Intelligent Systems, Control and Automation: Science and Engineering

Martti Lehto
Pekka Neittaanmäki *Editors*

Cyber Security: Analytics, Technology and Automation

Springer

Intelligent Systems, Control and Automation: Science and Engineering

Volume 78

More information about this series at http://www.springer.com/series/6259

Martti Lehto · Pekka Neittaanmäki
Editors

Cyber Security: Analytics, Technology and Automation

Springer

Editors
Martti Lehto
Department of Mathematical Information Technology
University of Jyväskylä
Jyväskylä
Finland

Pekka Neittaanmäki
Department of Mathematical Information Technology
University of Jyväskylä
Jyväskylä
Finland

ISSN 2213-8986 ISSN 2213-8994 (electronic)
Intelligent Systems, Control and Automation: Science and Engineering
ISBN 978-3-319-18301-5 ISBN 978-3-319-18302-2 (eBook)
DOI 10.1007/978-3-319-18302-2

Library of Congress Control Number: 2015938724

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

Foreword

Global cyberspace is made of complex, multi-layered information networks which encompass the communication networks of the public sector, business community, security authorities and control and monitoring systems used by industry and critical infrastructure which, by means of the Internet, create a worldwide network.

Imaginative data processing and utilization, arising from the needs of citizens and the business community, are some of the most important elements of a thriving society. Information and know-how have become key 'commodities' in society, and they can be utilized all the more efficiently through information technology. Different interactive electronic services are ubiquitously available, irrespective of time and place. While the public sector, the economy, the business community and citizens benefit from globally networked services, the digital IT society contains inherent vulnerabilities which can generate security risks to citizens, the business community or the vital functions of society.

Society is gradually becoming an information-based service culture providing, to an ever greater extent, both public and commercial digital services to citizens. Electronic ICT networks and digital services are vital to the functioning of society. Along with the general trends of change, the advancements in technology and the utilization of the Internet, the operating environment is heavily influenced by the global nature of this increasingly expanding sector and the changing habits among users as well as the challenges associated with reliability and security.

Cyber security risks have become more and more commonplace. Risks which were once considered improbable are now appearing all the more regularly. This trend epitomizes the new forms of instruments and methods being used in attacks, as well as ever-increasing vulnerabilities and the higher motivation of the attackers. The growing impact of cyber-attacks calls for new, creative and innovative solutions so as to mitigate the risks. In the past, individual persons or small hacker groups were the attackers, nowadays, however, various state-run organizations employing state-of-the-art cyber weapons are, in particular, carrying out targeted attacks. These so-called Advanced Persistent Threats (APTs) focus on carefully selected targets; their development requires sophisticated expertise and ample resources.

One global trend is that services are being moved to the cloud. Public officials, companies and citizens alike are increasingly switching to cloud storage and cloud computing. As a concept, cloud computing illustrates a change in the paradigm, one in which services are provided within a 'cloud' whose technical details remain opaque to and beyond the control of the users of the service. Cloud computing exhibits a new model of generating, using and providing ICT services, which includes dynamically scalable virtual resources as services provided over the Internet. In accordance with the prevailing trend government organizations are increasingly moving critical IT infrastructure-related data to the cloud, which brings about new cyber security challenges. Cloud computing and cloud services are integrally linked with Big Data, which is being used on a platform for the creation of new services to end users. This, in turn, necessitates close cooperation between cloud service providers and cyber security solution providers. In the future cloud services will lay emphasis on the generation of specific cyber security solutions and the protection of identity and privacy as well as miscellaneous solutions associated with data encryption.

The Internet of Things (IoT) stands for the transformation of industry where industrial products and industrial production utilize the Internet, nanotechnology and the entire ICT sector. The IoT gives objects, or "things", recognizable identities and they communicate across the global ICT network. New equipment, such as different industrial and service robots as well as information-gathering sensors, are linking into such networks at an increasingly accelerated pace. The latest step in this development involves different kinds of vehicles, such as cars, trolleys and buses as well as different types of heavy machinery.

Modern cars are smart devices with most of their systems controlled by computers. Communication between other cars, traffic control systems and user devices (e.g. smart phones) is also increasing. While infotainment systems provide many services for the driver, they can also be a distraction. All of this information traffic poses the risk of technical or user errors, and even enables remote attacks against cars.

Cross-disciplinary and holistic cyber security research is needed to solve these new challenges. Due to the complexity of the field, research must meet the four basic paradigms of science: the theoretical, experimental, model-based and data-based computational approaches.

Computational science represents the third paradigm of science, one which uses computers to simulate phenomena or situations which may not yet exist in the real world. Rapid advances in IT technology and methodological competence facilitate the introduction of increasingly complex and realistic computational models for solving research related problems. The methods of computational science can also be successfully employed when seeking solutions to situations where traditional methods fail to generate sufficiently accurate results. A computational approach can increase awareness among those sectors of cyber security which are important to society. The computational approach does not only strengthen multi- and cross-disciplinary research, it also expedites and intensifies product development. Simultaneously, it helps lower the barriers between fields of research in both the

public and private sector. It also boosts innovation and generates new breakthroughs in research and product development.

In many cases, large-scale simulations are accompanied by challenges in data-intensive computing. Overcoming the challenges in data-intensive computing has required the optimization of data-movement across multiple levels of memory hierarchies. These considerations have become even more important as we are preparing for exascale computing.

The volume of information and recorded data in the digital world is vast. By intelligently combining real time information, compiled from different sources, it is possible to create entirely new kinds of information which can help break down barriers between sectors. Cyber security is vital to all Big Data-type applications and the integration of the morsels of information generated through data mining demands high-level software and ICT competence. The development of Big Data-research methods provides better opportunities for scientists in different fields to conduct research in different areas and also find solutions to their questions. In addition to development in Big Data methodology, it is important to pay attention to multidisciplinarity and promote cross-disciplinary cooperation inter alia between mathematicians, information technology scientists and social scientists.

Comprehensive security builds on the most effective elimination of all threats to the lives of individuals. These days ICT and concomitant cyber security solutions play a critical role in safeguarding comprehensive security. Security in its myriad forms, and especially cyber security, is a field which will only grow in terms of competence and business opportunities.

Cyber security competence cuts across the different sectors and spheres of education. Top-level expertise in cyber security is needed to generate and improve situational awareness in cyber security as well as effective contingency plans against cyber threats, create systems that defend critical infrastructures and to develop functional cyber security solutions.

Jyväskylä
October 2014

Pekka Neittaanmäki
Martti Lehto

Contents

Part I
Cyber World Today

Phenomena in the Cyber World

Martti Lehto

Abstract This chapter describes and evaluates the cyber world, including its phenomena, from a strategic perspective. As no universally accepted definitions for the cyber world exist, associated literature and publications address it in many different ways. A five-layer model is constructed for cyber threats, which include cybervandalism, cybercrime, cyber intelligence, cyberterrorism and cyberwarfare. This chapter depicts the standards-based risk model, cyber operations and cyber-weaponry, as well as the critical structures of society as the targets. Moreover, cyber security definitions are provided. Cyber world phenomena are addressed in more detail in other chapters of this book.

1 What Does 'Cyber' Mean?

The word *cyber* is generally believed to originate from the Greek verb κυβερεω (kybereo)—to steer, to guide, to control. At the end of the 1940s **Norbert Wiener** (1894–1964), an American mathematician, began to use the word *cybernetics* to describe computerised control systems. According to Wiener, cybernetics deals with sciences that address the control of machines and living organisms through communication and feedback. Pursuant to the cybernetic paradigm, information sharing and manipulation are used in controlling biological, physical and chemical systems. Cybernetics only applies to machine-like systems in which the functioning of the system and the end result can be mathematically modelled and determined, or at least predicted. The cybernetic system is a closed system, exchanging neither energy nor matter with its environment. (Porter 1969; Ståhle 2004)

The prefix cyber is often seen in conjunction with computers and robots. **William Gibson**, a science-fiction novelist, coined the term *cyberspace* in his novel

M. Lehto (✉)
Department of Mathematical Information Technology,
University of Jyväskylä, Jyväskylä, Finland
e-mail: martti.lehto@jyu.fi

M. Lehto and P. Neittaanmäki (eds.), *Cyber Security: Analytics, Technology and Automation*, Intelligent Systems, Control and Automation: Science and Engineering 78, DOI 10.1007/978-3-319-18302-2_1

Neuromancer (Gibson 1984). Science-fiction literature and movies portray the Gibsonian cyberspace, or matrix, as a global, computerised information network in which the data are coded in a three-dimensional, multi-coloured form. Users enter cyberspace via a computer interface, whereafter they can 'fly' through cyberspace as avatars or explore urban areas by entering the buildings depicted by the data.

