receiving APP lists by listening and analyzing system information. Therefore, detecting the system anomaly caused by a Covert Channel sender and receiver APP is our system's important preliminary task. Fortunately, there are classes in Android API which can be used to listen to system state changes, such as ContentObserver and FileObserver; we can apply those API classes to detect the system anomalies when they occur.

As soon as our system is aware of anomalous activity, it logs the time of the occurrence. Based on the category of the system anomaly, the system will acquire a running Android APP list and permissions of the APP, and put the suspect APPs into a Suspect Covert Channel Sender APP list. At the same time, considering modifying certain system settings doesn't require android system permissions, so the system will also monitor Android APPs whose installation times are close to the occurrence of the system anomaly and put them into a Suspect Covert Channel Sender APP list.

The Covert Channel Detection Module will disturb data transfer over the Covert Channel. When system anomaly detected, the system will determine which category the system anomaly falls into, and then find the shared resource of the Android Covert Channel on the Android system. After the system finds the shared resource, it will interfere with data transmission over the Covert Channel by randomly modifying the Android system settings. This design can make up for the disadvantage of the analysis time requirement for the system, as shown in figure 4.

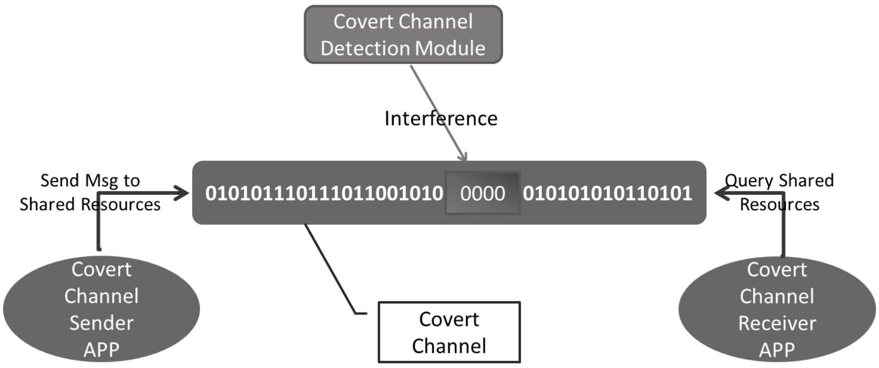

Covert Channel Detection Module
Interference
Send Msg to Shared Resources
010101110111011001010 0000 010101010110101
Query Shared Resources
Covert Channel Sender APP
Covert Channel
Covert Channel Receiver APP

Figure 4: Covert channel interference illustration diagram

As can be seen in Figures 2 and 3, we find that Covert Channel receiving APP usually has one essential task—transferring the received data to the outside world. Therefore, once the Covert Channel receives an Android APP and tries to send the received data from Covert Channel to a remote site, it will require Android Permissions such as SMS_SENT or INTERNET to carry out the action. The Covert Channel Detection Module, when it detects a system anomaly, will acquire the list of running Android APPs which have the permission to execute such a task. Both Suspect Covert Channel Sender's APP list and Suspect Covert Channel Receiver's APP list can be formed into a Covert Channel Analysis Matrix, as shown in Table 1. In our system, the Covert Channel Analysis Matrix is implemented with Android built-in SQLite database.

As long as the system anomaly is detected, the Covert Channel Analysis Matrix will be updated by both newly created Suspect Covert Channel Sender and Receiver lists. After several updates, the Covert Channel Detection Module will get the APP pair which has the maximum value in the Covert Channel Analysis Matrix. The system will give the smartphone user a warning about the Covert Channel threat and the suspect APP pair. The whole workflow of Covert Channel Detection Module is depicted in Figure 5.

Table 1: Illustration for covert channel analysis matrix

	APP_r_1	APP_r_2	APP_r_3	APP_r_4	APP_r_5
APP_s_a	1	1	1	1	1
APP_s_b	1	1	1	1	1
APP_s_c	3	1	1	1	1
APP_s_d	2	2	1	1	1
APP_s_e	1	1	1	1	1

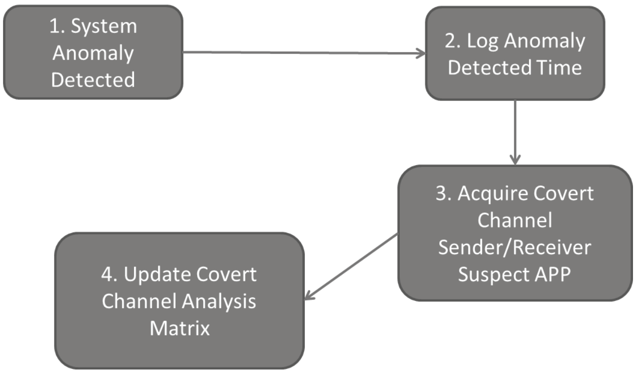

1. System Anomaly Detected
2. Log Anomaly Detected Time
3. Acquire Covert Channel Sender/Receiver Suspect APP
4. Update Covert Channel Analysis Matrix

Figure 5: Workflow for covert channel detection module

3.4 Eavesdropping behavior analysis module

An Eavesdropping Behavior Analysis Module will be activated under following circumstances: Circumstance 1—Start time which smartphone user sets up the module. Circumstance 2—The incoming or outgoing phone call established. Circumstance 3—When Covert Channel Detection Module detects a system anomaly, this module will be activated.

Once the Eavesdropping Behavior Analysis Module is activated, it runs tests to check if the system state meets any of the following conditions : Condition 1—Mobile microphone is not muted: Because the eavesdropping behavior is only possible when microphone is not muted.

Condition 2— The speed of the usage of internal and external storage is higher than the system threshold. We design this condition into our system because some malicious eavesdropping malware stores the recorded audio file in a local system, and then uses them for criminal purposes. Depending upon the mobile computing hardware used, the sampling rate varies. In our system, we setup the threshold at 64 kbps.

Condition 3—During the activation period of Eavesdropping Behavior Analysis Module, Android APPs are run which have permission, such as Audio_Recorder or Internet to execute remote or local eavesdropping.

In the system, it will give the user a warning messages and log the event with time, possible eavesdropping APPs and eavesdropping type based on whether the criteria satisfy certain conditions after the test. At the same time, the suspected malicious APPs list is generated in this module will also be used to update the Covert Channel Analysis Matrix in Covert Channel Detection Module. The workflow for Eavesdropping Behavior Analysis Module is illustrated in Figure 6.

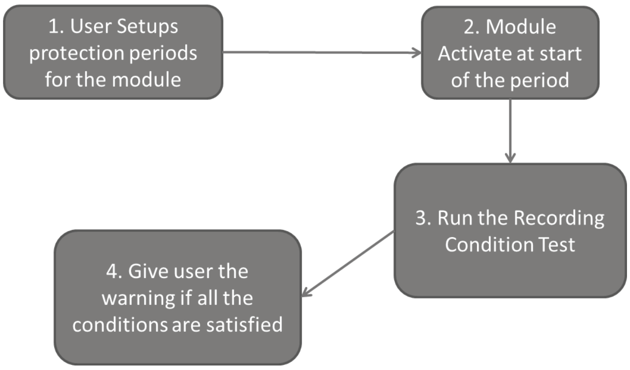

1. User Setups protection periods for the module
2. Module Activate at start of the period
3. Run the Recording Condition Test
4. Give user the warning if all the conditions are satisfied

Figure 6: Eavesdropping behavior analysis module

For Covert Channel based eavesdropping malware, our Anti Malicious Eavesdropping System includes the two modules mentioned above. These two modules can detect the known Covert Channels and analyze eavesdropping behavior, and warn the smartphone user. We test this system with Valware in section 4.

4. Experiment and result

In this section, we test this Anti-Malicious Eavesdropping System. At the beginning, a Covert Channel with three types of shared resources is implemented for experimental purposes. The shared resources used in the study are: 1. Updated Directory 2. Screen Brightness 3. System Volume. The characteristics of each Covert Channel are described in Table 2.

Table 2: The characteristics of each covert channel

Shared Resource used by Covert Channel	Characteristics
System Volume	1. If the media device is not used, then it is hard for the user to be aware of a problem. 2. It doesn't require Android Permission to modify the systems audio setting, but the range of the volume is from 0 to 15. Therefore the channel capacity is limited. 3. The possibility of being interfered with is high.
Screen Brightness	If it is not well designed, a change in screen brightness will raise the attention of the user. The range of screen brightness is from 0 to 255, the channel capacity is better than the system volume covert channel. The possibility of interference is high. Covert Channel Sending APP requires Android permissions to carry out the transmission.

Shared Resource used by Covert Channel	Characteristics
Updated Directory	It is hard to detect the Reading and Writing files into External storage. The capacity is not largely limited by the system setting. Covert Channel Sending APP requires Android permissions to carry out the transmission.

Figure 7 depicts the system CPU usage after Valware starts operating in the Android system. It indicates that it will be hard to detect mobile malware through the abnormal usage of the system resources due to advanced semiconductor technology.

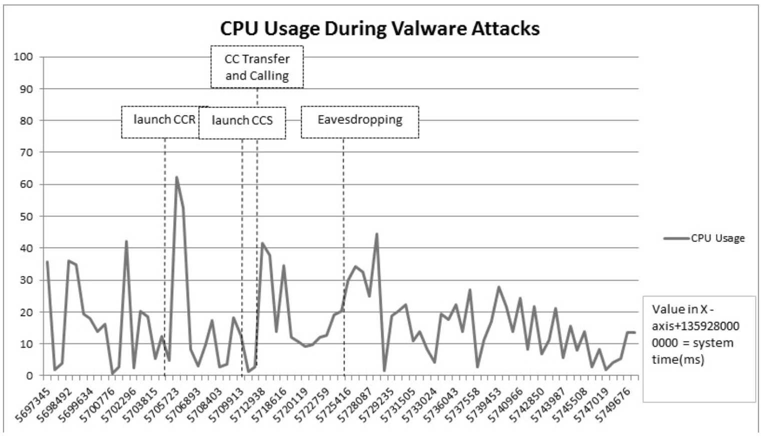


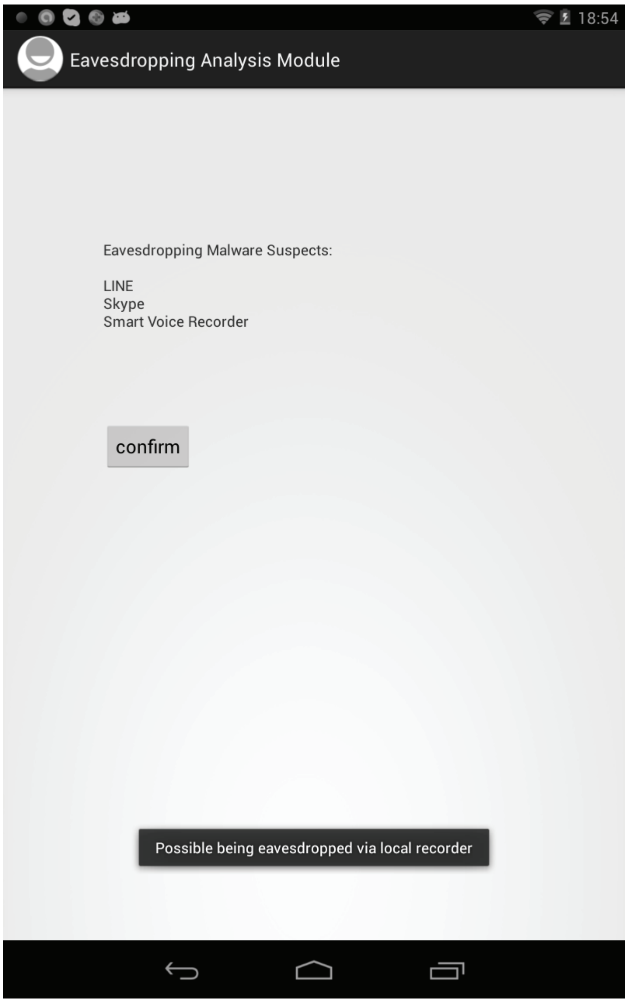
Figure 7: CPU Usage during valware attacks

Table 3: Responsiveness of each Antivirus APP

	Valware w/ Volume CC.	Valware w/ Brightness CC	Valware w/ Updated Directory CC
Lookout Premium	Scanned during installation. No countermeasure deployed during Valware attack.	Scanned during installation. No countermeasure deployed during Valware attack.	Scanned during installation. No countermeasure deployed during Valware attack.
Avast! Mobile Security	Scanned during installation. No countermeasure deployed during Valware attack.	Scanned during installation. No countermeasure deployed during Valware attack.	Scanned during installation. No countermeasure deployed during Valware attack.
AegisLab Antivirus Premium	Scanned during installation. No countermeasure deployed during Valware attack.	Scanned during installation. No countermeasure deployed during Valware attack.	Scanned during installation. No countermeasure deployed during Valware attack.
Anti-Malicious Eavesdropping System	Both modules provide smartphone users warning and log the event.	Both modules provide smartphone users warning and log the event.	Both modules provide smartphone users warning and log the event.

Android 4.2 Nexus7 is installed with Antivirus APPs, and then it is tested during a Valware attack. The responsiveness of each Antivirus APP during Valware attack is compared in Table 3. It shows that the Covert Channel and eavesdropping attack can be successfully launched under the current situation, and the Anti-Malicious Eavesdropping System will give smartphone users proper warnings for eavesdropping, and these types of covert channel attack. The system also interferes with the Covert Channel message and analyzes the Suspect Covert Channel based malware APPs.

After Valware starts its attack, our Anti-Malicious Eavesdropping System can detect system anomalies, as shown in Figure 8. This experiment proves this system can also get the list of current popular VoIP APPs (such as Line, Skype and 3CX) and local audio recording APPs (such as Smart Voice Recorder). If those APPs are detected during the important timing, the system will give smartphone users warning messages which inform users of the potential risk.



Eavesdropping Malware Suspects:

LINE
Skype
Smart Voice Recorder

Possible being eavesdropped via local recorder

Figure 8: Screenshot of eavesdropping analysis APP

Figure 9 shows that our Anti-Malicious Eavesdropping System is a lightweight solution for Android systems. That is, the system requires relatively small system resources and system permissions, and it provides smartphone users options to deal with threats like Covert Channel based malicious eavesdropping malware.

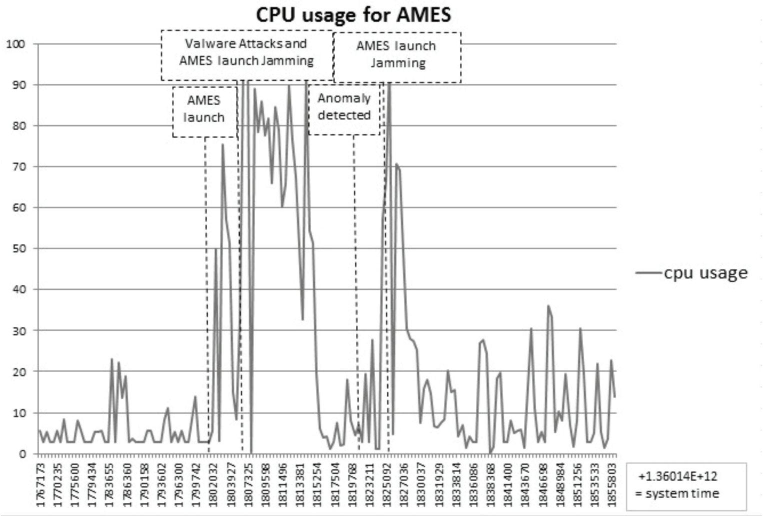


Figure 9: CPU Usage during anti-malicious eavesdropping system activation

5. Conclusion

This study mainly focuses on three types of Covert Channel based malicious eavesdropping malware, and provides a solution which can analyze and detect targeted threats. Android Covert Channels in our study include Covert Channels which utilize shared resources like Updated Directory, screen brightness and system audio volume.

Working under the premise that the Android Framework is not modified, we designed a system which takes system information such as system audio volume, screen brightness, and Android permissions as the weight in Covert Channel Analysis Matrix to analyze the possible malicious behavior. With Eavesdropping Behavior Analysis Module targeting possible malicious eavesdropping, smartphone users are given proper warnings and the data passing through Covert Channel types mentioned above can be wholly or partially blocked. In addition, this Anti-Malicious Eavesdropping System only uses limited system resources, as can be seen in the experiment.

The concept of a Covert Channel was proposed in 1973; Covert Channels have been one of the important topics in information security research. With the popularity of mobile devices, malware which is based in Covert Channels is likely to increase in the future. We understand the variety of the Covert Channel attacks, and we will work at providing proper detection approaches.

Acknowledgements

This work was supported by the Department of Electrical Engineering, National Cheng Kung University, under projects from the National Science Council (NSC 101-2219-E-006 -004) of Taiwan.

References

Butler W. Lampson (1973), “A note on the confinement problem,” Communications of the ACM, vol. 16, no. 10, pp. 613–615.

Hamed Okhravi, Stanley Bak, and Samuel T. King (2010), “Design, Implementation and Evaluation of Covert Channel Attacks,” in HST.

IDC (2012), Android Expected to Reach Its Peak This Year as Mobile Phone Shipments Slow, [Online], http://www.idc.com/getdoc.jsp?containerId=prUS23523812

NICHOLAS KULISH (2011),Germans Condemn Police Use of Spyware, [Online], http://www.nytimes.com/2011/10/15/world/europe/uproar-in-germany-on-police-use-of-surveillance-software.html?_r=0

Oliver Jung, M. Petraschek, T. Hoeher, I. Gojmerac (2008), "Using SIP identity to prevent man-in-the-middle attacks on ZRTP," Wireless Days, 2008. WD '08. 1st IFIP, vol., no., pp.1-5.

Roman Schlegel et al (2011), "Soundcomber: A Stealthy and Context-Aware Sound Trojan for Smartphones" in NDSS.

Xuxian Jiang, and Yajin Zhou (2012). “Dissecting android malware: Characterization and evolution,” In Proceedings of the 33rd IEEE Symposium on Security and Privacy.

Yi-Bing Lin, and M. H. Tsai (2007), "Eavesdropping Through Mobile Phone," IEEE T. Veh. Tech., vol 56, pp. 3596-3600.

PhD Research Papers

Technology, Violence and law: Cyber Attacks and Uncertainty in International law

Samuli Haataja
Griffith University, Gold Coast, Australia
s.haataja@griffith.edu.au

Abstract: In 2007 Estonia was faced with a new type of international violence that was difficult to conceptualise. Characterisations of the cyber attacks by Estonian officials at the time ranged from war, crime to terrorism. The technological makeup of cyberspace led to a range of problems for the traditional distinctions between these categories and hence international law was uncertain in its application to this new form of violence. These issues are among those generally discussed in literature on cyber attacks and international law. This literature also tends to follow a typical pattern of writing about law and technology, and arguably this does not result in a developed understanding of the relationship between law and technology. However, another body of literature exists which seeks to understand the intersection of law and technology better by looking at past events where technology created problems for the law, the socio-technical context of the law and the values that law seeks to protect. By adopting the insights from this body of literature, the uncertainties that cyber attacks (technology) creates for law will be explored. Accordingly, it will be shown that cyber attacks create a number of uncertainties for international law. On one level, this new type of violence has created uncertainties in the application of existing law and thus led to legal issues. These are centred around doctrinal issues on state responsibility (particularly attribution) and what constitutes an illegitimate use of force. On another level, they raise uncertainties about the compatibility of law premised upon a technological environment in which state sovereignty is central to regulate behaviour in an environment in which states lack a monopoly of violence and distinctions between the actors inflicting this violence is less clear. Exploring these uncertainties will lead to a more developed appreciation of how technology can shape the way we understand violence in international law.

Keywords: cyber attacks, technology, international law

1. Introduction

War, as a form of violence, has persisted throughout known history. Indeed it has been a defining feature of the modern state given the close links of notions of territory and violence to sovereignty. Besides the legal monopoly of violence afforded to states by virtue of their sovereignty, states have historically also held the exclusive technological capacity to monopolise violence. Initially this took place across land but technological advancement expanded it into the sea, air and space domains, though territory remained the physical and legal basis for sovereignty or sovereign rights within these domains. Recently however, technological change has given rise to a new form of violence complicating the traditionally held monopoly of violence by states. Cyber attacks constitute this new form of violence[1] and have become increasingly prominent in recent years. However, the prevalence of cyber attacks and the legal issues they raise has led to the general view that law is being outpaced by technology and thus unable to control this technologically enabled violence.

Using conceptualisations of the Estonia 2007 cyber attack as an example, this paper will explore why and how technology (cyber attacks) has resulted in this perception that law is falling behind technological developments. In doing so it will also explore the broader uncertainties that cyber attacks raise for international law. The first part of this paper will outline the different conceptualisations of the cyber attacks evident in Estonia's official responses to the attacks. It will also provide a brief account of how existing literature tends to differentiate these conceptualisations (war, crime and terrorism) and how these become problematic in the cyber context. The second part will review how existing literature on cyber attacks and international law tends to characterise the problems that technology creates for law and how the approach of this literature to the relationship between law and technology is rather narrow. It will also outline existing literature that seeks to provide more sophisticated understandings of the relationship between law and technology. Part three will then discuss the uncertainties in international law made evident by the new form of violence that cyber attacks embody. Collectively this paper will demonstrate that a broader understanding of the relationship between law and technology is needed in order to fully appreciate the uncertainties that

[1] For the purposes of this paper, the notion of violence is used to avoid context specific distinctions between different types of 'armed conflicts' and those that do not fit within these legal constructs. The term 'cyber attack' is also used broadly as a means of distinguishing this novel form of violence from the traditional types of violence that international law has been concerned with.

cyber attacks raise for international law and how technology can shape the way we understand violence in international law.

2. Estonia 2007

In 2007 Estonia was subject to a new type of violence which was vectored through cyberspace. Estonia was victim to a number of waves of cyber attacks following its government's decision to relocate a politically contentious war memorial statue.[2] Given the scale and duration of these attacks against a state heavily reliant on the internet, the events were widely described in the media as a cyber war (Landler and Markoff, 2007)(Farivar, 2007). The Estonian government's initial responses in early May reflected the rhetoric, as it was believed that the cyber attacks were part of a broader Russian attack against Estonia and hence that Estonia was under attack from Russia (Paet, 2007a)(Ansip, 2007).[3]

In the days and months that followed the attacks, Estonian officials described the cyber attacks as a part of a war, as crime and as terrorism. On 1 May, Estonia's Minister of Foreign Affairs declared that '[t]he European Union is under attack, as Russia is attacking Estonia' (Paet, 2007a). This was echoed the following day in a speech by the Prime Minister to the Riigikogu (the Estonian Parliament), who stated that the sovereign state of Estonia was 'under a heavy attack' (Ansip, 2007). On the other hand, on 11 May the Minister of Foreign Affairs spoke at a Council of Europe committee meeting about the ongoing incident, however he framed this within the language of cyber crime and the Convention on Cybercrime framework urging Russia to take measures against cyber criminals operating within its territory (Paet, 2007b). In the lead up to a meeting of European Union defence ministers on 14 May, Estonia's Minister of Defence made a statement raising issues regarding whether cyber attacks could constitute military action under the North Atlantic Treaty Organisation framework, thus allowing for the invocation of the North Atlantic Treaty's collective security provisions (Estonian Ministry of Defence, 2007). In this context, he stated that 'not a single NATO defence minister would define a cyber attack as a clear military action at present; however, this matter needs to be resolved in the near future' (ibid). A day after the meeting, the Estonian Minister of Defence compared the attacks to terrorism stating that, given the 'scale of damage and the way these cyber-attacks have been organised, we can compare them to terrorist activities' (AFP, 2007). Additionally, the following month the Minister of Defence explicitly identified the problem of classification stating that '[i]n our minds, what took place was cyber-warfare and cyber-terrorism' (Aaviksoo, 2007).

In November 2007, approximately six months after the attacks, Estonia's Minister of Defence gave a speech in which he backed away from describing the events as 'cyber war' stating that the term had 'no real content for the time being' (Aaviksoo, 2007). Instead, he said he 'tend[s] to term the events that took place in Estonia … as cyber-terrorism' (ibid). This is largely reflected in the impact and objectives of the attacks, described by the Minister of Defence to have been primarily of 'psychological nature' causing intimidation of the people and limiting their access to information online (ibid). However, despite the lack of long term consequences or physical harm or destruction caused by the attacks, they were nonetheless stated to have 'posed a serious threat to Estonian sovereignty' (ibid).[4]

The three different classifications of the cyber attacks evident in these responses show at least three legal regimes that are potentially applicable. These range from the general 'use of force' regime, the international laws seeking to tackle cyber crime and the international prohibition on terrorism. This also results in three different (though sometimes overlapping) conceptualisations of who was responsible for the attacks and what the legal and practical implications of that responsibility mean. If regarded as cyber war then Russia would have been the enemy creating the threat. If classified as cyber crime or terrorism, then cooperation would have been sought with Russia to end the attacks launched by the culprits. Thus the responses to the cyber attacks were mixed and the implications of each different characterisation varied in seriousness and in terms of who was responsible for the threat. As will be shown, this was a result of uncertainties in international law (especially at the time) about how to classify such an attack. Nonetheless, as the Minister of Defence's speech in November 2007 demonstrated, even six months after the attacks Estonia continued to believe that its sovereignty was threatened by this new form of violence and there remained uncertainty about how to

[2] For a detailed account of the events, see (Tikk, Kaska and Vihul, 2010).

[3] Despite the lack of concrete evidence to prove or disprove who was responsible for the attacks, some regard a degree of Russian government involvement as the only plausible explanation (Ottis, 2008).

[4] Domestically however, the attacks were treated as criminal acts (Czosseck, Ottis and Talihärm, 2011).

conceptualise it under international law. As the following sections will demonstrate, much of this uncertainty in how to conceptualise the attacks was a result of perceived ‘gaps’ in the international legal framework at the time dealing with cyber attacks.

The uncertainties faced by Estonia in conceptualising the cyber attack are also the focus of much of the literature on cyber attacks and international law. As demonstrated by Estonia’s official responses, cyber attacks can be classified as war, crime or terrorism and hence can fall under different legal regimes. A common way this classification is made is by looking at the intention and/or identity of the actors involved. Shackelford for example notes the difficulty of discovering the identity of those responsible for cyber attacks due to the speed of attacks and the attackers’ ability to maintain their anonymity (Shackelford, 2009, p. 232) This problem is exacerbated by the fact that boundaries between crime and terrorism are breaking down more generally, and in cyberspace states can encourage private actors to commit cyber attacks and hide behind a veil of plausible deniability (ibid, p. 233). Hollis makes a similar point, noting that the current architecture of the internet makes it ‘difficult to know which set of proscriptions—crime, war, or terrorism—applies’ (Hollis, 2011, p. 378) to a cyber attack.

Indeed, even in the real world there are problems with classifying an act into these categorises, because, according to Osler, ‘[t]he system breaks down when acts of crime look like war, and acts of war look like crime’(Osler, 2003, p. 604). Brenner highlights the added difficulties of making these distinctions in cyberspace, noting that ‘these threat dichotomies break down when attacks are vectored through the virtual world of cyberspace’ (Brenner, 2009, p. 70). Thus, as will be demonstrated, technological change is further complicating these distinctions between different types of violence (be they traditionally characterised as war, crime or terrorism), highlighting challenges to existing law and leading to the view that law is being outpaced by technology.

3. Existing literature

Besides discussing the need to determine an actor’s identity and intentions in order to classify a cyber attack, existing literature on cyber attacks and international law also tends to follow a similar pattern of writing about the relationship between law and technology. In fact, this pattern of writing is not unusual for people writing about legal problems created by technology. This pattern is described by Tranter as the ‘law and technology enterprise’ (Tranter, 2011) and usually begins with a technological crisis or event that reveals gaps in the law or creates challenges to existing law (ibid, p. 32). It then moves on to describe new, generally value-free law which is needed to fill these gaps. Given a positivistic approach to the law, the values that law embodies are not debated and law is simply reduced to a means of implementing policy (ibid, p. 70). Further, the historical relationship between developments in technology and law is often neglected (ibid, p. 72). Effectively this narrows the ways in which the relationship between law and technology is explored and it is also assumed that law has the capability to ‘control the impacts of technological change’ (ibid, p. 70).

The literature on cyber attacks and international law largely follows this pattern of writing about law and technology. A technological crisis or event is identified (the attacks against Estonia are commonly used as an example) or the destructive potential of cyber attacks is emphasised. For example Hollis notes the ‘enormity of the cyberthreat problem’ (Hollis, 2011, p. 390) and that ‘[i]n terms of effects, cyberthreats can be merely annoying or apocalyptic’ (ibid). Similarly Shackelford maintains that ‘cyber attacks represent a threat to international peace and security that is potentially as daunting and horrific as nuclear war’ (Shackelford, 2009, p. 198). The picture painted is mostly grim as cyber attacks can have unfathomable and apocalyptic consequences, and in scale can be comparable to nuclear attacks. This crisis event then reveals or highlights ‘gaps’ in the law as it is unable to ‘keep up’ with technology. For example Hathaway et. al. argue that ‘[c]yber-attacks present a new and growing threat—one that current international and domestic laws are not yet fully prepared to meet’ (Hathaway et al, 2012, p. 877) and that new international law is needed to ‘fill the gaps in existing law’ (ibid). Even the *Tallinn Manual on the International Law Applicable to Cyber Warfare* (Schmitt, 2013) expressly acknowledges that given the speed of development of these technologies, ‘cyber practice may quickly outdistance agreed understandings as to its governing legal regime’ (Schmitt, 2013, p. 3).

This literature also presents a common solution to addressing the problems. The solution is the creation of new law designed specifically for cyber attacks or one that extends existing law to this technology (Hathaway et al, 2012)(Hollis, 2011)(Hoisington, 2009, p. 441)(Shackelford, 2009, pp. 197-198). Finally, the literature tends to focus on the applicability of international legal doctrine to cyber attacks and consequently provides a rather positivistic account of law. Specifically, the law on the use of force (*jus ad bellum*) and law in war (*jus in bello*) are common areas of international law discussed by authors engaging in debates about whether, how and why these bodies of law can be extended to cyber attacks (Waxman, 2011)(Schmitt, 1998-1999)(Schmitt, 2012)(Jensen, 2002)(Hoisington, 2009)(Brown, 2006)(Hollis, 2007)(Shackelford, 2009)(Hathaway et al, 2012).

Therefore this literature generally regards cyber attacks as a novel technological development creating new challenges to international law. The way to respond to these challenges is the creation of new law. However, the approach towards law is primarily positivistic, as the focus of much of this literature is on international legal doctrine with the political, social and technological context of law left marginalised. Despite some exceptions, this is largely the pattern of writing about law and technology that the cyber attacks and international law literature reflects. This is ultimately a narrow way of understanding the relationship between law and technology ignoring the historical relationship between the two, the technological environment in which law operates and the values that the law seeks to protect.

4. Law and technology

Another body of scholarship that focuses on law and technology more broadly seeks to address these concerns. Scholars within this body of literature seek to provide a more sophisticated understanding of the relationship between law and technology. They recognise that law and technology issues are not new and acknowledge other (now) historical events in which the two have intersected. They also consider the values that law protects and the socio-technological environment in which the law exists in order to provide a more fruitful account of the relationship between law and technology (Tranter, 2011, pp. 72-73). For example Mandel looks at the 'lessons that can be learned from past responses to once-new legal issues created by technological advance' (Mandel, 2007, p. 552) which help to understand the 'current and future law and technology issues' (ibid) that arise. Cockfield in turn argues that 'technological developments can undermine important interests and values that the law seeks to protect' (Cockfield, 2003-2004, p. 338) and that one needs to consider the 'broader context of technological changes that may affect important interests and values' (ibid, p 384). Gifford seeks to address the ways in which law and technology interact, one being 'when society decides that technology produces undesirable results and employs legal rules to contain or modify those results' (Gifford, 2007, p. 572). Therefore these authors stress the importance of history, the values that law protects and understanding the socio-technological environment in which law exists in order to better understand the relationship between law and technology.

Bennett Moses is another prominent 'law and technology' scholar who notes that the relationship between law and technology is often described as a race, with law always falling behind the pace of technology (Bennett Moses, 2007a). However, she seeks to provide a more detailed account of why and under what conditions law is outpaced by technology, in order to better understand the relationship between the two. According to Bennett Moses, technological change is change in what actors are practically or 'technically capable of doing' (Bennett Moses, 2007b, p. 598). This occurs when new forms of conduct are made possible that adjusts the limitations placed on actors by the previous technological state of affairs (Bennett Moses, 2007b, p. 594). Sometimes this leads to moments in time that the law has problems in dealing with this new form of conduct revealing 'gaps' in the law (Bennett Moses, 2007a, p. 241). Additionally, technological change can reveal the socio-technical context that previously underpinned the law. A situation in which 'gaps' in the law may be revealed occurs when law is uncertain in its applicability to new forms of conduct made possible by technological change (Bennett Moses, 2007a, p. 248).

While a degree of legal uncertainty can exist prior to technological change, this uncertainty tends to be amplified by technological change (ibid, pp. 250-253). Whether a new form of conduct is permissible or not will depend on whether it fits into an existing legal category. However, some forms of new conduct made possible by technological change cannot be easily classified as they do not fit easily into the existing legal categories (ibid, pp. 235-255). At other times, the issue is not with the new form of conduct, but instead with the legal category itself as '[s]ome legal categories and concepts become ambiguous in light of technological change' (ibid, p. 257). Thus these uncertainties that result from technological change are unique and more problematic than simply legal uncertainty as they arise from the technology itself (ibid, p. 257-258).

5. Cyber attacks and uncertainty in international law

Uncertainty is a particularly prominent reason for why legal issues are raised by cyber attacks and the 'technological change' that they embody. While there is no specific point in time at which cyberspace technologies 'changed' and led to legal problems, the wide scale cyber attacks in 2007 against a state so reliant on the internet marked a significant event that led to increased discourse about these issues. The new forms of conduct made possible, as reflected by the Estonia incident included: the ability of non-state actors to exploit legitimate internet technologies and protocols and launch large scale cyber attacks against a state; the potential of governments to support such actors while maintaining plausible deniability and avoiding existing international legal obligations; the ability to undermine a state's sovereignty in its cyberspace in previously inconceivable ways (especially given the increasing reliance of states on the internet); and the ability of actors to do this with relative anonymity. Thus broadly, the technological change resulted in a wider range of actors (including those whose identity remains unknown) with the ability to engage in a unique type of violence in cyberspace. As the abovementioned literature on cyber attacks and international law demonstrates, the technological change that cyber attacks embody have led to challenges for the law or revealed 'gaps' in the law. Here the law and technology literature, specifically the writings of Bennett Moses, helps us understand why these problems arise.

As demonstrated by Estonia's mixed responses to the cyber attacks, it was faced with a number of uncertainties at the time. These uncertainties were largely doctrinal uncertainties. They revolve around questions of attributing state responsibility and whether a cyber attack can constitute a use of force. While there existed a degree of legal uncertainty in both of these contexts, these uncertainties were amplified by technological change. Instead of physical acts of violence committed by clearly identifiable actors, the attacks Estonia faced were vectored through cyberspace and were difficult to successfully trace due to technological structure of the internet and use of attacks that exploited this structure. Further, as a kind of violence in cyberspace, the cyber attacks also highlighted the uncertainty in the 'use of force' as a concept and whether it could include non-traditional uses of force made possible by technological change.

However, the uncertainties that this new form of technologically enabled violence creates are not limited to legal doctrine, which is largely the focus of existing literature on cyber attacks and international law. Instead, there is also uncertainty on another level about the compatibility of law, premised on the monopoly of violence of states, to regulate behaviour in a different technological environment. Historically, international law was concerned with ways in which to protect a state's sovereignty over its territory and developed throughout a time when the nature of violence was primarily physical and the primary actors were states. Further, the technological ability to engage in large scale violence (war) was regarded as solely within the realm of states and at times this ability was regarded as a defining feature of sovereignty. Only states held this right and technological capacity and international law on the use of force was developed by states around this technological context. International law's history with the regulation of violence has also been mixed, though state sovereignty and their consequent monopoly of violence have remained central in this history. While changing over time, the overriding purpose of the *jus ad bellum* for example was to provide restrictions and prescriptions on the violent conduct of states. However, these rules were developed over time to deal with different forms of physical violence by one state that threatened the sovereignty of another state. In this regard, international law was built around a certain technological environment – one in which states were the only actors who held this ability to engage in wide scale violence, both legally and practically (or technologically).

Due to technological change, there has been a diffusion of power in cyberspace and states are no longer able to maintain a monopoly of violence in this realm. Thus power relationships are changing, though international law remains premised on a different technological environment in which state power is supreme. As the technological change reflected in the Estonia incident demonstrates, a new technological environment has been made apparent in which states are no longer the sole actors capable of engaging in an effective form of wide scale violence. Nor is it always clear when a state is involved in acts of violence as the current technological environment makes it easy to mask one's identity and thus blurs traditional legal distinctions of acts for which states may be responsible for. Consequently, there is not just uncertainty about how existing legal doctrine on state responsibility should apply or what legal regime cyber attacks should be categorised into. There is also uncertainty as to the ability of international law, structured around the centrality of sovereign states and based on the assumption that states are the only actors technologically capable of maintaining a monopoly of violence, to regulate behaviour in cyberspace where power is more diffused, violence is no longer solely physical, distinctions between violence inflicted by state and non-state actors is less clear, and physical territory is less fundamental. This raises questions about the significance of territoriality, physicality and violence to sovereignty and international law, and reinvigorates one of the central concerns of international law – its ability to control international violence.

6. Conclusion

Collectively, this paper has sought to demonstrate how cyber attacks, as a new form of violence made possible by technological change, create uncertainties for international law leading to the belief that law is being outpaced by technological change. While the traditional conception of international law is inherently state based with a close connection to physicality and territoriality, the current technological environment is characterised by different power structures. Cyber attacks represent technological change reflective of this new technological environment, hence challenging some of these traditional assumptions. However, these uncertainties are not limited to legal doctrine as is the focus of much of existing literature and instead raise more fundamental questions about the relationship between technology, violence and international law. This paper has shown that a broader account of the relationship between law and technology is needed to better understand how technology shapes our understandings of violence in international law. This in turn will help us obtain a better understanding about how and why international law is being outpaced by technological change, an understanding of the past and present technological environment of international law and how technology can shape the way we understand violence in international law.

References

Aaviksoo, J. (2007) "Cyber-Defense: Estonia's Recent Experience of this Unnoticed Third World War", Paper presented at 24th International Workshop on Global Security, Paris, http://www.csdr.org/2007book/aaviksoo07.htm.

AFP. (2007) "Estonia urges firm EU, NATO response to new form of warfare: cyber-attacks", *The Sydney Morning Herald* [online], 17 May, http://www.smh.com.au/news/Technology/Estonia-urges-firm-EU-NATO-response-to-new-form-of-warfarecyberattacks/2007/05/16/1178995207414.html.

Ansip, A. (2007) "Prime Minister Andrus Ansip's speech in Riigikogu", [Press Release], 2 May, http://valitsus.ee/et/uudised/pressiteated/majandus-ja-kommunikatsiooniministeerium/13183.

Bennett Moses, L. (2007a) "Recurring Dilemmas: The Law's Race to Keep Up with Technological Change", *University of Illinois Journal of Law, Technology & Policy,* Spring, No. 1, pp 239-286.

Bennett Moses, L. (2007b) "Why Have a Theory of Law and Technological Change?", *Minnesota Journal of Law, Science & Technology*, Vol 8, No. 2, pp 589-606.

Brenner, S. W. (2009) *Cyberthreats: The Emerging Fault Lines of the Nation State*, Oxford University Press, New York.

Brown, D. (2006) "A Proposal for an International Convention To Regulate the Use of Information Systems in Armed Conflict", *Harvard International Law Journal,* Vol 47, No. 1, pp 179-221.

Cockfield, A. J. (2003-2004) "Towards a Law and Technology Theory", *Manitoba Law Journal*, Vol 30, No. 3, pp 383-416.

Czosseck, C., Ottis, R. and Talihärm, A. M., (2011) "Estonia after the 2007 Cyber Attacks", *International Journal of Cyber Warfare and Terrorism*, Vol 1, No. 1, pp 24-34.

Estonian Ministry of Defence. (2007) "Minister to attend meeting of EU defence ministers in Brussels", [Press Release], 13 May, http://www.kmin.ee/en/1429.

Farivar, C. (2007) "Cyberwar I", *Slate* [online], 22 May, http://www.slate.com/articles/technology/technology/2007/05/cyberwar_i.html.

Gifford, D. J. (2007) "Law and Technology: Interactions and Relationships", *Minnesota Journal of Law, Science & Technology*, Vol 8, No. 2, pp 561-588.