Cyber, as a concept, can be perceived through the following conceptual model (Kuusisto 2012):

- Cyber world: the presence of human post-modern existence on earth.
- Cyber space: a dynamic artificial state formed by bits
- Cyber domain: a precisely delineated domain controlled by somebody,
- Cyber ecosystem: systems of a cyber-community and its environment
- Cyber environment: constructed surroundings that provide the setting for human cyber activity and where the people, institutions and physical systems with whom they interact,
- Cyber culture: the entirety of the mental and physical cyberspace-related achievements of a community or of all of humankind.

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 defense, in addition to protecting the critical infrastructures that permeate and drive the 21st century global economy.

The Australian cyber security strategy defines cyberspace on the foundation of Australia's digital economy and the importance and benefits of ICT to the entire national economy. In accordance with the strategy "Australia's national security, economic prosperity and social wellbeing are critically dependent upon the availability, integrity and confidentiality of a range of information and communications technologies (ICT). 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 (Lehto 2013).

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 (Ibid).

Finland's cyber security strategy succinctly states: "Cyber domain means an electronic information (data) processing domain comprising of one or several information technology infrastructures. Representative to the environment is the utilisation of electronics and the electromagnetic spectrum for the purpose of storing, processing and transferring data and information via telecommunications networks" (Ibid).

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" (Ibid).

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 (Ibid).

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 (Ibid).

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.

South African researchers have created a model of cyberspace's most important structures. These are the economy, politics, the armed forces, psychology and information. According to this model economic structures are a significant target for cyber threats. Political structures are responsible for maintaining national security and the viability of an open society. The armed forces are tasked to maintain national security and to protect society against the measures of cyberwar. The psychological dimension plays an important role in the cyber world; psychological operations can influence human thinking and behaviour. The revolutions in North Africa demonstrated the influence of the media on people's opinions. Information plays the most important part in each cyber threat situation. The western information societies are dependent on the existence, credibility and availability of information (Ibid).

Using the Martin C. Libicki's structure for the cyber world has created a five-layer cyber world model: physical, syntactic, semantic, service and cognitive. The physical layer contains the physical elements of the communications network. The syntactic layer is formed of various system control and management programs and features which facilitate interaction between the devices connected to the network. The semantic layer is the heart of the entire network. It contains the information and datasets in the user's computer terminals as well as different user-administered functions. The service layer contains all those public and commercial services which

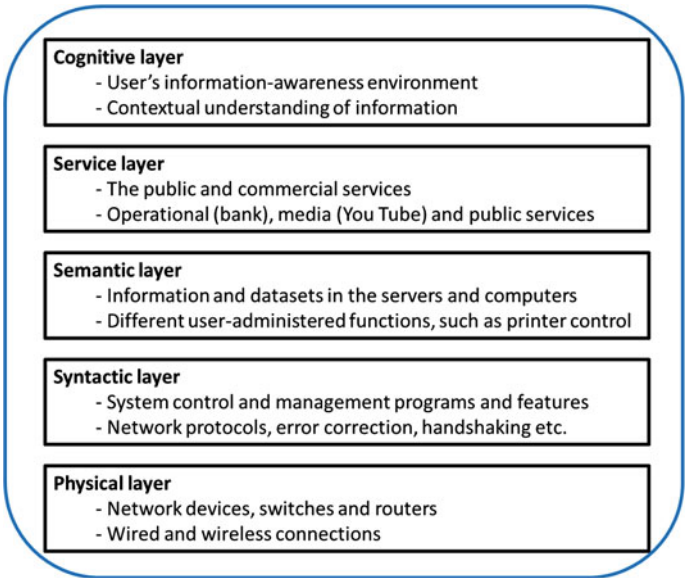

Fig. 1 The five layers of the cyber world

the users use in the network. The cognitive layer portrays the user's information-awareness environment: a world in which information is being interpreted and where one's contextual understanding of information is created. The cognitive layer can be seen from a larger perspective as the mental layer; including the user's cognitive as well as emotional awareness. Concepts related to emotions, such as trust, acceptance, and experience, are central to emotional awareness (Libicki 2007).

Figure 1 shows the five-layer cyber world model

Cyberspace is more than the internet, including not only hardware, software, data and information systems, but also people and social interaction within these networks and the whole infrastructure. The International Telecommunication Union (ITU 2011) uses the term to describe the "systems and services connected either directly to or indirectly to the internet, telecommunications and computer networks." The International Organisation for Standardisation (ISO 2012) defining cyber as "the complex environment resulting from the interaction of people, software and services on the internet by means of technology devices and networks connected to it, which does not exist in any physical form."

The US Joint Publication 3-13 (Information Operations 2012) says that "Cyberspace is a global domain within the information environment consisting of the interdependent network of information technology infrastructures and resident data, including the Internet, telecommunications networks, computer systems, and embedded processors and controllers."

To sum up, the cyber world can be defined as a global and multidimensional ICT network, into which the user (man or machine) can connect via fixed or mobile data

terminals, and virtually move about within it. In other words, the cyber world is an amalgamation of the Internet, other physical networks, digital services and virtual reality: it is a multi-user virtual environment.

2 Drivers of Change in the Cyber World

Time, data, network and intelligence have driven change in the cyber world.

In just a single minute on the web 216,000 photos are shared on Instagram, a total of $83,000 sales take place on Amazon, there 3 days worth of video is uploaded to YouTube. Google performs 2 million searches each minute and 72 h worth of video is uploaded to YouTube within the space of 60 s. 70 new domains are registered and 571 new website are created within a minute online, at the same time there are 1.8 million likes on Facebook, 204 million emails sent and 278,000 tweets posted. There are almost four times more Google searches than a year ago, and 180,000 more tweets are sent. YouTube video uploads have increased from 25 h up to 3 days. Relatives and friends now spend the equivalent of 1.4 million minutes chatting over Skype, compared to 370,000 min year 2012 (Woollaston 2013).

The Slammer computer worm is one example of change as regards the temporal factor. Slammer is a computer worm that caused a denial of service on some Internet hosts and dramatically slowed down general Internet traffic, starting at 05:30 UTC on January 25, 2003. It is estimated that it reached its full level of global internet infection within 10 min of release. At its maximum approximately 120,000 individual computers worldwide were infected and those computers generated an aggregate of over 1 terabit/second of infection traffic. At the point of maximum infection traffic the worm caused a loss of approximately 15 % of hosts on the internet where losses defined by a lack of reachability to the host (Travis et al. 2003).

Google CEO Eric Schmidt have said in the October 2010: "There was 5 Exabyte's of information created between the dawn of civilization through 2003, but that much information is now created every 2 days, and the pace is increasing." We produce 2.5 quintillion (2.5×1018) bytes of information every day. One example of Big Data is the Square Kilometre Array (SKA) planned to be constructed in South Africa and Australia. When the SKA is completed in 2024 it will produce in excess of one exabyte of raw data per day, which is more than the entire daily internet traffic at present. The Large Hadron Collider, at the European Organisation for Nuclear Research (CERN), which has 150 million sensors and is creating 22 petabytes of data in 2012 (Research Trends Issue 2012).

The utilisation of information technology has grown exponentially and it has become an inseparable element in the life of communities and individuals alike. While in 1988, 3.6 % of the world's population were on the Internet, at present there are approximately 2 billion Internet users (30 % of the world's population). More than 210 billion e-mail messages, 0.5 billion blog entries and 2.2 billion Google searches are made on a daily basis. Even though Facebook was launched as late as 2004, its users already number approximately 0.5 billion people. More than 4

billion people across the world now use mobile phones and 675 million smart phones were sold in 2012. Nearly 90 % of innovation in automobiles is related to software and electronics systems. Soon, there will be 1 trillion connected devices in the world, constituting an internet of things.

Imaginative data processing and utilisation, arising from the needs of citizens and the business community, are some of the most important elements in a thriving society. Information and know-how have become key 'commodities' in society, and they can be utilised all the more efficiently through information technology. Different interactive electronic services are ubiquitously available, irrespective of time and place. While the public sector, the economy and the business community as well as citizens benefit from globally networked services, the digital IT society contains inherent vulnerabilities which may generate security risks to citizens, the business community or the vital functions of society.

3 Cyber Threats and Vulnerabilities

3.1 Cyber Threats

Threats to society's vital functions may directly or indirectly target national systems and/or citizens, from within or outside the national borders. 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. The threats to society's vital functions can be divided into three entities which are: physical threats, economic threats and cyber threats.

Physical threats include:

- Natural disasters (e.g. earthquake, tsunami, volcanic eruption, flood).
- Environmental disasters (e.g. nuclear fallout, oil spill, toxic chemical discharges).
- Widespread technical disruptions (especially those in ITC systems).
- Conventional warfare with kinetic weapon systems.
- Terrorist strikes with kinetic weapon systems, and
- Civil unrest (violence, sabotage).

Economic threats include:

- Deep national depression.
- Deep global depression.
- Disruption in national or global financing markets, and
- Sudden global shortage of goods and services.

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.

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. 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 (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" (Commission of the European Communities 2007).

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.

The threats to society's vital functions can also simultaneously occur in each of the three abovementioned dimensions. For example, cyber operations and action aimed at collapsing an adversary's economy can be included in conventional warfare. When it comes to terrorism, different operations in the cyber world and the economic system can be included in strikes that cause physical destruction.

Disruptions can impact and escalate across the dimensions. For instance, a natural disaster can cause widespread disruptions in the power grid, which may adversely affect the operation of payment systems and the food distribution chain. When prolonged, they may result in civil disturbances.

3.2 Cyber Activism

Cybervandalism and hacking saw the light of day in January 1985 when two Pakistani brothers released *Brain*, the first computer virus developed for the pc environment. Hacking was the pursuit of amateurs until 2000, when professionally coded malware began to pop up in the network environment. The first spyware appeared in the mid-2000s, targeting the weapons industry, governments and NGOs, among others. The discovery of the computer worm *Stuxnet* in 2010 heralded a new dawn as regards malware. Stuxnet, co-created by the United States and Israel, was discovered as it was spreading in Europe, India and the Middle East. It was rumoured to contain up to 20 zero-day exploits. Within 25 year hacking, originally an amateur activity, matured into state-run information warfare in global networks and systems (Hyppönen 2010).

Hactivism stands for the different forms of computer and online activism, mostly on the Internet. The term was coined by conjoining the words *hacker* and *activism*. Whereas hacktivism has become a specific field of research in activism, the term itself has yet to become fully established. The reason for this is that, on the one hand, activism can tap into a range of instruments developed by hackers and, on the other hand, hackers can advance their own agenda (Hintikka 2013).

Hactivism often refers to social movements which either independently or assisted by hackers seize and utilise the possibilities offered by networks (McCaughey 2003). Jordan (2008) defines hactivism as an activity which is only possible on the Internet and exploits the manipulation of technology. In other words it relies on technological expertise. Correspondingly, hackers themselves view hacking as activism that opposes the use of technology to limit civil rights, such as Internet censorship.

Vegh (2003) divides online activism into two main categories: Internet-enhanced and Internet-based. According to him, the former concerns activism in which the Internet is mostly used as an extra communications channel or for the purpose of spreading awareness. The latter is only achievable on the Internet, just as Jordan posits.

Mobile technology and the social media offer entirely new vistas for modern cyber swarming. Harbingers of the activists' new modus operandi were in the air as early as 1998 in London and in 1999 in Seattle when groups of activists were mobilised over the Internet and led by mobile phones. Cyber swarming has assorted forms and motives. The so-called 'Botellón' gatherings, where young Spaniards socialise while drinking alcohol, are on the most benign side of the spectrum. The 2011 riots in Britain and the events associated with the Arab Spring were more serious in nature. During the British riots social media was used in organising looting and disturbances. Therefore, the UK is presently considering limitations on the use of social media in areas where riots are taking place.

In Egypt, on 25 January 2011, approximately 15,000 people gathered in the centre of Cairo for an anti-government demonstration. The organising has taken place in the social media. The next day the Egyptian leadership blocked access to

Twitter and Facebook, and Internet services were almost entirely disabled on the night of Friday, 28 January. On Friday, mobile phone services were altogether discontinued in certain areas. By breaking off social communications the government aimed to prevent people from organising, preclude their situational awareness and coordination through cyber swarming. However, the measures were half-hearted and, at the end of the day, the President lost his power. Cyber swarming claimed its first notable victory.

3.3 Cybercrime

Commission of the European Communities defines that cyber-crime is understood as "criminal acts committed using electronic communications networks and information systems or against such networks and systems". The cyber-crime is applied to three categories of criminal activities. The first covers traditional forms of crime such as fraud or forgery, though in a cyber-crime context relates specifically to crimes committed over electronic communication networks and information systems. The second concerns the publication of illegal content over electronic media (i.e. child sexual abuse material or incitement to racial hatred). The third includes crimes unique to electronic networks, i.e. attacks against information systems, denial of service and hacking. These types of attacks can also be directed against the crucial critical infrastructures in Europe and affect existing rapid alert systems in many areas, with potentially disastrous consequences for the whole society (Commission of the European Communities 2007).

Up until a few years ago virus coding was still a young men's hobby through which they sought pleasure and acclaim among their peers. Nowadays online crime is a professional activity aimed at achieving financial gain. While the criminals rarely operate on their own, they do not necessarily form a close-knit organisation. Cooperation that resembles outsourcing is the most common practice, in which the criminals take on specific roles. A skilful programmer may code malware and sell it to a botnet operator. The operator, in turn, will sell his network services to spammers or cyber blackmailers that threaten companies with denial of service attacks. In addition, those who peddle credit card or bank account information normally prefer to sell their information, rather than use the data themselves. These complex chains make it extremely difficult to solve crime, especially when the perpetrators can be spread across the globe. Many a time the traces lead to countries whose authorities lack the will, resources or powers to solve such cases. Since the risk of being caught is negligible, online crime is an extremely lucrative business. The vast number of potential victims more than makes up for the low rate of success, or marginal profit per unit (Kääriäinen 2010).

The number of cyber-attacks has dramatically increased in recent years: it has more than doubled within the past 3 years. At the same time, the financial consequences have risen by nearly 40 %. In 2011 the average annual cost for an American organisation amounted to USD 8.9 million. These days, the annual losses

caused by cybercrime are close to USD 400 billion. According to forecasts, the value of solutions used in thwarting denial of service attacks keeps growing at an annual rate of 18.2 %, expected to reach USD 870 million by 2017.

3.4 Cyber Espionage

Cyber espionage can be defined as action aimed at acquiring secret information (sensitive, proprietary or classified) from private citizens, competitors, groups, governments and adversaries for political, military or financial gain by using illicit methods on the Internet or in networks, programs or computers (Liaropoulos 2010).

Cyber espionage goes on continuously and different countries' cyber organisations continually hack their potential adversaries' systems. This is not a question of vandalism; rather it is professional espionage which seeks to pinpoint an adversary's Achilles' heels, attempting to uncover strategic level-information. Whereas the USA, China and Russia are the major, established actors, Israel, France, Iran and North Korea are also active in this area. India, Pakistan and South Korea are presently developing their cyber capacities. The newcomer in this field is Brazil. China set up its first cyberwar unit as a regular part of its armed forces in 2003. North Korea has four cyberwar units, and it is believed to engage in close cooperation with China (Bergqvist 2010).

3.5 Cyberterrorism

Cyberterrorism uses cyber-attacks against critical IT systems and their control systems. The goal of such attacks is to cause harm and spread fear among people. Cyberterrorism aims to create an impact at the national and international level alike (Beggs 2006).

Cyberterrorism is a new form of terrorists' capabilities made available through new technologies and networks. It makes it possible for terrorists to carry out strikes almost without any physical risk to themselves. It is difficult to define cyberterrorism: there is an ongoing debate whether it is a distinct phenomenon or just a form of information warfare conducted by terrorists (Arquilla and Ronfeldt 2001).

Cyberterrorism differs from other terrorist technologies because it includes offensive IT capabilities, used on their own or together with other means of attack. The focus of cyberterrorism is to create physical harm to IT systems or personnel and equipment by means of information technology. This being the case, for the cyber terrorist the network is a medium/vehicle that facilitates the cyber-attack. Such cyber warfare entails that the various cyber weaponry (assorted malware) can be delivered online to their intended target. This means that the terrorists must carefully consider the scale of physical or technological damage they intend to

achieve within the network, lest it cease to operate as a medium for propagating cyber weapons when needed (Ibid).

IT networks other than computer networks can also become targets. Attacks against them can significantly degrade the operations of a government and the armed forces, so long as they rely on commercial networks, service providers and the Internet. Cyber-attacks which are designed for creating harm and disruptions are probably made to support operations aimed at achieving physical damage. Network hijacking or seizing the network under one's own control may be done in support of another cyber operation, or for achieving some other, independent goal. Cyberterrorism may also target TV broadcasting systems in which case the terrorists can demonstrate their authority, advance their goals in the media, create chaos and show off their power (Ibid).

The US Federal Bureau of Investigation (FBI) defines cyberterrorism as criminal activity that uses computers and network tools in a manner that results in confusion, uncertainty and/or damage to digital services. It is done by creating chaos and uncertainty among the civilian population so as to intimidate or coerce a government or its people in furtherance of political, social or ideological objectives (Tereshchenko 2013).

A group of researchers at Boston University maintain that the goal of cyber terrorists is to foster death anxiety and cause physical harm, and by doing so increase people's willingness to acquiesce to their political or financial demands. A cyber terrorist only regards information, communications and infrastructure as targets; by attacking them he can generate fear and terror or use the attack as a force multiplier together with other means of attack (Jacobs et al. 2010).