Hathaway, O. A., Crootof, R., Levitz, P., Nix, H., Nowlan, A., Perdue, W. and Spiegel, J. (2012) "The Law of Cyber-Attack", *California Law Review*, Vol 100, No. 4, pp 817-886.

Hoisington, M. (2009) "Cyberwarfare and the Use of Force Giving Rise to the Right of Self-Defense", *Boston College International & Comparative Law Review*, Vol 32, No. 2, pp 439-454.

Hollis, D. B. (2007) "Why States Need an International Law for Information Operations", *Lewis & Clark Law Review*, Vol 11, No. 4, pp 1023-1062.

Hollis, D. B. (2011) "An e-SOS for Cyberspace", *Harvard International Law Journal*, Vol 52, No. 2, pp 373-432.

Jensen, T. E. (2002) "Computer Attack on Critical National Infrastructure: A Use of Force Invoking the Right of Self Defense", *Stanford Journal of International Law*, Vol 38, No. 2, pp 207-240.

Landler, M. and Markoff J. (2007) "After Computer Siege in Estonia, War Fears Turn to Cyberspace", *The New York Times*, 29 May.

Mandel, G. N. (2007) "History Lessons for a General Theory of Law and Technology", *Minnesota Journal of Law, Science & Technology*, Vol 8, No. 2, pp 551-570.

Osler, M. (2003) "Capone and bin Laden: The Failure of Government at the Cusp of War and Crime", *Baylor Law Review*, Vol 55, No. 2, pp 603-616.

Ottis, R. (2008) "Analysis of the 2007 Cyber Attacks Against Estonia from the Information Warfare Perspective" in Remenyi D. (ed) *Proceedings of the 7th European Conference on Information Warfare and Security*, Academic Publishing Limited, Reading.

Paet, U. (2007a) "Declaration of the Minister of Foreign Affairs of the Republic of Estonia. E. Government", [Press Release], 1 May, http://valitsus.ee/et/uudised/pressiteated/majandus-ja-kommunikatsiooniministeerium/13634.

Paet, U. (2007b) "Address by Minister of Foreign Affairs of Estonia Urmas Paet", [Press Release], 11 May, http://www.vm.ee/?q=en/node/3665.

Schmitt, M. N. (1998-1999) "Computer Network Attack and the Use of Force in International Law: Thoughts on a Normative Framework." *Columbia Journal of Transnational Law*, Vol 37, No. 3, pp 885-938.

Schmitt, M. N. (2012) "Classification of Cyber Conflict", *Journal of Conflict & Security Law*, Vol 17, No. 2, pp 245-260.

Schmitt, M. N. (ed) (2013) *Tallinn Manual on the International Law Applicable to Cyber Warfare*, Cambridge University Press, Cambridge.

Shackelford, S. J. (2009) "From Nuclear War to Net War: Analogizing Cyber Attacks in International Law" *Berkeley Journal of International Law*, Vol 27, No. 1, pp 192-252.

Tikk, E., Kaska, K. and Vihul, L. (2010) *International Cyber Incidents: Legal Considerations,* Coperative Cyber Defence Centre of Excellence, Tallinn.

Tranter, K. (2011) "The Law and Technology Enterprise: Uncovering the Template to Legal Scholarship on Technology", *Law, Innovation and Technology*, Vol 3, No. 1, pp 31-83.

Waxman, M. C. (2011) "Cyber-Attacks and the Use of Force: Back to the Future of Article 2(4)", *Yale Journal of International Law*, Vol 36, No. 1, pp. 421-459.

Modelling Cyber Warfare as a Hierarchic Error Effect of Information

Harry Kantola[1] and Juhani Hämäläinen[2]
[1]Finnish National Defence University, Helsinki, Finland
[2]Finnish Defence Forces Technical Research Centre (PVTT), Riihimäki, Finland
harry.kantola@mil.fi
juhani.hamalainen@mil.fi

Abstract: The use of Cyber domain in military concept has become more and more attractive, since both the society and the military itself are gradually more dependent on the functionality of related systems and resources. However, the evaluation of the impact of a possible malfunction in these systems is difficult and time consuming and there is a need for creating simplified analysis of effectiveness in Cyber Warfare. We shall introduce a description for cyber Warfare as an error source for information studied at hierarchic levels. In particular, we observe the relationship between the errors and their implications on the information. The cumulative appearance of errors generated by Cyber Warfare on different levels may render up to strategic effects. For instance, one bug in a code can terminate large scale system completely, i.e., technical error has yield to operational damage. Are these effects of core value or does the ambitious actions taken eventually result only in minor discomfort? In order to estimate the measure of effectiveness within Cyber Warfare, extensive simulations should be done and a large amount of variables have to be identified and determined. Instead of simulating effects on all levels, we propose a modelling framework step by step by generating a path from technical knowledge to strategic level, similarly to military configuration as technical, tactical, operational and strategic levels. In this study, we consider Cyber Warfare as an error source on all levels of framework by introducing and estimating it as an error of information. However, we shall not consider possible Cyber Warfare techniques or sources, but instead we model Cyber Warfare as reliability factor for given information. This yields to a hierarchic model construction of Cyber Warfare, emphasised with examples, where lower level models provide information and in particular disinformation from the technical details to the strategic entities. This study tries to identify and understand causality between large-scale problems on operational and strategic level and details in systems on technical and tactical level in the context of Cyber Warfare.

Keywords: information, cyber warfare, modelling, military, hierarchy, causality

1. Introduction

Modern society is dependent on information, information systems and their applications. Cyber domain has become hype in linguistics and therefore we have chosen to use the term Information sphere domain instead for better understanding. Allen and Gilbert (2009) have proposed the definition Information Sphere Domain for usage concerning the information in cyber for creating a better understanding of the matter in interest. According to their definition, the information in the system can be primarily altered and modified instead of the cyber itself. Therefore, when debating about cyber warfare, the issue is to interfere, corrupt or destroy information in the system or systems of systems.

Information can be defined and measured in terms of bits and it is usually denoted by I (Shannon 1948). We shall consider information errors due to cyber warfare formally and demonstrate this approach with illustrative examples. We shall not study special techniques how the interference can be generated but how the errors can evolve from the details to wider entities.

By using simple models, there exist possibilities to describe and "think" complex systems in order to focus on the essential part of the problem (Cutts, 2009). As Box and Draper (1987) states "All models are wrong; the practical question is how wrong do they have to be to not be useful." enables the simplification of the examination of the problem. This enables the possibility to create a general model for estimation of the impact of information error added in different layers by approaching the problem with simplified models.

The aim of the simplified model is to establish a method for evaluating the impact of information error in different layers as well as the causalities of the errors between the layers. Other goals of the model are to assist in finding out actions that do not render in additional value in the other layers and to estimate the cumulative effects for upper layers.

This paper focuses on introducing a hierarchic information error model and joining that with the categorization of computer network attacks (CNA) (Kantola, 2011). More precisely we search for the correlation between introduced errors at different military levels and CNA activities at different networks.

2. Cyber warfare

A classical approach to interfere with an adversary organization requires activities on three different levels, on the physical, the information and the social layer (Kott 2007, Beynon-Davies 2002; Libicki 2007).

In context for cyber warfare this requires a possibility to establish contact with the system in the physical level. On the information level the requirement is that there exist understanding and knowledge of the information system itself and how do systems interact with other systems and their environments. Finally, on the social layer, the focus is on knowledge of using principles of information systems. The cognitive understanding about the systems is fundamental for the entire network to be able to use it effectively. (Libicki 2007). This means that the effect of actions will always culminate to upper levels regardless on which level the actual action has taken place. By this we can state that information is in focus for all organization and military installation and therefore vulnerable for cyber attacks. It is no difference on what the type of information is interfered, the information on the technical level or the analyzed database information on the strategic level.

3. Information model, errors and hierarchy

We shall consider information in hierarchic levels and in particular how the cyber warfare can change information. The levels of our construction are technical, tactical, operational and strategic. We denote the amount of information affected with cyber warfare by I_{CW} and information at different levels of the information hierarchy with upper indices “technical (TE), tactical (TA), operational (OP) and strategic (ST)” yielding to double index notation for affected information. By using these notations, information is given by I^{TE}, I^{TA}, I^{OP} and I^{ST} and affected information by $I_{CW}{}^{TE}$, $I_{CW}{}^{TA}$, $I_{CW}{}^{OP}$ and $I_{CW}{}^{ST}$. Let us finally denote the remaining undisturbed information with *italic* notation, i.e., by I^{TE}*,* I^{TA}*,* I^{OP} and I^{ST}*.* This is emphasized in figure 1.

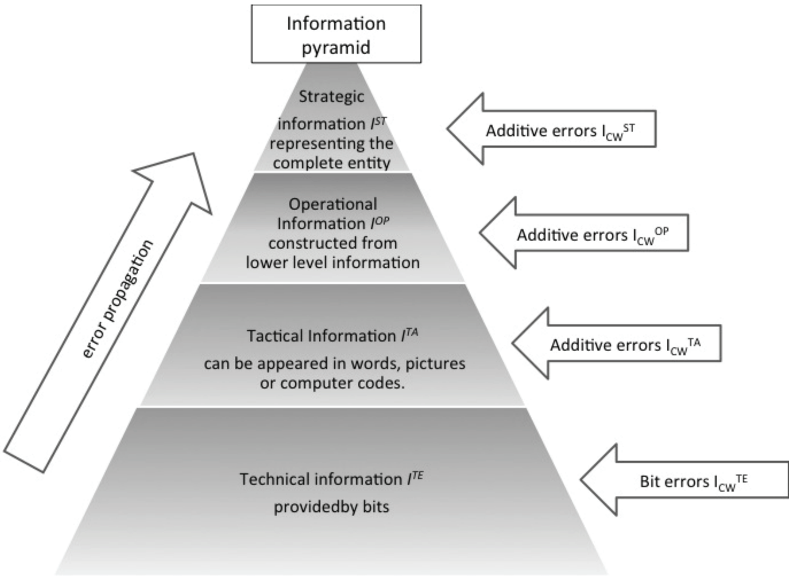

Information pyramid
Strategic information I^{ST} representing the complete entity
Operational Information I^{OP} constructed from lower level information
Tactical Information I^{TA} can be appeared in words, pictures or computer codes.
Technical information I^{TE} providedby bits
error propagation
Additive errors $I_{CW}{}^{ST}$
Additive errors $I_{CW}{}^{OP}$
Additive errors $I_{CW}{}^{TA}$
Bit errors $I_{CW}{}^{TE}$

Figure 1: The pyramid visualizes the information levels in hierarchic structure. The arrow in the left hand side represents the information and error flows from bottom to the top and arrows at right cumulative errors interfering information. Information expressions in the middle consist of complete information available at each level

Model for information is now constructed in a functional form as follows. We begin with technical information I^{TE} given by bits. If we consider a closed system, for instance closed software built up in a hierarchic manner, the hierarchic information description is proposed to be in a form $I^{TA}=I^{TA}(I^{TE})$, $I^{OP}(I^{TA})$, $I^{ST}=I^{ST}(I^{OP})$=> $I^{ST}=I^{ST}(I^{OP}(I^{TA}(I^{TE})))$. In order to consider open information system, all level shall include additive information denoted by $I_0{}^{TA}$, $I_0{}^{OP}$and $I_0{}^{ST}$, which may be affected by CW as well. For instance, in decision making it can present the specialists existing knowledge and it can be disturbed by cyber warfare as well. By using functional notation this provides following expressions for information.

$$
\begin{aligned}
&I^{TE}=\mathrm{I}^{TE}-\mathrm{I}_{CW}^{TE}\\
&I^{TA}=\mathrm{I}^{TA}(I^{TE})+\mathrm{I}_0^{TA}-\mathrm{I}_{CW}^{TA}=\mathrm{I}^{TA}(\mathrm{I}^{TE}-\mathrm{I}_{CW}^{TE})+\mathrm{I}_0^{TA}-\mathrm{I}_{CW}^{TA}\\
&I^{OP}=\mathrm{I}^{OP}(I^{TA})+\mathrm{I}_0^{OP}-\mathrm{I}_{CW}^{OP}=\mathrm{I}^{OP}(\mathrm{I}^{TA}(\mathrm{I}^{TE}-\mathrm{I}_{CW}^{TE})+\mathrm{I}_0^{TA}-\mathrm{I}_{CW}^{TA})+\mathrm{I}_0^{OP}-\mathrm{I}_{CW}^{OP}\\
&I^{ST}=\mathrm{I}^{ST}(I^{OP})+\mathrm{I}_0^{ST}-\mathrm{I}_{CW}^{ST}=\mathrm{I}^{ST}(\mathrm{I}^{OP}(\mathrm{I}^{TA}(\mathrm{I}^{TE}-\mathrm{I}_{CW}^{TE})+\mathrm{I}_0^{TA}-\mathrm{I}_{CW}^{TA})+\mathrm{I}_0^{OP}-\mathrm{I}_{CW}^{OP})+\mathrm{I}_0^{ST}-\mathrm{I}_{CW}^{ST}.
\end{aligned}
$$

These are the proposed models for information at different hierarchic levels.

However, this model can be used to estimate the cyber warfare effects only if the functional forms are known. Interference I_{CW} can be realized due to several reasons. The bits might be intentionally delayed, changed or destroyed by malwares, viruses etc. It is possible to prevent the functioning of software under interests or physically damage critical components with CW. However, these changes might be also due to software errors or wrong using principles. Information errors at tactical level follow either from the technical level or can appear in tactical level. Similarly errors in operational and strategic levels propagate from lower levels or are added at these levels. This shows the causality between the errors. In principle, one bit error can cause the damage of whole information in the “worst case” situation. On the other hand, technical bit error can be harmless and does not appear in application errors. An example of the hierarchic description of effects in electromagnetic context is provided in Järviö & Hämäläinen 2011, where electromagnetic threats are considered in terms of hierarchic system.

We can estimate the amount or ratio of interfered information if it is given by the number of bits by I_{CW} and $(\mathrm{I}_{CW})/\mathrm{I}$. However, these do not tell anything on the “value” of missed information, which is situation dependent and should be evaluated case by case.

4. Examples

We shall now consider simple examples, where we connect our model interpretation to realistic questions. First, we take an example of an excel sheet where one extra empty space can destroy the functionality of the calculation sheet by making one function (here “EXACT”) functioning wrong.

Table 1: Example of 8 bits error (empty space in ASCII bits) breaking the calculation sheet

	A	B	Function EXACT(A;B)
Comparable texts	20H	20H	TRUE
Extra empty space after 20H in B column	20H	20H	FALSE

This example shows the difficulty of seeing the source of error. This kind of error can easily lead to bigger errors, if this part of the sheet provides input information for further calculation and it can be easily propagated through all levels in hierarchy.

The second example is a small physical damage, i.e. broken pin, in VGA-connector causing errors of transformed visual information like errors in colour. These errors may yield to wrong interpretation, for instance, in maps.

5. Characterization of cyber events

Computer Network Attacks (CNA) has been categorized according to the level of technology and the purpose of the attack and according to on what level on the hierarchy it has had its effect (Kantola, 2011, p 62.). This study did not take into account if the attack or threat was aimed at a civilian or state organization. Kantola’s findings of the categorization of different methods and attacks are shown in figure 2. Note that the figure 2 emphasizes the applicability of categorization method, but does not present all possible attack methods. Information interaction between the levels has been represented in this figure by arrows.

The category in the lower left hand corner represents methods which main goal is to disrupt or block usage of networks. These kinds of attacks are rather easy to realise, although the achieved effect are relatively low. Conducting these kind of attacks are also fairly safe for the attacker, since their origin is difficult to track. (Kantola 2011).

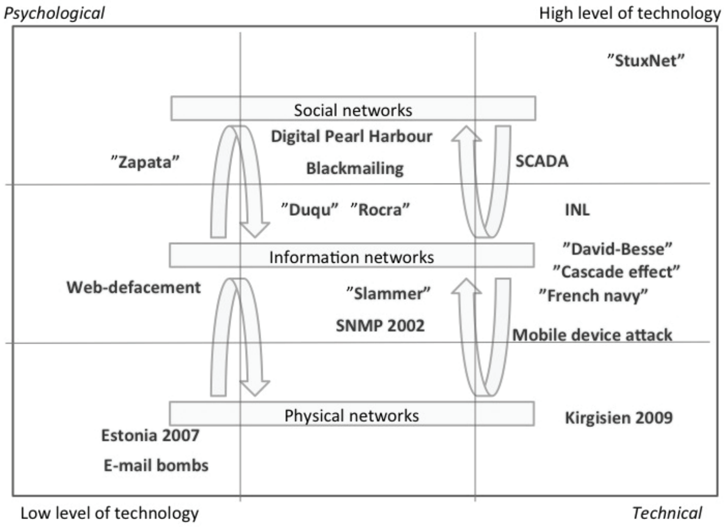


Figure 2: Categorizing of CNA- incidents by technology and achieved effect

The upper left hand corner represents methods with higher effect, but also dependent on several factors. These types of attacks are most often aimed at the socio-psychological aspect of the user and not only on the network and equipment itself. These types of attacks are rarer, since the attack method requires high level of coordination and understanding about the social network and the cultural aspects concerning the target group. Typical for this category is that the CNA activity can assist in psychological actions (operations) and activity in the social network and social media. These methods have a lot in common with psychological warfare, but the distinction is that it is the existing information that has been altered, disrupted from usage or spread in technical networks by technical means. In this category there are a large risk for attacker included, since it is often possible to identify the attacker and carry through counterattack using the altered information.(Kantola, 2011).

The category in the lower right side represents activities that have a low risk but can have a high payback when succeeded. These are technical means focused on interrupting or altering technological systems. These methods utilize the physical network topology and acts upon possibilities these enable. There are several restrictions in this category and they are largely dependent on what kind of technology is used and how the environment itself looks like. (Kantola, 2011).

Methods and practices that yield high effect, but include also high risk, are placed in the upper right hand corner. These types of actions include a combination of physical, informational and social activities. These types of actions are highly specialized and require thorough study of the object, which is going to be attacked (Kantola, 2011). In the middle, we find combinations of these categories and methods that are not that fully developed or utilized as the once in the corners.

If we consider the information flows to the established network hierarchy of known incidents, we would find that the information sphere domain would act accordingly to figure 3.

As physical networks function mostly on technical level, these are comparable to technical level on information hierarchy. Previous study (Kantola, 2011) shows that events in information networks are mostly represented by activities that have tactical and operational aspects yielding to tactical or operational error level. Finally, activities in social networks affect the decision making parties and are thus related to strategic level. Due to this, we shall combine the hierarchic information error description and CNA categorization.

The hierarchic error description and the CNA categorization correlate and present therefore a possibility to evaluate the impact of error effect in information created by CNA-activities. The introduced mathematical formulas might be applicable with further development. By choosing what kind of information error is inserted in the system, it is possible to find out possible consequences on the other (higher) levels. These values are not absolute since parts of the package are dependable on psychological values, which are not possible to give a value that is for example how much does decision maker interprets given information and how much he or she will count into uncertainties before acting on the given information. The combination of these two findings is presented in figure 4.

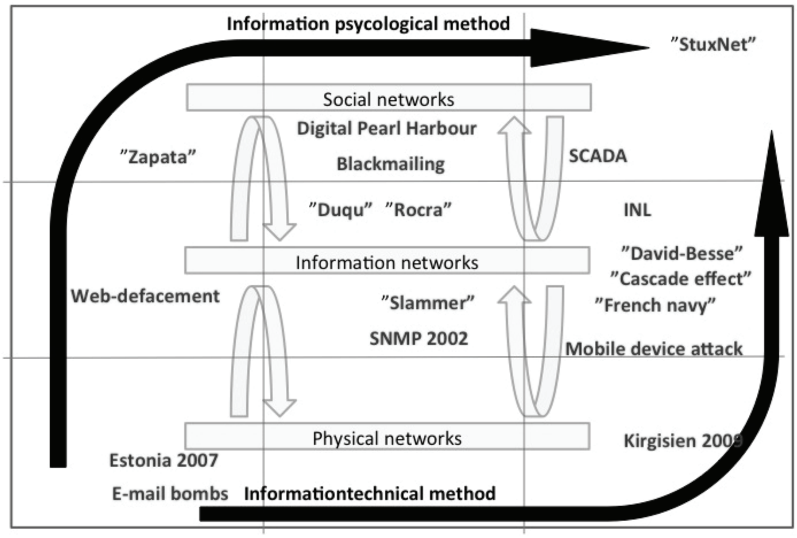


Figure 3: Information flows in different network hierarchy related to the presented categories of CNA-actions

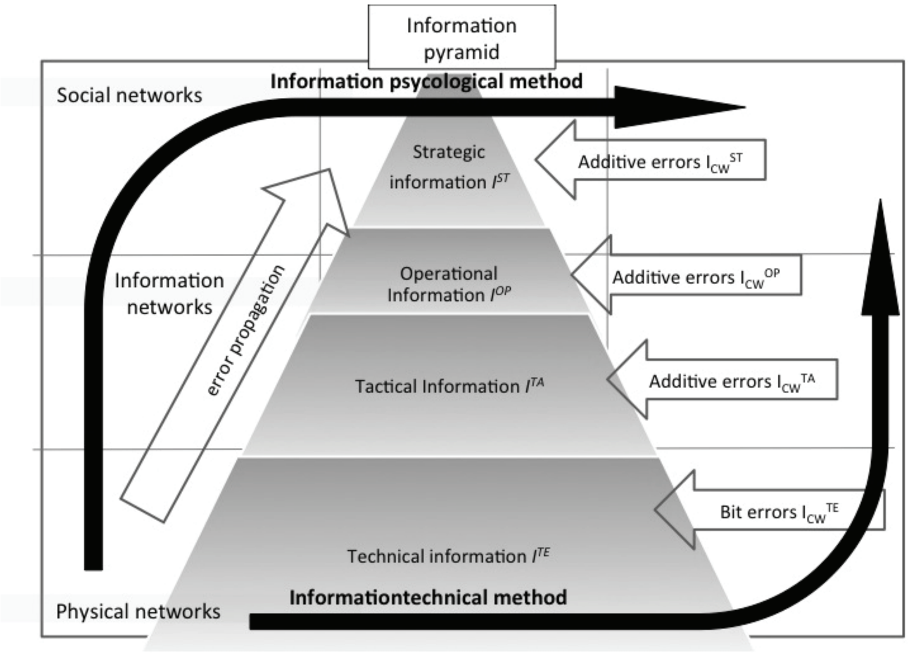

Information pyramid
Social networks
Information psycological method
Strategic information I^{ST}
Additive errors I_{CW}^{ST}
Information networks
error propagation
Operational Information I^{OP}
Additive errors I_{CW}^{OP}
Tactical Information I^{TA}
Additive errors I_{CW}^{TA}
Bit errors I_{CW}^{TE}
Technical information I^{TE}
Informationtechnical method
Physical networks

Figure 4: Combined picture of hierarchic and CNA categorization

The introduced hierarchic information error description shows different level errors appearances in different networks. The error propagation arrow describes how errors inherits between the levels and in this sense represents the causality as appeared error develops within time.

6. Examples on the known cyber events related to this proposal

Examples that support this study can be drawn from history. For example the"stuxnet"-case can be implemented in this scrutiny. The information was altered on the technical level so that the centrifuges were not functioning properly. This rendered in malfunctionalities in tactical level, the enrichment of uranium was not working and eventually in the operational level, when the enrichment plan had to be rethought and the industry complex had to be partially rebuilt with new centrifuges. On the strategic level was this played out as a strategic loss in international reliability and also the complete plan on how to use enriched uranium was postponed by a couple of years.

Another example would be the case with the DOWNAUP –virus that aggressively spread in networks using server vulnerabilities. Variants of this virus are known as Conficker and Kido (F-Secure, 2009). A similar worm is believed to be have used when the Canadian state network was attacked and users were denied access to their computers while the attacker tried to download information from the network (Austen, 2011). These kinds of attacks are on the information network level and act directly on the operational aspect accordingly to the hierarchy in the information pyramid. Separate persons or groups can still do their work or thing but they cannot interact and cooperate with the rest of the community or organization.

7. Conclusions

We have demonstrated the usability of hierarchic and previously mentioned categorization of CW events. We began by considering theoretically information at different hierarchic levels but interpreted in bits as it is its fundamental form. We proceeded by proposing functional forms of CW-disturbed information at different levels and presented schematic examples of effects of small errors. However, we liked to consider also the hierarchy of organization theory taken into account in the matter of CW attacks.

Since error of information cumulates when proceeding upward in the hierarchy it seemed logical to finally combine these two approaches in order to propose dimensions for cyber events. By combining these two approaches, the variety and categorization of errors in information is clarified. The systematic modelling of information errors supports the exploitation of vulnerabilities and causalities between different CW effects. Theoretical structure in terms of bits would provide more comprehensive description of the CW actions and also support the estimation of the effectiveness of CW. Since both descriptions include dimensional aspects, the proposed methodology would provide different impacts for defining dimensions of the CW. Estimates of error ratios or absolute error values can be given provided that functional forms of mathematical formulas are known.

Further studies might be focused on developing criteria for the mathematical formula in order to establish critical values of bit error rates. However, the level of the bit error rates does not include the consideration of the quality of interfered information and this should be elaborated. The studies of formal CW dimensions would also be a subject of interest.

References

Allen P. and Gilbert D. in Czosseck C. (ed.) and Geers K. (ed.) (2009) “The Virtual Battlefield Perspectives on Cyber Warfare”, Amsterdam, IOS Press. Pp 132-142.

Austen, Ian. (2011) “*Canada Hit by Cyberattack*”. The New York Times, http://www.nytimes.com/2011/02/18/world/americas/18canada.html?_r=3. Last read 27 January, 2013.

Beynon-Davies, Paul. (2002) “*Information Systems, An Introduction to Informatics in Organisations*”,Hampshire : Palgrave, ISBN: 0-333-96390-3.

Box G. E. P. and Draper N. R. (1987) “*Empirical Model-Building and Response Surfaces*”, Wilay, New York. pp 74 and 424.

Cutts A. “*Warfare and the Continuum of Cyber Risks: A Policy Perspective*.” in Czosseck C. (ed.) and Geers K. (ed.) (2009) “The Virtual Battlefield Perspectives on Cyber Warfare”, IOS Press, Amsterdam p. 67.

F-Secure, http://www.f-secure.com/v-descs/worm_w32_downadup_a.shtml, Last read 27 Jan 2013.

Järviö, P. and Hämäläinen, J. (2012) “A Hierachical System Approach to Electromagnetic Threats”, *EUROEM 2012 Book of Abstracts”* p.127, Toulouse, France.

Kantola H. (2011) “*Datanätverksattacker, trend eller nödvändighet*? – Ur ett småstatsperspektiv”. Finnish National Defense University, Helsinki.

Kott, Alexander (editor) (2007) “*Information Warfare and Organizational Decision-Making*”, Artech House, Inc, Norwood, Maine.

Shannon, C., E.,(1948) “A Mathematical Theory of Communication” *The Bell System Technical Journal,* Vol. 27, pp. 379-423 and 623-656, 1948.

Resolving Terminology Heterogeneity in Digital Forensics Using the Web

Nickson Karie[1, 2] **and Hein Venter**[1]
[1]**Department of Computer Science, University of Pretoria, South Africa**
[2]**Department of Computer Science, Kabarak University, Kenya**
menza06@hotmail.com
hventer@cs.up.ac.za[†]

Abstract: Frequency generators are devices that are widely used in the field of medicine in an attempt to effect chemical changes in the human body for the purpose of curing certain diseases. In addition, it is believed by many medical practitioners and researchers that frequencies can be generated, using sympathetic resonance to stimulate organ function or even to physically vibrate offending bacteria, viruses and parasites, resulting in their elimination from the human body. Digital forensics, which is rather a relatively new branch of forensic science, has attracted a wider array of people, such as computer professionals, law enforcement agencies and practitioners, that always need to cooperate in this profession. However, this has inflicted new problems with terminology heterogeneity within the domain, analogous to the offending bacteria, viruses and parasites in the human body, which needs to be resolved and/or eliminated. This paper therefore, proposes a new method that uses the Web in an attempt to resolve terminology heterogeneity in the digital forensic domain. The proposed method is based on 'terminology frequency' which is automatically generated by the Web each time new information is added. We refer to this frequency as the Web Terminology Frequency (WTF) in this paper. Web search engines are however employed as tools that have the ability to capture the terminology frequencies. Finally, we show how the computed WTF is used to deduce a method coined as the 'Terminology Heterogeneity Resolver' (TE-HERE) in this paper. Experiments conducted using the proposed TE-HERE method focuses on the digital forensic domain terminologies. However, in the authors' opinion, the TE-HERE which is technically simple and does not require any human-annotated knowledge can as well be applied in other domains. This is because the findings presented in this paper indicate that the TE-HERE method produce influential results. TE-HERE is a novel approach to resolving terminology heterogeneity in digital forensics and constitutes the main contribution of this paper

Keywords: digital forensics, digital forensic terminologies, terminology heterogeneity, web terminology frequency, terminology heterogeneity resolver (TE-HERE), web search engines

1. Introduction

The Web which dates back to 1989 (Gribble 2012, CERN 2012, WWW Foundation 2012) has grown to be such a vast entity where an astronomical amount of information is amassed. In addition, it is the largest semantic electronic database in the world. This "database" is available to all and can be queried using any Web search engine that can return aggregate hit count estimates for a large range of search queries (Cilibrasi and Vitànyi, 2007). New information is also added to the Web on a daily basis. To tap into this rich bank of information, Web search engines are the most frequently used tools to query for information related to a particular term (Karie and Venter, 2012). To the authors' knowledge, there is so far no better or easier way to search for information on the World Wide Web than simply using Web search engines like Google. However, we do not dispute the existence of other techniques that can be used to search for and extract information from the Web. Therefore, in this paper we use the Web as a live and active electronic text corpus in an attempt to resolve terminology heterogeneity in digital forensics.

Heterogeneity and specifically, semantic heterogeneity is a problem that is not well understood in many domains. In addition, there is not even an agreement regarding a clear definition of this problem (Sheth and Larson, 1990). However, according to (Merriam-Webster Dictionary, 2012), heterogeneity refers to the state of being heterogeneous. Anything that is heterogeneous lacks uniformity. Heterogeneity can occur in a domain, for example, when the people involved have differences in perceptions. The use of different terminologies to describe exactly the same thing/object in digital forensics for example, causes terminology heterogeneity.

As a matter of fact, in the case where different digital forensics experts have to testify in a court of law, if evidence presentation is uniformly and correctly done using uniform terminologies, it is much more useful in apprehending criminals, and stands a much greater chance of being admissible in the event of a prosecution. Lack of uniformity in the usage of terminologies to communicate exactly the same evidence that convicts criminals might create loopholes for the attackers to evade responsibility. Therefore, resolving terminology heterogeneity can help in the above mentioned problems. This can further help create uniformity in communication, understanding, presentation and interpretation of domain knowledge and information.

As for the remaining part of this paper, section 2 discusses related work. In section 3 some technical background is explained, followed by a discussion of the proposed TE-HERE method in section 4. Experimental results are considered in Section 5, while conclusions are drawn in section 6 and mention is made of future research work

2. Related work

Heterogeneity problems have been widely addressed at different levels by different researchers (Prasenjit and Gio, 2002). Various methods and models for resolving terminology heterogeneity have also been proposed. However, much of these methods and models are found in the domain of information sharing and particularly in interoperating databases (Xu and Lee, 2002). Existing research have also managed to identify different types of heterogeneity some of which include: structural semantic heterogeneity discussed by (Colomb, 1997) and schematic heterogeneity discussed by (Bishr, 1998). The biggest problem as discussed by (Colomb, 1997) lies in what can be called the fundamental conceptual heterogeneity. Fundamental conceptual heterogeneity occur when the terms used in two different ontologies have meanings that are similar, yet not quite the same (Sheth and Larson, 1990).

The concept of resolve terminology heterogeneity in ontologies is discussed by (Prasenjit and Gio, 2002). They propose a method that looks up in a dictionary or semantic network like WordNet (Miller, 1995) to determine similarities of words based on word similarity compound from a domain-specific corpus of documents. Their method however, focuses more on the use of information theory and hierarchical taxonomy such as the WordNet to resolve terminology heterogeneity in ontologies while in this paper we use the Web as a live and active text corpus (Xu et al, 2011) to resolve terminology heterogeneity in digital forensics.

In another paper (Sansonnet and Valencia, 2005) propose a method to solve semantic heterogeneity between software agents in an open world with a formalism for knowledge representation based on simplicial complexes. They propose an algorithm for solving terminological heterogeneity between knowledge bases formalized with the generalized simplicial notations. They further introduce the notion of an approximated mapping between heterogeneous terms. However, their work focuses more on software agents and knowledge bases while this paper focuses on digital forensic terminologies.

In their paper (Larab and Benharkat, 1996) introduces a schema integration method for federated databases based on a terminological reasoning approach. There method deals with the integration of terminologies that translate the export schemas (parts of DB^2 schemas which participate to the federation). However, the main focus of their method is in schema integration for federated systems based on a terminological reasoning approach while the current paper focuses on resolving terminology heterogeneity in digital forensics based on WTF.

In a paper by Serafini and Tamilin (2007) they address the problem of reasoning with instances of heterogeneously formalized ontologies. They build their approach upon the capability of mappings to enforce a propagation of concept membership assertions between ontologies. Their approach is grounded on a distributed description logic framework, which encodes ontologies as description logic knowledge bases and mappings as bridge rules and individual correspondences. They focus more on reasoning with instances of heterogeneously formalized ontologies while the current paper focuses on terminological heterogeneity in digital forensics.

In another paper Hakimpour and Geppert (2001) presents an approach to integrate schemas from different communities, where each such community is using its own ontology. Their approach focuses on merging ontologies based on similarity relations among concepts of different ontologies. They further present formal definitions of similarity relations based on intensional definitions. However, their approach concentrates on schema integration from different communities while we focus on terminology heterogeneity in digital forensics.

Another effort by Muresan et al, (2003) presents a two-step approach which they use to build a terminological database. Their work addresses obstacles to understanding results across heterogeneous databases brought about by the lack of ability to determine conceptual connections between differing terminologies. However, the problem with this approach is that it demands the construction of a terminological database while in our approach we use the Web which is considered the largest semantic electronic database in the world (Cilibrasi and Vitànyi, 2007).

There also exist other related works on resolving terminology heterogeneity; however, neither those nor the cited references in this paper have attempted to use WTF to resolve terminology heterogeneity in the way that is introduced in this paper. Our approach uses the Web and the Web search engines to resolve terminology heterogeneity in digital forensics. However we acknowledge the fact that the previous work on terminology heterogeneity has offered useful insights toward the development of the TE-HERE method in this paper. In the section that follows we present a detailed explanation of the technical background to aid in the understanding of the TE-HERE method discussed later in section 4.

3. Technical background

Much of the theory explained in this paper is based on the WTF. Thinking of WTF for the first time can be very interesting. However, the theory of frequency has been used in many different domains for various purposes. Frequency generators for example, are widely used in the field of medicine in an attempt to effect chemical changes in the human body for the purpose of curing certain diseases (Frequencyrising, 2012).

Moreover, according to Kilgarriff (1997) frequency lists exist and are very useful representations of meaning for information retrieval, text categorisation, and numerous other purposes. In addition, frequency lists act as a representation of the full text which is susceptible to automatic objective manipulation. However, the full text which is very rich in information cannot be readily used to make for example, similarity judgments (Kilgarriff, 1997). Therefore, WTF is introduced in this paper in an attempt to resolve terminology heterogeneity in digital forensics.

According to (Cilibrasi and Vitànyi, 2007) Google events capture all background knowledge about the search terms concerned available on the Web. The Google event *x*, consists of a set of all Web pages containing one or more occurrences of the search term *x*. Thus, it embodies, in every possible sense, all direct context in which *x* occurs on the Web. This constitutes the Google semantics of the term *x*. For this reason, in all our experiments, the Google search engine was used.

Hit counts reported by Web search engines for a specified terminology *x* are useful information sources for this study and, as such, are used as input for computing the WTF. The hit counts of a search term query *x* is defined as the estimated number of Web pages containing the queried term *x* as reported by a Web search engine (Bollegala et al, 2011). However, the hit count may not necessarily be equal to the exact terminology frequency, because the queried term *x* may appear many times on a single Web page. However, for the purpose of this study, assuming that each Web page returned by the Web search engine contains a single instance of the searched term *x,* then it becomes clear that the total number of web pages (hit counts) reported by a search engine, can be equated to the estimated terminology frequency of the term *x* on the Web.

Note that, frequency is a general term used in different fields to define the number of occurrences of a repeating event *x* per unit time *t* (Bakshi et al, 2008). Calculating the frequency of any repeating event *x* can be accomplished by counting the number of times that event *x* occurs within a specific time period *t*, then dividing the count by the length of the time period as shown in equation 1.

$$Frequency\ (f) = \left(\frac{\textbf{Number of times an event } x \textbf{ occurs}}{\textbf{Length of the time period}(t)}\right) \quad (1)$$

Using equation 1 to compute the WTF, we replace the number of times an event x occurs with the hit counts reported by a Web search engine for the specified search term *x.* On the other hand, the length of the time period *t* is replaced by the time (in milliseconds) taken by the Web search engine to successfully execute the search query for the specified search term *x*. Substituting these values in equation 1 gives equation 2.

$$Frequency\ (f) = \left(\frac{\textbf{Hit counts for the specified search term } x}{\textbf{Length of the time period } (t) \textbf{ in milliseconds}}\right) \qquad (2)$$

The result obtained from equation 2 therefore gives a generalized WTF of the search term *x* with respect to the time (in milliseconds) taken by the Web search engine to successfully execute the search query for the specified search term *x.* Re-writing equation 2 in the context of WTF gives equation 3.

$$WTF = \left(\frac{\textbf{Hit counts for the specified search term } x}{\textbf{Length of the time period } (t) \textbf{ in milliseconds}}\right) \qquad (3)$$

However, for the purpose of this study, the following notations are adopted.

f(x) denotes the estimated hit counts returned by a Web search engine for any specified search term *x* on the Web.
f(t) denotes the time period *t* in milliseconds taken by the Web search engine to successfully execute a search query for the given search term *x*.

Replacing these notations in equation 3, gives equation 4.

$$WTF = \left(\frac{f(x)}{f(t)}\right) \qquad (4)$$

Equation 4 therefore is used in this study as the basis for computing the WTF for any given digital forensics terminology *x* on the Web. The proposed TE-HERE method also utilizes the computed WTF in resolving terminology heterogeneity and is explained in the section that follows.

4. The proposed TE-HERE method

Cooperation among different people working in a domain is inevitable and often involves handling of information from diverse sources, analysing it and further presenting it to other stakeholders. Information sources however, cannot be predetermined because they may be autonomously created and maintained (Prasenjit and Gio, 2002). The owners of the information sources at times may as well prefer to maintain their autonomy.

However, effective cooperation among different people in any domain presupposes that information from different sources should be harmonised in such a way as to create uniformity and common understanding in the domain. The harmonisation process however, can be very costly and close to impossible if done manually; especially when the people involved have differences in background and perceptions on the meanings and usage of certain domain terminologies.

Therefore, we propose in this paper a method coined as the **T**erminology **He**terogeneity **Re**solver (TE-HERE) which utilizes WTFs to resolve terminological heterogeneity.

With reference to equation 4, many of the Web search engines deliver search results within fractions of a second (milliseconds). In addition, the search query execution time even for the same search term *x* using different search engines might in many cases be different. Some search engine deliver search results faster than others. Therefore, to eliminate on these variations in the search query execution time, the authors introduce an assumed search query execution time of one second (1000 milliseconds) across all search engines. Thus, for computing the WTF, the search query execution time used is one second irrespective of the search engine used. This is done by multiplying the result of equation 4 by 1000 milliseconds equivalent to 1 second.

As a concrete example, let the specified search term *x* be 'Digital Evidence' using the Google search engine as of 14th June 2012, the search query executed returned 103,000,000 hits in 0.16 seconds (160 milliseconds). Using equation 4 to compute the WTF gives a value of **643,750**. However, assuming an execution time of one second, the results of equation 4 is further multiplied by 1000 as shown in equation 5.

$$WTF = \left(\frac{f(x)}{f(t)}\right) * 1000 \qquad (5)$$

From equation 5, the WTF changes to **643,750,000**. For this reason infer from this calculation that irrespective of the search engine used for the specified search term *x,* if all other factors remain constant, then the computed WTF would be **643,750,000**. Equation 5 is therefore used in this study to define the proposed TE-HERE method and can be re-written as equation 6.

$$\text{TE-HERE} = \left(\frac{f(x)}{f(t)}\right) * 1000 \qquad (6)$$

Equation 6 therefore, defines the TE-HERE, a new method for resolving terminology heterogeneity in digital forensics using the Web and Web search engines. The experimental results obtained using the proposed TE-HERE method was found to be influential and are discussed in the section that follows.

5. Experimental results

As mentioned earlier, the theory of frequencies has been used in different domains for different purposes. However, in this study, WTF is introduced and used to resolve terminology heterogeneity in digital forensics. Terminologies which produce the highest TE-HERE values are deemed to have the highest frequency. Therefore, it should be possible to use the realized frequency to influence domain members to adopt the use of certain terminologies during communication and presentations of domain knowledge and information. This also implies that, terminologies with the highest TE-HERE values can be adopted for use, for example, in a court of law or civil proceedings where uniformity in the understanding and interpretation of evidence information by all stakeholders is a priority.

While the theory discussed in this paper is rather intricate, the resulting method is simple enough. Knowing that there exists terminology heterogeneity in digital forensics, the computed TE-HERE values of the terms in question as defined by equation 6 can be used as a quick guide to resolve the terminology heterogeneity.

Given any two digital forensic terms, for example, x_1 and x_2 used to refer to the same thing; we find the number of hit counts for the terms, denoted as $f(x_1)$ and $f(x_2)$. We also note the time period in milliseconds $f(t_1)$ and $f(t_2)$ taken by the Web search engine to successfully execute the search term queries.

As a concrete example, let the search term x_1 be ‘Digital evidence’ and search the term x_2 be ‘Electronic evidence’. Using the Google search engine, the hit counts reported for the term x_1 and x_2 as on 14 June 2012, it follows that:

“Digital evidence” $f(x_1)$ =103000000 and $f(t_1)$ = 160 milliseconds (0.16 sec)
“Electronic evidence” $f(x_2)$ =34900000 and $f(t_2)$ = 330 milliseconds (0.33 sec)

Substituting these values in equation 6 gives the following TE-HERE values ‘Digital Evidence’ = 643,750,000 and ‘Electronic evidence’ = 105,757,575.8. Since Digital Evidence has the highest TE-HERE value, it simply means that it has the highest frequency of usage and thus preferred by many as compared to Electronic evidence. It can also mean that, in case of a digital forensics investigation, the term ‘Digital Evidence’ can be used in the place of ‘Electronic Evidence’ without misleading the receivers of such information.

To further analyse the performance of the proposed TE-HERE method, we conducted two sets of experiments. First we use a data set proposed by Prasenjit and Gio (2002). Secondly, the proposed TE-HERE method is tested using digital forensics domain terms to measure its performance against selected terms. These two experiments are discussed in the two sub-sections that follow respectively.

5.1 The Mitra and Wiederhold data set

To test the performance of the proposed TE-HERE method, we use a data set proposed by Prasenjit and Gio (2002) in Table 1. This data set was used in two different ontologies. However, an analysis of the different terms showed that they were used to refer to the same thing. The input to the TE-HERE method is therefore the reported Google hit counts *f*(*x*) and the time period *f*(*t*) for any of the terms in question. The TE-HERE method works by computing the TE-HERE value for each of the terms (see Table 1) using equation 6. Given any of the terms pairs x_1 and x_2, the associated computed TE-HERE values determine their frequency as seen from table 1.

Table 1: Mitra and Wiederhold data set, *Original source:* (Prasenjit and Gio 2002*).*

Term (x_1)	TE-HERE Values	Term (x_2)	TE-HERE Values
Passenger	885185185.2	Passenger	885185185.2
Cargo	1920833333	Payload	92727272.73
Departure Time	24565217.39	Time	38066666667
Arrival Time	38391304.35	Time	38066666667
Arrival City	9148148.148	Destination	2495652174
Name	27242424242	Location Name	25250000
Departure City	17842105.26	Origin	3052000000
Airport	2637931034	Airforce Base	2182352.941
Flight	2496296296	Sortie	665217391.3

The TE-HERE Values shown in Table 2 depicts the estimated frequency of the digital forensic terms in question. Table 1 on the other hand, was used mainly for the purpose of testing the proposed TE-HERE method with terminologies not necessarily originating from digital forensics. This was done to provide a clear picture of the performance and accuracy of the TE-HERE method.

From Table I, column 1 and 3 shows the terms used while column 2 and 4 indicate the TE-HERE Values computed using equation 6. For example, the terms ‘Cargo’ and ‘Payload’ (See Table 1) with TE-HERE Values of 1,920,833,333 and 92,727,272.73 respectively, depicts the performance of the TE-HERE method. In this case, because Cargo has the highest TE-HERE value, it can be adopted in place of Payload without distorting information. This is also depicted in the other columns of Table 1. From this experiment, it is clear that using the TE-HERE method; we can generate satisfactory terminology frequency results. However, it is also possible to get false positives; though in a significant number of cases the TE-HERE values were satisfactory. Therefore, we suggest that for any application, if the results are not satisfactory the domain expert can decide based on the computed TE-HERE Values which of the terms in question can be adopted for use. For example, in the case of ‘Departure Time’ and ‘Time’ (see Table 1) a decision can be made to decide on which of the terms can be adopted for use.

5.2 Digital forensics terminologies

In Table 2, a part of the experimental findings is presented using selected digital forensics domain terms. Each term enclosed in double quotes is used as a single Google search term with the reported hit counts denoted in Table 2 as *f*(*x*) and the associated execution time period denoted as *f*(*t*). The computed TE-HERE value shows the measures obtained to ascertain the performance of the TE-HERE method. The selected digital forensics terms used are: 'Digital Evidence', 'Electronic evidence', 'Multimedia evidence' and 'Digital and multimedia evidence’.

The authors found that these terms are mostly used in discussions that involve the digital forensics investigation and evidence presentations in either a court of law or civil proceedings, hence the motivation for the experiment indicated in Table 2. In all the experiments conducted, the TE-HERE method produced influential results as can be seen from Table 1 and 2 respectively.

To determine the TE-HERE value of the search terms in Table 2, the proposed TE-HERE method was used. The first column of the table show the sampled digital forensics terminologies and their corresponding hit counts values *f(x)* shown in column two, the time taken in milliseconds by the search engine to execute the search query *f(t)* is shown in column three and finally the TE-HERE values are presented in the last column.

In the case of a need to resolve terminology heterogeneity in the digital forensics domain and/or any other domain for example, the proposed TE-HERE method can be used. Since the term ‘digital evidence’ has the highest TE-HERE value of **643,750,000** it simply means that, ‘digital evidence’ is widely used and probably preferred by many stakeholders.

Therefore, this value taken as the global estimated terminology frequency can be used to influence other members to adopt the use of the term ' digital evidence ' in all communications and presentations involving digital forensic evidence either in a court of law or civil proceedings. This will as well create uniformity in communication, understanding and interpretation of domain knowledge and information.

Table 2: Experimental findings of digital forensic terminologies using the TE-HERE method

Digital Forensics Terminologies	Hit Counts *f(x)*	Time in milliseconds *f(t)*	TE-HERE Values
Digital and multimedia evidence	17300000	240	72083333.33
Multimedia evidence	61400000	210	292380952.4
Digital Evidence	103000000	160	643750000
Electronic evidence	34900000	330	105757575.8

The performance of the proposed TE-HERE method is further backed up by a graphical representation of the Table 2 results shown in Figure 1.

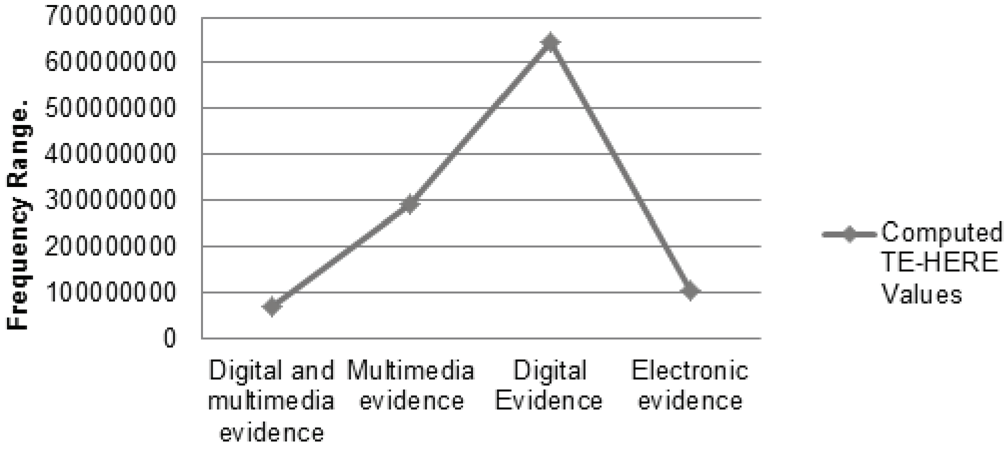


Figure 1: Graphical representation of the TE-HERE values from Table 1.

In this paper, we have adopted a relatively technical simple and new method for resolving terminology heterogeneity in digital forensics coined as the “Terminology Heterogeneity Resolver” (TE-HERE). This method uses the Web as a live and active text corpus, comprising automatic generated WTF. In addition, Web search engines are employed in this study as tools that have the ability to capture the terminology frequency.

To the best of the authors' knowledge, there exists no other experiments in digital forensics similar to the one explained in this paper, that can be used as a baseline to judge the performance of the proposed TE-HERE method thus found it hard to benchmark the measures produced by the TE-HERE method for the specified digital forensic terminologies. This is, therefore, a novel approach of using the Web and the Web search engines to resolve terminology heterogeneity in digital forensics.

The advantage of using the TE-HERE method is that, there is so much of the data involved (The entire Web) thus; TE-HERE values can be generated with respect to the thousands of data points available on the Web. The TE-HERE method is also economical and technically simple because it utilises the freely available information on Web and the Web search engines. In the case of digital forensics, for example, the TE-HERE method can be used to influence members to adopt the usage of certain domain terminologies.

6. Conclusion

The problem addressed in this paper was that of terminology heterogeneity in the digital forensics domain where different stakeholders use different terminologies to describe or refer to exactly the same thing/object, resulting in the lack of uniformity in communication, understanding and interpretation of domain knowledge and information. Knowing that digital forensics is relatively a new field, addressing terminology heterogeneity at an early stage can help resolve both the present and future terminology heterogeneity problems, thereby creating a lasting uniformity in the domain.

Finally, the proposed TE-HERE method was found to generate influential results making it a method to consider as a quick guide for resolving terminology heterogeneity. This is backed up by the fact that the TE-HERE method is technically simple and economical. Though the initial experiments done were based on the digital forensics domain terminologies, the authors believe that the TE-HERE method can be applied in other domains as well. As part of the future work, the authors are now planning to use the TE-HERE method to build an automated model for resolving terminology heterogeneity and even more as a way towards resolving semantic disparities in the digital forensics domain. However, more research needs to be done so as to improve on the performance and accuracy of the TE-HERE method.

References

Bakshi, U.A., Bakshi, A.V., and Bakshi, K.A. (2008), Time and Frequency Measurement, Technical Publications, (Electronic Measurement Systems.) First Edition:2008 pp. (4–1). [online], http://books.google.co.za/books?id=jvnI3Dar3b4C&pg=PT183&hl=en#v=onepage&q&f=false

Bishr Y.A., (1998), Overcoming the Semantics and Other Barriers to GIS Interoperability. International Journal of Geographic Information Science, Vol. 12, No. 4, pp 299-314.

Bollegala, D., Matsuo, Y. and Ishizuka, M. (2011), A Web Search Engine-Based Approach To Measure Semantic Similarity Between Words, IEEE Transactions on Knowledge and Data Engineering, Vol. 23, No. 7, pp 977-990.

CERN, (2012) How the web began [online], http://public.web.cern.ch/public/en/about/webstory-en.html

Cilibrasi, R.L. and Vitànyi, P.M.B. (2007), The Google Similarity Distance, IEEE Transactions on Knowledge and Data Engineering, Vol. 19, No 3, pp. 370–383.

Colomb, R.M. (1997) Impact of Semantic Heterogeneity on Federating Databases, The Computer Journal, Vol. 40, No. 5 pp 235-244.

Frequencyrising, (2012) Bio Frequency Generator, [online], http://www.frequencyrising.com/frequency-generator.html

Gribble, C. (2012), History of the Web Beginning at CERN, [online], Hitmill, http://www.hitmill.com/internet/web_history.html

Hakimpour, F. and Geppert, A. (2001), Resolving Semantic Heterogeneity in Schema Integration: an Ontology Based Approach, Proceedings of the of the International Conference on Formal Ontologies in Information Systems, pp 297—308.

Karie, N.M. and Venter, H.S. (2012), Measuring semantic similarity between digital forensics terminologies using web search engines, Proceedings of the Annual Information Security for South Africa (ISSA), pp.1-9,Sandton, Johannesburg

Kilgarriff, A. (1997), Using Word Frequency Lists to Measure Corpus Homogeneity and Similarity between Corpora. Proceedings of the AISB Workshop, Falmer.

Larab, O. and Benharkat, A., (1996), Resolving Semantic Heterogeneity in Databases with a Terminological Model: Correspondence Refinement. Proceedings of the International Workshop on Description Logics, Cambridge, MA, USA. pp.150-154.

Merriam-Webster Dictionary, (2012) [online], http://www.merriam-webster.com/dictionary/heterogeneity

Miller, G.A. (1995), “WordNet: A Lexical Database for English”, [online], PrincetonUniversity, http://wordnet.princeton.edu/

Muresan, S., Popper, S.D., Davis, P.T., and Klavans, J.L. (2003), Building a Terminological Database from Heterogeneous Definitional Sources. Proceedings of the annual national conference on Digital government research 2003.

Prasenjit, M. and Gio, W. (2002), Resolving terminological heterogeneity in ontologies, Proceedings of the ECAI-02 Workshop on Ontologies and Semantic Interoperability. Lyon, France, July.

Sansonnet, P. and Valencia, E. (2005), Terminological Heterogeneity Between Agents Using a Generalized Simplicial Representation, EUMAS 2005- Proceedings of the Third European Workshop on Multi-Agent Systems, Brussels, Belgium, pp 363-374.

Serafini, L. and Tamilin, A., (2007), Reasoning with Instances of Heterogeneous Ontologies. [online], Data & Knowledge Management Group, http://ceur-ws.org/Vol-314/52.pdf

Sheth, A.P. and Larson, J. (1990), Federated database systems for managing distributed Heterogeneous and autonomous databases. ACM computer Surveys, Vol 22 No. 3, pp 183 – 236.

WWW Foundation, (2012), History of the Web, http://www.webfoundation.org/vision/history-of-the-web/

Xu, Z. and Lee, Y.C. (2002), Semantic heterogeneity of geodata, Proceedings of the Symposium on geospatial theory, processing and application, Ottawa.

Xu,Z., Luo, X., Yu, J. and Xu, W. (2011), Measuring semantic similarity between words by removing noise and redundancy in web snippets. Concurrency and Computation: Practice and Experience , Vol. 23 No. 18, pp 2496-2510.

Protecting Critical Infrastructure Services in the Cloud Environment

Áine MacDermott, Qi Shi, Madjid Merabti, and Kashif Kifayat
PROTECT: Research Centre for Critical Infrastructure Computer Technology and Protection, School of Computing and Mathematical Sciences, Liverpool John Moores University, UK
a.mac-dermott@2008.ljmu.ac.uk
q.shi@ljmu.ac.uk
m.merabti@ljmu.ac.uk
k.kifayat@ljmu.ac.uk

Abstract: Due to the scalability of resources and performance, as well as improved maintainability it is apparent that cloud computing will eventually reach IT services that are operating critical infrastructures. Since IT infrastructures have become an integral part of almost all organisations, cloud computing will have a significant impact on them. Protecting sensitive critical infrastructure data in the cloud environment is the explicit focus of our work. The scale and dynamic nature of cloud computing cause challenges for their management, including investigating malicious activity and/or policy failure. Sufficient security metrics needs to ensure the confidentiality, integrity, and availability of the data on the cloud. Hosting critical infrastructure services in the cloud brings with it security and resilience requirements that existing cloud services are not well placed to address. The protection of critical infrastructure and data has been particularly highlighted with the increase of advanced persistent threats being designed to specifically target the systems that control these infrastructures. There had been a belief in the industry that even if these control systems had vulnerabilities, they were not actively being targeted, but this is now believed to be unrealistic. The severity of the control system specific malware and cyber attacks has led to an increased awareness and understanding from security vendors, as well as the government, both in the United Kingdom and the United States of America. Gaining a deeper understanding of the infrastructure security needs is of utmost importance as there is currently a paradigm shift in assessing the extent of risks and protecting against zero-day vulnerabilities. To continue to offer safe and reliable services for public consumption, infrastructure providers need to recognise they have a problem and lack adequate tools, processes and experience to deal with it. It is inevitable that cloud computing will be introduced into their computing paradigm, and effective protection metrics need to be developed to deal with this revolution. Multiple intrusion detection activities have been introduced to address the issue of intrusion detection within cloud computing environments. Security, risk management, control, and trust are important issues that deter and holdback organisations from fully adopting cloud computing and profiting from its many advantages. Our research aims to develop a framework for the protection of critical infrastructure data in the cloud computing environment. We aim to develop a model that can be tailored to the different cloud environments, creating anomaly based intrusion detection techniques tailored to the specialist nature of the cloud computing environment. Our framework should provide monitoring of the network and have an improved detection efficiency based on the efforts in literature. It will be composed of network-based IDS, with a distributed architecture so that there is no single point of failure. The use of unsupervised learning methods will observe the behaviours of users within the network.

Keywords: critical infrastructure protection, security, control system, cloud computing, anomaly based intrusion detection, risk management, trust

1. Introduction

Critical infrastructures, such as the power grid and water distribution facilities, include a high number of devices over a large geographical area. These infrastructures face significant threats as the growth in the use of industrial control systems such as SCADA (supervisory control and data acquisition) systems and their integration to networks in order to coordinate and manage the actions of these devices. While this provides great capabilities for operation, control, and business, this augmented interconnectedness also increases the security risks due to cyber related vulnerabilities they possess. The importance of protecting these infrastructures has been particularly highlighted by the increase in advanced persistent threats (APTs), such as 'Stuxnet' and 'Duqu', which were designed to target these control systems and disrupt their functionality. Effective protection of SCADA systems is thus crucial, as these are important components of critical infrastructures and it is apparent that existing methods do not meet the security requirements of such interconnected infrastructures.

As more sectors adopt cloud services in their computing environment, the trend will also reach IT services operating critical infrastructure. There needs to be an assurance that the cloud computing environment can provide proficient protection of the sensitive critical infrastructure data. The focus of our research is on intrusion detection techniques for infrastructure services in the cloud computing environment. The reality of today's advanced malware and targeted attacks is that 100% protection is not realistic; reducing attack vectors and marginalising the impact of an attack is the ideal goal. Infrastructures will inevitably fully grasp the benefits being offered by cloud computing. Once their services are in the cloud environment resourceful functionality is essential. Resources of a private cloud could be used to detect attacks on a control system by monitoring the data and logs it produces.

This paper provides an overview of the protection problem faced by critical infrastructure vendors. The structure of this paper is as follows: Section 2 provides background on critical infrastructures and their vulnerabilities, cloud computing and its security issues. Section 3 details existing approaches currently attempting to resolve this protection problem. In Section 4, we detail the migration of infrastructure services to the cloud environment how our proposed method can offer provision to this. Section 5 illustrates our conclusions.

2. Background

The European Commission defines critical infrastructure as "the physical and information technology facilities, networks, services and assets that, if disrupted or destroyed, would have a serious impact on the health, safety, security or economic well-being of citizens or the effective functioning of governments in EU countries" (Commission 2006). These infrastructures are physical or mechanical processes mostly controlled electronically by systems, usually called supervisory control and data acquisition (SCADA) or process control systems (PCSs), composed of computers and interconnected by networks.

Traditionally, these infrastructures were inherently secure systems as they were largely based on dedicated communication links. Nowadays, modern critical infrastructures largely make use of IT technologies, where wireless sensor networks (WSNs) with open access have become an integral part of virtually any critical infrastructure. Examples of these infrastructures include telecommunication, energy grid, water grid, and transportation, to name a few.

SCADA is the standard approach to computer-based management of industrial processes. SCADA systems are widely used in industrial installations to control and maintain field sensors and actuators. Orders are communicated from the SCADA level downwards to production lines. Low level sensors are used to gather basic data to provide a view of the situation at the lowest level. This is further propagated upwards to the information system. Most industrial plants now employ networked process historian servers for storing process data and other interfaces.

Table 1 illustrates these levels of order.

Table 1: Levels of order

	Order	***Managed by:***
↑	Production planning and control	SCADA
	Micro controllers	Processor based devices
Low Level	Production lines	Sensors gather basic data

Concern in the industry is how many of these infrastructures are dependent upon each other for functioning (Brewer 2012). Interdependencies among computers, communication, and power infrastructures have increased security risks due to the complexity of the integrated infrastructures. Disruptions in one part of the infrastructure may spread out through the system and have cascading effects on other sectors (Ten et al. 2010). Critical infrastructure control systems may vary depending on the different environments and infrastructure. However, they possess the same basic components and characteristics, namely for supervisory control and acquisition of data. Securing control systems is difficult as the usual security assumptions and practices applied for protecting IT systems are not sufficient. The risks associated to the highlighted cascading effects that could occur have increased the support for the implementation of risk management and risk assessment policies.

2.1 Control system weaknesses

Control systems are created with different design goals compared to traditional IT systems. Each point in the SCADA network is a potential entry point. Little work has been done to enhance their security as it had been believed there was no problem prior to the increase in threats targeting these systems (Byres et al. 2004). The connectivity of SCADA networks increases the risk of cyber attacks and the critical need to improve the security of these networks. When a security breach occurs in a SCADA system, the results are often different to those on traditional IT systems.

Open communication protocols such as Modbus and DNP3 are increasingly used to achieve interoperability. Many SCADA protocols use TCP/IP (Transmission Control Protocol/Internet Protocol) and provide no additional authentication or protection. Vulnerabilities in the TCP/IP protocol include IP spoofing and man-in-the-middle attacks. Additionally, the standardisation of software and hardware used in SCADA systems potentially makes it easier to mount SCADA-specific attacks, as was evident in the case of Stuxnet (Miyachi et al. 2011).

‘Stuxnet’, discovered in June 2010, is an example of malware that was created to specifically target control systems. Stuxnet was a computer worm that attacked Iran’s uranium enrichment program. It targeted the SCADA system in the infrastructure and exploited the fact that the Sieman’s PLC (programmable logic controller) does not require authentication to upload rogue ladder logic, making it easy for the attackers to inject their malicious code into the system (Zetter 2012). ‘Duqu’ is a computer worm that was discovered in September 2011, and was related to the Stuxnet worm. It was believed to have been designed to steal sensitive data in order to launch further cyber attacks (Brewer 2012).

Detailed knowledge is required for control system attacks. These systems are highly customised, and their configuration and functionality can differ depending on their environment. As subsystems of critical infrastructure, it is of utmost importance that protection measures are in place as they are essential for production planning and control. Security measures that can be applied to the existing complex infrastructure and systems, and address the security requirements have to be developed. Due to the complexity of these systems and their differentiating environments, this has proven difficult. These systems were designed for the purpose of control and data acquisition and have low memory resources. Therefore ‘updating’ these systems is not an option. Embedded control systems are often subject to strict timing requirements when passing data in a network. Given these small timing windows, introducing a small amount of overhead could be catastrophic (Reeves et al. 2012).

2.2 Cloud computing and critical infrastructure

Hosting critical infrastructure services in the cloud brings with it security and resilience requirements that existing cloud services are not well placed to address. Due to virtually unlimited scalability of resources and performance, as well as significant improvement regarding maintainability, it is inevitable that cloud computing will be introduced into their computing paradigm. Cloud computing will eventually reach the IT services that are operating critical infrastructures (OTE 2012). Vendors are growing increasingly anxious about protecting the sensitive data their infrastructures handle daily and desire sufficient security metrics to be developed.

With the technical development and market growth in cloud computing, organisations that provide, operate and maintain IT systems for critical infrastructure are making the decision as to when they should make the computing paradigm shift. Cloud services can offer efficient access to large IT infrastructures that benefit from the economy of scale. Therefore, it would be highly desirable to maintain irrecoverable and valuable data obtained from critical infrastructure within secure cloud infrastructures.

Cloud providers usually build up large scale data centres and provide cloud users with computational resources in three delivery models (Annapureddy 2010). Software as a Service (SaaS) is a software application delivery model in which enterprises hosts and operates their applications over the internet so the customers can access it. One benefit of this model is customers do not need to buy any software licences or any additional equipment for hosting the application. Platform as a Service (PaaS) is a model that provides a platform for building and running customer applications. Enterprises can build applications without installing any tools on their local systems and can deploy them without many difficulties. Infrastructure as a Service (IaaS) provides a convenient option for organisations by migrating the IT infrastructure to the cloud provider. This means it is the responsibility of the cloud provider to tackle the issues of IT infrastructure management, such as configuring servers, routers, firewalls, and so on.

There are security issues at each level of the cloud computing paradigm. These levels are application level, virtual level and physical level. For the purpose of our analysis, we are focusing on issues on the application level (Saas) and the virtual level (PaaS and IaaS) as this is where vendors have most concerns. Physical security issues could include hardware security, hardware reliability, and network security to name a few. Table 2 presents an overview of the security requirements and threats associated with each delivery model.

Table 2: Security requirements and threats associated with each delivery model

Level	Service level	Security requirements	Threats
Application level	Software as a Service (SaaS)	Access control Communication protection Data protection from exposure Privacy in multitenant environment Service availability Software security	Data interruption Exposure in network Interception Modification of data at rest and in transit Privacy breach Session hijacking Traffic flow analysis
Virtual level	Platform as a Service (PaaS) Infrastructure as a Service (IaaS)	Access control Application security Cloud management control security Communication security Data security (Data in transit, data at rest, remanence) Secure images Virtual cloud protection	Connection flooding DDoS Defacement Disrupting communications Exposure in network Impersonation Programming flaws Session hijacking Software interruption Software modification Traffic flow analysis

Based on the problem at hand, it is evident that sufficient security metrics need to be developed for protecting the secure data being stored in the cloud environment. The ability to clearly identify, authenticate, authorise, and monitor who or what is accessing the assets of an organisation is essential to protecting an information system from threats and vulnerabilities. Intrusion detection systems (IDSs) are the current security measures protecting the service as a whole.

3. Existing protection methods

Security, risk management, control, and trust are issues that deter organisations from fully adopting cloud computing and profiting from its many advantages. To address these issues, multiple intrusion detection activities have been developed for the cloud environment. The work of (Hamad & Al-Hoby 2012) take an innovative approach to tackling this security problem. They designed and implemented the Cloud Intrusion Detection Service (CIDS), which aims to be a scalable intrusion detection framework that can be deployed by cloud providers to enable clients to subscribe with the IDS in a service-based manner. Their system consists of three separate layers: User Layer, System Layer, and Database Layer. It is a re-engineered version of Snort. The model outperforms currently used solutions for service- based IDS but at the same time provides minimal overhead to the case of traditional IDS deployment for single network protection.

In (Xin, Ting-lei, & Xiao-yu 2010) they focus on the problem of IDS scalability in the cloud environment. To overcome the limitations of using traditional IDSs a new intrusion detection mechanism based on cloud computing (IDCC) is proposed. This architecture is designed to be scalable to support wide networks and to be highly available. It is composed of four components: Local data collectors (LCs), Local analysers (LAs), Remote data collectors (RCs), and Cloud computing data centre (CCDC). Monitoring the security activity in a multi-site network is the objective of these components. It proves its ability to compact similar alerts and to correlate alerts coming from heterogeneous platforms on several sites to detect intrusions that are more complex.

(Dhage et al. 2011) conveys that when there is only one IDS in the entire network, the load on it increases as the number of hosts increases. It is difficult to keep track of different kinds of attacks or intrusions, which are acting on each of the host present in the network. In order to overcome this limitation, they propose an architecture in which mini IDS instances are deployed between each user of cloud and the cloud service provider. As a result, the load on each IDS instance will be lesser than that on single IDS and hence that small IDS instance will be able to do its work in a better way. For example, the number of packets dropped will be less due to the lesser load which single IDS instance will have.

Intrusion detection systems are one of the most popular devices for protecting cloud computing systems from various types of attack (Mahmood & Agrawal 2012; Shelke et al. 2012; Taghavi Zargar et al. 2011; OTE 2012; Annapureddy 2010; Chen et al. 2011). IDSs can observe the traffic from each virtual machine (VM) and generate alert logs and can manage cloud computing globally. A key problem is the management of logs. As IDSs generate large amounts of logs, administrators have to decide what to analyse first. In addition, it is difficult to analyse logs because communication between many systems and many consumers generates large amounts of logs. Effective log and resource management is desired, as an administrator may miss important alerts and events, thus endangering their system.

The work of (Lee et al. 2011) propose a multi-level IDS and log management method based on consumer behaviour for applying IDS effectively to the cloud system. They assign risk level to user behaviour based on analysis of their behaviour over time. Applying differentiated levels of security strength to users based upon the degree of anomaly increases the effective usage of resources. Their method proposes the classification of generated logs by anomaly level. This is so that the system administrator analyses logs of the most suspected users first.

Since cloud infrastructure have enormous network traffic, traditional IDSs are not efficient to handle such a large data flow. There needs to be a strong balance between IDS security level and system performance. Due to the large data sets, classification techniques require a huge amount of memory and CPU usage. In (Mahmood & Agrawal 2012) the focus is on techniques to reduce the number of computer resources, both memory and CPU time required to detect an attack. Feature reduction is used to remove worthless information from the original high dimensional database of cloud traffic data. A back propagation algorithm is applied on reduced cloud traffic data for classification. Their original contributions show that dimensional reduction techniques help compact similar alerts and correlate alerts coming from heterogeneous platforms on several sites to detect intrusions that are more complex.

It is important to note that the existing approaches in this area do not tackle the protection of services migrating to the cloud environment efficiently. In current solutions, they provide theoretical intrusion methods for the cloud infrastructure, or for network protection, but these are not sufficient for our protection problem. We believe the development of a secure private cloud network where the analysis of logs generated from the historian in the control system could provide efficient data processing and extraction of the system behaviour. This will overcome the challenges associated with processing the massive data sets generated by the control systems.

4. Critical infrastructure in the cloud environment

Infrastructure vendors will inevitably take advantage of the benefits cloud computing have to offer (Khorshed et al. 2012). Concerns over protecting sensitive data and services in this environment remain. Nonetheless, utilising the cloud environment is a natural extension of remote access as it removes the requirement for the user to be in the same location as the infrastructure. Remote access to critical infrastructure is already common place. We analysed the range of cloud environments and services on offer, and developed a risk assessment of what benefits and risks could be associated with deploying each service (M. Zhou et al. 2010). There is a strong relationship between cloud storage and critical infrastructure, as they are both distributed systems and may possess the same underlying issues. The natural progression of this utilisation is determining what services could be shifted from a critical infrastructure environment to a cloud environment. In addition, how the functionality of these services can be improved, and determining how the protection issues differ from the traditional critical infrastructure environment.

This perception has led us to the current progression of our work. Most industrial plants employ networked process historian servers storing process data and other possible business and process interfaces. For example, direct file transfer from programmable logic controllers (PLCs) to spreadsheets (Zhu & Sastry 2010). PLCs in the control system generate a huge amount of data and logs. Logs of communication are stored in the historian databases. These databases possess historical data that is being logged 24/7 from over 6,700 data points so that it could be easily accessible by both operators and engineers (Fovino et al. 2010; Verba 2008). These historian servers receive data from field control processors or front-end processors, which issue control commands to and poll data from devices in field networks.

The control network typically contains assets such as the human-machine interface (HMI) and other workstations, which run control system applications on conventional computer platforms. The field network devices directly monitor and control a physical process, such as refining, manufacturing, or electric power generation/transmission/distribution (Briesemeister et al. 2010). Figure 1 illustrates the arrangement of critical infrastructure components and communication links.

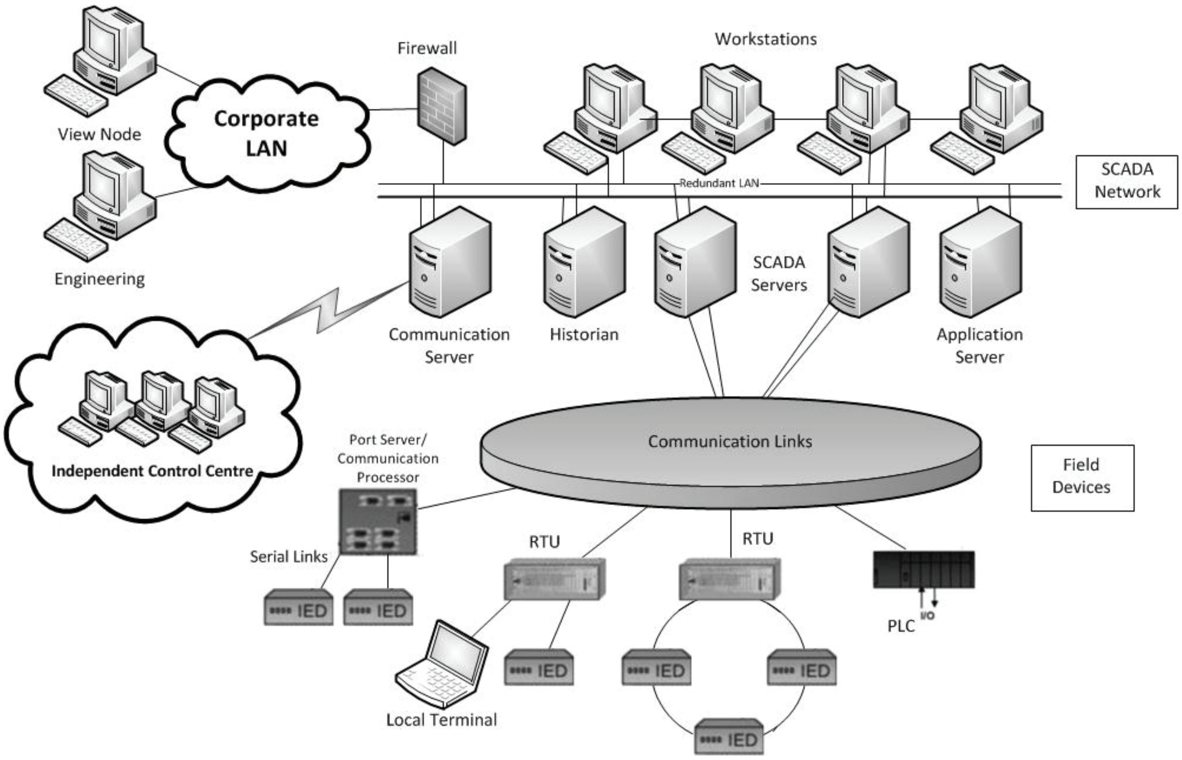


Figure 1: Critical infrastructure components

We propose, in order for critical infrastructure to utilise the cloud environment, is that if the historian database sent this information to a Private Cloud. If critical infrastructure migrates to the cloud, the auditing infrastructure must also be moved to the cloud. The use of a private cloud to audit the data from the system and process it more effectively would be valuable. This will overcome the challenges associated with processing the massive data sets generated by the control systems. The cloud environment is suitable as it has massive storage and computational capabilities, which would be beneficial for the processing of large amounts of data. Private clouds grant complete control over how data is managed and what security measures are in place. To utilise this environment with critical infrastructure services we will need to test our premise against historian data and determine how we can improve upon the processing of the produced logs. The private cloud environment is distributed and elastic, which can offer improved processing rates and efficiency compared to current methods.

This is two-folds though. We could also use this collection of data to perform behavioural analysis and modelling of this information flow. Looking for trends and subtle changes in the data would be beneficial in achieving state awareness. Behaviour modelling can take place without affecting the system in any way. We would aim to avoid the real time detection, as we would be looking for the effect of attacks rather than the cause, i.e. long term effects of attacks as they can slowly modify the system over time, and this behavioural analysis technique would be a great way of overcoming these problems. By monitoring the evolution of the plant process states, and tracking down when the industrial process is entering into a critical state, it would be possible to detect these attack patterns (known or unknown) aiming at putting the process system into a known critical state by using state of commands. In control system architectures, the major cyber attack vector is the flow of network commands (A. Carcano et al. 2003).

Figure 2 represents how our proposed method may be applied to the current infrastructure illustrated in Figure 1.

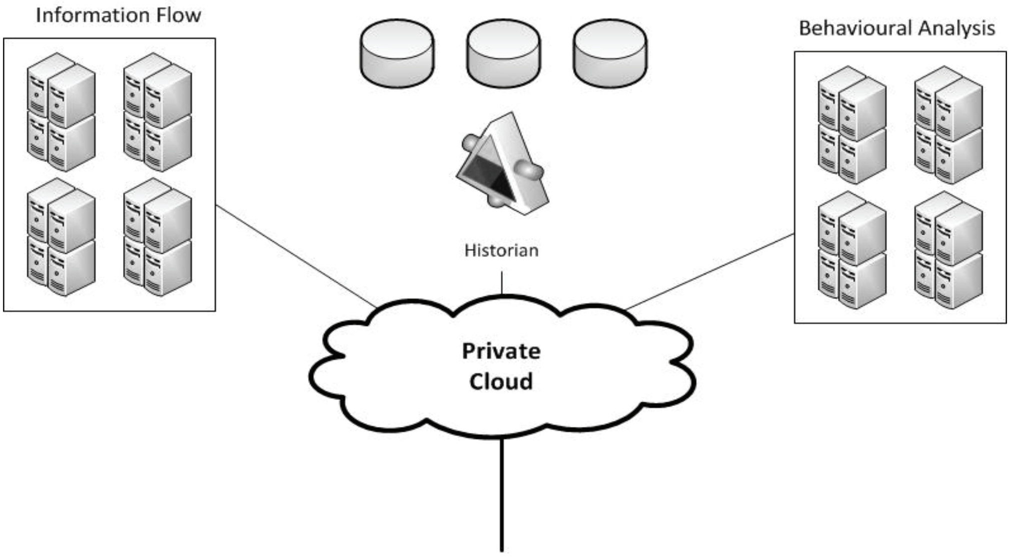


Figure 2: Proposed private cloud attributes

By processing the sensor data from the historian in a private cloud environment, we can analyse the behaviour of the infrastructure and use the critical state metric as a trigger for logging a chain of packets. It is possible to discriminate between critical states due to cyber attacks and critical states due to faults/physical attacks. We aim to build a reference behaviour model, with the use of Bayesian classification procedure associated to unsupervised learning algorithms to evaluate the deviation between current and reference behaviour. A reference audit data set representing the normal system behaviour will be used to create the model (Scott 2004). By training with unsupervised learning algorithms, namely Bayesian classification, the log analysers perform well in discovering the inherent nature of the dataset and clustering similar instances into classes.

5. Conclusions

Critical infrastructures have become the central nervous system of the economy in all countries. Their failure has a high socioeconomic impact, as these infrastructures are dependent upon each other. Disruption in a single sector is likely to have cascading effects on other sectors. The need for critical infrastructure protection has been emphasised with the increase of threats being designed specifically targeting the systems that control these infrastructures. With the technical development and growth, it is clear that cloud computing will eventually reach the IT services that are operating critical infrastructure. In cloud environments, network perimeters will no longer exist from a cloud user's perspective, which render traditional security protection methods such as firewalls not applicable to cloud applications. There has been increased interest in developing security mechanisms and intrusion detection systems that will protect the services and sensitive data they possess. This report has shown our plans to build upon our initial research into protecting critical infrastructure services in the cloud environment. By using private cloud environments and behavioural analysis, our work can achieve this, and tackle the scalability issues and log management issues associated with current proposals.

References

Annapureddy, K., 2010. Security Challenges in Hybrid Cloud Infrastructures. In *Aalto University, T-110.5290 Seminar on Network Security*.

Brewer, R., 2012. Protecting critical control systems. *Network Security*, 2012(3), pp.7–10.

Briesemeister, L. et al., 2010. Detection, correlation, and visualization of attacks against critical infrastructure systems. In *2010 IEEE Eighth Annual International Conference on Privacy Security and Trust (PST)*. pp. 15–22.

Byres, E.J., Franz, M. & Miller, D., 2004. The Use of Attack Trees in Assessing Vulnerabilities in SCADA Systems. In *IEEE International Infrastructure Survivability Workshop (IISW'04)*. Lisbon, Portugal.

Carcano, A. et al., 2003. A Multidimensional Critical State Analysis for Detecting Intrusions in SCADA Systems. *IEEE Transactions on Industrial Informatics*, 7(2), pp.179–186.

Chen, S., Nepal, S. & Liu, R., 2011. Secure Connectivity for Intra-cloud and Inter-cloud Communication. *2011 40th International Conference on Parallel Processing Workshops*, pp.154–159.

Commission, E., 2006. European Programme for Critical Infrastructure Protection.