The motive separates the cyber criminal from a cyber warrior or cyber terrorist. Where the cyber criminal strives for financial gain, the cyber warrior fights for his military objectives and the cyber terrorist pursues his own agenda. According to a US Congressional Research Service report all such groups use analogous tactics and techniques. The motives and aims make the distinction between these groups. (US-CRS 2008)

3.6 Cyber Warfare

As there is no generally accepted definition for cyber warfare it is quite liberally used in describing events and action in the digital cyber world. The concept of cyber warfare became extremely popular in 2008–2010, partly superseding the previously used concept of information warfare which was launched in the 1990s. For some, cyber warfare is war which is conducted in the virtual domain. For others, it is the counterpart of conventional 'kinetic' warfare. According to the OECD's 2001 report, cyberwar military doctrines resemble those of so-called conventional war: retaliation and deterrence. Researchers agree with the notion that the definition of cyberwar should address the aims and motives of war, rather than the forms of cyber operations. They believe that war is always widespread and

encompasses all forms of warfare. Hence, cyber warfare is but one form of waging war, used alongside kinetic attacks (OECD 2001).

In the 1990s cyber warfare was associated with the concept of information warfare (IW) as its subset. Libicki (1995) defined the sectors of IW as follows:

- Command-and-control warfare, C2 W
- Intelligence-based warfare, IBW
- Electronic warfare, EW
- Psychological operations, PSYOPS
- Hackerwar
- Information economic warfare, IEW
- Cyberwar

The United States defines information warfare as a range of actions taken during a conflict or war by means of information operations (IO) to achieve information superiority over an adversary. The US doctrine includes cyber operations as part of information operations. Air Force Doctrine Document 2-5 (2005) defines information operations as follows:

1. Influence Operations
    a. Psychological operations, PSYOPS
    b. Military deception, MILDEC
    c. Operations security, OPSEC
    d. Counterintelligence (CI) operations
    e. Counterpropaganda operations
    f. Public affairs (PA) operations
2. Network Warfare Operations
    a. Network attack, NetA
    b. Network defense, NetD
    c. Network warfare support, NS
3. Electronic Warfare Operations
    a. Electronic attack
    b. Electronic protection
    c. Electronic warfare support

The concept of Network Centric Warfare (NCW) emerged in American discourse at the end of the 1990s: in NCW the network gained prominence over information. The NCW concept was launched in 1998 in the US Naval Institute's publication "*Network-Centric Warfare: Its Origin and Future*", written by Vice Admiral **Arthur K. Cebrowski** (1942–2005) (Director for Space, Information Warfare, and Command and Control on the U.S. Navy staff) and **John J. Garstka**. They maintained that "For nearly 200 year, the tools and tactics of how we fight have evolved with military technologies. Now, fundamental changes are affecting the very character of war" (Cebrowski and Garstka 1998; Senenko 2007).

They went on to say that Network-centric warfare and all of its associated revolutions in military affairs grow out of and draw their power from the fundamental changes in American society. These changes have been dominated by the co-evolution of economics, information technology, and business processes and organizations, and they are linked by three themes (Cebrowski and Garstka 1998):

- The shift in focus from the platform to the network
- The shift from viewing actors as independent to viewing them as part of a continuously adapting ecosystem
- The importance of making strategic choices to adapt or even survive in such changing ecosystems.

Later the concept was published in the book *Network Centric Warfare* written by, in addition to John Gartska, **David S. Alberts** (Director, Research OASD-NII), and **Frederick P. Stein** (MITRE Corporation). According to their definition network centric warfare is "an information superiority-enabled concept of operations that generates increased combat power by networking sensors, decision makers, and shooters to achieve shared awareness, increased speed of command, higher tempo of operations, greater lethality, increased survivability, and a degree of self-synchronization" (Alberts et al. 2000).

"The term network centric warfare broadly describes the combination of strategies, emerging tactics, techniques, procedures and organisations that a fully or even a partially networked force can employ to create a decisive warfighting advantage" (Garstka 2003).

All of the abovementioned sectors need to be analysed from the offensive and defensive perspective. When it comes to IW and information operations, information is at the core of thinking. Information is seen as the fourth operational factor which glues together the three accepted operational factors: force, space and time. In IW information is understood to be data accumulation, present in any format or system, which can be utilised in communication and interaction. Furthermore, IW encompasses the following concepts: information systems, information environment, information functions and information superiority (STAE 2008).

Cyber warfare, in its present form, can be understood to incorporate both IW and EW, thereby establishing a modus operandi that complies with network centric warfare. Cyber-thinking hopes to bring the structures of cyberspace, i.e. the critical infrastructure, alongside information that is at the core of the information environment. All vital functions of society are more or less networked. Being 'networked' refers to action which is not fixed to any time or place and the management of functions. Network structures, along with information, are gaining in prominence. Yet another significant paradigm shift is the fact that while information warfare is generally perceived to occur during conflicts and war, nowadays cyber threats—in all their different forms—have become a part of everyday life for people and institutions.

Cyber warfare can be divided into strategic and operational-tactical warfare, depending on the role assigned to cyber operations in the different phases of war. State actors launch offensive cyber operations in situations where the states are not

at war with each other. In this case, the cyber-attacks constitute a cyber conflict in a low intensity conflict, as was the case with Estonia in 2007.

In the spring of 2007 Estonia was subjected to a three-week long series of cyber-attacks which targeted, among others, the government, the police, the banking system, the media and the business community. The cyber campaign mainly used denial of service (DOS) attacks targeting among other things web servers, e-mail servers, DNS servers and routers (Ottis 2008).

The Russo-Georgian War, also known as the South Ossetia War, was fought during the first week of August, 2008 between Georgia and the Russian Federation, and the army of the Republic of South Ossetia. In this short-lived war cyberwarfare was used as a part of conventional 'kinetic' operations. As early as 8 August several Georgian and South Ossetian websites experienced DOS attacks. The campaign against Georgian websites began on the night of August the 9th. The attacks targeted the websites of Georgia's government and President, and *Georgia-online*. On 11 August the Georgian authorities decided to fight the 'disinformation' and stopped all Russian TV broadcasts in the country. Caucasus Online, Georgia's leading Internet service provider, prevented access to all pages that had a.*ru* Internet domain suffix. The Russian *RIA Novosti* news agency's website was attacked and went down for a few hours on 10 August. The website of Russia's English-speaking TV channel *RussiaToday* was attacked on 12 August and remained inoperative for approximately 24 h. Hackers gained access to the web pages of Georgia's Central Bank and the Ministry of Defence and tampered with some media footage in them.

Libicki (2011) argues that a cyber-attack used in lieu of kinetic methods creates more ambiguity in terms of effects, sources, and motives. The cyber-attacks change the risk profile of certain actions, and usually in ways that make them more attractive options. He presents four hypothetical uses of cyber-attacks. One, cyber-attacks may be used by a victim of small scale aggression to indicate its displeasure but with less risk of escalation than a physical response would entail. A state rich in cyber warriors may also use the threat of cyber war to deter the potential target against support proxy war fighters. Cyber-attacks can be used by one state to affect the outcome of conflict in another state without having to make any sort of visible commitment, even an implied one. Cyber-attacks do not need to be directed towards adversaries, although the risks of making new enemies if the source of the cyber-attacks is discovered are obvious.

3.7 Cyber World Vulnerabilities

Threat, vulnerability and risk form an intertwined entirety in the cyber world. First, there is a valuable physical object, competence or some other immaterial right which needs protection and safeguarding. A threat is a harmful cyber event which may occur. The numeric value of the threat represents its degree of probability. Vulnerability is the inherent weakness in the system which increases the probability of an occurrence or exacerbates its consequences. Vulnerabilities can be divided

into those that exist in human action, processes or technologies. Risk is the value of the expected damage. Risk equals probability times the loss. It can be assessed from the viewpoint of its economic consequences or loss of face. Risk management consists of the following factors: risk assumption, risk alleviation, risk avoidance, risk limitation, risk planning and risk transference. Countermeasures can be grouped into the three following categories: regulation, organisational solutions (management, security processes, methods and procedures and the security culture) and security technology solutions.

According to the ISO 27005 definition, risks emerge from the "potential that a given threat will exploit vulnerabilities of an asset or group of assets and thereby cause harm to the organization". The risk depends on:

- The **Asset** covering its business importance, existing vulnerabilities or weaknesses and level of protection implemented through controls;
- The **Threat** consisting of a threat agent who—depending on their capabilities—utilizes an attack vector to compromise an asset or set of assets. The effectiveness of an attack depends on the capability of the threat agent and the sophistication of the attack;
- The **Impact** that takes into account the value that the asset represents for the business and the consequences when the confidentiality, integrity, availability or privacy of that asset is compromised though the threat (ENISA 2012b).

Figure 2 shows the interaction between threats, vulnerabilities, risks and countermeasures as per the ISO 15408:2005 standard (ENISA 2012b).