Dhage, S.N. et al., 2011. Intrusion detection system in cloud computing environment. *Proceedings of the International Conference & Workshop on Emerging Trends in Technology - ICWET '11*, pp.235–239.

Fovino, I.N. et al., 2010. Modbus/DNP3 State-Based Intrusion Detection System. *2010 24th IEEE International Conference on Advanced Information Networking and Applications*, pp.729–736.

Hamad, H. & Al-Hoby, M., 2012. Managing Intrusion Detection as a Service in Cloud Networks. *International Journal of Computer Applications*, 41(1), pp.35–40.

Khorshed, M.T., Ali, a. B.M.S. & Wasimi, S. a., 2012. A survey on gaps, threat remediation challenges and some thoughts for proactive attack detection in cloud computing. *Future Generation Computer Systems*, 28(6), pp.833–851.

Lee, J., Park, M. & Eom, J., 2011. Multi-level Intrusion Detection System and log management in Cloud Computing. *2011 13th International Conference on Advanced Communication Technology (ICACT)*, (1), pp.552–555.

Mahmood, Z. & Agrawal, C., 2012. Intrusion Detection in Cloud Computing environment using Neural Network. *International Journal of Research in Computer Engineering and Electronics*, 1(1), pp.1–4.

Miyachi, T. et al., 2011. Myth and reality on control system security revealed by Stuxnet. In *2011 Proceedings of SICE Annual Conference (SICE)*. Tokyo, Japan: IEEE, pp. 1537–1540.

OTE, 2012. *Discussion on the Challenges for the Development of a Context for: Secure Cloud Computing for Critical infrastructure IT*.

Reeves, J. et al., 2012. Intrusion detection for resource-constrained embedded control systems in the power grid. *International Journal of Critical Infrastructure Protection*, 5(2), pp.74–83.

Scott, S.L., 2004. A Bayesian paradigm for designing intrusion detection systems. *Computational Statistics & Data Analysis*, 45(1), pp.69–83.

Shelke, M.P.K., Sontakke, M.S. & Gawande, A.D., 2012. Intrusion Detection System for Cloud Computing. *International Journal of Scientific & Technology Research*, 1(4), pp.67–71.

Taghavi Zargar, S., Takabi, H. & Joshi, J., 2011. DCDIDP: A Distributed, Collaborative, and Data-driven Intrusion Detection and Prevention Framework for Cloud Computing Environments. *Proceedings of the 7th International Conference on Collaborative Computing: Networking, Applications and Worksharing*, pp.332–341.

Ten, C.W., Manimaran, G. & Liu, C.C., 2010. Cybersecurity for critical infrastructures: attack and defense modeling. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans,* 40(4), pp.853–865.

Verba, J., 2008. Idaho national laboratory supervisory control and data acquisition intrusion detection system (SCADA IDS). *International Conference on Technologies for Homeland Security (HST)*, (208), pp.469–473.

Xin, W., Ting-lei, H. & Xiao-yu, L., 2010. Research on the Intrusion detection mechanism based on cloud computing. *Intelligent Computing and Intelligent Systems (ICIS)*, pp.125–128.

Zetter, K., 2012. Researchers Release New Exploits to Hijack Critical Infrastructure. *Wired.com*. Available at: http://www.wired.com/threatlevel/2012/04/exploit-for-quantum-plc/ [Accessed April 10, 2012].

Zhou, M. et al., 2010. Security and Privacy in Cloud Computing: A Survey. *2010 Sixth International Conference on Semantics, Knowledge and Grids*, pp.105–112.

Zhu, B. & Sastry, S., 2010. SCADA-specific Intrusion Detection/Prevention Systems: A Survey and Taxonomy. *Proceedings of the 1st Workshop on Secure Control Systems (SCS).* 2010.

Information Security Management System Standards: A gap Analysis of the Risk Management in ISO 27001 and KATAKRI

Riku Nykänen and Mikko Hakuli
University of Jyväskylä, Jyväskylä, Finland
riku.t.nykanen@student.jyu.fi
mikko.s.hakuli@student.jyu.fi

Abstract: An information security management system (ISMS) provides controls to protect organizations their most fundamental asset, information. Risk management is an essential part of any ISMS. ISO27001 is a widely adopted ISMS standard that sets specific information security requirements for the management system. Organizations that claim to have adopted ISO27001 can be formally audited and certified to comply with the ISO27001 standard. KATAKRI is a Finnish national security auditing criteria that is based on several ISMS standards and best practices. It was initially intended to be used by public sector to audit private sector service providers, but it has been adopted also as a baseline of requirements for private sector security standards. Since many organizations have claimed ISO27001 certification, it is beneficial to analyse the gaps between ISO 27001 and national KATAKRI certifications. This paper explores structures of ISO 27001 and KATAKRI and presents results of gap analysis of risk management requirements between ISO 27001 controls for information security management and KATAKRI requirements.

Keywords: information security management system (ISMS), risk management, ISO 27001, KATAKRI

1. Introduction

Risk management is an essential part of all major information security management systems. One of the key objectives of risk management is to identify and secure key assets to enable business operations and their continuity. The information technology causes a number of risks in performing operational activities and these risks are expected to continue to escalate as new technologies emerge (Pereira and Santos, 2010).

Information security helps to mitigate the various risks through the application of a suitable range of security controls (Posthumus and von Solms, 2004). Each industry operates in different risk environment. In addition to common risks each organization has its own unique risks. Hence organizations continuously struggle to choose and implement the cost efficient set of security controls that mitigates the risks to acceptable level. (Baker and Wallace, 2007)

Many organizations apply certification for their ISMS to convince their stakeholders that security of organization is properly managed and meets regulatory security requirements (Broderick, 2006). Security aware customers may require ISMS certification before business relationship is established (KATAKRI, 2011). As there is a variety of different ISMS approaches available, organizations may even be requested to have multiple certifications.

ISMS standards are not the silver bullet and they possess potential problems. Usually guidelines are developed using generic or universal models that may not be applicable for all organizations. Guidelines based to common, traditional practices take into consideration differences of the organizations and organization specific security requirements. (Siponen and Willison, 2009)

In this study we compare the internationally widely used ISO/IEC 27001 to Finnish national ISMS approach called KATAKRI. Comparison is limited to risk management requirements of ISMS. The paper is structured as follows: in the section 2 an overview of risk management as part of ISMS and overview of selected standards are presented. In section 3 we briefly present need for gap analysis and present a model of how the requirements were divided into phases for analysis; section 4 presents summary of the results of the gap analysis; conclusions of the results of the gap analysis are presented in section 5; discussion and future work are presented in section 6.

2. Risk management as part of information security management

2.1 Risk components in security ontology

Area of security involves people with different roles within organizations. This emphasizes the role of common understanding of the used terminology. Comprehensive study of security ontologies (Blanco et al., 2011) denotes that security community, including risk analysis community, lacks common ontology thus there exist many domain specific ontology definitions.

Risk components should be identified in Certification and Accreditation (C&A) process requiring risk management (Gandhi and Lee, 2007). ISO/IEC definitions are commonly used for terms asset, vulnerability, threat and control. Assets are something having value for the organization and what needs to be protected. Countermeasures can mitigate or reduce vulnerabilities to acceptable level. Control (countermeasure) is a mean of managing risk, including policies, procedures, guidelines, practices or organizational structures. Threat a potential cause of an unwanted incident, which may result in harm to a system or organization. Vulnerability is a weakness of an asset or group of assets that can be exploited by threats. (ISO/IEC 27002) In this paper we use previous ISO/IEC definitions unless otherwise stated.

2.2 Definition of requirements for ISMS

Desirable and “complete” security requirements cover seven facets: who, where, what, when, why, which and how? Structured requirement definitions with well-designed requirement attributes provide clearer, concise, and informative requirements compared to natural language requirement definition. (Lee et al., 2006)

C&A requirements are generally written in natural language instead of structured requirements (Gandhi and Lee, 2007). According to Lee et al. (2006) natural language requirements suffer from range of problems related to, for example, consistency, completeness and redundancy. Natural language requirements are often long and verbose, but decomposing a requirement may change the meaning or context of the requirement. However, decompositions ease requirement compliance evaluation. Another problem is varying requirement abstraction levels. Decomposition and restructuring is a solution for this problem also. The third addressed problem in the natural language requirements is that requirements suit to multiple requirement categories. The last of the presented problems is having redundant requirements. The same requirement may be expressed even within same document using different terminologies.

2.3 ISO/IEC 27000 standards family

ISO/IEC 27001 is an information security standard published by the ISO/IEC standardization organization in 2005. It specifies the requirements for establishing, implementing, operating, monitoring, reviewing, maintaining and improving a documented Information Security Management System. ISO/IEC 27001 specifies requirements for the management of the implementation of the security controls. The controls and implementation guidelines than an organization may use are presented in ISO/IEC 27002. Controls represented in appendix of ISO/IEC 27001 and in ISO/IEC 27002 are normative. Organization defines which of the controls it shall implement. Organization may request certification against ISO/IEC 27001 for implemented ISMS. ISO/IEC 27001 contains definition of the term and definitions. Definitions refer to other ISO/IEC standard documents. Hence all ISO/IEC 27000 family standards share a common ontology.

ISO/IEC 27001 describes four-phase cyclic process known as “Plan-Do-Act-Check” (PDCA).

- Plan: establish security policy, objectives, processes and procedures.
- Do: implement the security policy and relevant procedures.
- Check: assess and measure the process performance.
- Act: take corrective and preventive actions.

Applying PDCA model, organization adopts a process approach for establishing, implementing, operating, monitoring, reviewing, maintaining and improving an ISMS. ISO/IEC 27000 Information Security Management System standards family includes also ISO/IEC 27005 standard for risk management. Its purpose is equal to ISO/IEC 27002 as it provides implementation guidance that can be used when planning risk management activities.

Boehmer (2009) claims that ISMS based on ISO 27001 is equivalent to risk management, which again is equivalent cost/benefit management. Risk approach is in the interest of organizations that want to avoid wasting investments in information security, and to find cost-efficient, risk mitigating controls.

2.4 KATAKRI – Finnish national security auditing criteria

Another approach to manage corporate security is Finnish national security auditing criteria, KATAKRI. It is published by the ministry of defence, but Confederation of Finnish Industries, Finnish Communications Regulatory Authority, ministry of foreign affairs and ministry of the interior have also participated in the preparation of the criteria. Initial version was published in 2009 and the updated version II in 2011.

The first goal of national security auditing criteria is to harmonize official measures while assessing organization security level. The second defined goal is "to support companies and other organizations as well as authorities with their service providers and subcontractors to work on their own internal security". Therefore criteria contain unofficial recommendations to help users to apply useful security practices. (KATAKRI, 2011)

KATAKRI is organized as requirements compliance questionnaire. It defines a number of requirements in form of questions. Each question consists of a tripartite classification of criteria, corresponding to the security level concepts: the base level (level IV), the increased level (level III) and the high level (level II). For KATAKRI certification the organization shall select the pursued security level. Based on selection, every criterion defined for the selected security level must be complied in each question. The questions and criteria are defined in natural language.

Criteria are divided into four main areas:

- administrative security
- personnel security
- physical security
- information security

Areas are not meant to be used independently. It is instructed to take all four areas into account when performing accreditation audit using KATAKRI. (KATAKRI, 2011)

KATAKRI does not include definition of terminology that is used. Each question contains, in addition to requirements to all security levels, two columns; "recommendations for the industry" and "source/additional information". For the questions having sources defined, definitions of terms can be derived from defined requirement sources. Lack of the common ontology can be seen as major weakness of KATAKRI compared to other ISMS standards.

3. Risk management compliance gap analysis

In this research we focus on ISO/IEC 27001 and KATAKRI risk management requirements. Organization may request certification for implemented ISMS against both standards. They both define their own specific set of requirements that ISMS must fulfill to be compliant.

In the preface of KATAKRI it is stated that "the criteria have been created from the perspective of absolute requirements and they do not include a marking system which is used in some criteria". Also, ISO/IEC 27001 states that "excluding any of the requirements specified in Clauses 4, 5, 6, 7, and 8 is not acceptable when an organization claims conformity to this International Standard." As both approaches present absolute prerequisite to meet all requirements with yes/no satisfaction criteria, results are comparable by comparing requirements as results are in same scale. Both KATAKRI and ISO/IEC 27001 use the scale of being full compliance or non-compliance. Partial compliance is not accepted. As Karabacak and Sogukpinar (2006) state that the official certification can be difficult as it is "all-or-nothing" design.

The main research question was to analyze is the ISO/IEC audited risk management process compliant with KATAKTRI requirement for risk management. As result of the analysis we expected to see gap analysis of risk management requirements of ISO/IEC 27001 and KATAKRI to both directions. We hope to see that results of this analysis will help organization having either of the certifications to evaluate easier amount of actions required to pursue the other certification.

The risk management requirements are covered in ISO/IEC 27001 in section 4.2.1. There are six main requirements. Three of these requirements contain ten more specific requirements for the corresponding main requirements.

In KATAKRI, risk management requirements are covered in the first part, administrative security. In this part there is subdivision A400, "Identifying, assessing, and controlling risks". This part contains 12 questions, which each contain several requirements. Risk management requirements are not only limited to section A400, but there are risk management requirements also in other subdivisions of the administrative security main part.

Fenz and Ekelhart (2011) have analyzed five commonly used ISRM methodologies and derived a generic ISRM view out of the selected methodologies. They have created five phases for risk management. Phases and their outputs are represented in table 1.

Table 1: Information security risk management phases and their outputs by Fenz and Ekelhart (2011)

ISRM phases and outputs	
Phase	Output
System characterization	Inventory list of assets to be protected, including their acceptable risk level.
Threat and vulnerability assessment	List of threats and corresponding vulnerabilities endangering the identified assets.
Risk determination	Quantitative or qualitative risk figures and levels for identified threats.
Control identification	List of potential controls that can mitigate the risks to an acceptable level.
Control evaluation and implementation	List of cost-efficient controls that have to be implemented to reduce the risk to an acceptable level.

We identified the risk management requirements from ISO/IEC 27001 and KATAKRI and used qualitative requirement analysis to categorize them into the ISRM phases. The content of the each category was analysed to find the gaps between the requirement definitions within context of an ISRM phase. Within the context of each ISRM phase, the goal of the requirement was determined for all requirements (Jingbai et al., 2008). The gaps in the requirements were detected analysing differences in the goals of the requirements.

Both ISO/IEC 27001 and KATAKRI define requirements to establish risk assessment procedure, which is outside of the scope of ISRM phases. Hence these requirements were analysed as separate set of requirements.

4. Results

This chapter represents key results of the requirement analysis. In the following tables 2 and 3, requirement criteria without corresponding criteria in other specification is presented in *italic* style. Tables don't include all requirements for clarity, but the most important requirements for all phases are included.

Requirements outside of the scope of ISRM phases set the prerequisites to implement risk assessment methodology, which shall implement requirements categorized into phases. Table 2 represents identified requirements for risk assessment procedures.

Table 2: Risk assessment procedure requirements mapping

Risk assessment procedure requirements mapping	
KATAKRI	ISO/IEC 27001
Define a risk assessment procedure (A401.0) Results of the risk assessment procedure are documented (A401.0) *Measure risk assessment process (A407.0)* *Risk assessment is performed annually or when significant changes occur (A403/level III) or risk assessment is part of management process (A403/level II)* *Results of risk assessment are considered when setting goals of the security work (A404.0)*	Identify a risk assessment methodology suited to requirements (4.2.1c1) Develop criteria for accepting risks (4.2.1c2)

Identified risk management requirements from ISO/IEC 27001 and KATAKRI were mapped to the presented ISRM phases. Results of the mapping are presented in table 3. Corresponding security level is presented in KATAKRI requirements. In addition table includes ISO/IEC 27005 mapping (Fenz 2011).

Table 3: Information security risk management phase mapping

Information security risk management phase mapping			
Phase	KATAKRI	ISO/IEC 27001	ISO/IEC 27005 (Fenz 2011)
System characterization	Asset identification (A401.1) Identify owners of assets (A401.1)	Identify acceptable levels of risk (4.2.1c2) Asset identification (4.2.1d1) Identify owners of assets (4.2.1d1)	Asset identification
Threat and vulnerability assessment	Threat assessment (A401.1) Identify vulnerabilities (I706.0)	Identify threats (4.2.1d2) Identify vulnerabilities (4.2.1d3)	Identify threats Identify vulnerabilities
Risk determination	Assess risks (A401.2) Risks are prioritised (A405.0) Likelihood risk estimation (A405.0/level II) Risk assessment covers at least security management and personnel, information and premises security (A402.0) *Risks relating to external actors are identified (A402.0, A409.0)* *Risk assessment influences to security training (A405.0)*	Identify impact (4.2.1d4, 4.2.1e1) Assess threat likelihood (4.2.1e2) Assess vulnerability (4.2.1e2) Likelihood risk estimation (4.2.1e4)	Identify impact Assess threat likelihood Assess vulnerability Likelihood risk estimation
Control identification	(No requirements)	*Identify and evaluate options for the treatment of risks (4.2.1f)*	Evaluate existing and planned controls
Control evaluation and implementation	Controls are proportioned to the assets and the relevant risks (A401.1) Management approved chosen controls (A401.2) Management approval for residual risks (A401.2)	Select control objectives and controls (4.2.1g) Management approval for residual risks (4.2.1h)	Information security risk treatment (risk avoidance, risk transfer, risk reduction, or risk retention)

As seen from table, KATAKRI does not explicitly require identify and evaluate possible options to mitigate the risks. Rationale for this can be found from the other sections of KATAKRI documentation. Criteria itself contains mandatory controls for each defined security level. Therefore it is not mandatory for organization to evaluate other possible risk treatment options or controls. As ISO/IEC 27001 does not set any specific controls, but only defines normative controls, it is mandatory for organization itself to identify and evaluate appropriate options for risk treatment.

5. Conclusions

Comparing natural language requirements has exposed variety of problems. Many of the analyzed requirements have problems with the completeness. KATAKRI also contains several redundant requirements. Mutual ontology between compared standards facilitates analysis. While KATAKRI is lacking definition of terms, its definitions must be extracted from referred documents. In subdivision A400, “Identifying, assessing, and controlling risks” both ISO/IEC 27001 and 27002 are among the referred documents. Hence risk management terminology is coherent in both documents, but problems exist in other parts of the KATAKRI.

Gap analysis indicates that the KATAKRI certified ISMS implements the most of the risk management requirements of ISO/IEC 27001, but some exceptions exist. As presented in previous chapter, KATAKRI does not have requirement to evaluate and identify possible options for risk treatment. Rationale for this is that

KATAKRI itself defines minimum set of controls for each defined security level. ISO/IEC 27001 does not define any mandatory controls, but all controls defined in ISO/IEC 27002 are under considered as normative. The second ISO/IEC 27001 requirement missing from KATAKRI is risk likelihood analysis, which is required by the KATAKRI only on the high security level (level II). KATAKRI requires grouping risks by the importance, but this is not exactly same requirement as likelihood analysis, because risk importance may comprise other risk attributes such as impact. The third difference is the identification of the vulnerabilities. KATAKRI does not require risk management process to identify vulnerabilities, but has requirement to identify the technical vulnerabilities in section of information assurance.

ISO/IEC 27001 certified ISMS does not automatically fulfill all KATAKRI risk management reguirements. Following requirements from KATAKRI are not included in ISO/IEC 27001:

- Risk management process is measured.
- Risk assessment is performed annually or when significant changes occur (A403/level III) or risk assessment is part of management process (A403/level II).
- Risk assessment results drive security work.
- Management has approved chosen controls.
- Risk assessment is also required, when relevant, from external actors like subcontractors and service providers.
- Risk assessment influences to security training.

When organization implements ISMS using PDCA model, the requirements for measurement, periodic assessment, results driving security work and management approval for security controls, should be fulfilled. These are part of “check” and “act” phases of PDCA model to measure results and achieve continuous improval of ISMS.

The other two deviating requirements, “assessing external parties” and “assessment influence to security training” are covered by normative controls in ISO/IEC 27002. Requirement assessing external parties is analogous to “Addressing security in third party agreements”. In ISO/IEC 27002, control “Information security awareness, education, and training” has guideline to include known threats in security training. If this control is implemented, ISMS procedure should also fulfill the KATAKRI requirement.

In this study our target was to compare contents of risk management requirements between ISO/IEC 27001 and KATAKRI. As results show, some deviations between requirements exists to both directions and requirements are not completely overlapping. Major deviation between models is the identification possible options for the risk treatment. Where ISO/IEC 27001 requires organizations to implement a process to identify potential options, KATAKRI defines itself a minimum set of controls for each of the three security levels. Most of the KATAKRI requirements missing from ISO/IEC 27001 are fulfilled when ISMS is implemented using PDCA model. Other deviations in the risk management are minor and a well implemented ISMS should cover these requirements.

6. Discussion

This research was limited to analyzing KATAKRI and ISO/IEC 27001 requirements for risk management. For organizations having either of certifications, it would be meanful to have analysis of complete requirement definitions. Comparison structure should compare each security level from KATAKRI to combination of ISO/IEC 27001 and 27002. As we have seen that some of the KATAKRI requirements are covered in the normative controls of ISO/IEC 27002, which should be included in comparison even it is normative document.

In this study we have identified some problems that KATAKRI currently comprises. One of them is the lack of common ontology over the document. This leaves possibility for interpretation instead of having exact requirements for ISMS. Another identified problem is the natural language requirements. As long as KATAKRI is structured as requirements compliance questionnaire, the problem can only be mitigated enhancing requirement definition quality.

Future research is continued on evaluating existing risks for IT companies and how current ISMS certification models correlate to existing risks. One of the goals is to study if the ISMS certificate will help organizations to find cost-efficient, risk reducing security controls or does certification just cause additional costs for the organization that doesn't reduce actual risks at all.

References

Baker, W.H. & Wallace, L. 2007, "Is Information Security Under Control?: Investigating Quality in Information Security Management", Security & Privacy, IEEE, vol. 5, no. 1, pp. 36-44.

Blanco, C., Lasheras, J., Fernández-Medina, E., Valencia-García, R. & Toval, A. 2011, "Basis for an integrated security ontology according to a systematic review of existing proposals", Computer Standards & Interfaces, vol. 33, no. 4, pp. 372-388.

Boehmer, W. 2009, "Cost-Benefit Trade-Off Analysis of an ISMS Based on ISO 27001", Availability, Reliability and Security, 2009. ARES '09. International Conference on, pp. 392.

Broderick, J.S. 2006, "ISMS, security standards and security regulations", Information Security Technical Report, vol. 11, no. 1, pp. 26-31.

Fenz, S. & Ekelhart, A. 2011, "Verification, Validation, and Evaluation in Information Security Risk Management", Security & Privacy, IEEE, vol. 9, no. 2, pp. 58-65.

Gandhi, R.A. & Lee, S. 2007, "Discovering and Understanding Multi-dimensional Correlations among Certification Requirements with application to Risk Assessment", Requirements Engineering Conference, 2007. RE '07. 15th IEEE International, pp. 231.

ISO/IEC 27001:2005 2005, Information technology – Security techniques – Information security management systems – Requirements, ISO copyright office, Geneva, Switzerland.

ISO/IEC 27002:2005 2005, Information technology – Security techniques – Information security management systems – Code of practice for information security management, ISO copyright office, Geneva, Switzerland.

Jingbai, T., Keqing, H., Chong, W. & Wei, L. 2008, "A Context Awareness Non-functional Requirements Metamodel Based on Domain Ontology", Semantic Computing and Systems, 2008. WSCS '08. IEEE International Workshop on, pp. 1.

Karabacak, B. & Sogukpinar, I. 2006, "A quantitative method for ISO 17799 gap analysis", Computers & Security, vol. 25, no. 6, pp. 413-419.

KATAKRI 2011, National Security Auditing Criteria version II, Ministry of Defence, Finland.

Lee, S., Gandhi, R., Muthurajan, D., Yavagal, D. & Ahn, G. 2006, "Building problem domain ontology from security requirements in regulatory documents", Proceedings of the 2006 international workshop on Software engineering for secure systemsACM, New York, NY, USA, pp. 43.

Pereira, T. & Santos, H. "A Conceptual Model Approach to Manage and Audit Information Systems Security", Proceedings of the 9th European Conference on Information Warfare and SecurityAcademic Conferences Limited, , pp. 360.

Posthumus, S. & von Solms, R. 2004, "A framework for the governance of information security", Computers & Security, vol. 23, no. 8, pp. 638-646.

Siponen, M. & Willison, R. 2009, "Information security management standards: Problems and solutions", Information & Management, vol. 46, no. 5, pp. 267-270.

Time Correlated Anomaly Detection Based on Inferences

Abimbola Olabelurin[1], Georgios Kallos[2], Yang Xiang[3], Robin Bloomfield[4], Suresh Veluru[1] and Muttukrishnan Rajarajan[1]
[1]School of Engineering and Mathematics, City University London, UK
[2]BT Technology Service and Operations, Adastral Park, Ipswich, UK
[3]School of Information Technology, Deakin University, Australia
[4]School of Informatics, City University London, UK
olalexb@yahoo.com

Abstract: Anomaly detection techniques are used to find the presence of anomalous activities in a network by comparing traffic data activities against a "normal" baseline. Although it has several advantages which include detection of "zero-day" attacks, the question surrounding absolute definition of systems deviations from its “normal” behaviour is important to reduce the number of false positives in the system. This study proposes a novel multi-agent network-based framework known as *Statistical model for Correlation and Detection (SCoDe)*, an anomaly detection framework that looks for time-correlated anomalies by leveraging statistical properties of a large network; monitoring the rate of events occurrence based on their intensity. *SCoDe* is an instantaneous learning-based anomaly detector, practically shifting away from the conventional technique of having a training phase prior to detection. It does acquire its training using the improved extension of *Exponential Weighted Moving Average (EWMA)* which is proposed in this study. *SCoDe* does not require any previous knowledge of the network traffic, or network administrators chosen reference window as normal but effectively builds upon the statistical properties from different attributes of the network traffic, to correlate undesirable deviations in order to identify abnormal patterns. The approach is generic as it can be easily modified to fit particular types of problems, with a predefined attribute, and it is highly robust because of the proposed statistical approach. The proposed framework was targeted to detect attacks that increase the number of activities on the network server, examples which include Distributed Denial of Service (DDoS) and, flood and flash-crowd events. This paper provides a mathematical foundation for *SCoDe*, describing the specific implementation and testing of the approach based on a network log file generated from the cyber range simulation experiment of the industrial partner of this project.

Keywords: time-correlated anomaly detection, intrusion detection system, denial of service attack, time-series analysis, exponential weighted moving average

1. Introduction

Intrusions into information systems are security violations, or attempts that compromise the confidentiality, integrity or availability (CIA) of a system. These intrusions are attacks that result in anomalous behaviour and activity in the system, threatening the reliability and quality of service (QoS) of such system. Intrusion Detection Systems (IDS) are deployed to monitor hosts, networks or other resources using several techniques to detect security violations. It is very important that the detection of these intrusions is accurate, timely and rapid with low-risk rate of false positives.

Anomaly detection aims at finding the presence of anomalous patterns or activities in a network by comparing traffic data activities against a "normal" baseline. Using numerous techniques for IDS (Patcha and Park, 2007; Gupta and Kotagiri, 2010; Ye et al, 2001), anomaly detection systems have several advantages over other detection techniques, which include detection of "zero-day" attacks. However, the question surrounding the absolute definition of systems deviations from its “normal” behaviour is important to reduce the number of false positives triggered by these systems.

In the recent past, some research approaches (Alarcon-Aquino and Barria, 2001; Brutlag 2000) to anomaly detection assume that an anomaly is an abrupt change in some features extracted from the traffic, so the reference is simply the previous-to-current traffic window. Other approaches (Manikopoulos and Papavassiliou, 2002; Oliner, Kulkarni, and Aiken, 2010) assume that anomaly occurs when a traffic window is sufficiently different from a reference window that have been previously chosen and approved by network administrators. These approaches have several disadvantages. The problem with the first approach is that it only detects abrupt changes from one traffic window to another while the second approach requires an expert to have previous knowledge of the traffic in order to determine the reference (normal) traffic window. Additionally, due to the non-stationary nature of network traffic (Simmross-Wattenberg et al, 2011), having just one reference window can be problematic, as network traffic exhibits cyclic-stationary behaviour and using a single reference window will not be appropriate under all network conditions.

Consider some of the worms, malwares, or malicious codes that do not allow time for human interventions (Moore et al, 2003). The primary breach is usually undetected and the existence of the exploit remains unknown, so that malicious code may continue to run indefinitely, stealing computing resources (as in a zombie network), spoofing content, denying service, and many more. According to a recent Kaspersky report (RT, 2012), a typical example is the recently discovered Flame virus that does not cause any physical damage but it is the most sophisticated malware, capable of stealing different types of information with the help of its spyware tools. This virus can record audio, capture screen and transmit visual data at the same time. The anomalous activities by this malicious code are expected to lead onto an increase in the event frequency.

Furthermore, consider a typical Denial of Service (DoS) attack, an overwhelming number of service requests can be sent to a particular server over a short period of time and thus deny the service to normal users. These DoS attacks suddenly and significantly increase the number of events and activities on the server.

It has been shown that anomalous behaviour triggered by a worm, malicious code and DoS attacks can be detected because of temporal consistency, similarity (low temporal variance) in system calls; as they exhibit time-correlated attribute behaviours (Malan and Smith, 2006). Though many of these behaviours can be individually insignificant, but they provide important and powerful information when aggregated. This study proposes a novel 'multi-agent' network-based framework known as *Statistical model for Correlation and Detection (SCoDe)*, an anomaly detection framework that looks for time-correlated anomalies by leveraging statistical properties of a large network and monitoring the rate of events occurrence based on their intensity. These agents can mutually recognise each other's activities, instantaneously learn and dynamically adapt to the system. Each agent sense, communicate and generate responses regarding their monitored attributes. Attributes may include different networked system activities at user levels (type of software/program use, login/logout period and location), system level (amount of free memory, cumulative and per user CPU usage), process level (system process time, idle time, relationship among processes) and packet level (type of connection, average number of packet sent/received, protocol and port used).

Hence, in the context of the proposed framework, *an anomaly is defined as an unexpected 'state change' in time across multiple attributes in a network system.* The proposed model was targeted to detect attacks that increase the number of events and activities on the network server, example of which include Distributed Denial of Service (DDoS) and flood and flash-crowd anomalies.

SCoDe as shown in Figure 1 deployed each agent on its host, using the modified and improved variation of Exponential Weighted Moving Average (EWMA) to separately model the activity and behaviour of a relevant system attribute. The ability of EWMA to monitor the rate of event occurrence, its sensitivity to small variation in process mean, its ability to customize the detection of small shift and large shift in the process and ease of parameter modification makes it useful for anomaly detection in event intensity.

SCoDe monitors each agent's behaviour and reports a *SCoDe score*, which quantifies the divergence of recent event from the EWMA, which represent the smoothed event frequency model. The model then computes the numerical average of all agents' scores and checks whether this *network score* exceeds a threshold, statistically defined from the traffic data to give an arbitrarily high probability of perfect detection (False positive = false negative = 0). By doing these computations systematically, it is shown in this paper that the proposed model overcomes noise and can detect system anomaly.

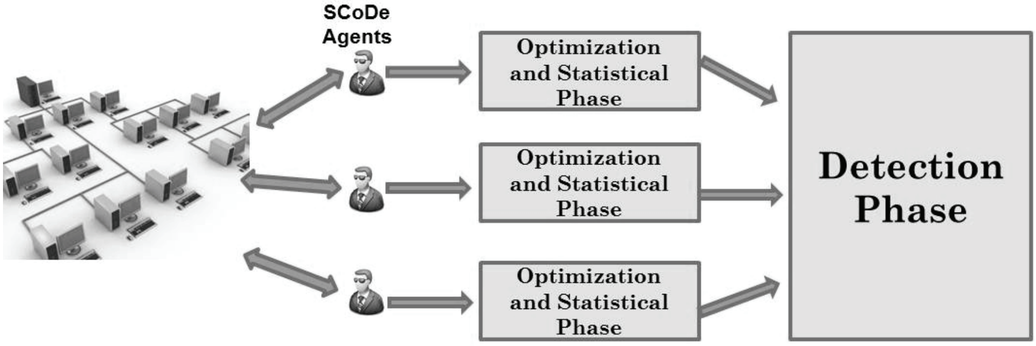

SCoDe Agents
Optimization and Statistical Phase
Optimization and Statistical Phase
Optimization and Statistical Phase
Detection Phase

Figure 1: *SCoDe* approach to time-correlated anomaly detection

The proposed model assumes that the aggregated network traffic follows Gaussian distribution (Papoulis, 1984), intuitively expect anomalies on individual agent in the network to be relatively common, but do not expect irregular high anomaly score from multiple agents to be strongly correlated in time.

The approach is generic as it can be easily modified to fit particular types of problems, with a defined attribute, and it is highly robust because of the proposed statistical approach. It requires no prior knowledge about normal events in order to set thresholds; instead, it learns what constitutes the normal from its observations. The thresholds depend on the instantaneous observed events, ensuring that what constitute a normal for one network is considerably different from another.

To validate the proposed framework, the result obtained in this work is compared to those reported in (Ye, Vilbert and Chen, 2003), which employed the first order EWMA technique in detecting intrusions. The section that follows describes and analyses some of the related work and Section 3 discusses the proposed approach to intrusion detection, which include the methodology used, data acquisition process and result obtained.

2. Related work

2.1 Intrusion detection approach

Anomaly-based intrusion detection has a long history of detecting `Zero-day attacks' (Patcha and Park, 2007; Ye, Chen and Borror, 2004). However, its operation has been fundamentally limited by the poor quality of the models by generating excessive and high rate of false positive alarms (Oliner, Kulkarni, and Aiken, 2010).

In order to solve some of the above mentioned issues, different approaches have been discussed. Gupta and Kotagiri (2010), Ji and Ma (1997), and Tombini et al (2004) used combination of either "strong" classifier, "weak" classifier or both, in a stacked or serial layered approach for better classification of analysed events. A major drawback is that these techniques are expensive with regards to model training, processing time and decision making. Event or alert correlation was also proposed as a solution in Valdes and Skinner (2001), while cumulative trigger was investigated in Huang et al (2007). However, most of these collaborative detection efforts raise alarm only on individual clients.

In general, harshly simplifying, higher level correlation, taking multiple alerts and additional information from heterogeneous sources into consideration tends to reduce alert volume, improve alert content and track attacks spreading across multiple packets (Viinikka et al, 2006). *SCoDe* builds on the principle of Syzygy model (Oliner, Kulkarni, and Aiken, 2010) that appropriately uses the aggregate behaviour of a community or groups of network events to decide whether to raise an alarm for the community and not individual clients, making strong guarantees and better efficient anomaly detector even with a noisy model.

While the model in Viinikka et al (2006) monitors the behaviour of *n* clients in a community in order to detect deviation from the previously observed "normal" behaviour, the propose generic model uses the *SCoDe* agent to monitor *n* number of client behaviours, employing EWMA to analyse and find deviation from each "immediate" client's behaviour in real time. Thereafter, the model analyse the aggregate of these deviations to look for time-correlated deviations in the network and raise an alarm.

2.2 Exponential weighted moving average (EWMA)

Statistical Process Control (SPC) techniques have typically been used for monitoring and controlling quality of manufacturing processes. They can be univariate or multivariate, and they have been extensively used in forecasting and time series analysis, detecting changes in process mean, process variance, and relationship among multiple variables (Montgomery and Johnson, 1976).

This method has been deployed and utilized for intrusion detection in Ye et al (2001), and Cisar and Maravic (2007*).* It extracts trends by forecasting the “normal” operations and highlighting any anomalies from alert information provided by sensors performing pattern matching.

According to Roberts (2000), the first order EWMA is simply defined as:

$$z_i = \lambda x_i + (1-\lambda) z_{i-1} \qquad (1)$$

and its extended variation

$$z_i = \lambda \sum_{j=0}^{i-1} (1-\lambda)^j x_{i-j} + (1-\lambda)Z_{i-j} \tag{2}$$

where λ is the smoothing constant ($0 < \lambda \leq 1$), ($1 - \lambda$) is called smoothing factor, x_i is the observation (event frequency in this model), z_i is the smoothed observation at time *i*.

As the monitored statistic in (1) is the process mean, the estimated standard deviation of the EWMA statistics is approximately

$$\sigma_z = \sqrt{\frac{\lambda}{2-\lambda}}\, \sigma_x \tag{3}$$

where σ_x is given by $\frac{\sigma}{\sqrt{n}}$ and σ is the standard deviation of x, supposed to be known a priori. Because of the recursive equations discussed above, the equation (3) can be written as

$$\sigma_z = \sqrt{\frac{\lambda\,[\,1-(1-\lambda)^{2i}]}{2-\lambda}}\, \sigma_x \tag{4}$$

One of the key factors that influence the performance of EWMA described above is the choice of the smoothing factor λ, which has a value range between 0 and 1 (Montgomery and Johnson, 1976). This factor determines the rate at which older (past) events are weighted into in the EWMA equation and can be made sensitive to a small or gradual process change by the choice of its value. For instance, a value of λ = 1 gives more weight to recent data and less weight to older data and vice-versa.

The choice of λ was left to the judgement of the analyst and it was arbitrarily set at 0.2 ± 0.1 by Hunter (1986) due to his experience with econometric data. Following the same trend, Lucas et al (1990) set the value of λ to 0.25 because this value was used for a similar sample data by Lucas and Crosier (1982). Instead of selecting a particular value of λ, (Ye, Chen and Borror, 2004) tested a wide range of values from 0.001 to 0.9. However, Cisar and Maravic (2007*)* showed that the optimal value for λ is the value which resulted in the smallest mean squared errors (MSE) when the decay factor was iterated between 0.1 and 0.9.

3. The *SCoDe* model

3.1 Proposed methodology

The goal of *SCoDe* is to identify anomalous events in the network within a specified detection window of time period *i*, and its simple implementation is as follows:

Step 1: Use each *SCoDe* agent to obtain, track and update the network of client's event frequency modelled as *SCoDe score* x_i, at a regular specified time interval period ***i***.

Step 2: For each x_i, use equation (2) to compute EWMA statistics z_i, setting the initial value of EWMA z_o as the estimated mean of the events μ_x.

Step 3: For each x_i and z_i, generate the network *anomaly score* s_i ($0 \leq s_i \leq 1$) according to a set of rules and produce the *SCoDe* matrix.

Step 4: Compute the average score among anomaly score s_i within a time window to generate the network score ***C*** that represents the state of the network. If ***C*** > ***V*** (a network threshold), then the model reports an anomaly. Subsection B shows how to compute the threshold V, given a desired false positive rate and the training data.

3.2 Anomaly score

For each of the behaviour and corresponding *SCoDe* agent, the model defines a *t*-dimensional feature vector

$$A_m = (s_1, s_2, s_3, \dots, s_t) \tag{5}$$

to represent an agent *m* update for its attribute and each A_m can be plotted as a point in a dimensional feature space. The instantaneous value s_i is set to 1 if the $x_i > z_i +$ δ where δ is the defined attribute threshold, otherwise x_i is set to 0. The larger the value of δ, the better, though any positive value will be sufficient.

In order to significantly improve the performance of the model in detecting an off target process immediately after the EWMA is started and after smoothed high frequency event that usually result in false alarm, equations (2) and (4) were substantially exploited leading to a constantly changing attribute threshold δ, with upper and lower limit UCL_z and LCL_z, set as

$$\begin{aligned} UCL_z &= \mu_z + L\sigma_z \\ LCL_z &= \mu_z - L\sigma_z \end{aligned} \tag{6}$$

where *L* is a value set to obtain desired s-significant level (usually 3 and known as the “three-sigma rule”).

For the total *n* set of *SCoDe* agents and attributes, each dimensional feature vector is treated as an independent point in a set and represented in a *SCoDe* matrix as follows:

$$A_{n,t} = ((s_{1,1} \quad s_{1,2} \quad s_{1,3} \quad \cdots \quad s_{1,t}@s_{2,1} \quad s_{2,2} \quad s_{2,3} \quad \cdots \quad s_{2,t}@| \quad | \quad |@s_{n,1} \quad s_{n,2} \quad s_{n,3} \quad \cdots \quad s_{n,t})) \tag{7}$$

In order to minimize the classification error and loss of information that could result from the binary vector (0 and 1), *SCoDe* applied a five levels scoring method to the anomaly score as described below

$$s_1 = \left\{ \begin{array}{ll} 0 & if\ \delta = 0 \\ 0.25 & 0 < \delta < \sigma \\ 0.50 & \sigma < \delta < 2\sigma \\ 0.75 & 2\sigma < \delta < 3\sigma \\ 1.0 & \delta > 3\sigma \end{array} \right\} \tag{8}$$

This novel approach enables *SCoDe* to be very sensitive to slight variations in parameter, classify and detect any small deviations from the smoothed profile in order to achieve anomaly detection.

Smoothing factor optimization

SCoDe model optimizes the smoothing factor by defining the one-step-ahead prediction errors as

$$e_i = x_i - z_{i-1} \tag{9}$$

and iterates λ values from 0.01 to 0.99. Thereafter, the sum of the squared errors SSE_λ and the mean squared error MSE_λ for each λ is obtained by using the following equations:

$$\begin{aligned} SSE_\lambda &= \sum_{i=0}^{n-1} e_i^{\,2} \\ MSE_\lambda &= SSE_\lambda \frac{\square}{\square}(n-1) \end{aligned} \tag{10}$$

where *n* is the total number of events. The λ value that result in the least MSE_{λ} is chosen for the model and this process significantly prove to improve the sensitivity and accuracy of the model.

3.3 Detection phase

Given a detection window, *SCoDe* computes the average score among all attributes anomaly signal to generate the *network score* ***C*** that represents the state of the client. If ***C*** > ***V*** (network threshold), then the model reports an anomaly.

Consider a network of *n* attributes and let $s_i \sim X$ where *X* is a random variable with finite mean and finite positive variance. By Central Limit Theorem (Papoulis, 1984), as n → ∞, the network scores are distributed normally with mean μ_X and variance σ_x^2/n:

$$C = \mathbf{average}_i\ (\ [\![s]\!]_{\downarrow i}) = \frac{1}{n}\sum_i (X) \sim Norm\ (\mu_{X}, \frac{\sigma_X^2}{n}) \qquad (11)$$

When ($E(|X|^{\uparrow 3}$) = ρ < ∞ where $E(\ )$ denotes the expected value, convergence happens at a rate on the order of 1/$\sqrt{n}$ (Berry-Esseen theorem). Concretely, let C' = C - μ_X and let F_n be the cumulative distribution function *(cdf)* of C'$\frac{\sqrt{n}}{\sigma_X}$ and ϕ the standard normal cdf. Then, there exist a constant *B* > 0 such that $\forall x, n, |F_n(x) - \phi(x)| \le \frac{B_\rho}{\sigma_X^3\sqrt{n}}$

Now consider when some numbers of the attributes *d* ≤ *n* of the network have been exploited. The anomaly score, as *n, d* → ∞ will be

$$C = \frac{1}{n}\left(\sum_{i=1}^{n-d}(X) + \sum_{i=1}^{d}(Y)\right) \sim Norm\left(\frac{(n-d)\mu_X + d\mu_Y}{n}, \frac{(n-d)\sigma_X^2 + d\sigma_Y^2}{n^2}\right) \qquad (12)$$

The rate of convergence guarantee the asymptotic behaviour at relatively small values of *n and d*, and even when $d \ll n$.

According to (Oliner, Kulkarni, and Aiken, 2010), we can choose any positive V between $\frac{\sigma_X^2}{n}$ and $\frac{\sigma_X^2}{\square}n + \delta$ and guarantee that there exist *n* and *d* that give arbitrarily high probability of perfect detection (False Position = False Negative =0). Without knowing δ however, the best strategy is to pick the lowest value of *V* such that the false positive is acceptable.

In line with (Oliner, Kulkarni, and Aiken, 2010), *V* threshold is generated by

$$V = \mu_H + 2\sigma_H^{\square} \qquad (13)$$

where *H* is the distribution of anomaly score for network with attribute *n*.

3.4 Data acquisition process

The data used to test the proposed *SCoDe* framework for intrusion detection were collected from the Cyber Range experiment performed by one of the industrial partners. The simulated experiment emulated the traffic of a large complex network illustrated in Figure 2, and evaluated the resilience of the network to Cyber-attacks.

The experiment comprised of a standard setup of:

- A main enterprise network consisting of 200 XP workstations, a Demilitarised Zone (DMZ), 2 Intrusion Detection Systems (Demilitarised IDS and Internal IDS) and 1 firewall;
- A branch enterprise network consisting of 10 XP workstations and a DMZ VLAN;
- 8 routers for the interconnection of the branches;

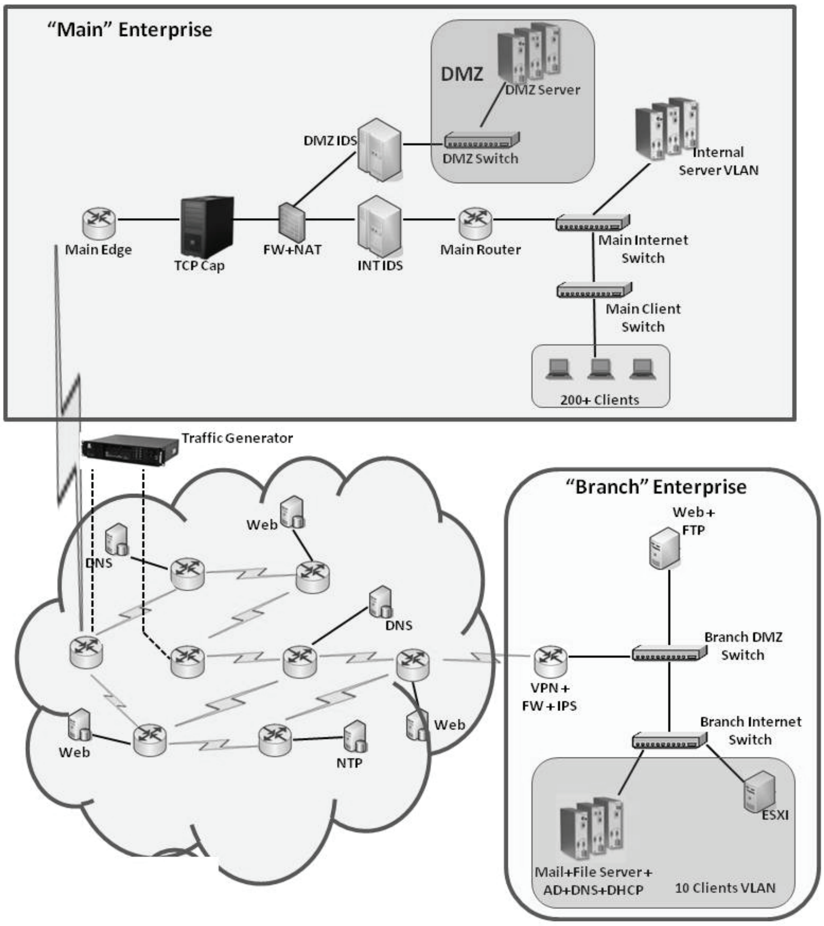


Figure 2: Schematic architecture of the experiment network diagram

A traffic generator simulated mixed regular traffic and injected malware internet traffic between the branch and the main networks. The injected malicious traffic consisted of attacks such as intrusive port scanning, Nexpose Vulnerability assessment, Bruteforce attacks, and DDoS attacks.

The experiment was run for over 140 hours between March 1, 2012 and March 7, 2012 and was focused on in-band monitoring and logging. The traffic generator was used to produce a total steady-state traffic volume of circa 12 Mbits/s in the main network and the different IDS namely DMZ IDS and INT IDS were used to monitor and detect malicious events. While the INT IDS monitors the traffic to the intranet, the DMZ IDS monitors the traffic to the DMZ. The data used for the evaluation of *SCoDe* consists of log files alert from the DMZ IDS.

In the first stage of the analysis, *SCoDe* was used to analyse the DMZ IDS SNORT alerts containing 8428 entries lasting 141 hours in time. A typical SNORT intrusion alert from the experiment is shown in Appendix A. Each alert has features which include Time, Type, Description, Priority level, Protocol Type, Source IP and Port, Destination IP and Port, and Sensor Information.

In the data acquisition stage, the frequency of events is taken at specific intervals which can be in seconds, minutes or hours. One-hour periods was chosen to be reasonably stationary for the data as the event frequency ranges from 0 to 1332 events as shown in Figures 4 to 6.

3.5 Results

The frequency of events occurring each hour is modelled as the observation data x_i and the available three different priority levels of the log data serves as the three attributes used in the proposed model. The graphs in Figures 4 to 6 show the event analysis of the three attributes that was model and Figure 7 shows the graph of the generated network score of the *SCoDe* model at each time interval. Out of the total 141 time intervals, attack with malicious traffic occurred 11 times.

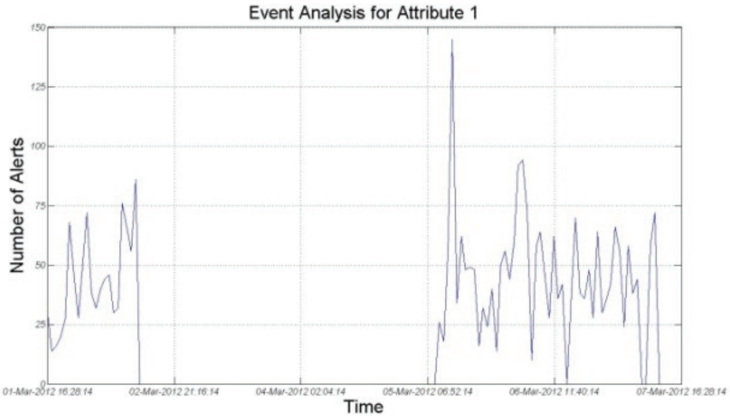


Figure 4: Analysis of attribute (priority) 1 events

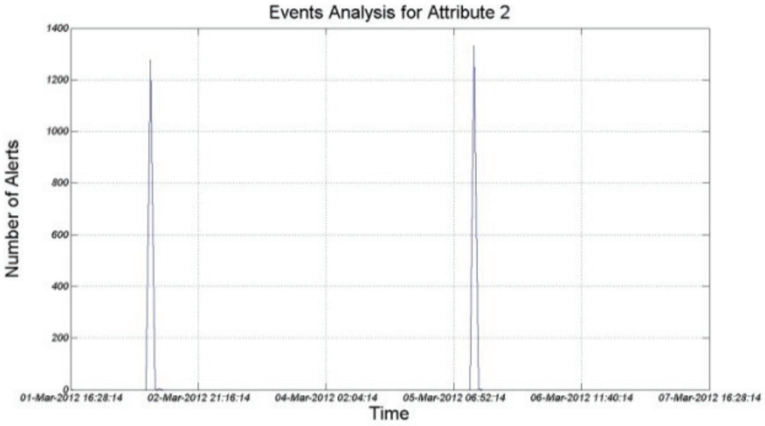


Figure 5: Analysis of attribute (priority) 2 events

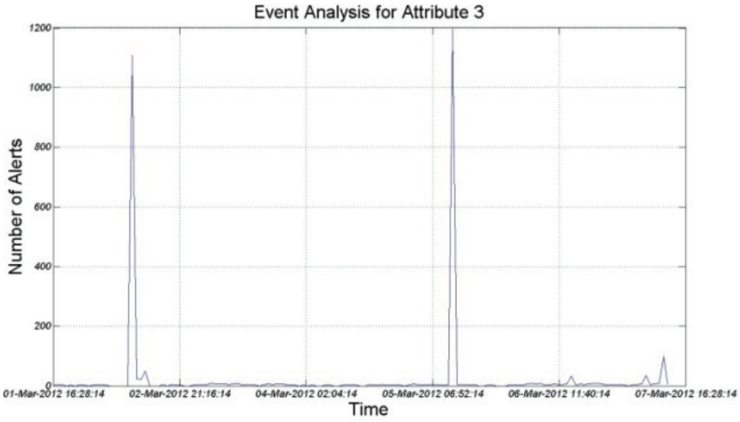


Figure 6: Analysis of attribute (priority) 3 events

SCoDe appropriately signal anomalous behaviours across the network at 10 different time intervals which corresponds to 90.9% detection rate, and the type of attack detected within those time intervals include intrusive port scanning, Nexpose vulnerability assessment, Nexpose exhaustive penetration audit, Bruteforce audit, and DDoS with false positives rate of 0.57%.

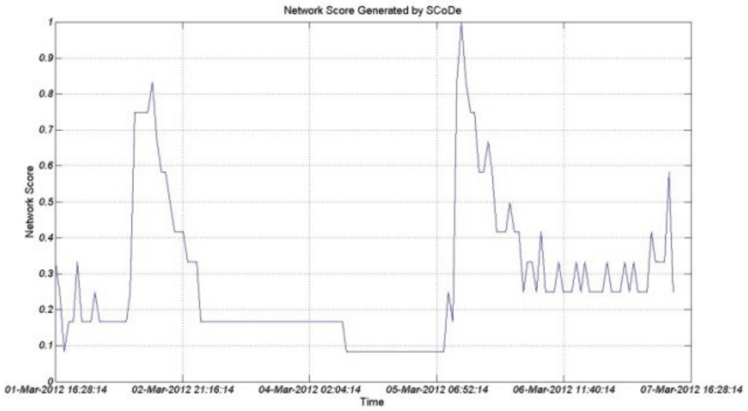


Figure 7: The generated network score

To validate the *SCoDe* model, the proposed model was compared with similar EWMA techniques used in (Ye, Vilbert and Chen, 2003), where λ values of 0.2 and 0.3 was experimented with two different significant levels *L* of 1.96 and 3.

For each technique with a given parameter combination, the number of hits and the number of false alarms were computed as show in Table 1. A hit refers to an anomaly detection signal on an intrusive event, and false alarm is a signal on usual event (steady-state traffic). The results show that when *L* is 1.96, the number of false alarms is very high irrespective of the λ value and the hit rate was between 80% and 90%. However, a better result is obtained when *L* is 3; with the hit rate still standing between 80% and 90% but a very low false alarm rate between 0 and 1 was achieved. However, the proposed *SCoDe* model achieved 90.9% hit rate and one false alarm using the optimal value of λ (0 < λ < 1) and its five-level scoring system for the time-series intrusion data, an improvement of around 1% on the detection rate. The difference was SCoDe ability to detect several short DoS attacks that went undetected with other techniques considered.

Table 1: Comparison of the proposed model against other EWMA based models

Technique	λ	L	Hits	False Alarm	Summary
$EWMA_{0.2,\ 3}$	0.2	3	9	1	Raised alarm on individual event
$EWMA_{0.2,\ 1.96}$	0.2	1.96	9	9	Raised alarm on individual event
$EWMA_{0.3,\ 3}$	0.3	3	8	0	Raised alarm on individual event
$EWMA_{0.3,\ 1.96}$	0.3	1.96	8	8	Raised alarm on individual event
SCoDe	0 -1	0-3	10	1	Raised alarm on aggregate event

Although other EWMA methods achieve the result above on alarm raised on individual attributes, *SCoDe* computes the aggregate attribute events before raising the alarms, as event aggregation has been significantly proved in Oliner, Kulkarni, and Aiken (2010), and Malan and Smith (2006) to reduced system noise.

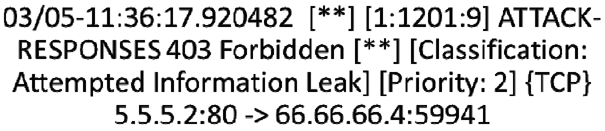
4. Conclusion and future work

This paper presented an approach for detecting anomalous events by leveraging the statistical properties of a large network. The proposed approach explored the aggregate system attribute changes and its effect in detecting attacks that increase the number of events and activities in a network such as Nexpose vulnerability assessment, Nexpose exhaustive penetration audit, Bruteforce audit, and DDoS. The prototype implementation, called *SCoDe* used in this study proposed a network of agents to address the dual problem of accuracy and efficiency for building robust and efficient intrusion detection systems. *SCoDe* raised alarms on aggregated events and the evaluation results show improved attack detection rate of about 1% compared to other similar EWMA techniques. This relatively small but significant improvement enables short DoS attacks to be quickly detected before it degenerate to full-scale DDoS. The proposed model can detect attacks that increased the frequency of events occurring in information systems. However, other anomalies such as port scanning attack should be detectable with minor modification to the current approach. Further work will also include evaluation of the model in different attack scenarios. Progress work also includes detecting time-correlated anomalies using different attributes from user levels, system level, process level and packet level activities.

Acknowledgements

Thanks to BT and Northrop Gumman UK, the industrial partners of this project for their valuable contributions and supports.

Appendix A: Alert sample from the experiment



```
03/05-11:36:17.920482  [**] [1:1201:9] ATTACK-
RESPONSES 403 Forbidden [**] [Classification:
Attempted Information Leak] [Priority: 2] {TCP}
5.5.5.2:80 -> 66.66.66.4:59941
```

References

Alarcon-Aquino V. and Barria J. (2001) *Anomaly detection in communication networks using wavelets*, Communications, IEE Proceedings, vol. 148, pp. 355 –362, Dec.Brutlag J. D. (2000) *Aberrant Behavior Detection in Time Series for Network Monitoring*, Proceedings of the 14th Systems Administration Conference *(LISA 2000)*, Dec.

Cisar P. and Maravic Cisar S. (2007*) EWMA statistic in adaptive threshold algorithm*, in Intelligent Engineering Systems, 2007 (INES 2007) 11th International Conference on, pp. 51 –54, July.

Gupta K. K., Nath B. and Kotagiri R. (2010) *Layered approach using conditional random fields for intrusion detection*, IEEE Trans. Dependable Secur. Comput., vol. 7, pp. 35–49, Jan.

Huang L., Nguyen X., Garofalakis M., and Hellerstein J. M., *Communication-efficient online detection of network-wide anomalies*," in IEEE Conference on Computer Communications (INFOCOM), pp. 134–142, IEEE, 2007.

Ji C. and S. Ma (1997) *Combinations of weak classifiers,* IEEE Trans. on Neural Network, vol. 8, no.1 pp. 32–42, Jan.

Lucas J. M. and Crosier R. B., (1982) *Fast initial response for CUSUM quality control schemes: give your CUSUM a head start,* Technometrics, vol. 42, no. 1, pp. 102–107.

Lucas J. M., Saccucci M. S., Baxley R. V. (Jr.), Woodall W. H., Maragh H. D., Faltin F. W., Hahn G. J., Tucker W. T., Hunter J. S., MacGregor J. F., and Harris T. J. (1990) *Exponentially weighted moving average control schemes: properties and enhancements*, Technometrics.

Malan D. J. and Smith M. D. (2006) *Host-based detection of worms through peer-to-peer cooperation*, *IEEE Security and Privacy*, Harvard University, USA

Manikopoulos C. and Papavassiliou S. (2002) *Network intrusion and fault detection: a statistical anomaly approach*, *Communications Magazine, IEEE*, vol. 40, pp. 76 – 82, Oct.

Montgomery D. and Johnson L. (1976) *Forecasting and time series analysis,* McGraw-Hill

Hunter J. S. (1986) *The exponentially weighted moving average,* J Quality Technology, vol. 18, no. 4, pp. 203–207.

Moore D., Paxson V., Savage S., Shannon C., Staniford S., and Weaver N. (2003), *Inside the Slammer Worm*, IEEE Security and Privacy, vol. 1, no. 4, pp. 33 – 39.

Oliner A. J., Kulkarni A. V., and Aiken A. (2010) *Community epidemic detection using time-correlated anomalies*, in *Proceedings of the 13th international conference on Recent advances in intrusion detection*, RAID'10, (Berlin, Heidelberg), pp. 360 – 381, Springer-Verlag, 2010.

Papoulis A. (1984) *Probability, Random Variables, and Stochastic Processes*, Mc-Graw Hill.

Patcha A. and Park J. M. (2007) *An overview of anomaly detection techniques: Existing solutions and latest technological trends*, *Computer Network*, vol. 51, pp. 3448 – 3470.

Roberts S. W. (2000) Control chart tests based on geometric moving averages, *Technometrics*, vol. 42, pp. 97–101, Feb.

RT (2012) *Flame virus explained: How it works and who's behind it*, tech. rep.,[Online], Available: http://rt.com/news/flame-virus-cyber-war-536/ [May 2012].

Simmross-Wattenberg F., Asensio-Perez J., Casaseca-de-la Higuera P., Martin-Fernandez M., Dimitriadis I., and Alberola-Lopez C. (2011) *Anomaly detection in network traffic based on statistical inference and alpha-stable modeling*, *Dependable and Secure Computing, IEEE Transactions on*, vol. 8, pp. 494 –509, July-Aug.

Tombini E., Debar H., Me L., and Ducasse M. (2004) *A serial combination of anomaly and misuse IDSs applied to http traffic*, in *Proceedings of the 20th Annual Computer Security Applications Conference*, ACSAC '04, (Washington, DC, USA), pp. 428–437, IEEE Computer Society.

Valdes A. and Skinner K. (2001) *Probabilistic alert correlation*, in *Proceedings of the 4th International Symposium on Recent Advances in Intrusion Detection (RAID 2001)*, pp. 54–68, 2001.

Viinikka J., Debar H., M´e., L., and S´eguier R. (2006) *Time series modelling for IDS alert management,* in *Proceedings of the 2006 ACM Symposium on Information, computer and communications security*, ASIACCS '06, (New York, NY, USA), pp. 102–113, ACM, 2006.

Ye N., Chen Q., and Borror C. (2004) *EWMA forecast of normal system activity for computer intrusion detection*, Reliability, IEEE Transactions on, vol. 53, pp. 557 – 566, Dec.

Ye N., Li X., Chen Q., Emran S., and Xu M. (2001) *Probabilistic techniques for intrusion detection based on computer audit data*, Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, vol. 31, pp. 266 –274, July.

Ye N., Vilbert S., and Chen Q. (2003) *Computer intrusion detection through EWMA for autocorrelated and uncorrelated data*, Reliability, IEEE Transactions on, vol. 52, pp. 75 – 82, March.

Applicability of Cultural Markers in Computer Network Attack Attribution

Charmaine Sample
Capitol College, Laurel, Maryland, USA
charsample50@gmail.com

Abstract: Computer Network Attack (CNA) attribution presents on going challenges for information security professionals. The distributed nature of the Internet combined with various anonymizing technologies contributes to making the attribution problem more difficult, especially when traversing hostile networks. What is needed is a new way to assist in attribution performance; this method must be technology independent. Culture offers a technology independent vector for analysing CNAs. The human mind uses both conscious and unconscious thought, and both of these processes are culturally influenced. This researcher seeks to determine if those cultural influences leave traces in CNA choices and behaviours. Geert Hofstede's cultural dimensions provide a framework for evaluating and understanding various behaviours. Hofstede's framework has been used in academia and business for research in order to better understand other cultures. Hofstede avails his data for researchers in all disciplines. The goal of this study is to determine if Hofstede's framework can be applied to the cyber environment in order to understand CNAs with the hope of greater understanding of cyber adversary choices and behaviours. The preliminary findings support the hypothesis: culture influences CNA choices and behaviours. Two sets of data were examined across all six cultural dimensions. The analysed data displayed statistically significant findings across three dimensions: power distance, individualism versus collectivism, and indulgence versus restraint. The tests performed were quantitative and included means comparison tests for the first data set, and group comparison tests in the second data set. The findings revealed valuable data in both the easily seen visible results, and in the areas that lacked data. These findings suggest that culture not only influences CNA choices and behaviours, but may also influence non-behaviours. The results of this research study suggest the need for additional research targeted toward specific cultural dimensions.

Keywords: computer network attack (CNA) choices and behaviours, Hofstede, cultural dimensions, automatic thought

1. Introduction

Computer Network Attack (CNA) attribution, "determining the identity or location of an attacker" (Wheeler & Larsen, 2003, p. 1) continues to challenge security professionals due to the various "stepping-stone" (Zhange, Persaud, Johson & Guan, 2005, p.1) and anonymizing techniques and products. These anonymizing solutions have resulted in a game of "cat and mouse" between security professionals and attackers where no significant progress is made in solving the CNA attribution problem. New approaches are needed in order to change the dynamics of attribution. Attack attribution beyond IP addresses offers a paradigm change. This study examines the CNA attribution problem from a different perspective; culture.

Hofstede et al., (2010) defines culture as "software of the mind" or "the mental programming" that defines a group of people (Hofstede et al., 2010). Hofstede's work is widely used by researchers in various industries. Most recently Doctor Dominick Guss (2011, 2004) used Hofstede's data in order to determine the role of culture in complex problem solving and dynamic decision-making. Yu & Yang (2009) extended the model for technology innovation. This researcher seeks to build the foundation for additional research in the role of culture in CNA by examining some well-known attack behaviours through the prism of Hofstede's cultural dimensions. Hofstede et al., (2010) quantified culture and operationalized data in six different dimensions: power distance (pdi), individualism versus collectivism (ivc), masculine/feminine (m/f), uncertainty avoidance (uai), long term orientation versus short term orientation (LTOvSTO), and indulgence versus restraint (ivr).

Two attack behaviours that are examined across the various dimensions are the aggressive, nationalistic, patriotic themed website defacements and an examination of individual hackers. The results of the collected data are compared with the values Hofstede associates with the general population in order to determine if statistically speaking, results correlate with certain ranges in various dimensions.

2. Literature review

Dijksterhuis (2004) observes, "a little introspection reveals that the processing capacity of consciousness is limited. People are not able to concentrate consciously on two different things simultaneously" (Dijksterhuis, 2004, p. 587). The nature of CNA attacks requires rapid complex thought; therefore, the attacker must rely on both conscious and unconscious thought.

Bargh and Morsella (2008) observed that culture permeates thought both conscious and unconscious. Baumeister & Masicampo (2010) blend the unconscious and conscious distinction, “conscious thought is for incorporating knowledge and rules for behaviour from culture. Over time, automatic responses then come to be based on that new input “ (Baumeister & Masicampo, 2010, p.948). Thus, many of the cultural influences over thought become engrained even as a part of conscious thought.

Buchtel and Norezayan (2008) observed the differences between eastern and western cultures along with the role of education in developing automatic behaviour. Buchtel & Norezayan (2008) noted the difference in contextualization between eastern and western cultures. Furthermore, Buchtel & Norezayan (2008) observed that culture influenced thought patterns. “The cultural differences are best conceptualized as difference in habits of thought, rather than differences in the actual availability of information processing” (Buchtel & Norenzayan, 2008, p. 219).

Guss (2004) also observed in microworlds simulations that culture played a significant role in problem solving. Guss (2004) stated “culture can influence the perception of the problem, the generation of strategies and alternatives, and the selection of one alternative” (Guss, 2004, p.6). Guess used Hofstede’s cultural dimensions in order to define culture.

Minkov (2013) says “culture has an independent existence, … [culture] can be studied independently of its carriers: the human beings” (Minkov, 2013, p. 15). This frees up the researcher to study culture as it relates to any issue. Hofstede’s dimensions make possible for this researcher to quantitatively determine the relationship between culture and CNA choices and behaviours. The study examines some general CNA behaviours in the context of all six cultural dimensions. An explanation of each dimension follows; with an emphasis on education, and technology use as this behaviour apply to each dimension.

2.1 The power distance index (PDI) dimension

The pdi dimension measures the measure of equality within a society. “Power distance can therefore be defined as the extent to which less powerful members of institutions and organizations within a country expect and accept that power is distributed unequally” (Hofstede et al., 2010, p. 61). Problem in high power distance societies are resolved by a show of power and in egalitarian societies problems are solved by flexibility (Hofstede et al., 2010, p. 63). Minkov provides the following observation on power distance: “Generally speaking, power distance is about treating people differently, depending on their group membership” (Minkov, 2013, p. 414).

When pdi is examined in terms of hacker behaviour nationalistic, patriotic hacking has been observed by Chinese Hackers (London, 2011, Chan 2005, Qiu 2003). In a high pdi society the hackers depend more on each other than themselves (Chan 2005). Characteristics of high pdi societies are loyalty and protection. Hofstede et al. (2010) acknowledges the nation of this high power distance relationship, “the junior partner owes the senior respect and obedience, while the senior partner owes the junior protection and consideration” (Hofstede et al., 2010, p. 80). In cyber terms this implies that the hackers act out of a sense of loyalty against opposing nations while acting with the knowledge that their government will provide them with protection.

Searches on nationalistic, patriotic themed website defacements through scholar.google.com resulted in close to 20 countries being represented. Hofstede et al., provide pdi values for 78 countries. Of the 20 countries some had to be eliminated due to not being found in Hofstede’s list, and others due to a lack of supporting reports, such as academic studies, or even news stories. The following countries were identified as having participated in nationalistic patriotic themed website defacements, Bangladesh, China, India, Indonesia, Iran, Israel, Malaysia, Pakistan, Philippines, Portugal, Russia, Singapore, Taiwan, Turkey.

2.2 The individualism versus collectivism (IVC) dimension

Individualism versus collectivism deals with how the individual relates to the larger group known as the society. “Individualism pertains to societies in which the ties between individuals are loose; everyone is expected to look after him- or herself and his or her immediate family” (Hofstede et al., 2010, p. 92). In the individualist society, the individual is responsible for his or her own personal growth and success.

In the collectivist society the needs of the group, or the collective are always considered first and above the needs of the individual. “Collectivism as its opposite pertained to societies in which people from birth onward are integrated into strong, cohesive in-groups” (Hofstede et al., 2010, p. 92). Not only is the individual supposed to consider the group first, but also the individual must avoid direct confrontation. “In most collectivist cultures direct confrontation of another person is considered rude and undesirable. The word *no* is seldom used, because saying “no” *is* a confrontation; “you may be right” and “we will think about it” are examples of polite ways of turning down a request” (Hofstede et al., 2010, p. 106).

Collectivism provides a moderating influence over behaviours. “It seems clear that groups and individuals make different decisions in strategic games and, more often than not, group decisions are closer to the ‘rational’ solution (Bornstein et al., 2003, p. 604). While collectivism moderates behaviours individualism does not. “The research on experimental games has uncovered many instances in which individuals deviate systemically from the game-theoretic prediction” (Bornstein et al., 2003, p.604). Additionally, when forced to go against their cultural upbringing individualist perform worse. The American individualist participants performed best when operating individually and with their names marked but abysmally low when operating as a group and anonymously” (Hofstede et al., 2010, p.121).

In terms of technology, collectivism can be seen to interfere with creativity. “The Golden Mean value is advantageous for the construction of harmonious society. But it advocates maintaining present situation and denies transformation which seriously influences technological innovations” (Yu & Yang, 2009, p. 462). This relationship with creativity suggests that if innovation is a problem in collectivist countries, then improvement abilities by collectivist societies may be considered a positive outcome.

2.3 The masculine feminine (M/F) dimension

This dimension deals with gender roles in the society. “A society is masculine when emotional gender roles are clearly distinct: men are supposed to be assertive, tough, and focused on material success, whereas women are supposed to be more modest, tender, and concerned with the quality of life. A society is called feminine when emotional gender roles overlap: both men and women are supposed to be modest, tender and concerned with the quality of life” (Hofstede et al., 2010, p. 140).

This dimension can sometimes be misunderstood. Men from feminine countries are not effeminate, nor are women from masculine countries masculine. Instead, this dimension deals with how conflict is handled. “Masculine countries tend to (try to) resolve international conflicts by fighting; feminine countries by compromise and negotiation” (Hofstede et al., 2010, p. 173). This lack of negotiation along with escalation of attack activities might be viewed as masculine behaviour. Consider the on going cyberwar between the United States and China where neither side appears to be willing to negotiate and both countries share masculine scores, US 62, and China 66.

2.4 The uncertainty avoidance (UAI) dimension

The fourth dimension for examination is uncertainty avoidance. “Uncertainty avoidance can therefore be defined as the extent to which the members of a culture feel threatened by ambiguous or unknown situations” (Hofstede et al., 2010, p. 191). People in low uncertainty avoidance cultures view the new as curious in contrast to their counterparts in high uncertainty avoidance cultures that view the new as fearful.

One area where the dimensional differences can be clearly seen is in education. Hofstede compared learning in England to Germany. Germany scores in the middle to high range for this dimension and England has a relatively low score. German students preferred the learning environment more structured, and the British preferred open-ended (Hofstede et al., 2010). Hofstede further noted that one characteristic in the high uncertainty avoidance culture is precision. “Most Germans, for example, favoured structured learning situation with precise objectives, detailed assignments, and strict timetables. They liked situations in which there was one correct answer that they could find. They expected to be rewarded for accuracy” (Hofstede et al., 2010, p. 205).

This dimension offers some potentially interesting behaviour in the cyber environment. Consider the malware program Flame. Flame was allegedly a joint effort between the US and Israel (Zetter, 2012, Nakashima, 2012).

One distinguishing feature of Flame was the use of a collision. Collisions are categorized by Mitre (capec.mitre.org) as a type of probabilistic attack. Probabilistic attacks inherently have an element of uncertainty built into them and not surprising the US has a low uncertainty avoidance score of 46. Even more interesting is the precision also associated with Flame, not surprisingly Israel has a high uncertainty avoidance score of 81. This dimension offers many additional research opportunities in Cybersecurity that this researcher hopes to explore.

2.5 The long-term orientation versus short-term orientation (LTOvSTO) dimension

"Long-term orientation (LTO) stands for the fostering of virtues oriented toward future rewards—in particular, perseverance and thrift" (Hofstede et al., 2010, p. 239). LTO focuses on the distant time horizon. Short-term orientation (STO) deals with a preference for more immediate gratification or returns.

In terms of thinking the differences between LTO and STO exists and have been documented. Hofestede et al. (2010) note the difference between the analytical Western thought process and the synthetic, holistic Eastern thought process. "Western analytical thinking focused on elements, while Eastern synthetic thinking focused on wholes" (Hofstede et al., 2010, p. 250). This statement by Hofstede et al. (2010) is consistent with the work by Buchtel & Norezayan (2008). Buchtel & Norezayan (2008) attribute the difference between analytical and holistic thinking to culture. This cultural dimension deals primarily with strategies, and provides an interesting research area but will not be a primary focus area for this research effort.

2.6 The indulgence versus restraint (IVR) dimension

The final cultural dimension deals with indulgence versus restraint (IVR). "Indulgence stands for a tendency to allow relatively free gratification of basic and natural human desires related to enjoying life and having fun" (Hofstede et al., 2010, p. 281). The opposite pole of this spectrum deals with moderation or restraint. "The items that defined the positive pole of this dimension were 'moderation', 'keeping oneself disinterested and pure', and 'having few desires'" (Hofstede et al., 2010, p. 288). Cynicism and other negative emotions are often times associated with the restraint pole of this dimension.

Certain behavioural consistencies have been observed. One such area deals with math, science and logical reasoning. "Societies whose children are better in mathematics are also societies whose children are better in science, in logical reasoning, and in reading. Success in all these domains is closely associated with weak monumentalism and strong flexhumility, even after taking into account the role of national wealth" (Minkov, 2011, p. 102). Conversely, a negative correlation exists between monumentalism and math performance. "The more monumentalist a particular society is, the lower its achievement in mathematics" (Minkov, 2011, p. 101).

Minkov (2011) suggests that the desire to finish first and be considered the best may lead to goals that foster superficial learning. Unlike intrinsically motived learning, superficial learning is extrinsically motivated (Minkov, 2011). Minkov (2011) provides a simple distinction between indulgent Americans and restrained Asians and Eastern Europeans. "Americans like to receive compliments. But in Japan and China, just like Eastern Europe, personal praise often causes embarrassment" (Minkov, 2011, p. 95).

Not surprising, most attacks that are named after their designers also coincide with indulgent societies. For example, the Kaminsky bug and the Morris worm were both authored by Americans one of the more indulgent societies. Certain type website defacements may also contain an indulgent behavioural component. An example of indulgent behaviour can be seen on the MI6 attack on an al-Qaeda website, where bomb making instructions were replaced by baking instructions. (Gardham, 2011),

The use of the Hofstede defined dimensions allows for a widely recognized framework for evaluating the behaviours, and a set of metrics for quantitative analysis. In spite of globalization and widespread use of the Internet the users are still educated within the context of their cultures. The statement by Hofstede et al., (2010) "software of the machines may be globalized, but the software of the minds that use them is not" (Hofstede et al., 2010, p.391), provides the launching point for the research, for if this statement is true, then a new vector for attack attribution may become available.

3. Methodology

This study consists of experiments using two different data sets being quantitatively compared and analysed for statistical significance using Z testing or Mann-Whitney U testing to obtain p-values. In order to test this hypothesis, culture and behaviours are decomposed into specific research questions. Two CNA activities are examined across all six dimensions.

The data used for comparison consists of primary and secondary data. The primary data, raw data, is provided by Hofstede and consists of his scoring results for 78 countries across six dimensions. Secondary data is used from academic peer reviewed articles, periodicals, news sites and web sites. One other source of data is Internet population data obtained from the Internet World Stats website. Additional information on the methodology will be explained in each subsection.

3.1 Data set one

The website www.zone-h.org provides examples of website defacements and a starting point for collecting data on nationalistic, patriotic themed website defacements. The use of scholar.google.com as a search engine resulted in reports describing attacks by various countries. Because the reports were analysis of attack behaviours and choices, a country either engages in the behaviour or does not. Participation is only scored once. Testing for the p-value will rely on means tests.

The first research question asks, do nationalistic, patriotic-themed website defacements correlate with high power distance societies? The comparisons will be made to the mean score across each dimension. Resulting p-value scores will use the 0.05 rules for statistical significance measure. The following data will be used with the following equation $Z_n = (\mu - \overline{x})/(\sigma/\sqrt{n}$, unless the data is not normally distributed, then the Mann-Whitney U test will be used.

Table 1: Sample data set with Hofstede dimensional scores

Country	PDI	IVC	M/F	UAI	LTOvSTO	IVR
Bangladesh	80	20	55	60	47	20
China	80	20	66	30	87	24
India	77	48	56	40	51	26
Indonesia	78	14	46	48	62	38
Iran	58	41	43	59	14	40
Israel	13	54	47	81	38	Null
Malaysia	104	26	50	36	41	57
Pakistan	55	14	50	70	50	0
Philippines	94	32	64	44	27	42
Portugal	63	27	31	104	28	33
Russia	93	39	36	95	81	20
Singapore	74	20	48	8	72	46
Taiwan	58	17	45	69	93	49
Turkey	66	37	45	85	46	49
Population Mean/Std Dev.	59 21.25070	45 23.97152	49 19.32747	68 22.99296	45 24.2320162	45 22.29343

3.2 Data set two

The second data set deals with frequencies. A search was performed on the top 20 all time black hat hackers. Only those hackers from after 1993 were accepted because the Internet was not fully global before that time. The expected results are a representative sample created from the data found at www.internetworldstats.com Table 3 contains the actual results, actual frequencies and the cultural dimension scores.

Table 2: Representative sample of countries and frequency

Country	PDI	IVC	M/F	UAI	LTOvSTO	IVR	Frequency
China	80	20	66	30	87	24	3
India	77	48	56	40	51	26	2
Japan	54	41	95	92	88	42	1
Russia	93	39	36	95	81	20	1
Germany	35	67	66	65	83	40	1
UK	35	89	66	35	51	69	1
US	40	91	62	46	26	68	2
Iran	58	41	43	59	15	40	1
Brazil	69	38	49	49	44	59	1
Mexico	81	30	69	82	24	97	1
Africa West	77	20	46	54	9	78	1

Table 3: Actual top hacker's country and frequency

Country	PDI	IVC	M/F	UAI	LTOvSTO	IVR	Frequency
US	40	91	62	46	26	68	6
UK	35	89	66	35	51	69	4
Russia	93	39	36	95	81	20	1
Germany	35	67	66	65	83	40	1
Canada	54	73	45	73	Null	Null	1
Philippines	94	32	64	32	27	42	1
Greece	60	35	57	112	45	50	1

4. Results: Figures and tables

Results from the tests for the first data set are presented in Table 4. These results were achieved by comparing mean scores. The Mann-Whitney test was used most often, to determine the Z score and the p-value.