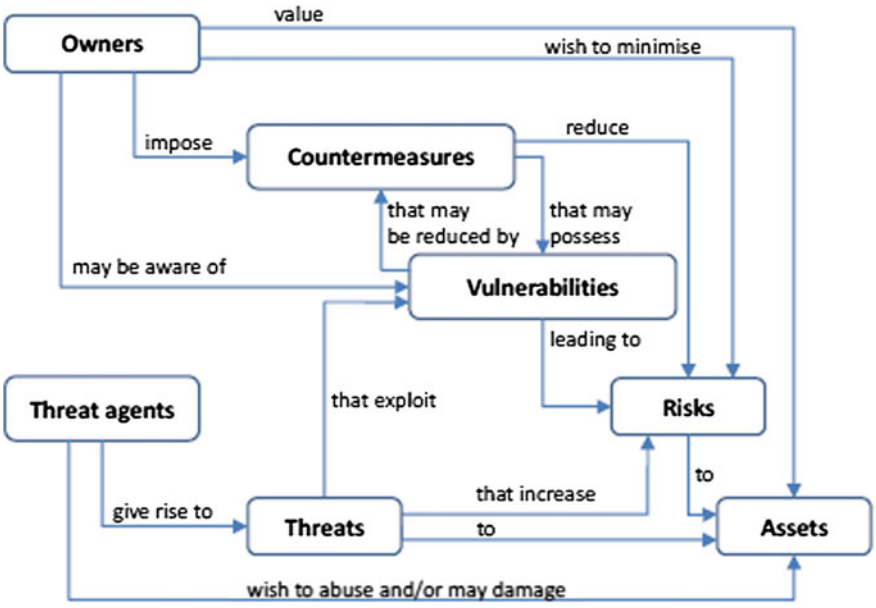

Fig. 2 Interaction model for cyber threats, cyber vulnerabilities, cyber risks and countermeasures

3.8 Cyber Operations

Cyber operations can be defined as a strategy for disrupting an adversary's IT and information-based systems while protecting one's own systems with defensive and offensive means. In cyberwar, cyber operations are not entirely independent operations, isolated from other warfare. Rather they are an integral part of the overall campaign.

Within the United States military domain, Computer Network Operations (CNO) is considered one of five core capabilities under Information Operations (IO) Information Warfare. CNO consists of computer network attack (CNA), computer network defense (CND) and computer network exploitation (CNE). The Joint Pub 3-13 (2012) uses a term Cyber Operations (CO). "CO is the employment of cyberspace capabilities where the primary purpose is to achieve objectives in or through cyberspace. Cyberspace capabilities, when in support of IO, deny or manipulate adversary or potential adversary decision making, through targeting an information medium (such as a wireless access point in the physical dimension), the message itself (an encrypted message in the information dimension), or a cyber-persona (an online identity that facilitates communication, decision making, and the influencing of audiences in the cognitive dimension)." JP 3-12, Joint Cyberspace Operations, is a new JP, signed 5 February 2013 and it is classified.

The National Security Agency (NSA) http://www.nsa.gov/careers/career_fields/ describes that Computer Network Operations mission involves three major functions (NSA 2013):

- Computer Network Attack (CNA): Includes actions taken via computer networks to disrupt, deny, degrade, or destroy the information within computers and computer networks and/or the computers/networks themselves.
- Computer Network Defense (CND): Includes actions taken via computer networks to protect, monitor, analyze, detect, and respond to network attacks, intrusions, disruptions, or other unauthorized actions that would compromise or cripple defense information systems and networks.
- Computer Network Exploitation (CNE): Includes enabling actions and intelligence collection via computer networks that exploit data gathered from target or enemy information systems or networks.

Rowe (2010) has defined techniques for reversible cyber-attacks:

1. Cryptographic attack
2. Obfuscating attack
3. Withholding-information attack
4. Resource-deception attack

Cyber operations can also be divided on the basis of the medium: whether one uses computers and the network as platforms or as targets (Seruga 2011).

3.9 Cyber Weaponry

A cyber weapon refers to a computer program which operates in computers other than those of the user, in the same vein as a computer virus. While a cyber-weapon, by its design, can be an independently mobile and spreading virus, mobility is not a necessary precondition. The most successful cyber weapons are phlegmatic by nature, being either nearly or totally inert in the local area network. In the latter case the cyber weapon must specifically be infiltrated into each target (Kiravuo et al. 2013).

The detected cyber weaponry such as Stuxnet and its kin, *Flame*, *Duqu* and *Gauss*, are all modular malware. The desired functionality of such malware is constructed from several process objects. Of these, the clearly identifiable ones include its 'warhead', the malware payload, and its platform, the delivery module (Ibid).

The platform is controlled by a *command, control and communication module* (C3) which operates independently or in contact with its command and control servers, receiving further commands from them. This module targets and activates the warhead components and it may also download new warheads from its command and control servers. The module also controls the mobility of the cyber weapon. Stuxnet was discovered because of a flaw in the command, control and communication module which made it spread more rapidly than originally intended (Ibid).

In order to break through to its target the weapon carries one or more *exploit modules* which are programmed to exploit system vulnerabilities. There are many kinds of vulnerabilities. For example, one may permit the installation of program code into network software in such a manner that it begins to execute the code. A different vulnerability may arise from a standard password in an automated system. By exploiting these vulnerabilities the cyber weapon seizes partial control over the targeted system and, having gained a toehold, manages to re-distribute itself across the system (Ibid).

With the help of the attack modules the *mobility and installation modules* implement the actual replication and mobility of the malware, installing the weapon into the target computer's operating system. The known cyber weapons can exploit the operating system manufacturers' certificates; this makes it possible for them to quite stealthily be installed as bona fide device drivers or library code. A cyber weapon may also contain an installable rootkit functionality which activates at this stage. It affects the operating system by obscuring the cyber weapon's ongoing processes and files in the computer's file system during system check (Ibid).

If the cyber weapon is designed to spread in the target organisation's LAN, one must first manage to insert it into the firewall-protected network. This can be achieved, for example, via infected USB flash drives or by e-mailing the cyber weapon to its target, as was the case with the Duqu Trojan. In such an instance the weapon burrows into its target by means of a *dropper* package which may outwardly appear to be a word processing document, for example (Ibid).

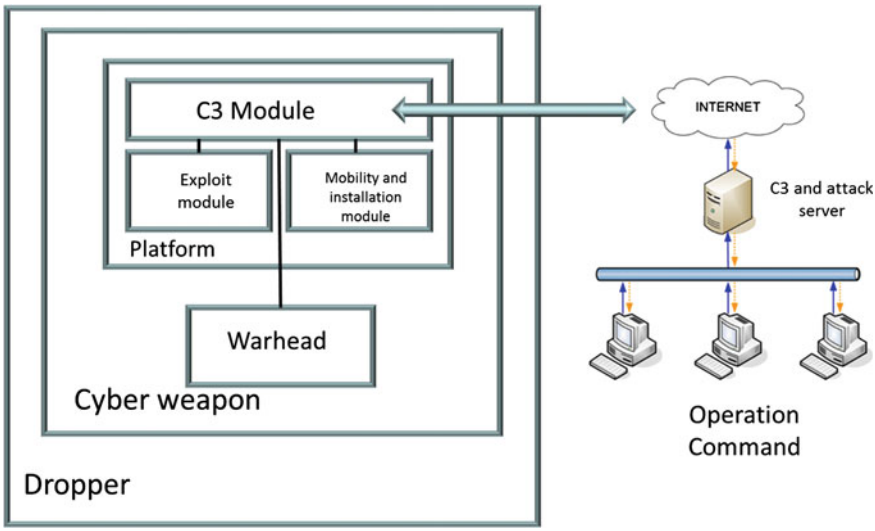

Fig. 3 A standard cyber weapon

Then the delivery vehicle, constructed of the abovementioned modules, will inject one or more actual warheads into the target. One warhead may carry out intelligence, seeking certain kinds of files from the target computer or network servers, hijacking typed passwords from keyboards or eavesdropping on the room through the computer's microphone, etc. Another warhead may cause harm, searching and destroying automated systems, disrupting databases and causing other such damage. (Ibid) Fig. 3 illustrates the design of a standard cyber weapon.

In other words cyber weapons comprise a group of extremely sophisticated computer programs whose functionality determines their targets. Cyberweaponry is used in cyber operations whose aim is to create a desired effect in the target through malware. Typical cyber weapons include (DCSINT 2005):

- Adware
- Backdoor
- Root-kit
- Scareware
- Sniffer
- Spyware
- Trojan Horse
  - Logic Bomb
  - Time Bomb
- Viruses
- Worms
- Zombie

ENISA (2012b) uses a cyber threat landscape model, consisting of threat scenarios. These scenarios include attack methods and techniques, malware and physical threats. The attack methods and techniques include the following:

- Abuse of Information Leakage
- Code Injection Attacks
- Compromising confidential information
- Botnets
- Denial of Service Attacks (DOS)
- Distributed denial of service attack (DDOS)
- Drive-by Exploits
- E-mail Spoofing
- Identity Theft
- IP Address Spoofing
- Keystore Logging
- Password Cracking
- Phishing
- Search Engine Poisoning
- Spamming
- Targeted Attacks

By merging the attack methods and techniques with the Libicki four-layer model one gets the synopsis illustrated in Fig. 4.