Table 4: Data set 1 results

Dimension	Z Score	p-value	Hypothesis	Alternative Hypothesis
PDI	2.42	0.0281	Reject	Accept
IVC	-2.35	0.015	Reject	Accept
M/F	0.5714	0.4247	Accept	Reject
UAI	-1.33	0.123	Accept	Reject
LTOvSTO	1.15	0.1251	Accept	Reject
IVR	-1.51	0.0655	Accept	Reject

Hypothesis: Nationalistic, patriotic themed website defacements correlate with low power distance societies. H_0: $u_0 < 59$ and Alternative Hypothesis: Nationalistic, patriotic themed website defacements correlate with high power distance societies: H_1: $u_0 >= 59$

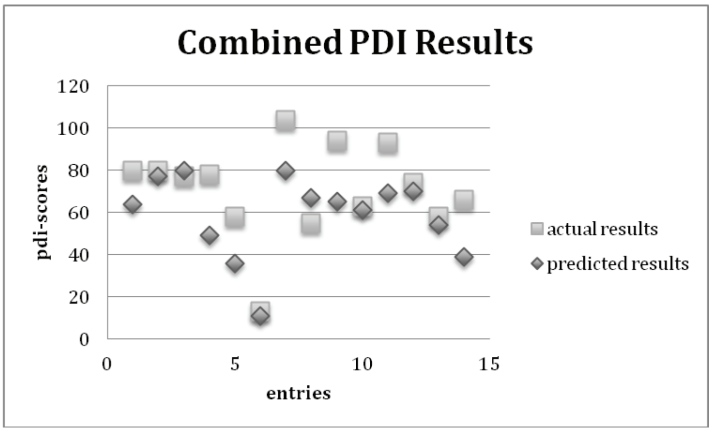


Figure 1: Actual and predicted PDI results

Hypothesis: Nationalistic, patriotic themed website defacements correlate with individualist societies. H_2: $u_0 > 45$ and Alternative Hypothesis: Nationalistic, patriotic themed website defacements correlate with collectivist societies. H_3: $u_0 <= 45$

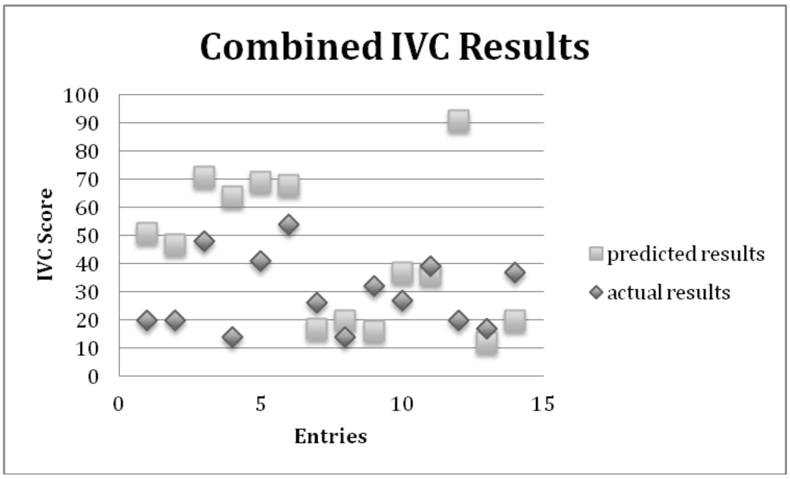


Figure 2: Predicted and actual IVC results

Hypothesis: Nationalistic, patriotic themed website defacements correlate with indulgent societies. H_4: $u_0 < 45$ and Alternative Hypothesis: Nationalistic, patriotic themed website defacements correlate with masculine societies. H_5: $u_0 >= 45$

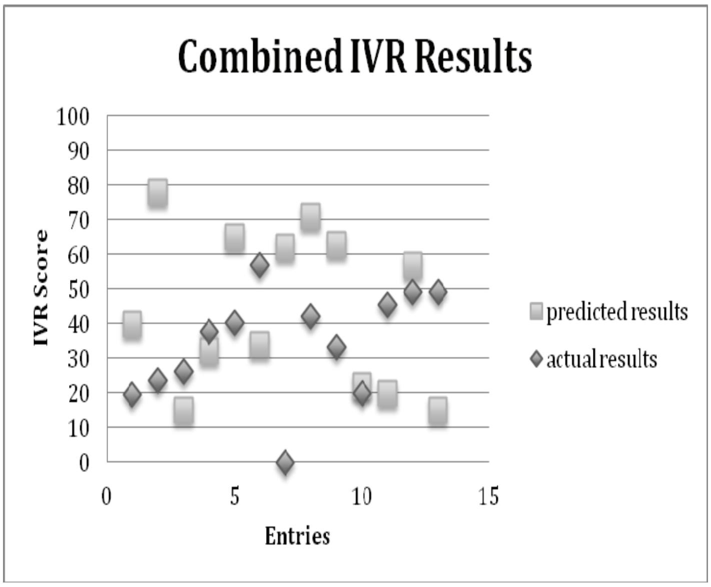


Figure 3: Predicted and actual IVR results

The second set of data deals with the frequency of individual "lone wolf" attacks, by examining the top hacker lists. The pooled estimate is represented by p̂, Table 5 contains the results for each dimension.

Table 5: Data set 2 results

Dimension	Z =	p-value	Hypothesis	Alternative Hypothesis
PDI	-2.03	0.0212	Reject	Accept
IVC	2.53	0.0057	Reject	Accept
M/F	0.04	0.484	Accept	Reject
UAI	0	0.5	Accept	Reject
LTOvSTO	0.85	0.1977	Accept	Reject
IVR	1.55	0.0606	Accept	Reject

5. Conclusion

The results indicate that culture does appear to play a role in CNA choices and behaviours. This initial analysis is simply a “proof of concept”. The PDI and IVC dimensions appear to support this initial hypothesis.

Often times what is not seen is equally as important as what is seen. This appears to be the case in this area of research. For example, when examining the IVC dimension with the first data set, not only did the results support the collectivist tendency, but also what was not seen was the lack of individualist results. The highest score in that dimension was below the cut-off for the top 1/3 range.

A similar finding also occurred with the second dataset and the IVR dimension. When both sets of scores a combined the finding suggests that lone wolf attackers are more likely to come from indulgent societies, *and* individual, “lone wolf” hackers are very unlikely to come from countries with strong restraint scores. When examining the first dataset, the restrained pole of the IVC dimension was dominant and representation the indulgent dimension was lacking. This is even more significant because the differences were close to opposite in response to the research questions.

The two simple CNA behaviours that were used in this study were chosen for expediency. This area of research holds a great deal of promise for both offensive and defensive uses. Specific attack types should be evaluated within Hofstede’s cultural dimensions framework in order to determine the validity of applying this framework on a much larger scale. Initial anecdotal evidence appears to support this suggestion, now a more formal study is needed.

The use of cultural markers in CNA attribution offers the promise of being able to augment existing solutions in order to find the true origination point of an attack. This research is in the very early stages and should not be viewed as a replacement to any of the existing attribution technologies. Instead, this research should be used as an additional piece in a very complex puzzle.

References

Bargh, J.A., and Morsella, E., (2008) “The Unconscious Mind”, *Perspectives on Psychological Science,* Volume 3, No. 1, pp.73-79.

Baumeister, R.F., and Masicampo, E.J., (2010) “Conscious Thought is for Facilitating Social and Cultural Interactions: How Mental Simulations Serve the Animal-Culture Interface”, *Psychological Review,* Volume 117, No. 5, pp. 945-971.

Buchtel, E.E. and Norenzayan, A., (2008). “Which Should You Use, Intuition or Logic? Cultural Differences in Injunctive Norms About Reasoning”, *Asian Journal of Social Psychology,* Volume II No.4, pp. 264-273. doi:10.1111/j.1467_839x.2008.00266.x.

Chan, B. (2005). “Imagining the Homeland: The Internet and Diasporic Discourse of Nationalism”, J*ournal of Communication Inquiry,* Volume 29 No.4, pp. 336-368.

Dijksterhuis, A. (2004). “Think Different: The Merits of Unconscious Thought in Preference Development and Decision Making, *Journal of Personality and Social Psychology,* 2005, Volume 7, No.5, pp. 586-598. doi:10.1037/0022-3514.87.5.586.

Gardham, D. (2011, June 2). “Mi6 Attacks al-Qaeda in ‘Operation Cupcake’”. *The Telegraph*. Retrieved from http://www.telegraph.co.uk/news/uknews/terrorism-in-the-uk/8553366/MI6-attacks-al-Qaeda-in-Operation-Cupcake.html on May 5, 2012.

Guess, C.D. (2004). “Decision Making in Individualistic and Collectivist Cultures”, *Online Readings in Psychology and Culture,* Volume 4.

Guss, C.D. (2011). “Fire and Ice: Testing a Model on Culture and Complex Problem Solving. *Journal of Cross-Cultural Psychology,* Volume 42, No. 7, pp. 1279 – 1298. doi: 10.1177/0022022110383320.

Hofstede, G., Hofstede, G.J., and Minkov, M. (2010). *Cultures and Organizations,* McGraw-Hill Publishing: New York, NY.

Internet World Stats website (2013) www.internetworldstats.com.

Minkov, M. (2013). *Cross-Cultural Analysis.* Thousand Oaks, CA: Sage Publications.

Minkov, M. (2011). *Cultural Differences in a Globalizing World.* WA, UK: Emerald Group Publishing Limited.

Nakashima, E., Miller, G., and Tate, J. (2012, June 19). “U.S., Israel Developed Flame Computer Virus to Slow Iranian Nuclear Efforts, Officials Say”, *The Washington Post.* Retrieved from http://articles.washingtonpost.com/2012-06-19/world/35460741_1_stuxnet-computer-virus-malware on July 2, 2012.

Qiu, J. L. (2003, October). “The Internet in China: Data and Issues”, In *Annenberg Research Seminar on International Communication,* Volume 16. Retrieved from: http://www.usc.edu/schools/annenberg/events/normanlearcenter/icbak/Papers/JQ_China_and_Internet.pdf on November 4, 2012.

Wheeler, D.A., & Larsen, G.N. (2003). *Techniques for Cyber Attack Attribution* (No. IDA-P-3792). INSTITUTE FOR DEFENSE ANALYSES ALEXANDRIA, VA. Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA468859 on January 7, 2012.

Yu, Y., and Yang, Q. (2009). "An Analysis of the Impact Chinese and Western Cultural Values Have on Technological Innovation", *Second International Workshop on Knowledge Discovery and Data Mining,* Jan 23-25, 2009, pp. 460-463 doi: 10.1109/WKDD.2009.149.

Zetter, K. (2012. May 28). "Meet 'Flame' the Massive Spy Malware Infiltrating Iranian Computers", *Wired.* Retrieved from http://www.wired.com/threatlevel/2012/05/flame/ on May 28, 2012.

Zhang, L., Persaud, A., Johnson, A., and Guan, Y. (2005, February). "Stepping Stone Attack Attribution in Non-Cooperative IP Networks", *Proceedings of the 25th IEEE International Performance Computing and Communications Conference,* (IPCCC 2006).

Zone-h website (2012). http://www.zone-h.org.

Zone-h mirror .website (2012). http://www.zone h.org/mirror/id/18480993, http://zonehmirrors.net/defaced/2012/09/01/razor-fragzone.uw.hu/ on

Securing Complex System-of-Systems Compositions

Nathan Shone, Qi Shi, Madjid Merabti and Kashif Kifayat
Liverpool John Moores University, Liverpool, UK
n.shone@2007.ljmu.ac.uk
q.shi@ljmu.ac.uk
m.merabti@ljmu.ac.uk
k.kifayat@ljmu.ac.uk

Abstract: In modern computing, the threat of cyber-warfare and the emergence of hacking syndicates have raised concerns regarding the security of complex systems. Systems under particular scrutiny include those involved in critical infrastructures and large public organisations. Recent high profile attacks and security breaches have highlighted the significance of security research. The increasing complexity of modern systems is largely due to the rising demand for functionality. However, systems are expected to retain their security, integrity and usability, which can often be affected by complexity. System-of-Systems (SoS) is a concept aiming to address these issues whilst increasing functionality beyond the capabilities of individual systems. It involves creating a large-scale decentralised super-system, composed of multiple independent component systems. Components voluntarily collaborate by pooling resources and sharing services, and are able to join, leave and change their contribution spontaneously. In complex environments, it is often necessary to lower security to achieve a compromise with functionality. This can lead to weaknesses being created or exposed. In an unpredictable SoS environment, which supports emerging behaviour, this is a major security risk, and is vulnerable to component misbehaviour. Real-time misbehaviour detection is a vital security feature of any SoS, requiring a solution that is lightweight and consumes minimal system resources, whilst providing adequate accuracy and indication of misbehaviour. However, the dynamic and uncertain nature of such systems, combined with their undefined boundaries, poses significant challenges. Many existing monitoring solutions and methods utilised in static systems are unable to function effectively in SoS environments. This paper presents our statistical anomaly detection based solution, to address the problem of detecting component misbehaviour on complex and uncertain systems in real-time. It is also able to make informed decisions regarding the degree of irregularity of anomalous behaviour. The paper also examines problems with existing solutions and presents some our promising initial results.

Keywords: component misbehaviour, system-of-systems, monitoring

1. Introduction

System-of-Systems (SoS) are increasingly characterised by large scale, decentralisation, complexity and uncertainty. A SoS can involve the voluntary collaboration of thousands of component systems with varying roles, contributions and abilities, thus producing diverse workload characteristics. The resultant system possesses both dynamic and uncertain structure and function, whilst lacking effective security and trust mechanisms. These facts make detecting behavioural anomalies on a SoS component system difficult (Agrawal 2009), considering the criticality of some linked applications.

SoS is an emerging technology involving the creation of a decentralised super-system, whose component systems voluntarily contribute to achieve a shared goal. It enables levels of functionality that are unachievable on stand-alone systems, as well as reducing complexity for end users. There is no traditional hierarchy, security or trust mechanisms in a SoS, and components can join, leave and change their contribution at any time. These features place greater reliance on an equilibrium of inter-component trust, which must be maintained. This trust is susceptible to abuse and can lead to catastrophic consequences, such as reduced performance, components leaving or even total SoS collapse. Despite this, the benefits of a SoS are abundant and applications have been linked to military (DiMario 2006), healthcare and aerospace (Jamshidi 2008) systems.

To facilitate connectivity, collaboration or functionality in complex and heterogeneous environments such as a SoS, it is often necessary to make system changes (Maier 2006). These changes often include a reduction in, or changes to security, despite the good intentions this often leads to complications. Such changes can create or expose weaknesses in the system, which can cause or permit misbehaviour, or can be exploited by emerging behaviour.

Misbehaviour raises many concerns in terms of SoS functionality and integrity, component security, reliability of contributed services and repercussions for system owners. However, most monitoring solutions lack the ability to discern between the dynamic behaviour in a SoS environment and genuine misbehaviour (Efatmaneshnik et al. 2012). Our work aspires to improve the accuracy of misbehaviour detection on SoS component systems by developing a real-time monitoring solution that (i) effectively detects misbehaviour whilst able to account for dynamics of the environment, (ii) assesses the irregularity and threat of the misbehaviour, (iii) adjusts thresholds to account for system changes.

A key area of research related to our work is anomaly detection, which aims to identify system behaviour not conforming to an expected norm. It is of particular importance to our work due to its capabilities of detecting novel threats. In an environment with undefined boundaries, that encourages the development of emerging behaviour, this is an essential characteristic. Our interests lie with methods suitable for continuous monitoring, rather than those using static system profiles or definitions of anticipated misbehaviour. Whilst designing a monitoring solution suitable for use in a SoS, there are several challenges to overcome. These include *scalability*, whereby methods must meet the demands of a constantly changing system. Methods must therefore be '*lightweight*' in terms of physical size, the number of metrics required to operate and the runtime resource requirements for the detection and analysis of misbehaviour. In addition, methods must be able to cope with the *dynamics* of a SoS environment, as there is no discernible normality or predictability. Dynamics are also evident in the fluctuation in system workload and changes in composition structure and functionality. They also need to be able to cope with alterations made to the system. Any methods considered must have the ability to utilise metrics from multiple levels of *abstraction* including hardware, operating system and applications. Methods must operate in *real-time,* as the criticality of linked applications cannot allow for any time delay, whilst providing protection. Importantly, the methods must have a high level of *accuracy*, with low false positive and false negative rates. Finally, our solution must establish *optimality* by balancing accuracy, resource requirements (e.g. storage), performance impact and reliability.

Modern approaches to anomaly detection involve using fixed thresholds for the groups or individual metrics being monitored. These thresholds are generally calculated offline using training data and remain static during the entire process. Derivative techniques such as Multivariate Adaptive Statistical Filtering (MASF) (Buzen & Shum 1995) maintain an additional threshold, detailing data according to time (e.g. day of week). However, the main problem with this approach is that it assumes the data distribution used to calculate thresholds is Gaussian; however, this can often be untrue. Also, this approach cannot adapt thresholds to reflect changes in the system or its behaviour, which leads to increased false readings and decreased accuracy and efficiency.

In this paper, we propose our novel solution, which overcomes the problems and limitations outlined above, hence improving the accuracy and reducing the number of false positives and false negatives. We make the following contributions:

- A novel framework to detect misbehaviour on SoS components.
- A statistical method to calculate thresholds for monitored metrics and can account for contribution to the SoS and the dynamics of the system.
- A statistical method to analyse anomalous behaviour and calculate a score indicating the degree of irregularity by observing behaviour of the target metric and metrics with proven relationships.
- A statistical method to analyse and refine thresholds, to account for behavioural changes in the system.

The remainder of this paper is structured as follows. Section 2 provides background information on existing techniques. Section 3 provides a brief overview of our framework, whilst Section 4 describes our proposed statistical techniques that overcome the problems outlined. Section 5 presents our experimental results and the paper concludes in Section 6.

2. Background

SoS is an emerging technology with no universally agreed definition (Xiao et al. 2011). Hence why many researchers have proposed methods to differentiate between a SoS and traditional systems (Boardman & Sauser 2006). The majority of existing research focuses on distinguishing a SoS from a traditional system as well as looking at the creation, management and maintenance of SoSs (Trivellato et al. 2011), (Gorod et al.

2008). There is limited research into SoS security, most of which focuses on theoretical security frameworks based on either semantics (Agrawal 2009) or trust management (Selberg & Austin 2008).

Peer-to-Peer (P2P) systems share some structural similarities with a SoS; therefore, we have also examined the applicability of solutions from this area of research. The majority of P2P monitoring solutions are based on the calculation of a score, indicating the reputation or trust of a particular node. Neighbouring nodes or higher ranking nodes are used to monitor and compute these scores based on interactions or observations (Jin & Chan 2010; Visan et al. 2011). There are several problems when applying this approach to a SoS. This method relies on prior interaction with the target node, but this poses questions of whether it is safe to assume the component is secure and if this could jeopardise critical SoS implementations. This approach relies on the integrity of components and assumes they are capable of assessing the dynamic behaviour (e.g. role or contribution changes) in a heterogeneous environment.

The term misbehaviour can be interpreted in different ways, but for our work, we define it as atypical behavioural anomalies, which are exhibited by a component system. Misbehaviour falls into two main classifications, deliberate and accidental. Deliberate misbehaviour is an event with malicious intentions, e.g. corruption of contributed services, excessive consumption of resources and theft of data. Accidental misbehaviour is an event unintentionally leading to irregular behaviour, e.g. components unable to function due to connectivity issues, malfunction or incorrect configuration. Misbehaviour can be identified in many ways but two of the most common approaches are deviation from a normal pattern of behaviour and changes in behavioural distribution over time. The *first* approach involves establishing a set of upper and lower thresholds, normal data will fluctuate within these limits but anomalous data will lie outside. The calculation of these thresholds is based on assumptions of the shape (or behaviour) of the data distribution and knowledge gained from studying historical data. This approach is known as parametric threshold creation. The *second* approach deals with the variability of data over an extended time period. An estimated variability of data is characterised and threshold limits calculated by using historical data. This approach avoids making assumptions regarding the shape of data distribution; this is known as non-parametric threshold creation.

Many approaches using statistical analysis of data often rely on Gaussian (normal) distribution as the probability distribution model (C. Wang et al. 2011). One such method prominent in anomaly detection is MASF (Buzen & Shum 1995), whereby thresholds are computed using the standard deviation of time-segmented data. Using Gaussian distribution, it is assumed 95% of data is within two standard deviations of the mean and 99% of data is within three standard deviations of the mean. The probability of data lying outside of three is 0.27%, thus deeming it a rare event and therefore anomalous. Overall, Gaussian assumptions are generally adhered to, however there can be exceptions. It is advantageous to use methods that do not rely on restrictive normality assumptions, such as those presented in Section 4.

3. Framework overview

In this section, we will provide a brief overview of the design of our proposed framework solution. To solve the problem of detecting misbehaviour on SoS components we have designed our novel solution, called Secure SoS Composition (SSC) Framework. The primary aim of our solution is to provide real-time protection to compositions and component systems against misbehaviour. SSC is a host-based solution, using a statistical anomaly detection approach to monitor for misbehaviour. Each component is independent and therefore responsible for monitoring its own behaviour. We propose that SSC is installed on component systems as a compulsory requirement. An overview of SSC is illustrated in Figure 1.

SSC calculates a threshold profile using our novel threshold calculation algorithm (Section 4.1), which combines parametric and non-parametric threshold calculation techniques. The threshold profile is a database containing extensive twenty-four hour profiles, detailing the expected behaviour of all the monitored system metrics. It consists of two separate sets of thresholds, S2T and DA. The training data used to calculate these thresholds is gathered by logging all monitored metrics before joining a SoS, over a ten-day period.

S2T: This is an average behaviour threshold set, integrated with the promised levels of SoS contribution. Components maintain a S3LA configuration file, which specifies the contribution and limitation of each observed metric (e.g. max / min RAM to contribute). This results in a behavioural profile detailing the expected normal system use and SoS contribution.

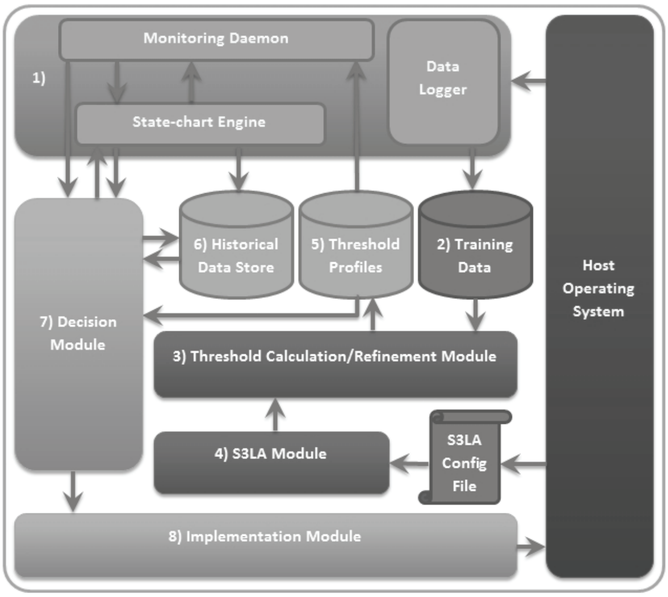


Figure 1: Overview of SSC framework

Dynamically adjusted (DA): This threshold set is an evolution of the S2T set, as it is adjusted to compensate for the expected data variability. Our novel threshold adjustment algorithm (Section 4.2) routinely adjusts this threshold set, by reviewing trending behavioural patterns and adjusting thresholds to account for them. This method works by observing metric readings that fall into Zone 1 (Between S2T Max and DA Max) and Zone 2 (Between S2T Min and DA Min) as illustrated in Figure 2.

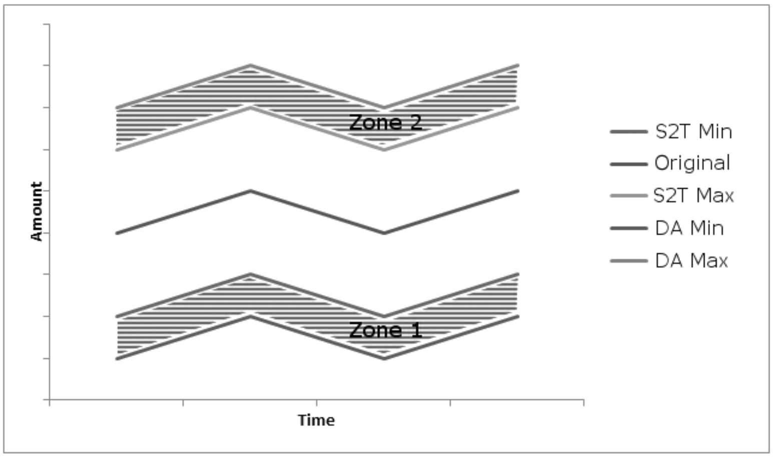


Figure 2: Illustration of threshold profile

Live monitoring data from the component system is monitored against the DA threshold set, by observing metrics using both categorical measures (activity distribution over metric categories) and ordinal measures (e.g. CPU usage) to identify problems in real-time.

Due to the complexity of the environment, any behaviour not conforming to the threshold profile is not automatically treated as misbehaviour. Instead, anomalous behaviour detected is sent to our novel decision algorithm (Section 4.4). This calculates a certainty score based on a scale from 0 (abnormal) to 1 (normal) to indicate the degree of irregularity for each reported behavioural event. To make an accurate decision regarding behaviour on a dynamic system, multiple sources of evidence are utilised. Our algorithm performs several statistical tests on the reported metric and behaviourally related metrics. Behavioural similarity is determined using weighting scores calculated using Kendall's Tau to represent the strength of the relationship. The certainty score is then utilised to determine relevant action based on the severity of the certainty score (e.g. limiting services or component disconnection).

SSC has an integrated state-chart engine, which helps to reduce unnecessary load, system footprint, required storage and improves monitoring performance. The state-chart engine uses four states (Normal, Low, High and Disconnected) to assess the current risk level of the system, as illustrated in Figure 3. The engine measures the categorical distribution of both high and low risk certainty scores, over a period of time. Each categorical group is assigned a limit of how many high and low risk certainty scores it can possess; if either of these limits are reached, the risk level is then raised accordingly. Each group is also assigned a time-out value; this is the period of time the scores recorded against the group are kept. Once this period has elapsed, and providing there is no re-occurrence, the score is deleted. This automated approach to threat level management helps to account for the dynamic behaviour encountered.

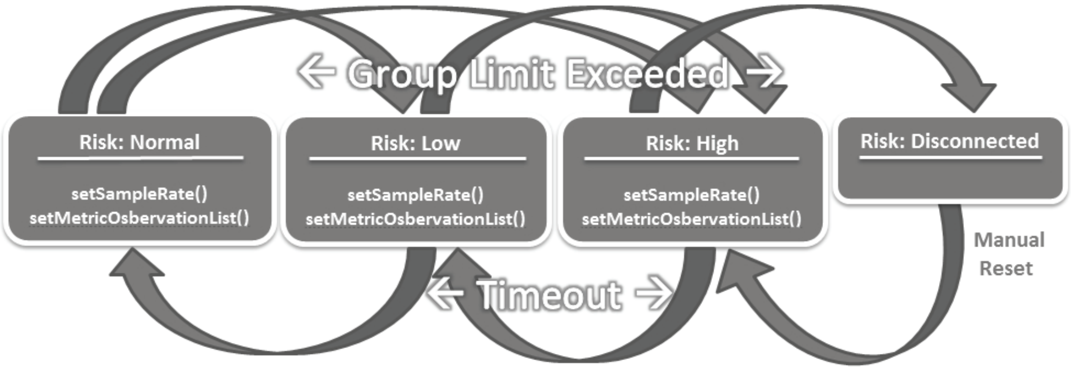


Figure 3: SSC monitoring state-chart

SSC is lightweight and efficient, consumes minimal resources, is able to function in real-time and cope with the dynamic environment and changes to the system over time, which many existing solutions are unable to do.

4. Proposed techniques

In this section, we will explain our novel algorithms and the main contributions of this work. In order for our solution to function successfully and efficiently in the uncertain and dynamic environment of a SoS, the development of these algorithms was paramount.

4.1 Threshold calculation

To monitor for behavioural abnormalities on the component, we must first create a behavioural profile to monitor against. The complexity and uncertainty of the behaviour encountered means traditional parametric approaches of profile establishment are ineffective. We have therefore created our own novel algorithm to formulate this profile, which combines both parametric and non-parametric techniques. Using the training data, we create our profile, which provides a behavioural estimate for each metric at regular intervals over twenty-four hours. This profile is an estimate of the expected levels of behaviour whilst involved in the SoS (also accounting for the dynamics of the system). The following formulae and explanations will provide an overview of the algorithms function.

In (1) we use the collected training data *P* to establish the mean value *A* for each metric *m* at each sample point *i* and *n* is the total number of sample points. This will provide us with an average of the system's normal metric usage taking into account any off-peak activities such as updates.

$$A_{m_i} = \frac{\sum_{j=1}^{n} P_{m_{i,j}}}{n} \quad (1)$$

We then calculate the maximum S2T threshold (*xst*) $xst_{m_i} = A_{m_i} + S_m + \sigma_{m_i}$ and the minimum S2T threshold (*nst*) $nst_{m_i} = A_{m_i} - T_m - \sigma_{m_i}$ for each metric *m* at each sample point *i*. The mean value *A* was produced in (1), *S* is the maximum and *T* is the minimum specified levels of SoS contribution (as specified in the S3LA configuration file) for each metric and *σ* is the standard deviation. This produces our S2T threshold set, which details the theoretical expected static behaviour of the system whilst contributing to the SoS. The deviation applied to the thresholds reduces the effect any discrepancies in the averaging process would have.

To calculate the DA threshold profile, we must first create two groups (*MU* in equation (2) and *ML* in (3)), these contain lists of sample points *i* for each metric *m* outside of the S2T threshold set. *MU* contains sample points above the S2T maximum threshold (*xst)* and *ML* contains sample points below the S2T minimum threshold (*nst)*. Data used in this process is from the training data *P.*

$$MU_m = \{ P_{m_i} \mid \forall i, 1 \leq i \leq n \wedge P_{m_i} > xst_{m_i} \} \quad (2)$$

$$ML_m = \{ P_{m_i} \mid \forall i, 1 \leq i \leq n \wedge P_{m_i} > nst_{m_i} \} \quad (3)$$

In (4) and (5), we calculate the maximum *xdt* and minimum *ndt* thresholds for the DA profile. We calculate the mean absolute deviation and standard deviation of values in both the *ML* and *MU* groups, which are added or deducted respectively. This method allows us to ascertain the expected levels of variation of each metric at each sample point. These are the finalised DA thresholds and are set as the default values used to monitor the behaviour of the live system.

$$xdt_{m_i} = \left(xst_{m_i} + \frac{\sqrt{\sum \left(P_{MU_{m_i}} - xst_{m_i} \right)^2}}{n} \right) + \sigma_{MU_m} \quad (4)$$

$$ndt_{m_i} = \left(nst_{m_i} - \frac{\sqrt{\sum \left(P_{ML_{m_i}} - nst_{m_i} \right)^2}}{n} \right) - \sigma_{ML_m} \quad (5)$$

4.2 Threshold adjustment

As the base system, SoS functionality and SoS roles are prone to change, so we have devised our method to adjust the threshold automatically to adapt to changes on the system. It uses a combination of live and historical data to examine the distribution of data lying between the S2T thresholds and the DA thresholds (i.e. values greater than S2T maximum but less than DA maximum). This process occurs on a regular basis and if required, adjustments to the thresholds are calculated. The gap between the S2T and DA thresholds is split into three quartiles Q_1, Q_2 and Q_3. This part of the process assumes the quartile distribution is Gaussian, therefore 25% of data is below Q_1 and 75% is below Q_3. The calculated adjustments are based on the changes required to the quartile limits to ensure the Gaussian distribution of values over the quartiles. However, thresholds cannot be adjusted above DA minimum or below DA maximum. The following explanations and formulae outline this process for calculating the maximum threshold adjustment; however, the same principle applies to calculate the minimum threshold adjustment.

In (6) we divide the difference between the *xst* (S2T) and *xdt* (DT) thresholds into quartiles (Q_1, Q_2 and Q_3) and calculate the quartile limits. All values *D* (these values are from the live system and from recent historical data) for metric *m* lying between these quartile values are split into representative groups *LQ* (lower quartile), *MQ* (mid quartile) and *UQ (*upper quartile).

$$\begin{aligned} LQ_m &= \{ D_{m_i} \mid \forall i, 1 \leq i \leq n \wedge D_{m_i} > xst_{m_i} \wedge D_{m_i} \geq Q_1 \} \\ MQ_m &= \{ D_{m_i} \mid \forall i, 1 \leq i \leq n \wedge D_{m_i} > Q_1 \wedge D_{m_i} \geq Q_3 \} \\ UQ_m &= \{ D_{m_i} \mid \forall i, 1 \leq i \leq n \wedge D_{m_i} > Q_3 \wedge D_{m_i} \geq xdt_{m_i} \} \end{aligned} \quad (6)$$

In (7) we calculate the upper quartile adjustment *uqa* where $p = \left(\frac{c(LQ_m) + c(MQ_m)}{tc} 100 \right)$, *tc* is the total count of values between the DA and S2T thresholds and *c()* is a function to return the number of values in a particular group. This method is based on the fact that in a Gaussian set of data, the upper quartile should define the lowest 75% of the data; therefore, 75% of all *D* values should be below Q_3. We calculate the adjustment required to the DA threshold, to ensure 75% of the values are below Q_3.

$$uqa_m = \begin{cases} \frac{(75-Q3)xdt_{m_i}}{100} & if\, c(UQ_m) > 0 \wedge p < 75 \wedge c(UQ_m) > c(MQ_m) \wedge c(UQ_m) > c(LQ_m) \end{cases}$$
(7)

In (8) we calculate the lower quartile adjustment *lqa* where $p = \left(\frac{c(LQ_m)}{tc}100\right)$ and *tc* is the total count of values between the thresholds. Again, this is based on the fact that the lowest 25% of the data set is below Q_1. We calculate the level of adjustment required to the DA threshold, to ensure 25% of the values are below Q_1.

$$lqa_m = \begin{cases} \frac{(Q1-25)ndt_{m_i}}{100} & if\, c(LQ_m) > 0 \wedge p > 25 \wedge c(LQ_m) > c(MQ_m) \wedge c(LQ_m) > c(UQ_m) \end{cases}$$
(8)

In (9) we adjust the DA threshold by applying the largest adjustment value (*lqa* or *uqa*) as this will provide the balance where it is needed most. If neither is larger, then no adjustment is made.

$$xdt_i = \begin{cases} xdt_i + uqa_m & if\, uqa_m > 0 \wedge uqa_m > lqa_m \\ xdt_i - lqa_m & if\, lqa_m > 0 \wedge lqa_m > uqa_m \\ xdt_i & else \end{cases} \quad (9)$$

4.3 Weighting calculation

The metric weight is a numerical representation of the strength of the relationship and behavioural similarity between two metrics. To calculate this value, we use Kendall's tau-b coefficient, which measures the association between two metrics and tests for statistical dependence. This method is considered more statistically relevant than similar methods, such as Spearman's rank coefficient. The Kendall's tau coefficient τ_B equation is shown in (10). Here, n_c is the number of concordant pairs, n_d is number of discordant pairs, $n_1 = \frac{\sum_i t_i(t_i-1)}{2}$, $n_2 = \frac{\sum_j s_j(s_j-1)}{2}$, *t* is the number of ties in the first metric *i* and *s* is the number of ties in the second metric *j*.

$$\tau_B = \frac{n_c - n_d}{\sqrt{(n(n-1)/2) - n_1)(n(n-1)/2) - n_2)}} \quad (10)$$

4.4 Decision calculation

If a monitored metric does not conform to its thresholds, then the event is passed to the decision algorithm to calculate a certainty score representing the degree of behavioural irregularity. The decision algorithm uses multiple sources of information when calculating the score. It uses a series of statistical observations of the metric and other metrics with which there is an established relationship. In this algorithm, there are two types of events (*evt)*, *2* represents an event where the minimum threshold has been exceeded and *3* an event where the maximum threshold has been exceeded. The following formulae and explanation detail how the algorithm works.

In (11) we calculate the threshold difference *td* between the metric *M* at sample number *j* and its respective threshold (*ndt* or *xdt)*, which were detailed in Section 4.1.

$$td = \begin{cases} \frac{M_j - ndt_{i,j}}{ndt_{i,j}} & if\,evt = 2 \\ \frac{M_j - xdt_{i,j}}{xdt_{i,j}} & if\,evt = 3 \end{cases} \quad (11)$$

In (12) we calculate the difference between the current and last metric value *ld*.

$$ld = \frac{M_j - M_{j-1}}{M_{j-1}} \quad (12)$$

In (13) we calculate the difference between the current metric value and the average metric value *ad,* using historical data.

$$ad = \frac{M_j - \bar{M}}{\bar{M}} \quad (13)$$

In (14) we calculate the frequency *f* that the metric exceeds its corresponding threshold. Here, *tc* is the total number of examined historical and current data samples, *nc* is the number of samples below the minimum threshold and *xc* is the number of samples above the maximum threshold.

$$f = \begin{cases} \frac{nc}{tc} & if\,evt = 2 \\ \frac{xc}{tc} & if\,evt = 3 \end{cases} \quad (14)$$

In (15) we calculate *sd*, which is the difference between the current metric value and the average of the values exceeding their corresponding threshold. Here, *en* is the metric values below the minimum threshold and *ex* is the metric values above the maximum threshold.

$$sd = \begin{cases} \frac{M_j - \frac{\sum_{i=1}^{n}(en_i)}{n}}{\frac{\sum_{i=1}^{n}(en_i)}{n}} & if\,evt = 2 \\ \frac{M_j - \frac{\sum_{i=1}^{n}(ex_i)}{n}}{\frac{\sum_{i=1}^{n}(ex_i)}{n}} & if\,evt = 3 \end{cases} \quad (15)$$

In (16) we calculate the metric score *ms* using current and historic data *P*, metric *i*, weighting value *W* and the *c()* count function. The metric score is a numerical representation of the similarity of anomalous behaviour across other monitored metrics. We calculate the probability that the reported occurrence is true misbehaviour, by assessing the number of other metrics also exceeding their respective thresholds. Weighting values are applied to the other metrics in accordance with their level of proven relationship, as determined by the weighting algorithm. The idea behind this approach is behaviourally similar metrics would provide a good indicator as to the severity and scale of the anomalies identified on the metric.

$$ms = \begin{cases} \frac{\sum_{i=1}^{n}\left(\frac{c(P_i < ndt_i)}{total}\right) + W_i}{n} & if\,evt = 2 \\ \frac{\sum_{i=1}^{n}\left(\frac{c(P_i < xdt_i)}{total}\right) + W_i}{n} & if\,evt = 3 \end{cases} \quad (16)$$

Finally, (17) calculates the final certainty score *C* by combining the values of the previous statistical tests. This value is returned for any necessary action to be taken. The algorithm uses multiple statistical observations that are all of equivocal value in determining the behavioural irregularity. This enables the solution to make an informed decision regarding the reported behaviour, rather than relying on the fact that the threshold had been exceeded. Thresholds are calculated behavioural estimates, heavily based on averages and as such it is possible for thresholds to be exceeded legitimately especially on a SoS.

$$C = \frac{(1-td)+(1-ld)+(1-ad)+(1-f)+(1-sd)+(1-ms)}{n} \qquad (17)$$

5. Experimentation and initial results

Our implementation runs on Ubuntu 11.10 and is written in C, as this allows increased interactivity with lower level operating system functions in real-time. Our databases are implemented as SQLite databases to keep the solution lightweight and reduce system footprint. The metric data used by our solution is gathered from the /proc virtual file system, callable system functions (i.e. sysinfo), port monitor, JMX proxies and file system monitor. The file system monitor watches for changes to important files and directories or their meta-data. Monitored metrics are measured using the scale by which they are ordinarily measured (e.g. bandwidth as MB/s). As no suitable SoS testing framework exists, we have used web services as an abstraction, as they facilitate a realistic simulation of a SoS interface, dynamic component interaction, service contribution and service utilisation. Our experiments were conducted on a virtual machine, which was allocated 1GB RAM, one core of an Intel Core i7 3.2GHz processor and setup with Apache Tomcat 6 and Axis 2. Ten different Java web services were installed to simulate the varying loads, complexity and usage of web services. A separate virtual machine running on the same virtual network used loadUI and soapUI to simulate dynamic component interaction by creating a random number of connected hosts and requests to the web services. To simulate misbehaviour on the component, we created a custom python script, which randomly causes metrics to exceed either of their calculated thresholds by random amounts and for varying periods. Table 1 outlines the setup for each experiment, detailing the simulated SoS load and the misbehaviour present on the system (simulated using our python script).

Table 1: Experiment setup

Experiment No.	Simulated SoS Component Load	Misbehaviour Risk Level
1	None	None
2	Static	None
3	Dynamic	None
4	Static	Low
5	Static	High
6	Dynamic	Low
7	Dynamic	High

5.1 Results

Table 2 shows the results obtained from our initial experiments. Our results have shown that SSC has low false positive rates and low false negative rates. The false positive and false negative rates were at their highest when dynamic component load and high-level risk misbehaviour were simulated. They also show that it has quick detection times, thus enabling instant reaction to a problem. They also show that SSC consumes low CPU and RAM whilst in operation, thus providing a minimalistic system footprint. Our promising initial results have indicated that SSC has the capability to detect misbehaviour in complex dynamic and uncertain environments.

Table 2: Performance statistics

Experiment No.	Avg. Detection Time (sec)	Consumed RAM (%)	Consumed CPU (%)	False Positive Rate (%)	False Negative Rate (%)
1	N/A	0.8	3	0	N/A
2	N/A	0.8	3	0	N/A
3	N/A	0.8	3	0	0
4	0.13	1.2	7	0	0
5	0.20	1.8	9	0	0
6	0.18	1.5	8	0.07	0
7	0.24	2.0	10	0.12	0.01

6. Conclusion and future work

This paper has outlined our proposed novel solution (SSC) which offers protection to SoS compositions by monitoring for internal component misbehaviour. We have also outlined our novel algorithms that enable SSC to function effectively, whilst accounting for the dynamic environment that most solutions would fail to account for. Our initial results are promising and show that our solution and algorithms are performing effectively. Our immediate future work will focus on refining our solution and algorithms. We believe that our initial results show it is possible to improve on the performance of existing solutions from similar research areas. As we are still in the preliminary experimentation phase of our research, we will be looking to test our solution against existing methods and solutions from similar research areas. We will then be able to produce an effective security monitoring solution that will maintain the security and integrity of future SoS compositions.

References

Agrawal, D., 2009. A new schema for security in dynamic uncertain environments. *2009 IEEE Sarnoff Symposium*, pp.1–5.

Boardman, J & Sauser, B, 2006. System of Systems - the meaning of of. In *2006 IEEE/SMC International Conference on System of Systems Engineering*. Los Angeles, CA: IEEE, pp. 118–123.

Buzen, J.P. & Shum, A.., 1995. MASF - Multivariate Adaptive Statistical Filtering. In *CMG Conference*.

DiMario, M.J., 2006. System of Systems Interoperability Types and Characteristics in Joint Command and Control. In *2006 IEEE/SMC International Conference on System of Systems Engineering*. Los Angeles, CA: IEEE, pp. 222–227.

Efatmaneshnik, M., Nilchiani, R. & Heydari, B., 2012. From complicated to complex uncertainties in system of systems. In *2012 IEEE International Systems Conference*. IEEE, pp. 1–6.

Gorod, A., Sauser, Brian & Boardman, John, 2008. System-of-Systems Engineering Management: A Review of Modern History and a Path Forward. *IEEE Systems Journal*, 2(4), pp.484–499.

Jamshidi, M., 2008. System of systems engineering - New challenges for the 21st century. *IEEE Aerospace and Electronic Systems Magazine*, 23(5), pp.4–19.

Jin, X. & Chan, S.-H.G., 2010. Detecting malicious nodes in peer-to-peer streaming by peer-based monitoring. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 6(2), pp.1–18.

Maier, M.W., 2006. Research Challenges for Systems-of-Systems. In *2005 IEEE International Conference on Systems, Man and Cybernetics*. IEEE, pp. 3149–3154.

Selberg, S. & Austin, M., 2008. Toward an evolutionary system of systems architecture. In *Proceedings of Eighteenth Annual International Symposium of The International Council on Systems Engineering*. Ultrect, The Netherlands, pp. 1–14.

Trivellato, D., Zannone, N. & Etalle, S., 2011. A Security Framework for Systems of Systems. *2011 IEEE International Symposium on Policies for Distributed Systems and Networks*, pp.182–183.

Visan, A., Pop, F. & Cristea, V., 2011. Decentralized Trust Management in Peer-to-Peer Systems. *2011 10th International Symposium on Parallel and Distributed Computing*, pp.232–239.

Wang, C. et al., 2011. Statistical techniques for online anomaly detection in data centers. *12th IFIP/IEEE International Symposium on Integrated Network Management and Workshops*, pp.385–392.

Xiao, B. et al., 2011. Reasearch on The History and Perspective of System of Systems. *Reliability, Maintainability and Safety, 2011 9th International Conference on*, pp.1262–1266.

Non Academic Papers

Proactive Cyber Defense: Understanding and Testing for Advanced Persistent Threats (APTs)

Anna-Maija Juuso [1], Ari Takanen[2] and Kati Kittilä[2]
[1]Codenomicon, Singapore
[2]Codenomicon, Oulu, Finland
Anna-Maija.Juuso@codenomicon.com
Ari.Takanen@codenomicon.com
Kati.Kittilä@codenomicon.com

Abstract: Government and critical infrastructure networks are increasingly reliant on cyberspace. This reliance is in stark contrast to the inadequacy of cybersecurity: Many of these networks used to be closed and lack the robustness needed to withstand cyber-attacks. As a result, vulnerabilities in network protocol implementations can be exploited with Internet hacking tools to disrupt operations or to steal sensitive information. Fuzzing is a black-box testing technique originally used by blackhat hackers to find exploitable vulnerabilities. In this paper, we demonstrate how specification-based fuzzers can be used to discover exploitable vulnerabilities proactively to make networks and devices more robust against cyber-attacks. Advanced Persistent Threats (APTs) typically utilize previously unknown, zero-day vulnerabilities. Thus, proactive vulnerability discovery is an essential part of effective cybersecurity.

Keywords: security testing, proactive cyber defence, fuzzing

1. Introduction

The security landscape is changing: Governments, critical infrastructure providers and defense organizations increasingly rely on the Internet to perform mission-critical operations. At the same time, cyber-attacks have become more professional with attackers investing more time and money into creating detection evasion techniques and developing sophisticated, targeted attacks exploiting zero-day vulnerabilities. Zero-day exploits are the biggest threat to security, because there are no defenses against them and the attacks can go unnoticed. Most organizations rely largely on signature-based security solutions, which only defend against known threats and require continuous rule updates to stay up-to-date on cyber-attacks.

Advanced Cyber Threat (APT) refers to sophisticated Internet abuse performed by highly-motivated and well-resourced groups, such as organized cyber criminals, hostile nation states and hacktivists. These attacks frequently utilize unknown, zero-day vulnerabilities. Zero-day vulnerabilities pose the greatest threat to network security, because there are no defenses for attacks against them (Codenomicon, 2010). The attacks can go unnoticed and once discovered it takes time to locate the vulnerabilities and to create patches for them (Codenomicon, 2010). Advanced attacks, like Stuxnet, can utilize multiple zero-days making them extremely difficult to defend against.

Vulnerabilities are flaws in software or software components in hardware, which enable cyber adversaries to exploit a system. Vulnerabilities are not created when a system is being attacked. They are design and implementation errors that are introduced into the code during development. Cyber adversaries need to find a vulnerability in the protocol implementation in order to devise an attack against a target system. By removing potential zero-day vulnerabilities proactively, you can make it significantly harder for cyber adversaries to devise attacks. Thus, the best way to prevent zero-day attacks is to get rid of exploitable vulnerabilities proactively. Fuzzing enables you to find previously unknown, zero-day vulnerabilities by triggering them with unexpected inputs.

In this paper, we present two types of fuzzing: mutation and generation-based fuzzing. We will examine generation-based fuzzing in detail and demonstrate how protocol specifications can be used to create specification-based fuzzers, a type of generation-based fuzzer. We describe fuzzer building techniques developed by our own researchers and demonstrate how specification-based fuzzers can be used to discover vulnerabilities, which could be exploited in cyber-attacks. This paper also includes two case studies. We used generation-based fuzzers to examine the robustness of two widely used communication technologies: Voice over IP (VoIP) and DVB/MPEG2-TS. VoIP technologies are widely used in command and control systems, whereas DVB and MPEG2-TS are used as transports in satellite communications. We use consumer electronics as test targets. However, the same code-base, codecs, and protocol stacks are also used in military communication.

2. Cybersecurity risks

Security threats have been growing in scale and sophistication for decades. Twenty years ago, cyber-attacks were primarily the domain of hobbyists. Then, as the opportunity for profiting from stolen information grew, criminals started taking a larger role. More recently spies working for government and corporate espionage are leading some of the most technically advanced and resource-intensive attacks to date. 2011 saw a significant increase in the activity of "hacktivist" groups like Anonymous and LulzSec (HP 2012). The motivation for these groups is retaliation for perceived wrongdoing. Rather than financial gain, their main goal is embarrassing their victims. However, their attacks are not without financial consequences.

2.1 Critical infrastructure

Critical infrastructure networks are no longer isolated. They are all connected to the cyberspace, the global network of interdependent information technology infrastructures and communication networks, and they depend on common commercial off-the-shelf software. When SCADA systems were first implemented in the 1960's, it would have been hard to image that critical control system could be attacked remotely or that printers in adjacent corporate networks could be used as weapons to attack them. Nobody could envision such threats, so no measures were taken to make the networks resilient against cyber-attacks. Newer SCADA devices communicate using Internet protocols, sometimes over the public Internet (Sommer 2011). This helps reduce the cost of dedicated communication links, but at the same time it makes the networks more open to outside attacks (Sommer 2011).

Power grids, telecommunication and transportation networks are exactly the type of infrastructure that has been the target of traditional warfare, only the weapons have changed (Washington Post 2012). There have only been a limited number of reported cyber-attacks against critical infrastructure (Washington Post 2012). However, it would be naïve to assume that hostile nation states, terrorists and hacktivists are not aware how vulnerable these networks are and how much havoc such an attack could cause (Washington Post, 2012). Stuxnet was the first publicly admitted act of cyber warfare. It demonstrated that cyber-attacks can cause significant physical damage to a facility (Kroft 2012). Stuxnet was a very sophisticated attack carefully selecting its victims and remaining undetected (Kroft 2012). It utilized four different zero-day vulnerabilities. However, many critical infrastructure networks are so vulnerable that it does not take a sophisticated attack like Stuxnet to exploit them. In many countries, large parts of the critical infrastructure is privately owned and extremely vulnerable (Washington Post 2012).

2.2 Corporate networks

A network is only as strong as its weakest element (Juuso 2010). If attackers manage to compromise a laptop or a smartphone of a remote user working over a VPN, then they have direct access to your network. Stuxnet was carried into the plants on a corrupted laptop or thumb drive (Kroft 2012). Corporate networks connected to critical networks are full of equipment, like VoIP phones, printers and storage devices. Nobody thought that these devices could be used to attack the networks, so the developers did not try to make this difficult or impossible to do (Varpiola 2012).

When sourcing equipment for critical networks or networks connected to critical networks, security or robustness testing should be used as an acceptance criterion. The challenge with outsourcing is that you lose visibility over the security and quality of the software development. Often buyers have been surprised to find that the middleware they have purchased has an open source core.

Networks for distributed organizations often include site-to-site, branch office, and remote access networks. There might also be additional network security layers such as VPNs and LANs. All these add to the complexity of a network making it more difficult to secure and increasing the importance of proactive measures.

2.3 Cloud security

In recent years the use of virtualization technologies and cloud services has increased dramatically. Cloud services and virtualization can help government agencies connect with citizens, improve efficiency and reduce costs. However, like any new technology, cloud services and virtualization introduce new security concerns. New technologies are not necessarily inherently less secure than old ones. They just have not been tested and used for as long. Also, the threats can be different. The potential security risk in cloud technologies is the hypervisor, which controls all the clients within a virtual cloud. If the implementations of PHYP or another hypervisor protocol contain vulnerabilities, these could be exploited to inject malicious code or to otherwise control client clouds.

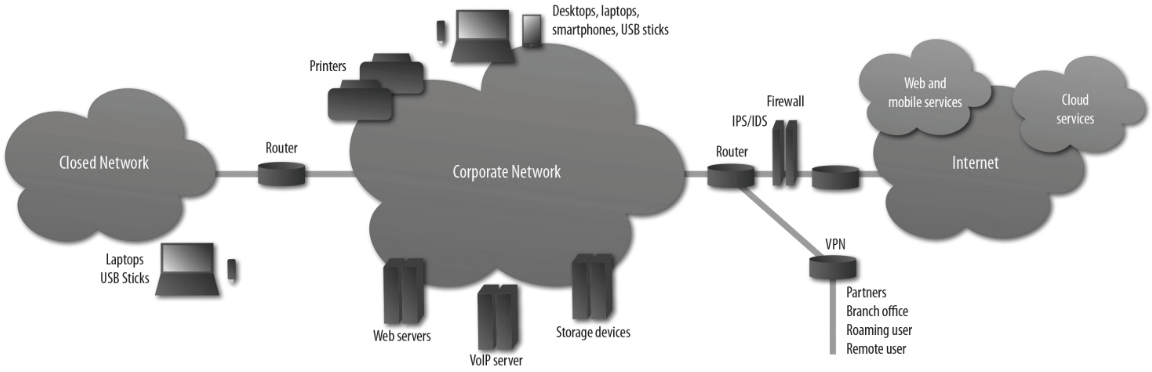


Figure 1: Closed and corporate networks

2.4 eGov and mGov

eGov and mGov services make information sharing between citizens, businesses and government more seamless: less paperwork, less bureaucracy and you can use the services whenever, wherever. Such initiatives should be applauded, as they improve the efficiency of government services and improve the quality of service experienced by the users. However, in developing services with external user interfaces to handle private and confidential information, the robustness of the services should be thoroughly tested before deployment. Any security incidents could erode user confidence and set back the development of the services.

3. Vulnerability exposure

Vulnerabilities are flaws in software or software components in hardware, which enable cyber adversaries to exploit a system. Vulnerabilities are not created when a system is being attacked. They are design and implementation errors that are introduced into the code during development (Rontti 2011). The errors become vulnerabilities once the software is released, and it gets exposed to outside attacks (Takanen 2011). Security researchers, security companies and hackers discover some of the vulnerabilities, and if they choose to report the findings, they can enable software developers to create patches for the found vulnerabilities (Takanen 2011). After the patch release the vulnerability becomes public knowledge (Takanen 2011).

3.1 No exposure, no publicity

Figure 2 categorizes vulnerabilities based on exposure. The exposure of a vulnerability depends firstly on whether the vulnerability can be accessed by outside attackers, and secondly, on how public the vulnerability is. During development, new vulnerabilities have zero exposure to attacks: nobody knows that they exist and they cannot be exploited by outsiders (Takanen 2011). After release the vulnerabilities have limited exposure: they are open to attacks, but the attackers first have to find them (Takanen 2011). After a patch is released, the exposure is full: the attackers have both the possibility to attack and the information they need (Takanen 2011). Public exposure can be avoided by deploying patches in a timely manner.

3.2 Known and unknown vulnerabilities

Figure 3 divides vulnerabilities into four groups: known and unknown vulnerabilities with and without patches. If you have vulnerabilities within your systems for which patches already exist, then clearly you should be doing better vulnerability research and be more vigilant about patch updates (Wang 2011). Most organizations do a good job employing various technologies like anti-virus, firewall, IPS/IDS to defend against known attacks and keep up-to-date with software updates (Wang 2011). During the small window when a vulnerability has been discovered but there is no patch yet, a workaround needs to be implemented.

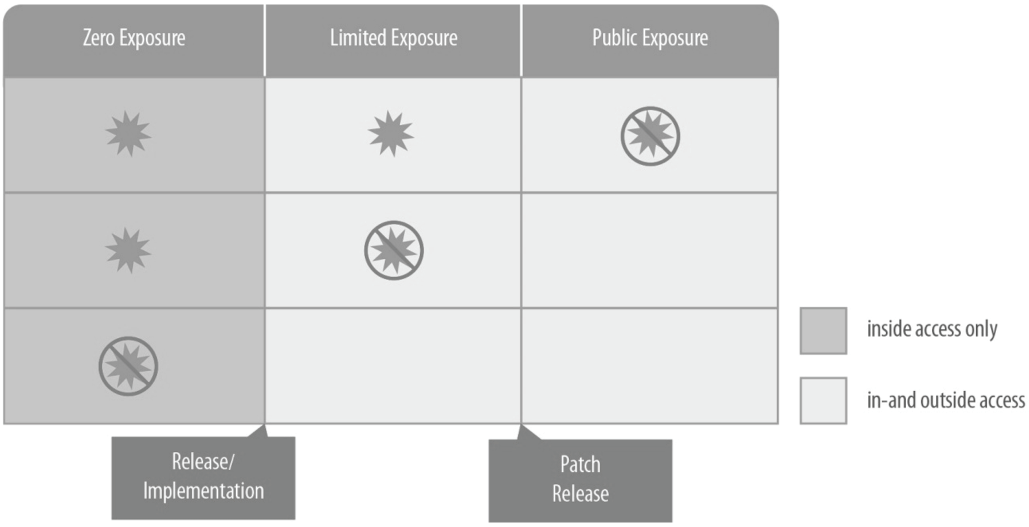


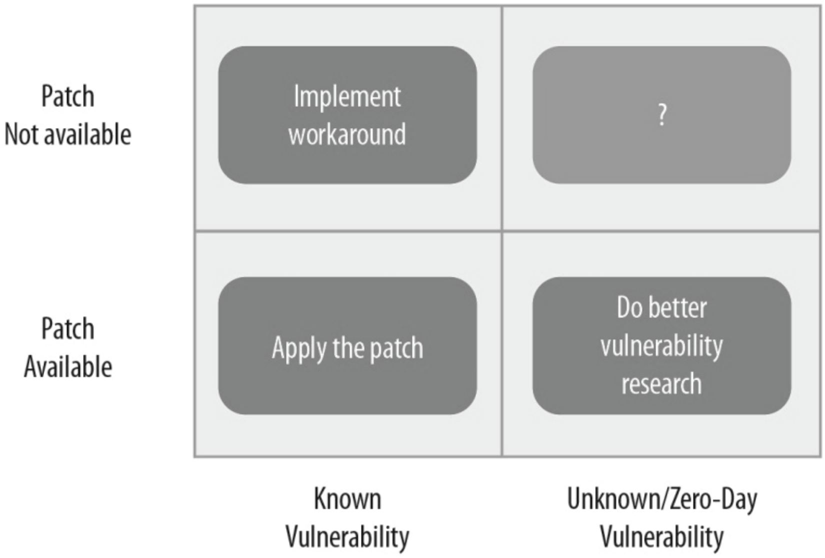
Figure 2: Vulnerability exposure. Based on (Takanen 2011)

3.3 Known and unknown vulnerabilities

Figure 3 divides vulnerabilities into four groups: known and unknown vulnerabilities with and without patches. If you have vulnerabilities within your systems for which patches already exist, then clearly you should be doing better vulnerability research and be more vigilant about patch updates (Wang 2011). Most organizations do a good job employing various technologies like anti-virus, firewall, IPS/IDS to defend against known attacks and keep up-to-date with software updates (Wang 2011). During the small window when a vulnerability has been discovered but there is no patch yet, a workaround needs to be implemented.

3.4 Zero-day vulnerabilities and APTs

In this paper, we focus on the fourth quadrant, the unidentified zero-day vulnerabilities. These vulnerabilities are the biggest threat to an organization's security (Juuso 2011). Their existence is unknown, there are no defenses for attacks against them and an attack can go completely unnoticed (Juuso 2011). If an attacker finds a zero-day vulnerability in a network, service or application, they can do what they want with it from website defacing, obstructing operations to stealing confidential information. It is unlikely that targets of high profile attacks are not keeping their patches up-to-date. It is the zero-day vulnerabilities in their systems that make APTs possible.



	Known Vulnerability	Unknown/Zero-Day Vulnerability
Patch Not available	Implement workaround	?
Patch Available	Apply the patch	Do better vulnerability research

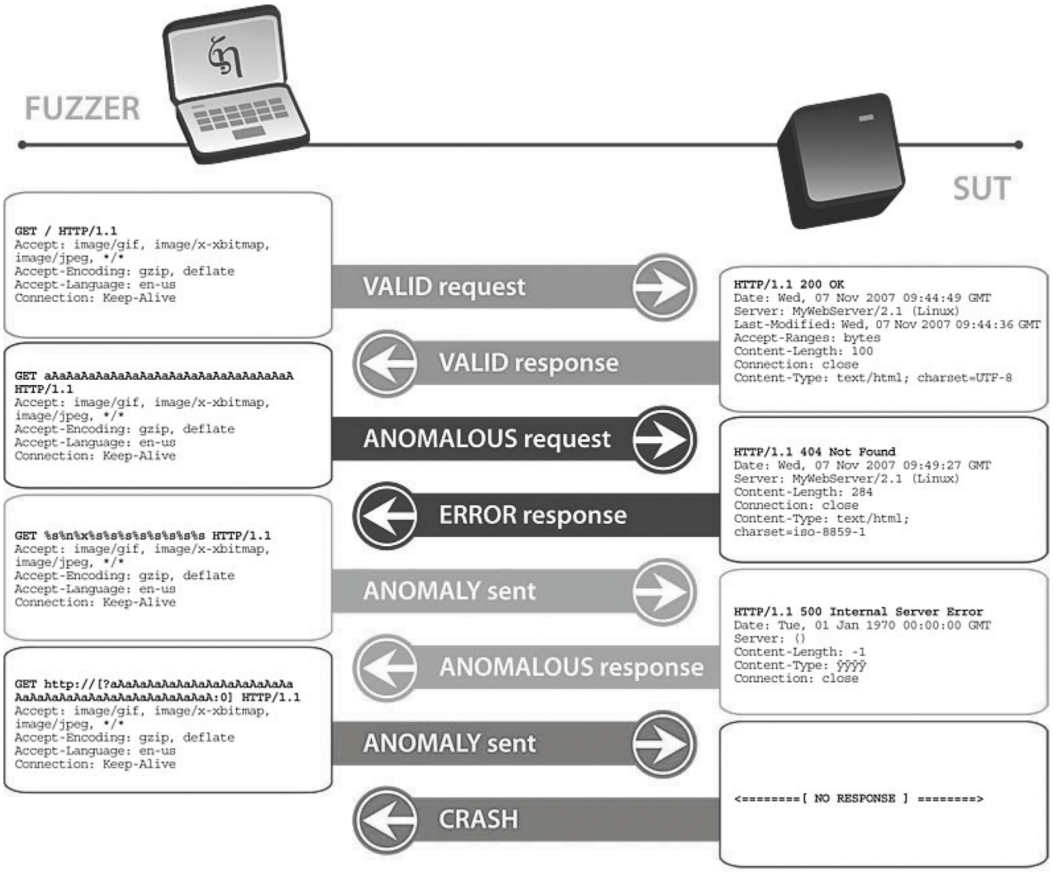
Figure 3: Known and unknown vulnerabilities (Forrester 2009)

4. Fuzzing

Fuzzing is a black-box robustness testing technique used to reveal zero-day vulnerabilities by triggering them with unexpected inputs. Basically, unexpected data in the form of modified protocol messages are fed to the inputs of a system, and the behavior of the system is monitored (Takanen 2008). If the system fails, e.g., by crashing or by failing built-in code assertions, then there is an exploitable vulnerability in the software (Takanen 2010).

Fuzzing simulates outside attacks. Code review techniques, like static code analysis, are very effective in finding reported vulnerabilities and variations of them from software. Fuzzing completes these methods by using unexpected inputs to find completely new vulnerabilities. Fuzzing provides a very good representation of potential attacks: it can find zero-day vulnerabilities that cyber adversaries are also looking for. By finding zero-day vulnerabilities proactively, networks can be made more robust against attacks reducing the risk of DoS disruptions to network services.

Fuzzing can be used to test all software and software components with hardware, so you can use it to test all types of applications, services and network equipment like routers, switches and servers, and also security software like anti-virus and firewalls (Wang 2011). The cost of not fuzzing can be high. It can be costly, if an important network element is offline or services unavailable, never mind the consequences of an advanced cyber-attack. Figure 4 depicts the flow of fuzzing tests.



```
GET / HTTP/1.1
Accept: image/gif, image/x-xbitmap,
image/jpeg, */*
Accept-Encoding: gzip, deflate
Accept-Language: en-us
Connection: Keep-Alive
```

```
HTTP/1.1 200 OK
Date: Wed, 07 Nov 2007 09:44:49 GMT
Server: MyWebServer/2.1 (Linux)
Last-Modified: Wed, 07 Nov 2007 09:44:36 GMT
Accept-Ranges: bytes
Content-Length: 100
Connection: close
Content-Type: text/html; charset=UTF-8
```

```
GET aAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaA
HTTP/1.1
Accept: image/gif, image/x-xbitmap,
image/jpeg, */*
Accept-Encoding: gzip, deflate
Accept-Language: en-us
Connection: Keep-Alive
```

```
HTTP/1.1 404 Not Found
Date: Wed, 07 Nov 2007 09:49:27 GMT
Server: MyWebServer/2.1 (Linux)
Content-Length: 284
Connection: close
Content-Type: text/html;
charset=iso-8859-1
```

```
GET %s%n%x%s%s%s%s%s%s%s%s HTTP/1.1
Accept: image/gif, image/x-xbitmap,
image/jpeg, */*
Accept-Encoding: gzip, deflate
Accept-Language: en-us
Connection: Keep-Alive
```

```
HTTP/1.1 500 Internal Server Error
Date: Tue, 01 Jan 1970 00:00:00 GMT
Server: ()
Content-Length: -1
Content-Type: ÿÿÿÿ
Connection: close
```

```
GET http://[?aAaAaAaAaAaAaAaAaAaAaAaAa
AaAaAaAaAaAaAaAaAaAaAaAaAaA:0] HTTP/1.1
Accept: image/gif, image/x-xbitmap,
image/jpeg, */*
Accept-Encoding: gzip, deflate
Accept-Language: en-us
Connection: Keep-Alive
```

```
<========[ NO RESPONSE ] ========>
```

Figure 4: The flow of fuzzing. (Takanen 2008)

A survey conducted by a large independent software vendor found that every single unique vulnerability found had been discovered by fuzzing (Wang 2011). The Internet is full of fuzzing kits, like the Phoenix Exploit Kit, Blackhole and Crimepack, favored among cyber adversaries to find exploitable vulnerabilities in networks and applications (Wang 2011). Industry leading companies are already using fuzzing to protect their networks against zero-day attacks. By finding zero-day vulnerabilities proactively, networks can be made more robust against attacks reducing the risk of advanced cyber-attacks (Rontti 2011).

5. Automating fuzzing

In fuzzing, thousands and even millions of misuse-cases are created for each use-case, thus most robustness testing solutions contain at least some degree of automation. There are two popular ways to automate fuzzing: mutation-based and generation-based fuzzing (Kaksonen 2009). In mutation-based fuzzing real-life inputs, like network traffic and files, are used to generate test cases by modifying the samples either randomly or based on the sample structure. In generation-based fuzzing, the process of data element identification is automated by using protocol models. Specification-based fuzzing is a form of generation-based fuzzing, which uses protocol and file format specifications to provide the fuzzer with protocol or file format specific information, e.g., on the boundary limits of the data elements (Kaksonen 2009).

Mutation-based fuzzing is dependent on the quality of samples used: Elements that are not included in the samples cannot be properly tested without additional models. Specification-based test generation achieves excellent coverage of the features included in the specification. However, new features and proprietary features not included in the specification are not covered (Varpiola 2012). Specification-based fuzzing provides better assurance level for the target system, but it requires an interface specification, whereas mutation-based fuzzing can even be used to test proprietary secret interfaces. Figure 5 shows the results of a fuzz test, in which both mutation- and generation-based fuzzers were used to test the same target system (Juuso 2012). The figure depicts the differences between mutation- and generation-based fuzzing in terms of vulnerabilities found and time needed.

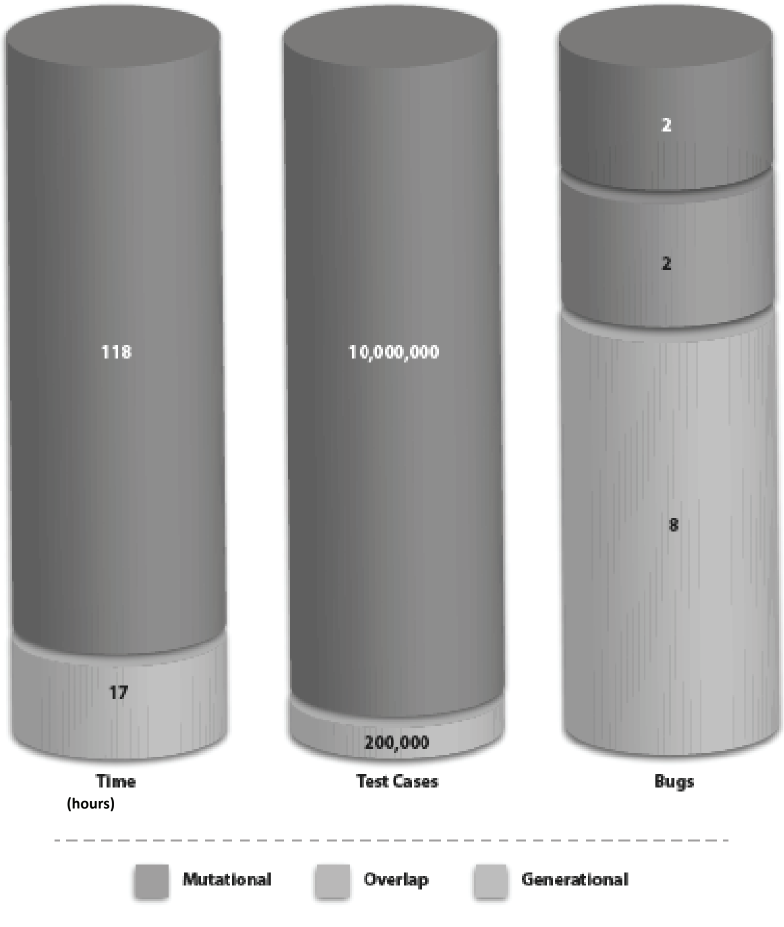


Figure 5: Mutation- and generation-based fuzzing. (Juuso 2012)

6. Building a specification-based fuzzer

The starting point for creating a specification-based fuzzer is acquiring a model of the protocol in a description language, such as BNF or ASN.1. Figure 6 shows an excerpt of a SIP protocol definition from the RFC3261 standard described in ABNF (Augmented Backus-Naur Form, RFC5234) format. (Rontti 2012)

SIP-message	=	Request / Response
Request	=	Request-Line *(message-header) CRLF [message-body]
Request-Line	=	Method SP Request-URI SP SIP-Version CRLF
Request-URI	=	SIP-URI / SIPS-URI / absoluteURI
SIP-URI	=	"sip:" [userinfo] hostport uri-parameters [headers]

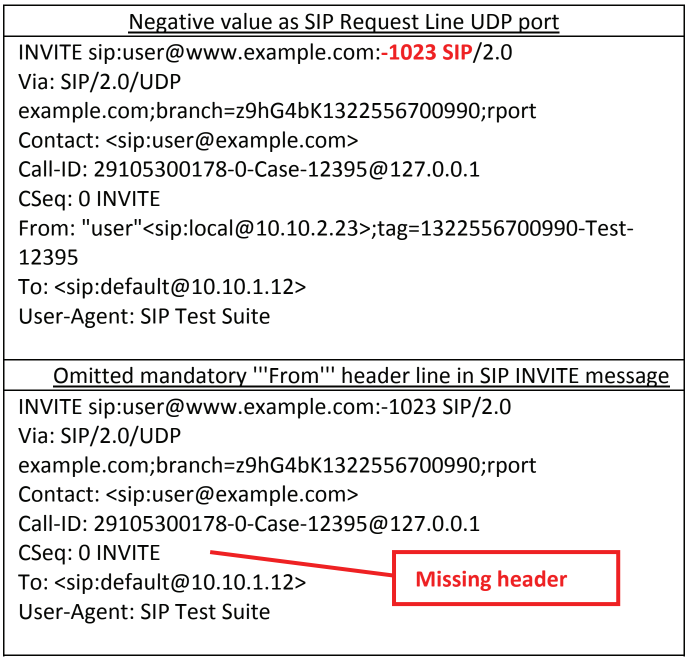
Figure 6: Excerpt of SIP in ANBF

6.1 Semantic rules

To create a fuzzer, the protocol model is augmented with semantic rules describing relationships between the protocol elements. Semantic rules make testing more efficient, because if test cases do not adhere to protocol semantics they will be dropped by the target system. This ensures that the anomalized packet hits the code deep in the state machine. (Rontti 2012).

6.2 Anomalies

Some fuzzers use completely random data as anomalies. However, such random anomalies are fairly inefficient. Intelligent fuzzers can use three levels of anomalies. Firstly, anomalies can be abnormal values in protocol fields. Such anomalies are used to test field level problems, like overflows. Secondly, fuzzers can also anomalize the message structure, for example by multiplying or deleting message elements. Field and structural level anomalies can be used, for example, to find vulnerabilities in message parsers, such as XML or ASN.1 parsers. Finally, the entire message sequence can be anomalized by multiplying or deleting entire messages. Such structural level anomalies can cause the state machine to dead-lock, or they can produce resource consumption issues such as out-of-memory situations. The examples in Figure 7 visualize two different types of anomalies in SIP messages. (Rontti 2012).



Negative value as SIP Request Line UDP port

```
INVITE sip:user@www.example.com:-1023 SIP/2.0
Via: SIP/2.0/UDP
example.com;branch=z9hG4bK1322556700990;rport
Contact: <sip:user@example.com>
Call-ID: 29105300178-0-Case-12395@127.0.0.1
CSeq: 0 INVITE
From: "user"<sip:local@10.10.2.23>;tag=1322556700990-Test-
12395
To: <sip:default@10.10.1.12>
User-Agent: SIP Test Suite
```

Omitted mandatory '''From''' header line in SIP INVITE message

```
INVITE sip:user@www.example.com:-1023 SIP/2.0
Via: SIP/2.0/UDP
example.com;branch=z9hG4bK1322556700990;rport
Contact: <sip:user@example.com>
Call-ID: 29105300178-0-Case-12395@127.0.0.1
CSeq: 0 INVITE
To: <sip:default@10.10.1.12>
User-Agent: SIP Test Suite
```

Missing header

Figure 7: Examples of anomalies in SIP messages

6.3 Test case execution

To see whether a test case has passed or failed some type instrumentation is needed. Valid-case instrumentation is a straightforward way of monitoring the target system via the injection vector. A valid test case, a message containing no anomalies, is sent to the target system, which should send a valid response. A new valid case is sent until a valid response is received. If no response is received, then the target system has failed. (Rontti 2012).

The strength of specification-based fuzzing comes from the fact that the fuzzer implements the protocol completely. Each message generated by the specification based fuzzer is based on the protocol model, and all messages received from the tested system are parsed back into the behavioral model. The parsing process is also used to check whether the received message can be handled with the rules described in the model. If they cannot be handled, the test case is terminated and the verdict of the test case is determined through external instrumentation. (Rontti 2012).

7. Case studies

In this paper, we relate the findings of two case studies. The first case study will look at VoIP. We relate the CVSS scores of VoIP vulnerabilities defined as a part of a technical report published by the European Telecommunications Standards Institute, ETSI. The second case relates the findings of Kuipers (2012). Kuipers and his team tested Smart TVs. In both case studies, the test targets were commercial off-the-shelf systems, but the same code-base, codecs, and protocol stacks are also used in military communications.

7.1 Voice over IP

Voice over IP (VoIP) is a technique where Internet protocols are used to perform voice signaling such as setting up calls and checking the availability of recipients. Voice is typically the only medium, but video is also used. The VoIP infrastructure contains multiple critical interfaces. In addition to VoIP protocols, such as SIP and RTP, attacks can also come through a wireless interface. Due to its complexity, IMS/VoIP infrastructures are a good case study for security analysis. (ETSI 2013).

Attack surface analysis is synonymous to studying the protocol interfaces and understanding the data paths inside the network implementing the service. This is a starting point for mapping threats and understanding the exposure of a system. With this information you can, for example, limit the scope of testing required to achieve sufficient confidence in a service. You will also understand how attackers can gain access to valuable information inside your network. There are several methods for conducting an attack surface analysis, including scanning, passive monitoring, deducting the physical properties of the system, and observing configuration. These methods are complementary and more than one of them should be used to map all the possible ways into a system's open interfaces. (ETSI 2013)

In the study, the CVSS metric is used to calculate the risk for each protocol in the case study. CVSS is a vulnerability scoring system that helps organizations prioritize and coordinate a joint response to IT security vulnerabilities (FIRST 2013). The CVSS score provides a simple metric from zero to ten for describing how easy an interface is to attack (FIRST 2013). While the primary use case of CVSS is to estimate the impact of realized vulnerabilities, the exploitability metric provides a convenient way of assigning numerical values to potential threats. Table 1 summarizes the results by listing the top-5 attack vectors into a VoIP service. In the table, each attack vector is identified based on the interface and the protocol used. A CVSS score is provided for each attack vector. The names of the interfaces have been simplified for easier understanding. The test results are confidential.

Table 1: Results

INTERFACE	PROTOCOL	CVSS score	NOTES
Terminal to Proxy	SIP	10	Direct unauthenticated connection
Terminal to Wi-Fi	802.11	8.6	Wireless
Terminal to AAA	IPv4/DHCP	8.6	Device registration, must be open
Terminal to Registrar	SIP	8	VoIP registration, often through proxy
Terminal to Web Services	HTTP	8	Application services for the terminal

The risks of VoIP attacks are apparent. According to a recent study funded by the Defense Advanced Research Projects Agency (DARPA), attackers could insert malware into VoIP handsets eavesdrop on confidential conversations (Sale 2013). Such VoIP devices are commonly used by government departments and corporations around the world (Sale 2013).

7.2 DVB/MPEG2-TS

Kuipers (2012) used generation-based Codenomicon Defensics fuzzing solutions to test modern Internet-enabled TV sets. The Defensics fuzzing tools are based on deep protocol models built from the protocol and file format specifications. The test cases were created automatically and sent to the target system. The TV sets' firmware were updated to the latest available version via the manufacturers' website or the devices' own update functionality. All available services in the TV sets were enabled in order to get the maximum test coverage. The TV sets were then scanned for open ports and services to find the attack surfaces. In addition, any accompanying documentation was consulted. All the discovered protocols were tested by fuzzing with the appropriate Defensics suite. The test results are shown in Table 2. (Kuipers 2012).

The tests provide a number of interesting results. Firstly, many of the TV sets failed standard Internet protocol tests, for example the IPv4 tests (Kuipers 2012). This is surprising, because most Internet-enabled TVs use either commercial or open source operating system, which should be relatively robust (Kuipers 2012). Secondly, the tests demonstrated that video data is clearly an easier attack vector than simple voice data (Kuipers 2012). However, from a military perspective, the most interesting results finding was that all the TV sets failed the DVB tests (Kuipers 2012).

Table 2 shows the results from our tests conducted with the Defensics suites.

Table 2: Results from our tests conducted with the Defensics suites

Protocol	Sony #1	Samsung	Panasonic	Sharp	Sony #2	LG
IPv4	pass	**fail**	**fail**	pass	pass	fail
SunRPC	n/a	n/a	n/a	n/a	n/a	n/a
DVB	**fail**	**fail**	**fail**	**fail**	**fail**	**fail**
UPnP	n/a	**fail**	pass	n/a	n/a	**fail**
Images	pass	**fail**	**fail**	n/a	n/a	**fail**
Audio	pass	pass	n/a	n/a	n/a	pass
Video	**fail**	**fail**	n/a	**fail**	**fail**	**fail**

The DVB protocol is not just used to transmit video/audio streams to TV, it is also used for video surveillance (CCTV), local communication (i.e. car to car transmission), navigational purposes, access to content on hand-held devices (DVB-H), and even to provide Internet access in remote locations where cable or mobile communication is not feasible (IP-over-DVB/MPEG). In addition to the numerous commercial applications of DVB, it is also used in satellites serving military and intelligence purposes (DVB-S and DVB-S2). The same code-base, codecs, and protocol stacks are used in both commercial and military applications of DVB.

Most common failure mode in DVB fuzzing was total reboot of the TV set. The corrupt data stream crashed a critical software component in the device, either in the stream parsing functionality or in some video or audio codec causing the entire system to fail. Other less fatal anomalies in the video streams and data elements of the Electronic Programme Guide (EPG) data resulted in data corruption or de-synchronization of the data producing incomprehensive video or audio. The failures were not analyzed to examine the exploitability of the found vulnerabilities, for example by injecting executable exploit code into the device.

DVB signals are mainly unencrypted, even if some of the data channels implement encryption. Therefore, anyone under the coverage of the satellite can use a sniffer to capture at least parts of the DVB data (Nve 2010) and create attacks based on the data streams. With sniffers attackers can gain access to sensitive information like user names and passwords, which enable them to launch further, more serious attacks against their target. Access to satellite controls allow an attacker to damage or destroy a satellite, to steal or corrupt satellite transmissions and even to compromise other networks connected to the satellite (U.S.-China Economic and Security Review Commission 2011).

8. Fuzzing best practices

If a protocol is not implemented properly it will contain vulnerabilities. Some protocols are just harder to implement than others. Especially, with new technologies it is important to reserve enough time for testing. It is better to find vulnerabilities in systems, before deployment or implementation, rather than to wait for the attackers to find them.

8.1 As an acceptance condition

Many vendors are in a hurry to push software onto the market, and often times it is the user who ends up doing the testing (Varpiola 2012). Actually, software products have the highest rate of defects of product sold today (Varpiola 2012). By insisting on using fuzzing as an acceptance condition, you can make vendors claim responsibility over the quality and security of their products (Wang 2011). A prominent US ISP already uses fuzzing as entry criteria for its network suppliers (Juuso 2011).