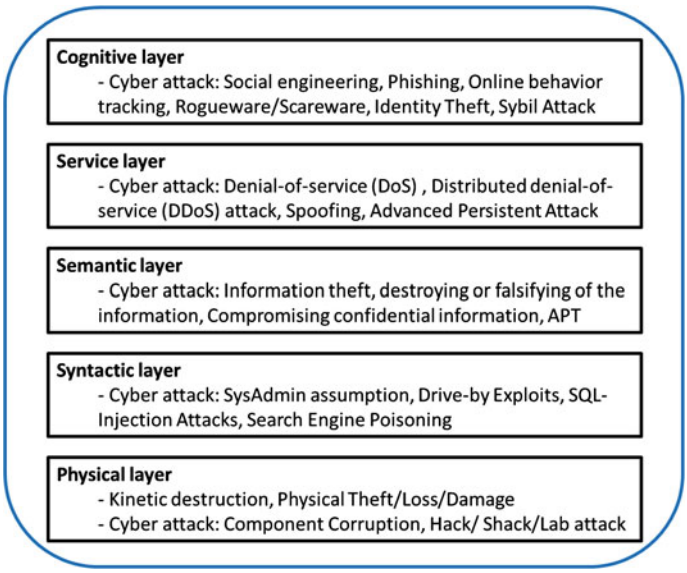

Fig. 4 Attack methods into the different layers of the cyber world

3.10 Society's Critical Structures as Targets

3.10.1 Critical Infrastructure

In the cyber world the most important threat focuses on critical infrastructure (CI). CI encompasses the structures and functions which are vital to society's uninterrupted functioning. It comprises physical facilities and structures as well as electronic functions and services. In order to secure them, one must identify and protect individual critical targets while constantly keeping an eye on the functioning of the infrastructure as a whole. (HVK 2013)

Most countries have a detailed definition regarding their critical infrastructure, including its importance to society, associated threats, its different parts and sectors, and often also the manner by which it is safeguarded. The definitions have normally seen the light of day in conjunction with new, internal security-related legislation.

In most countries, this definition has evolved over the years to include an ever-broader range of infrastructures. National definitions differ slightly in the criteria used to define the criticality of an infrastructure. Most countries and institutions use crosscutting criteria, which cover all infrastructures in all sectors. In Germany critical infrastructure are those "facilities and organizations of major importance to the community whose failure or impairment would cause a sustained shortage of supplies, significant disruptions to public order or other dramatic consequences" (KRITIS 2009).

In United States critical infrastructure means "systems and assets, whether physical or virtual, so vital to the United States that the incapacity or destruction of such systems and assets would have a debilitating impact on security, national economic security, national public health or safety, or any combination of those matters" (Executive Order 2013).

In United Kingdom the Critical National Infrastructure comprises "those assets, services and systems that support the economic, political and social life of the United Kingdom whose importance is such that loss could cause large-scale loss of life, have a serious impact on the national economy, have other grave social consequences for the community or be of immediate concern to the national government" (OECD 2008).

It is possible to identify three dimensions in safeguarding CI: political, economic and technical. The political dimension arises from different countries' shared interests in securing their CI systems and the ensuing increased mutual cooperation. The political dimension entails national legislation and national security needs as well as associated international cooperation around these two topics. International cooperation aims to achieve analogous solutions in countries whose needs are comparable. Uniform security legislation and security policies facilitate technical cooperation, especially when several countries have shared infrastructure. The economic dimension affects all companies and business actors which build, own and administer infrastructure systems and installations, and whose operations are driven by economic interests. The economic dimension also includes a fair

apportionment of security costs between the stakeholders. The technical dimension encompasses technological advances, including their utilisation, and all practical solutions and measures which states and businesses incorporate in securing the functioning of their critical infrastructure during possible disruptions (HVK 2013).

The Finnish point of view on critical infrastructure can be deduced from the threat scenarios which are defined in the strategy for securing the functions vital to society. They are (Security Strategy for Society 2010):

- The energy transmission and distribution network,
- The telecommunications and information systems network,
- The transport logistics system,
- The community technology network,
- The food supply network,
- The financial and payment systems network,
- The health care and welfare system, and
- The safety and security network.

The abovementioned networks are not isolated entities. Rather, they form the national critical infrastructure network within which many interdependencies exist. Should one system be paralysed or collapse, there would be knock-on effects elsewhere in the network. Therefore, the analysis of CI requires modelling so as to determine the interdependencies between the different elements in the network (Pye and Warren 2011).

According to the definition used by Finland's National Emergency Supply Agency critical infrastructure consists of the equipment and devices, services and IT systems which are so vital to the nation that their failure or destruction would degrade national security, the national economy, general health and safety and the efficient functioning of the central government (HVK 2013).

According to EU commission green book "critical infrastructure include those physical resources, services, and information technology facilities, networks and infrastructure assets which, 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." Critical infrastructure can be damaged, destroyed or disrupted by deliberate acts of terrorism, natural disasters, negligence, accidents or computer hacking, criminal activity and malicious behaviour. The goal of European Programme for Critical Infrastructure Protection (EPCIP) would be to ensure that there are adequate and equal levels of protective security on critical infrastructure, minimal single points of failure and rapid, tested recovery arrangements throughout the Union (EU Green Paper 2005).

3.11 Critical Information Infrastructure

From the viewpoint of cyber security an important subset has emerged inside CI: critical information infrastructure (CII).

According to EU commission "the ICT sector is vital for all segments of society. Businesses rely on the ICT sector both in terms of direct sales and for the efficiency of internal processes. ICTs are a critical component of innovation and are responsible for nearly 40 % of productivity growth. ICTs are also pervasive for the work of governments and public administrations: the uptake of eGovernment services at all levels, as well as new applications such as innovative solutions related to health, energy and political participation, make the public sector heavily dependent on ICTs. Last, not least, citizens increasingly rely on and use ICTs in their daily activities: strengthening CII security would increase citizens' trust in ICTs, not least thanks to a better protection of personal data and privacy" (EU Commission 2009).

Critical information infrastructure is a broad concept that designates both the information itself and the channels through which information is created and conveyed. In other words CII is ICT systems that are critical infrastructures for themselves or that are essential for the operation of critical. Critical Information Infrastructure Protection (CIIP) is usually understood as including both the protection of data (including issues of privacy) and the protection of information infrastructure (NATO 2007; EU Green Paper 2005).

3.12 Scada

When it comes to cyber security, various supervisory control and data acquisition (SCADA) systems of the industry have become noteworthy sectors in CI and CII. SCADA incorporates physical and software components which include sensors and measuring devices, telecommunications networks, drivers, communications equipment, the Human-Machine Interface (HMI) and applications. SCADA systems can be used in monitoring large or small systems. They are in use at nuclear power plants and oil and gas refineries, in telecommunications networks, in logistics systems, in premises control heating, cooling and electric systems as well as water distribution and wastewater management systems (SCADA 2004).

The cyber threat in SCADA networks hit the news with the discovery of the Stuxnet worm in 2010. Stuxnet is a computer worm discovered in June 2010 that is created by the United States and Israel to attack Iran's nuclear facilities. Stuxnet initially spreads via Microsoft Windows, and targets Siemens industrial control systems. It is the first discovered malware that spies on and subverts industrial systems, and the first to include a programmable logic controller (PLC) rootkit. The worm initially spreads indiscriminately, but includes a highly specialized malware payload that is designed to target only Siemens SCADA systems that are configured to control and monitor specific industrial processes. Stuxnet infects PLCs by subverting the Step-7 software application that is used to reprogram these devices (Cavelty 2011).

4 Cyber Security

Cyber security measures are associated with managing risks, patching vulnerabilities and improving system resilience. Key research subjects include techniques associated with detecting different network behavior anomalies and malware, and IT questions related to IT security. Since these research subjects mainly concentrate on the physical, syntactic and semantic layers, present research infrastructures are focused on studying phenomena in the aforementioned layers.

In short, cyber security can be defined as a range of actions taken in defence against cyber-attacks and their consequences and includes implementing the required countermeasures. Cyber security is built on the threat analysis of an organisation or institution. The structure and elements of an organisation's cyber security strategy and its implementation programme are based on the estimated threats and risk analyses. In many cases it becomes necessary to prepare several targeted cyber security strategies and guidelines for an organisation.

According to ITU the Cyber security is not an end unto itself; cyber security as a means to an end. The goal should be to build confidence and trust that critical information infrastructure would work reliably and continue to support national interests even when under attack. Therefore the focus of national cyber security strategies should be on the threats most likely to disrupt vital functions of society (ITU 2011).

In EU point of view "the borderless and multi-layered Internet has become one of the most powerful instruments for global progress without governmental oversight or regulation. While the private sector should continue to play a leading role in the construction and day-to-day management of the Internet, the need for requirements for transparency, accountability and security is becoming more and more prominent." The EU want safeguard an online environment providing the highest possible freedom and security for the benefit of everyone. The EU is presented five strategic priorities, which are (EU 2013):

- Achieving cyber resilience
- Drastically reducing cybercrime
- Developing cyberdefence policy and capabilities related to the Common Security and Defence Policy (CSDP)
- Develop the industrial and technological resources for cyber security
- Establish a coherent international cyberspace policy for the European Union and promote core EU values

Finland's Cyber security Strategy (2013) defines cyber security as follows: "Cyber security means the desired end state in which the cyber domain is reliable and in which its functioning is ensured." The strategy adds three notes to the definition:

> Note 1: In the desired end state the cyber domain will not jeopardise, harm or disturb the operation of functions dependent on electronic information (data) processing.

Note 2: Reliance on the cyber domain depends on its actors implementing appropriate and sufficient information security procedures ('communal data security'). These procedures can prevent the materialisation of cyber threats and, should they still materialise, prevent, mitigate or help tolerate their consequences.
Note 3: Cyber security encompasses the measures on the functions vital to society and the critical infrastructure which aim to achieve the capability of predictive management and, if necessary, tolerance of cyber threats and their effects, which can cause significant harm or danger to Finland or its population.

ITU (2011) has created the model of the national cyber security strategy, which has four parts: strategic context, ends, ways and means.

Strategic context consist factors influencing national cyber security activities. National interests flow from national values and guide political decisions. The realization of national interests achieves goals such as economic prosperity, security and stability of the State, protection of individual freedoms and a good international order. Nations use all tools of national power to realise their national interests. Cyber security strategies should focus on tackling threats most likely to prevent government agencies and businesses from carrying out critical missions (ITU 2011).

The Ends are the objectives that a national cyber security strategy seeks to accomplish. Just as national interests flow from national values, ends describe what a nation has to do to support national interests in cyberspace. Thus, cyber security strategies help focus efforts towards ensuring that cyberspace keeps a country secure and prosperous (Ibid).

The Ways identify the strategic activities to help countries govern the security areas. Governance defines how nations may use the resources to attain the outcomes that the ends envisage. In the multi-stakeholder domain of cyber security, the ways define how nations may allocate resources, coordinate and control the activities of all relevant stakeholders (Ibid).

The Means flow from the Ways. The means describe the resources available to achieve the stated ends. 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 a concrete action plans for each activity. The measures are dependent on the local conditions. Several national cyber security strategies have six priority areas (ITU 2011; Lehto 2013):

- 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

As the organisation's security solution is being constructed, one must strike a balance between cyber security, system functionality and ease of use. Inappropriate solutions selected for functionality and ease of use translate into massive vulnerability for the organisation.

References

Alberts DS, Garstka JJ, Stein FP (2000) Network centric warfare: developing and leveraging information superiority, 2nd revised ed. CCRP, Washington

Arquilla J, Ronfeldt D (eds) (2001) Networks and netwars: the future of terror, crime, and militancy. RAND, Santa Monica

Ashenden D (2011) Cyber security: time for engagement and debate. In: Ottis R (ed) Proceedings of the 10th european conference on information warfare and security (ECIW 2011, Tallinn). Academic Publishing, Reading, UK, pp 11–16

Beggs C (2006) Proposed risk minimization measures for cyber-terrorism and SCADA networks in Australia. In: Remenyi D (ed) Proceedings of the 5th european conference on information warfare and security (ECIW 2006, Helsinki). Academic Publishing, Reading, UK, pp 9–18

Bergqvist J (2010) Sinä olet sodassa: Sofistikoitujen algoritmien kamppailua kyberavaruudessa. Suomen sotilas 91(5):6–10

Career Fields, National Security Agency, http://www.nsa.gov/careers/career_fields/

Cavelty MD (2010) The reality and future of cyberwar, Parliamentary Brief. Accessed 30 Mar 2010

Cavelty MD (2011) Unraveling the stuxnet effect: of much persistence and little change in the cyber threats debate. Mil Strateg Aff 3(3):11–19

Cebrowski AK, Garstka JJ (1998) Network-centric warfare: its origin and future. Proc US Naval Inst 124(1):28–35

EU Commission (2009) Critical information infrastructure protection, vol 149. Brussels, COM (2009)

Executive Order (2013) Executive orders on national security, Order 13636 of February 12, 2013

Finland's Cyber Security Strategy (2013) Government resolution 24.1.2013, secretariat of the security and defence committee. Helsinki, Finland

Garstka JJ (2003) Network-centric warfare offers warfighting advantage, Signal, May 2003, pp 58–60

Gibson W (1984) Neuromancer. Berkley Publishing Group, New York

Green paper on a European programme for critical infrastructure protection, COM(2005) 576 final, Commission of the European Communities, Brussels. Accessed 17 Nov 2005

Grobler M, van Vuuren JJ, Zaaiman J (2011) Evaluating cyber security awareness in South Africa. In: Ottis R (ed) Proceedings of the 10th european conference on information warfare and security (ECIW 2011, Tallinn). Academic Publishing, Reading, UK, pp 113–121

Hintikka Kari A (2013) Haktivismi, Kansalaisyhteiskunnan tutkimusportaali. http://kans.jyu.fi/sanasto/sanat-kansio/haktivismi

HVK (2013) Huoltovarmuuskeskus, National Emergency Supply Agency, http://www.huoltovarmuus.fi/tietoahuoltovarmuudesta/kriittinen-infrastruktuuri-kasite/

Information Operations (2005) Air force doctrine document 2–5. http://www.iwar.org.uk/iwar/resources/usaf/afdd2-5-2005.pdf. Accessed 11 Jan 2005

ITU (2011) ITU national cybersecurity strategy guide, Geneva

Jacobs S, Chitkushev L, Zlateva T (2010) Identities, anonymity and information warfare. In: Demergis J (ed) Proceedings of the 9th european conference on information warfare and security (Thessaloniki, 2010). Academic Publishing, Reading, pp 120–127

Jordan T (2008) Hacking: digital media and technological determinism. Polity Press, Cambridge

Kääriäinen J (2010) Verkkorikollisuuden vaarat, Haaste 3/2010
Kiravuo T, Särelä M ja Manner J (2013) Kybersodan taistelukentät, Sotilasaikakauslehti 88(3):46–49
KRITIS (2009) National strategy for critical infrastructure protection (CIP strategy). Federal Ministry of the Interior, Berlin, (17th June 2009)
Kuusisto R (2012) KYBER—miten se voitaisiinkaan määritellä? lecture 10.10.2012 Helsinki
Lehto M (2013) The ways, means and ends in cyber security strategies. In: Kuusisto R, Kurkinen E (eds) Proceedings of the 12th european conference on information warfare and security (Jyväskylä, 2013), Academic Publishing, Reading, pp 182–190
Liaropoulos A (2010) War and ethics in cyberspace: cyber-conflict and just war theory. In: Demergis J (ed) Proceedings of the 9th european conference on information warfare and security, (Thessaloniki, 2010), Academic Publishing, Reading, pp 177–182
Libicki MC (1995) What is information warfare? strategic forum, institute for national strategic studies. National Defense University (Number 28; May 1995)
Libicki MC (2007) Conquest in cyberspace—national security and information warfare. Cambridge University Press
Libicki MC (2011) The strategic uses of ambiguity in cyberspace. Mil Strateg Aff 3(3):3–10
McCaughey M, Ayers MD (2003) Cyberactivism: online activism in theory and practice. Routledge, New York
NATO (2007) 162 CDS 07 E rev. 1—the protection of critical infrastructures, NATO Parliamentary Assembly
OECD (2001) IFP project on future global shocks, report: reducing systemic cybersecurity risk, 14.1.2001
OECD (2008) Protection of critical infrastructure and the role of investment policies relating to national security. OECD, May 2008
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 and security (Plymouth, 2008), Academic Publishing, Reading, pp 163–168
Porter A (1969) Cybernetics simplified. English Universities Press, London
Pye G, Warren M (2011) Analysis and modelling of critical infrastructure systems. In: Ottis R (ed) Proceedings of the 10th european conference on information warfare and security (Tallinn, 2011), Academic Publishing, Reading, pp 194–201
Research Trends (2012) Issue 30: special issue on big data. http://scienceofsciencepolicy.net/newsletter/research-trends-issue-30-special-issue-big-data
Rowe N (2010) Towards reversible cyberattacks. In: Demergis J (ed) Proceedings of the 9th european conference on information warfare and security (Thessaloniki, 2010), Academic Publishing, Reading, pp 261–267
Security Strategy for Society (2010) Government Resolution 16.12.2010, Ministry of Defence, Helsinki
Senenko CM (2007) Network centric warfare and the principles of war, Master thesis. Joint Forces Staff College, Joint Advanced Warfighting School, Norfolk, VA
Seruga J (2011) Information security awareness and enterprise security. Guest lecture at the University of Jyväskylä, 18 Aug 2011
Sotatekninen arvio ja ennuste, STAE (2008) Finnish defence forces: estimate and prediction of military technologies
Ståhle P (2004) Itseuudistumisen dynamiikka—systeemiajattelu kehitysprosessien ymmärtämisen perustana, verkkodokumentti
Supervisory Control and Data Acquisition (SCADA) Systems (2004) Technical information bulletin 04-1. National Communications System, Arlington, October 2004
Tereshchenko N (2013) US foreign policy challenges: cyber terrorism and critical infrastructure, e-International Relations, 12 June 2013
The European Network and Information Security Agency (ENISA) (2012b), ENISA threat landscape: Responding to the evolving threat environment
National Emergency Supply Agency (2013) http://www.huoltovarmuus.fi/tietoa-huoltovarmuudesta/