8.2 During SDL

Large software houses already include fuzzing as a part of their secure development lifecycles: Cisco's CSDL, Microsoft's SDLC and the Adobe Product lifecycles are good examples of this. Giants like IBM and Google also promote fuzzing (Varpiola 2012). The Microsoft secure development lifecycle (SDL) model endorses the use of fuzzing in the verification phase (Juuso 2011 B). However, fuzzing can be used throughout the development process from the moment the first software components are ready to even after the release (Juuso 2011 B). The earlier the vulnerabilities are found, the easier and cheaper it is to fix them (Codenomicon 2010). Indeed, by building security into your software you can avoid costly, critical and embarrassing software blunders.

9. Conclusion

Nation-sates are becoming increasingly aware of the reliance of their critical infrastructure on cyberspace. As result many countries have released their own cybersecurity strategies. As countries prepare for natural disasters they must also have plans and resources in place for cyber incidents. Fuzzing is already a part of secure development lifecycles in software companies. Governments should learn from their example and use fuzzing to verify the quality of in-house and third party software, before deployment.

Proactive cybersecurity is needed for three reasons. Firstly, we are so reliant on cyberspace that ad hoc responses to cyber-attacks are not enough. Secondly, the networks are now more open to attacks. Thirdly, defending networks against attacks exploiting zero-day vulnerabilities is getting increasingly difficult as networks get more complex. Thus, a proactive approach is needed.

Specification-based fuzzing can be used to test protocol implementations before deployment or integration in a systematic manner. The main benefit of specification-based fuzzing is that it can be used to find vulnerabilities based on the inherent qualities of a protocol without any prior knowledge of the vulnerability. By fuzzing their network equipment proactively, government and critical infrastructure providers can avoid costly security problems and improve their network uptime.

References

Codenomicon (2010), "How to Really Avoid Zero-Day Attacks – Build Security In, Don't Add it", [online], http://www.codenomicon.com/resources/whitepapers/codenomicon-wp-20100112.pdf

ETSI (2013), "IMS/NGN Security Testing and Robustness Benchmark" European Telecommunications Standards Institute, ETSI DTR/INT-00066, 2013. Stable draft.

FIRST (2013), "Common Vulnerability Scoring System (CVSS-SIG)", [online], http://www.first.org/cvss

Forrester (2009), "Next-Generation Marketing and Measurement", A commissioned study conducted by Forrester on behalf of Omniture.

HP (2012), "2011 Top Cyber Security Risks Report", [online], http://www.hpenterprisesecurity.com/collateral/report/2011FullYearCyberSecurityRisksReport.pdf

Juuso, A-M and Takanen, A. (2010), "Unknown Vulnerability Management", [online], http://www.codenomicon.com/resources/whitepapers/codenomicon-wp-unknown-vulnerability-management.pdf

Juuso, A-M and Takanen, A. (2011 A), "Unknown Vulnerability Management for Telecommunications, [online] http://www.codenomicon.com/resources/whitepapers/codenomicon-wp-telecommunications-20110204.pdf

Juuso, A-M and Takanen, A. (2011 B), "Building Secure Software using Fuzzing and Static Code Analysis", [online], http://www.codenomicon.com/resources/whitepapers/codenomicon-wp-fuzzing-and-static-code-analysis-20100811.pdf

Juuso, A-M and Varpiola, M. (2012), "Fuzzing Best Practices: Combining Generation and Mutation-Based Fuzzing", [online], http://www.codenomicon.com/resources/whitepapers/generation-and-mutation-based-fuzzing-registered.shtml

Kaksonen (2009) R. Kaksonen & A. Takanen, "XML Fuzzing Tool: Testing XML on Multiple Levels", *Testing Experience*, December 2009.

Kroft (2012) S. Kroft, "Stuxnet: Computer worm opens new era of warfare", 60 Minutes, July 2012.

Kuipers, R., Heikkinen, H. and Starck, E. (2012), "Smart TV Hacking: Crash Testing Your Home Entertainment", [online], http://www.codenomicon.com/resources/whitepapers/2012-smart-tv-hacking.shtml.

Miller, B.P. & al (1995), "Fuzz Revisited: A Re-examination of the Reliability of UNIX Utilities and Services", University of Wisconsin.

Nve, L. (2010), "Playing in a Satellite Environment 1.2", Presentation at the Blackhat DC Security Conference, Arlington, Virginia, January, [online], http://www.blackhat.com/presentations/bh-dc-10/Nve_Leonardo/BlackHat-DC-2010-Nve-Playing-with-SAT-1.2-slides.pdf

Rontti, T., Juuso, A-M. and Tirilä, J-M. (2011), "Securing Next Generation Networks by Fuzzing Protocol Implementations", Paper read at Technical Symposium at ITU Telecom World (ITU WT), Vienna, Austria, November.

Rontti, T., Juuso, A-M. and Takanen, A. (2012), "Preventing DoS Attacks in NGN Networks with Proactive Specification-based fuzzing", Communications Magazine, IEEE, Volume 50, Issue 9, September 2012, pp. 164 – 170.

Sale, R. (2013), "Cyber War Stakes Rising", [online], http://www.isssource.com/cyber-war-stakes-rising/.

Sommer, P. and Brown, I. (2011), "Reducing Systemic Cybersecurity Risk", OECD Project Future Global Shocks, [online], http://www.oecd.org/governance/risk/46889922.pdf

Takanen, A., Demott, J.D. and Miller, C. (2008), Fuzzing for Software Security Testing and Quality Assurance, Artech House.

Takanen, A. (2010), "Fuzzing: Helping to Avoid Zero-Day Attack", [online], http://www.continuitycentral.com/feature0754.html

Takanen, A. (2011), "Unknown Vulnerability Management and Testing", Fuzzing 101 Webinar, January, [online], http://www.codenomicon.com/resources/webcasts/20110120.shtml

U.S.-China Economic and Security Review Commission (2011), "2011 report to Congrss of the U.S.-China Economic and Security Review Commission, [online], http://origin.www.uscc.gov/sites/default/files/annual_reports/annual_report_full_11.pdf

Varpiola, M. (2012), "Embedded Device [fuzz] testing [against [A]PT]], Government and Defense perspective", presentation at Amphion Forum, Washington DC, United States, June.

Wang, C. and Takanen, A. (2011), "Fuzz your infrastructure - the blackhats are doing it, shouldn't you?", Codenomicon and Forrester, Fuzzing 101 Webinar, April.

Washington Post, The (2012), "Understanding cyberspace is key to defending against digital attacks", June, [online], http://www.washingtonpost.com/investigations/understanding-cyberspace-is-key-to-defending-against-digital-attacks/2012/06/02/gJQAsIr19U_story.html

Work
in
Progress
Papers

Efficient and Secure Remote Data Storing and Processing

Mohd Rizuan Baharon, Qi Shi, David Llewellyn-Jones and Madjid Merabti
School of Computing and Mathematical Sciences, Liverpool John Moores University, UK
M.R.Baharon@2011.ljmu.ac.uk
Q.Shi@ljmu.ac.uk
D.Llewellyn-Jones@ljmu.ac.uk
M.Merabti@ljmu.ac.uk

Abstract: Storing and processing data remotely is becoming a popular trend for people to deal with their data in order to overcome storage spaces and computing resources limitation. However, moving private data to untrusted third parties like clouds with limited digital control by data owners raises security concerns. Primitive encryption schemes seem ineffective to be used as such techniques require encrypted data need to be decrypted first before data can be processed. Thus, homomorphic encryption is believed to be one of the potential solutions as it allows arbitrary computation on encrypted data without decryption process. The first fully homomorphic encryption (FHE) scheme was introduced by Gentry, then followed by other researchers that produced similar and complex schemes like Gentry's one. Such schemes are suffering from poor efficiency as too much noise is produced by the process on encrypted data. Thus, this research work will deeply look at efficiency issue and propose a new scheme which minimizes the use of noise at the processing stage. The scheme implements the Elliptic Curve (EC) group as the underlying group since EC promises efficiency and strong security. The use of *n*-multilinear map along with the construction of the scheme is to enable the achievement of a FHE scheme. The key contribution of this work is to propose a FHE scheme with improved efficiency. Furthermore, an improve security of a Secure Sockets Layer protocol will be the second contribution as the used of the proposed FHE scheme enables data to be transmitted and processed securely and efficiently.

Keywords: cloud computing, data storage, homomorphic encryption, elliptic curve, *n*-multilinear map, secure sockets layer (SSL)

1. Introduction

Cloud Computing is becoming a popular IT technology that has changed the way people use IT technologies to run their businesses. This is due to cloud providing huge data storage and powerful computing resources to their clients who have less capability to store and process their data internally. Those benefits can be accessed with minimum requirements like desktop machines and Internet connection. As clouds provide services on a pay-as-you-use basis, it enables their clients to leverage their services based on client's needs. However, clients are still reluctant to adopt such a technology due to security concerns on their data (Subashini & Kavitha 2011).

Thus, research on securing data and their related processing by cloud-based applications is getting more attention from academia as well as enterprises working on or using cloud services. Clients, who are interested to use cloud services effectively, need to outsource their data to Cloud Service Providers (CSPs). However, outsourcing sensitive data into clouds with no physical and limited digital control by the clients raises serious concerns about data security. Due to the scale, dynamicity, openness and resource-sharing nature of cloud computing, addressing security issues in such environments is a very challenging problem (Zissis & Lekkas 2012).

To ensure privacy and integrity of the data is preserved, encryption techniques should be implemented. Primitive encryption schemes seem ineffective to be implemented in such an environment because the encrypted data cannot be processed without decryption (Mahmood 2011). Thus, a scheme that allows data to be processed in encrypted form like a FHE scheme is needed. However, existing FHE schemes are suffering from efficiency issues as they are computationally expensive. Ciphertexts generated through those schemes are "noisy" (Fan & Vercauteren 2012). Such limitations require an improved FHE scheme to be proposed. Thus, we propose a new FHE scheme based on a finite field that supports an *n*-multilinear map in this paper. The scheme is constructed based on an open problem raises by Boneh et al. (Boneh 2005), on their previous work on a bilinear map. Our scheme will be executed using SSL protocol to achieve high security between clients and cloud providers.

SSL is developed by Netscape (Chou 2002) providing a secure communication channel and mechanisms between two parties (a client and a server). The goals for SSL not only include security, but also interoperability, extensibility, and relative efficiency (Lee et al. 2007). An existing SSL provides a great solution for a secure communication in a cloud environment. However, only transmitted data that have been decrypted first can be processed by cloud-based applications. Thus, the implementation of our scheme in such a protocol enables data to be transferred and processed securely without the need of the decryption process.

This paper is structured as follows. We first describe our research contributions in Section 2 and briefly review some important concepts in Section 3. We then present some analysis, discussion and preliminaries results in Section 4. Finally, we conclude this paper in Section 5.

2. Our contributions

The expected key contributions of this paper are summarised in the following points:

- A new FHE scheme. We proposed a new FHE scheme based on a finite field that supports an *n*-multilinear map. We implement an EC group as the underlying group as EC promises high efficiency and strong security. Previous work based on a bilinear map allows arbitrary addition and one multiplication. Thus, to achieve arbitrary multiplication on encrypted data, an *n*-multilinear map will be implemented to the scheme subject to the existence of its generator in the map.
- An improve security of data processing using SSL protocol. SSL protocol uses symmetric and asymmetric encryption schemes to enable a huge amount of sensitive data to be transmitted securely and efficiently. However, such schemes do not allow data to be processed in cloud environments without a decryption process. Thus, a combination of symmetric encryption and the proposed scheme (asymmetric one) allows data to be transmitted and processed without to decrypt them first. Such protocol ensures sensitive data can be transmitted and processed securely and efficiently.

3. Background

This section describes fundamental concepts and definitions that have been used in our scheme.

3.1 An *n*-Multilinear map

Definition of n-Multilinear Map: (Papamanthou et al. 2010)

Let $G_1, G_2, \dots, G_n$ and G_t be cyclic groups of the same prime order *q* and $\mathbf{Z}_q^*$ be a finite field that is closed under multiplication operation. n -multilinear groups $G - G_1 - G_2 - \dots - G_n$ are all isomorphic to one another as they have the same order and are cyclic. An *n*-multilinear map is a function $e: G_1 \times G_2 \times \dots \times G_n \rightarrow G_t$ such that the following properties are satisfied:

- For all $a_1, a_2, \dots, a_n \in \mathbf{Z}_q^*$ and $g_1, g_2, \dots, g_n \in G$, $e(g_1^{a_1}, g_2^{a_2}, \dots, g_n^{a_n}) = e(g_1, g_2, \dots, g_n)^{a_1 a_2 - a_n} \in G_t$.
- The map is non-degenerate: If $g \in G$ generates G then $e(g, g, \dots, g) \in G_t$ generates G_t.

The Construction of n-Multilinear Groups of Order n Using Elliptic Curve Group:

An n -multilinear group $G = G_1 = G_2 = \dots = G_n$ of order *n* can be constructed as follows:

- Let $l = 2$, and $n = 10$ such that n is a square-free integer that is not divisible by 3. A square-free integer is one divisible by no square number, except 1. Then, $q = ln - 1 = 2(10) - 1 = 19.$
- Let $E_{(1,0)}(F_{19}): y^2 = x^3 + x$ defined over a finite field F_{19} be the groups of points. The curve has $q + 1 = ln = 20$ points in F_{19}. Thus, there exists a subgroup G in $E_{(1,0)}(F_{19})$ of order 10 since $n = 10$.
- Let G_t be the subgroup of a finite field that is closed under multiplication, $F_{19^2}^* = F_{361}^*$ of order *n*. Our aim is to have an *n*-multilinear map $e: G_1 \times G_2 \times \dots \times G_n \rightarrow G_t$ which includes the admissible *n*-multilinear map generator.

3.2 Elliptic curve over finite field F_q

Let $q > 3$ be an odd prime. An EC *E* over a prime field F_q is defined by an equation of the form:

$$y^2 = x^3 + ax + b$$

where $a, b \in F_q$, and $4a^3 + 27b^2 \not\equiv 0 \bmod q$. The set $E(F_q)$ consists of all points $(x, y), x \in F_q$, together with *O*. Furthermore, algebraic formula of adding distinct points and doubling a point on the curve are given as follows:

- $P + O = O + P = P$ for all $P \in E(F_q)$.
- If $P = (x, y) \in E(F_q)$, then $(x, y) + (x, -y) = O$. (The point $(x, -y)$ is denoted by $-P$, and is called the negative of *P*).
- (Point addition) Let $P = (x_1, y_1) \in E(F_q)$ and $Q = (x_2, y_2) \in E(F_q)$ where $P \neq \pm Q$. Then $P + Q = (x_3, y_3)$, where $x_3 = \left(\frac{y_2 - y_1}{x_2 - x_1}\right)^2 - x_1 - x_2$ and $y_3 = \left(\frac{y_2 - y_1}{x_2 - x_1}\right)(x_1 \quad x_3) \quad y_1$.
- (Point doubling) Let $P = (x_1, y_1) \in E(F_q)$, where $P \neq -P$. Then $2P = (x_3, y_3)$, where $x_3 = \left(\frac{3{x_1}^2 + a}{2y_1}\right)^2 - 2x_1$ and $y_3 = \left(\frac{3{x_1}^2 + a}{2y_1}\right)(x_1 - x_3) - y_1$.

4. Analysis, discussion and preliminary results

This section describes some analysis, discussion and preliminary results of our proposed scheme.

4.1 The proposed scheme

Suppose $m \in \{0,1\}$ be a plaintext and c be a ciphertext.

- KeyGen: Let G and G_t be finite fields of order n where $n = q_1 q_2$ such that q_1 and q_2 are two distinct random prime numbers. Pick two random generators $g, u \in G$ and set $\boldsymbol{h} = u^{q_2}$. The public key is $Pk = (n, G, G_t, e, g, \boldsymbol{h})$. The secret key is $Sk = q_1$.
- Encryption: $c(m) = g^m \boldsymbol{h}^r$.
- Decryption: $\llbracket m = c^{\dagger}(-1)\ (c(m)) = \llbracket log \rrbracket\ {}_{\downarrow}g^{\wedge}\ c\rrbracket\ ^{\dagger}(q_1 1)$ such that $\hat{g} = g^{q_1}$.

4.2 Scheme requirements

The scheme should hold such properties:

4.2.1 Homomorphic under $*$ operations

- $C - \sum_{i=1}^{n} c_i - \prod_{i=1}^{n} c_i\, \boldsymbol{h}^r$ is homomorphic under addition.
- $C = \prod_{i=1}^{n} c_i$ is homomorphic under multiplication.

4.2.2 Double layer encryption

- To ensure the privacy of the outsource data is preserved.

4.3 Basic description protocol

The protocol is illustrated in Figure 1, and its steps are explained below:

1. A creates raw data $V = v_i \in \{0,1\}$, $W = w_i \in \{0,1\}$ for $i = 1, 2, \ldots, n$ and secret key S_A . Then, *A* encrypts V and W using S_A, and encrypts S_A using P_B.
2. A sends $\bar{c}_{S_A}(V)$, $\bar{c}_{S_A}(W)$ and $c_{P_B}(S_A)$ to *C*.
3. B requests *C* to re-encrypt $\bar{c}_{S_A}(V)$ and $\bar{c}_{S_A}(W)$ using P_B. Then, *B* sends instructions to process on $c_{P_B}(\bar{c}_{S_A}(V))$ and $c_{P_B}(\bar{c}_{S_A}(W))$.
4. C re-encrypts $\bar{c}_{S_A}(V)$ and $\bar{c}_{S_A}(W)$ using P_B and run initial processes called partial decryption process on $\bar{c}_{S_A}(V)$ and $\bar{c}_{S_A}(W)$:

- $c_{P_B}(S_A) *_d c_{P_B}(\bar{c}_{S_A}(V)) - c_{P_B}(S_A *_d \bar{c}_{S_A}(V)) - c_{P_B}(V)$.
- $c_{P_B}(S_A) *_d c_{P_B}(\bar{c}_{S_A}(W)) = c_{P_B}(S_A *_d \bar{c}_{S_A}(W)) = c_{P_B}(W)$.

Then, $c_{P_B}(V)$ and $c_{P_B}(W)$ is computed to produce a result $c_{P_B}(V * W)$.
5. C sends the result $c_{P_B}(V * W)$ to *B*.
6. B decrypts the result using S_B to recover $V * W$.

Definition of $*_d$:

Let a and b be integers such that a is a secret key to encrypt a plaintext x and $b = c_a(x)$ is a ciphertext. Then, $a *_d b = c_a^{-1}(b) = x$.

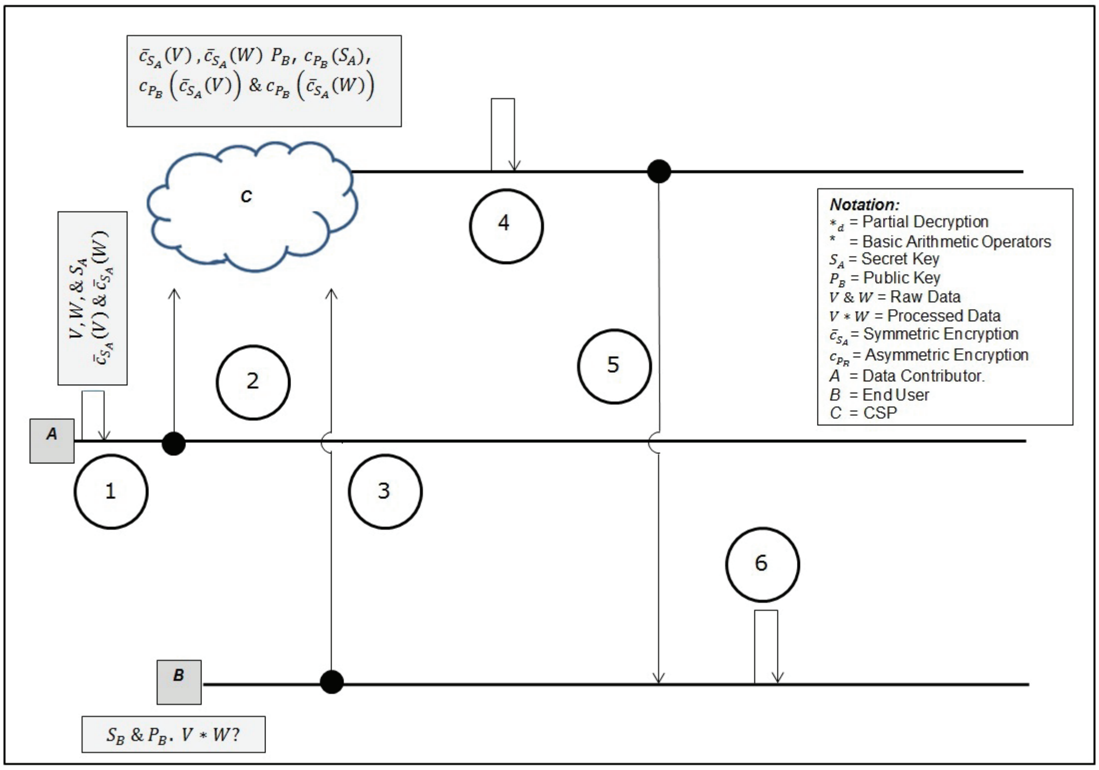

$\bar{c}_{S_A}(V)$, $\bar{c}_{S_A}(W)$ P_B, $c_{P_B}(S_A)$,
$c_{P_B}(\bar{c}_{S_A}(V))$ & $c_{P_B}(\bar{c}_{S_A}(W))$
C
$V, W, \& S_A$
$\bar{c}_{S_A}(V)$ & $\bar{c}_{S_A}(W)$
A
1
2
3
4
5
6
B
S_B & P_B. $V * W$?
Notation:
$*_d$ = Partial Decryption
$*$ = Basic Arithmetic Operators
S_A = Secret Key
P_B = Public Key
V & W = Raw Data
$V * W$ = Processed Data
$\bar{c}_{S_A}$ = Symmetric Encryption
c_{P_B} = Asymmetric Encryption
A = Data Contributor.
B = End User
C = CSP

Figure 1: The protocol

4.4 Preliminary results

Our preliminary results based on performance of encryption/decryption process have been summarised in table 1.

Table 1: Performance analysis for encryption/decryption process

Performer	Tasks	Method	Performance	Descriptions/Results
A	Encrypts raw data (*V* and *W*) using S_A	Symmetric encryption scheme	Fast	$\bar{c}_{S_A}(V)$, and $\bar{c}_{S_A}(W)$
	Encrypts S_A using P_B	The proposed scheme (asymmetric)	Fast	The key size is short. $c_{P_B}(S_A)$.
B	Decrypts $c_{P_B}(V * W)$ using S_B	The proposed scheme (asymmetric)	Fast/Slow	It is depends on the size of the output. $V * W$.
C	Re-encrypts $\bar{c}_{S_A}(V)$, and $\bar{c}_{S_A}(W)$ using P_B	The proposed scheme (asymmetric)	Fast	$c_{P_B}(\bar{c}_{S_A}(V))$, and $c_{P_B}(\bar{c}_{S_A}(W))$
	Run a partial decryption process		Fast	$c_{P_B}(S_A) * c_{P_B}(\bar{c}_{S_A}(V)) = c_{P_B}($. $c_{P_B}(S_A) * c_{P_B}(\bar{c}_{S_A}(W)) = c_{P_B}(S_A * \bar{c}_{S_A}(W)) = c_{P_B}(W)$.
	Computes $c_{P_B}(V * W)$	The proposed scheme (asymmetric)	Fast	*C* has a lot of computer resources power

5. Conclusion

We have described a secure and efficient FHE scheme for processing remote data in an encrypted form. Our technique implements EC as the underlying group as EC promises efficiency and an *n*-multilinear map to achieve a FHE. Our proposed scheme is implemented into the existing SSL protocol. The security of data that was encrypted, transmitted and processed using such a protocol is guaranteed in cloud environments due to no information disclosed at any stage. This work will be achieved if the generator of an *n*-multilinear map can be proved exists in the map. Thus, further work of this paper will be looking on various ways to prove that the generator exists in the map and can be computed efficiently. In addition, we will provide a proper security analysis of the scheme that will be based on a subgroup decision problem.

References

Boneh, D., 2005. Evaluating 2-DNF Formulas on Ciphertexts. In *Second Theory of Cryptography Conference, TCC 2005, Cambridge Proceedings*. pp. 325–341.

Chou, W., 2002. Inside SSL: The Secure Sockets Layer Protocol. *IT Professional*, 4, pp.47–52.

Fan, J. & Vercauteren, F., 2012. Somewhat Practical Fully Homomorphic Encryption. In *IACR Cryptology ePrint Archive*.

Lee, H.K., Malkin, T. & Nahum, E., 2007. Cryptographic Strength of SSL / TLS Servers: Current and Recent Practices. In *Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement*. pp. 83–91.

Mahmood, Z., 2011. Data Location and Security Issues in Cloud Computing. *2011 International Conference on Emerging Intelligent Data and Web Technologies*, pp.49–54.

Papamanthou, C., Tamassia, R. & Triandopoulos, N., 2010. Optimal Authenticated Data Structures with Multilinear Forms. *Procceeding of the 4th international conference on Pairing-based Cryptography*, pp.246–264.

Subashini, S. & Kavitha, V., 2011. A survey on security issues in service delivery models of cloud computing. *Journal of Network and Computer Applications*, 34(1), pp.1–11.

Zissis, D. & Lekkas, D., 2012. Addressing cloud computing security issues. *Future Generation Computer Systems*, 28(3), pp.583–592.

How to Improve Network Security Using Gamification

Anthony Keane and Jason Flood
Institute of Technology Blanchardstown, Dublin, Ireland
anthony.keane@itb.ie
jasoneflood@gmail.com

Abstract: Computer network systems have been shown in many surveys to be inherently vulnerable to breaches from both internal and external attacks. The majority of external network attackers are people with little real skills and those with better hacking skills often follow the same known sequence of repetitious attacks that largely rely on the opportunity of chance for success. Why computer networks are insecure at all often originates from the actual users and administrators of the systems. As IT Administrators depend on the manufacturers for security in their products, so does the hacker depend on the manufacturers for exploitable vulnerabilities in their products. The IT administrator looks for certain behavioral strengths in end-users while the hackers look for behavioral weaknesses in the same end-users. The totality in the security of the network lies in the balance between the IT Administrators knowledge/skills and the hacker's knowledge/skills. The margin of separation is often in the ability of each party to learn new tricks with time playing a crucial part. We propose a training system for IT Administrators to strengthen their knowledge and skills in the hope of tilting the security balance in their favour. Our system is based on providing a cycle of system security testing incorporated with training in Capture-The-Flag gamification via Cloud hosted virtual server systems. These are built and maintained by knowledge providers from voluntary organisations like the Honeynet Project, OWASP and many other more individuals at the forefront of network security in their international organisations.

Keywords: CTF, hacking, security training, network vulnerability, IT administrator training

1. Background overview

Computer Network insecurity stems from the many complex issues in design and development of the hardware and software components that make up a network. Firewalls, intrusion detection and prevention, virus scanning and filtering are traditionally utilised by administrators to protect their networks from the cyber threats coming from the Internet and other networks.

A summary of the current threats can be seen in many survey reports like the 2011 Symantec Internet Security Threat Report that provides an overview of the global threat activity. *The report is based on data from the Global Intelligence Network, which Symantec's analysts use to identify, analyze, and provide commentary on emerging trends in attacks, malicious code activity, phishing, and spam,* (Symantec 2012). Symantec recorded over 5.5 billion malware attacks, an 81% increase over 2010. Web based attacks increased by 36% with over 4,500 new attacks seen each day. 403 million new variants of malware created, a 41% increase over 2010 and 39% of all malware attacks via email used a link.

It has been shown that the vast majority of potential intruders are people with little real hacking skills and rely on the opportunity of chance to breach network systems. These script-kiddies use out-of-the-box tools and they have little understanding of how the tools work or what to do, once a breach has been made, (Honeynet 2012). As a result, these kind of attacks should be easy to defend against. While the IT Administrators depend on the manufacturers for security in their products, so does the hacker depend on the manufacturers for vulnerabilities. The security of the network hangs in the balance between the IT Administrators knowledge/skills being greater than the hacker's knowledge/skills. The deciding factor is often the speed at which each party learns new offensive and defensive tricks with time playing a crucial part in this cat-and-mouse game. The ultimate game changer is zero day attacks that take a long time to find and protect against but have a lower probability of occurrence but a higher impact risk.

In this paper, we are proposing a training system based on CTF games for IT Administrators to strengthen their knowledge and skill. Our system relies on a cloud hosting virtual server systems, built and maintained by highly skilled knowledge providers from international organisations such as Honeynet Project, OWASP and IRISS.

2. Identification of basic skill sets needed and why?

In 2012, the Verizon RISK Team in cooperation with the Australian Federal Police, Dutch National High Tech Crime Unit, Irish Reporting and Information Security Service (IRISS), Police Central e-Crime Unit, and United States Secret Service gathered data on system breaches and found that 81% of all compromised systems were hacked via attacks circumventing authentication through combining stolen or guessed credentials. 69% incorporated malware while 10% resulted from privilege misuse. Only 5% employed social tactics.

79% of victims had readily exploitable weakness in their systems. 96% of all attacks were considered simple and avoidable (in hindsight) without difficult countermeasures. This discovery fuels our hypothesis that better understanding of hacker techniques by IT Administrators could significantly reduce the amount of successful attacks. 94% of data compromised involved servers and 85% of breaches took 3 or more weeks to discover and often it was by a third party. The Verzion study states that *The challenge for the good guys lies in selecting the right tools for the job at hand and then not letting them get dull and rusty over time. Evidence shows when that happens, the bad guys are quick to take advantage of it, (DBIR 2012).*

The top threat actions against larger organizations are: Keyloggers and the use of stolen credentials, backdoors and command control, Tampering, Pretexting, Phishing, bruteforce and SQL injection.

The recommendations from Verzion are; implement a firewall or ACL on remote access services, change default credentials of systems and monitor/analyse the event logs. We would also add our own recommendations: Good preventive security practices are a must. Including installing and keeping firewall policies, antivirus and Operating Systems updated. Deploy an IPS. Have a well-planned incident-response measures. Limit connections to only those required for business needs. Testing of all these precautions is essential and this is where the CTF can play a role with IT Administrators.

Table 1: Eight common threat actions (DBIR 2012)

malware	Keylogger/form-grabber/spyware (capture data from user activity
malware	send data to external site/entity
malware	backdoor (allows remote access/control)
malware	Disable or interfere with security controls
hacking	exploitation of default or guessable credentials
hacking	exploitation of backdoor or command and control channel
hacking	brute force and dictionary attacks
hacking	Use of stolen login credentials

When global attack data is analysed, (Lindstrom 2012), the majority of attacks are seen to be different combinations of eight threat actions (see table 1), notice where basic hacking is combined with malware occurring frequently. Knowing more about attack sequencing is critical to recognizing the signature of an attack in development and helps to mitigate the likelihood of success.

3. Design of CTF environment

In computer security, Capture The Flag (CTF) is a computer network based competition that tests the typical skills needed to protect networks using known top-level hacks. There are two main styles of CTF; (1) In an **Attack/Defense** style competition; each team is given a networked machine(s) to defend. Teams are scored on both defending their machine and on attacking other machines; (2) **Jeopardy** style competitions usually involve multiple categories of problems, each of which contains a variety of questions of different point values.

Virtualization is often used to host a CTF as it has the capacity for making inventive leaps in content delivery. With this approach, players in the game can benefit from real-time feedback from their actions, the experience becomes more responsive and log file/packet analysis allows for rapid gathering of attack fingerprints. The CTF use of routers and switches allows for easy extension and scalability. Typically this kind of set up consist of approximately thirty individual challenge levels, meaning thirty separate virtual machines. Allowing players of the game root access without risking the integrity of the game or preventing other users from solving the challenge in parallel is a complex problem to solve. Attack/Defense game play is has a simpler technical challenge to be resolved where mainly ensuring that a player who successfully captures a machine does not bring the machine down and prevent others from completing the level is the central problem.

4. Metrics of skill proficiency and improvement

A primary challenge for the designers of CTFs is to cater for the skills difference between the players themselves, (Werther 2011). And the primary goal of any CTF is to encourage the participants to make new contacts and engage in social networking as this activity encourages learning pedagogy.

Results of our research carried out using numerous CTF challenges ran in Dublin in 2012 for OWASP Irish Chapter and IRISS, (Flood 2012) demonstrated that the typical security professional's experience was limited to "low-hanging fruit" knowledge.

In a traditional CTF event there is no mechanism for players to be assisted. This was recognized as a significant learning gap and limited the numbers willing to participate in the CTF challenges. With the introduction of the automated learner feedback module, learners had assistance throughout the games.

5. Conclusions and future work

Participation in CTFs increase security awareness by providing, in a cost effective way, detailed information on actual, exploitable security threats. This allows an organization to more intelligently prioritize re-mediation. Recovering from a security breach can cost an organization millions of dollars in IT re-mediation efforts; CTF's avoid these financial pitfalls by enabling the identification of vulnerabilities before exploitation. Improved security awareness protects an organization's reputation and trustworthiness. Many businesses now seek reassurance from partners concerning their commitment to security. Successes at CTFs are a key indicator in this regard.

Computer network systems can be made more secure if IT Administrators can become more aware of the old and new vulnerabilities in their systems and of the methodologies and tools used by the vast majority of hackers, which we have seen is limited to a small number of threat actions. The best and most interesting way of helping IT Administrators to fulfill their needs is for them to play CTF games with education features embedded in the challenges.

Our success in running such CTFs in Dublin has prompted us to look at expanding the number of challenges to a global audience via Worldwide organisations like OWASP and the Honeynet Project. This will enable others to contribute challenges and for us to get better statistics of what works with whom and how. Out goal is and will continue to be to enable network administrators to use available knowledge to best safeguard the networks under their control.

References

DBIR (2012), The Verzion 2012 Data Breach Investigations Report, [online] www.verizonbusiness.com/about/events/2012dbir/

Flood J, Denihan M., Keane A. and Mtenzi F. (2012), Black Hat Training of White Hat Resources: The Future of Security is Gaming, 7th International Conference for Internet Technology & Secured Transactions (ICITST-2012) London, UK

Honeynet (2012) Know Your Enemy: The Tools and Methodologies of the Script Kiddie [online] Honeynet Project: http://project.honeynet.org

OWASP (2010) OWASP Top Ten Web Application Security Risks for 2010 [online] https://www.owasp.org/index.php/Category:OWASP_Top_Ten_Project

Lindstrom G., (2012) Meeting the Cyber Security Challenge Publ. by Geneva Centre for Security Policy, [online] www.gcsp.ch/content/download/9408/113285/download

Stallings W. (2003) Network Security Essentials: Applications and Standards, 2nd Ed., Prentice Hall, New Jersey.

Symantec (2012) Symantec Internet Security Threat Report, Vol.17, April 2012, Mountain View, California, [online] http://www.symantec.com/threatreport/

Werther J., Zhivich M., Leek T., Zeldovich N., (2011) Experiences In Cyber Security Education: The MIT Lincoln Laboratory Capture-the-Flag Exercise, MIT Lincoln Laboratory, [online] http://people.csail.mit.edu/nickolai/papers/werther-llctf.pdf

Truly Protect Video Delivery

Nezer Zaidenberg and Asaf David
University of Jyvaskyla, Jyvaskyla, Finland
nezer@truly-protect.com
asaf@truly-protect.com

Abstract: Siwek (2006) has shown the cost of video piracy on video distribution is putting great strain on the video industry. In Averbuch et al(2011) one of the authors described a system for protected digital content distribution without hacker's ability to create illegal copies. In this paper we present a work in progress - enhancements to the aforementioned truly-protect system for video distribution.

Keywords: trusted computing, video delivery, digital rights management

1. Introduction

The Internet is becoming the leading platform for multimedia content distribution. The proper license scheme of such data is turning to be a major concern for media owners and providers. DRM (Digital Rights Management) solutions are the standard tool used to enforce proper terms of content use. While many DRM methods were proposed and implemented over the years. This was shown by MoRE and [dEZZY/DoD] (1999), Leyden (2007) and others. Hackers find ways to circumvent DRM solutions often by exploiting the ability of a malicious client to intercept the tool in runtime and obtain the encryption key. In this paper we introduce a new DRM solution, specifically aimed at streaming video.

Averbuch et al (2011) has introduced "Truly-Protect" system for content protection. This system relies on the encryption and creation of encrypted copy of the entire content for end-user. However, encryption of the entire content is not feasible in Video application (due to the sheer volume of the video data). This is especially true for applications that require different content for each user such as video conferencing etc. The volumes of video traffic with the requirement to encrypt every single stream will create too much strain on the server CPU.

The solution that we present tries to prevent this vulnerability. The decryption of the video and the rendering of frames happen atomically inside a virtual machine, which establishes a controlled and trusted environment. The machine-code instructions run by the VM are encrypted per client. The decryption of instructions happens only at runtime, using an encryption key that is held securely inside the CPU, inaccessible to any other process. Moreover, the genuineness of the system is verified prior to the key exchange. We believe this system achieves an improved level of security compared to existing DRM solutions

The system constructed can now distribute protected video and secondary ports for other codecs are currently being developed.

2. System design

The system requirements for video distribution are as follows

- Use modern encoding and viewing technology
- Maintain the original codec encryption rate
- Maintain the original codec quality
- Encrypt as little as possible
- Maintain viewer experience ("smoothness" of video)
- Remain resistant to attempts to view or temper with the system

These requirement were meant to allow the system provide acceptable viewing quality at acceptable bandwidth compared to the competition while reducing the stress on the streaming server.

In order to maintain requirement 1 we decided to use h264 codec.

Requirement 6 forces us to use kernel module because code running in user mode can very easily be inspected. (Code running on supervisor mode can still be inspected but “Truly-Protect” includes Kennell et al’s(2003) protocol that detects emulators) In order to meet requirements 2-5 we will follow the procedure described by Stuetz and Uhl(2010). In summary we will do the following in the client system

- Perform the “Truly-Protect” genuinely test the truly protect test ensure that the correct software is running and not an emulated copy. This is achieved by performing a computation on the memory blocks that holds the “Truly-Protect” code, checking the side effects as well as result are accurate within bounded time frame. (Stuetz and Uhl(2010) describe this method in greater detail)
- Only after the end user machine has been validated and a private key known only to end user CPU and stored in a protected CPU register we send our public key to the “Truly-Protect” server and receive the video decryption key encrypted by our public key. We decrypt the key and keep the video key saved only inside the End-User CPU where it is not accessible by tapping to the communication lines or buses or using any software on the End-User machine. Averbuch (2011) describes this in greater detail at el.
- Decrypt an h264 video motion vectors in the kernel after successful authentication.
- Plays the video normally on the user machine

We will also develop a server system that communicates with said client system.

3. Preliminary results

Our current system is able to decode and encrypt h264 video that remains protected to crackers attack while only encrypting 1% of the video traffic.

The operation system does all decoding (both h264 and audio as encryption decoding) algorithms by Linux kernel modules. The system offer improvement over current DRM solutions as does not rely on obfuscation methods (security by obscurity).

4. Attacks on the system

Like any video delivery system this system is vulnerable to attacks on the video output. (Such as attacks on the HDMI output as shown by Felton (2006) to simply recording the output using a camcorder.) While video can be grabbed and re-encoded in such method these methods degrade the output stream quality and add latency. Furthermore, they cannot be avoided. But the system is immune to all methods of tempering with the system or decoding the stream. Furthermore, the key exchange algorithms used in this system is not vulnerable to tampering, eavesdropping or man in the middle attacks.

References

Averbuch, A, Kiperberg, M, Zaidenberg, N. (2011) An Efficient VM-Based Software Protection. In NSS 2011.

Felton, Ed. (2006) Making and Breaking HDCP Handshakes [Online] available: https://freedom-to-tinker.com/blog/felten/making-and-breaking-hdcp-handshakes

Kennell, Rick and Jamieson Leah H.,. Establishing the genuinity of remote computer systems. In Proceedings of the 12th USENIX Security Symposium, 2003.

Leyden, J. (2007) [Online] Blu-ray DRM defeated available: http://www.theregister.co.uk/2007/01/23/blu-ray_drm_cracked/

MoRE and [dEZZY/DoD] (1999) [Online] The Truth about DVD CSS cracking available: http://www.cs.cmu.edu/~dst/DeCSS/MoRE+DoD.txt

Siwek S.E., (2006) The True Cost of Motion Picture Piracy to the U.S. Economy in *Policy Report 186*

Stuetz, T, Uhl,A. (2010) A Survey of H.264 AVC/SVC Encryption in *Technical Report*

ECCWS 2014

Piraeus, Greece

13th European Conference on Cyber-Warfare and Security
University of Piraeus, Greece
3-4 July 2014

For further information contact
info@academic-conferences.org
Or telephone
+44-(0)-118-972-4148

ISBN 978-1-908272-55-3