Towards a general policy on the fight against cyber crime (2007) COM (2007) 267 final, commission of the european communities, Brussels. Accessed 22 May 2007

Travis G, Balas E, Ripley D, Wallace S (2003) Analysis of the "SQL Slammer" worm and its effects on indiana university and related institutions, report. Advanced Network Management Lab, Cybersecurity Initiative, Indiana University, Bloomington, IN

Vegh S (2003) Classifying forms of online activism: the case of cyberprotests against the World Bank. In: McCaughey M, Ayers MD (eds) Cyberactivism: Online Activism in Theory and Practice. Routledge, New York, pp 71–95

Woollaston V (2013) Revealed, what happens in just ONE minute on the internet, Daily Mail Online, 30 July 2013

Cyber World as a Social System

Tuija Kuusisto and Rauno Kuusisto

Abstract The increasing applying of information and communication technology is transforming the society into an unknown ground of the continuously evolving cyber world. This challenges the actors of the society and has increased complexity. The purpose of this chapter is to form a hypothesis about how to identify patterns i.e., emergent phenomena about the cyber world for developing security in the society. The cyber world is considered as a complex adaptive system. A system modelling approach to complex adaptive systems is briefly outlined and a social system model of a society is introduced as a content analysis method to complex adaptive systems. The model is populated with a small set of empirical data to have a preliminary view on the content analysis of the cyber world. The results of the content analysis are patterns, i.e., emergent phenomena of the cyber world. They can be utilized for focusing the more detailed analysis of the cyber world on the most significant issues from the security planning and implementation point of view.

1 Introduction

The global digital business as well as the digitalization of the functions and structures of the society provides a lot of benefits to the government, business, customers and the citizens. However, the increasing applying of information and communication technology is transforming the society into an unknown ground of

T. Kuusisto (✉) · R. Kuusisto
National Defence University, Helsinki, Finland
e-mail: tuija.kuusisto@luukku.com

R. Kuusisto
e-mail: rauno.kuusisto@mil.fi

R. Kuusisto
Department of Mathematical Information Technology,
University of Jyväskylä, Jyväskylä, Finland

M. Lehto and P. Neittaanmäki (eds.), *Cyber Security: Analytics, Technology and Automation*, Intelligent Systems, Control and Automation: Science and Engineering 78, DOI 10.1007/978-3-319-18302-2_2

the continuously evolving cyber world. This transformation covers all the actors, organizations and functions of the society and influences on the way the systems and processes of the society are constructed and executed. This challenges the actors and has increased complexity. Already most of the functions and services of the society as well as the vital functions are automated and interconnected. The actors are dependent on the confidentiality, integrity and availability of information guiding the functions and services. Information in the cyber world is vulnerable to various security threats. Therefore, the security aspects of information as well as the security of the technology structures and systems holding and containing information have become a major issue.

The cyber world as an endlessly expanding domain offers splendid opportunities for new and in novel way manifesting actors, structures and activities. However, having an access to high level cyber potential is required to become a significant cyber world actor. This potential is typically formed only by the resources of nation states, because it is based on such things as existing prosperity, education system, science and research, as well as international business and trade. Some countries will develop more cyber potential than the others but even the major business organizations need a host nation with high enough resources to create the required potential.

However, due to the existing high cyber potential and interconnection, some of the new kinds of actors of the cyber world will emerge and attract people at least for a while. For example, some of the online shopping and social media providers have succeeded well and political and religious groups have been able to recruit followers by communicating their message in the cyber world. These emergent global actors will bring a lot of cultural issues with a variety and even divergence of values to compete in the minds of people. The interpretations of norms and their value will become more inconsistent and the requirements for developing regulations will remain contradictory. Deeper and wider understanding of the characteristics of the society transforming in the cyber world are needed for improving the security of the society.

The purpose of this chapter is to form a hypothesis about how to identify patterns i.e., emergent phenomena and further on simple rules about the cyber world for developing security in the society. These patterns and rules can be used for understanding and maybe organizing the complexity. First, the main concepts of the cyber world are shortly discussed and the complex adaptive system (CAS) theory (Holland 1996) is referred to increase understanding about the nature of the cyber world. A system modelling approach to complex systems is briefly outlined and a social system model of a society is introduced as a content analysis method to complex adaptive systems. The social system model is based on the theories of social systems (Parsons 1951; Habermas 1984, 1989; Luhmann 1999). The model is populated with a small set of empirical data to have a preliminary view on the content analysis of the cyber world. The results of the content analysis are patterns, i.e., emergent phenomena of the cyber world. They can be utilized for focusing the more detailed analysis of the cyber world on the most significant issues from the security planning and implementation point of view.

2 Concepts

2.1 The Main Concepts of the Cyber World

The concepts and models of the cyber world referred and outlined in this chapter are based on communication philosophy (Habermas 1984, 1989), sociology (Parsons 1951; Luhmann 1999), cognition philosophy (Bergson 1911; Damasio 1999; Merleau-Ponty 1969), organizational culture (Schein 1992; Hofstede 1984), information (Polanyi 1966) and complexity (Holland 1996; Kauffman 1995; Ball 2004). Merriam-Webster (2013) gives a common definition for cyber and defines it as 'of, relating to, or involving computers or computer networks (as the Internet)'. It can be noted that even if most of the cyber world is constructed of information, the explicit expressions about the electronic information processing aspect of cyber cannot be found on the common definitions of cyber.

This paper uses the concept of cyber world to emphasize the comprehensive nature of cyber. The cyber and the physical worlds are partly overlapping as depicted in Fig. 1. This follows the common human perception about the edge of these two worlds being obscure. For example, supply chain management systems and controlling systems as well as additive manufacturing known as 3D printing consist of constructs and items of the cyber and the physical worlds that are dependent on each other.

One of the definitions for world is 'the earth with its inhabitants and all things upon it' and for physical 'of or relating to material things' (Merriam-Webster 2013). Based on these definitions, the physical world is defined as the earth with its inhabitants and all things upon it relating to material things. The definitions of world and cyber give the cyber word the following definition: The earth with its inhabitants and all things upon it related to or involving computers and computer networks.

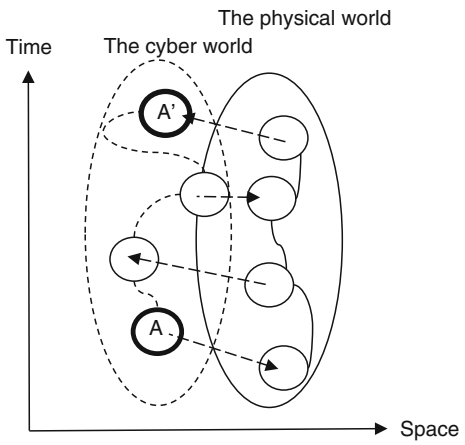

Fig. 1 The relationship between the cyber world and the physical world

As ITU (2011) states the terms cyberspace, cyber environment and critical information infrastructure are used interchangeably. The definition of cyber world is close to the definition of the cyber environment in ITU-T (2008), which says that the cyber environment 'includes users, networks, devices, all software, processes, information in storage or transit, applications, services, and systems that can be connected directly or indirectly to networks'. Hathaway and Klimburg (2012) present similar thinking when they argue that cyber space contains people and social interaction in the networks in addition to hardware, software and information systems of the internet.

The cyber world includes not only the computers and data and information networks, but also the complete and comprehensive system of human existence in those networks. This interpretation of the concept of cyber world allows to deal with the essential issues and phenomena that emerge from this novel domain. These issues include human social behavior supported by information technical solutions. In this chapter, information and communication technology is considered as an enabler rather than dominator of human existence. This approach allows the studying of the phenomena and characteristics of the cyber world without locking the study on the structural restrictions of any technology.

2.2 The Physical and the Cyber World Framework

A framework for classifying roughly activities of the cyber and the physical worlds is outlined in Fig. 2. The framework supports especially the modelling of activities of the overlapping parts of the cyber and the physical worlds. It can be applied for increasing understanding about how the cyber and the physical world activities are interconnected. In addition, it serves, e.g., as a modeling tool when producing

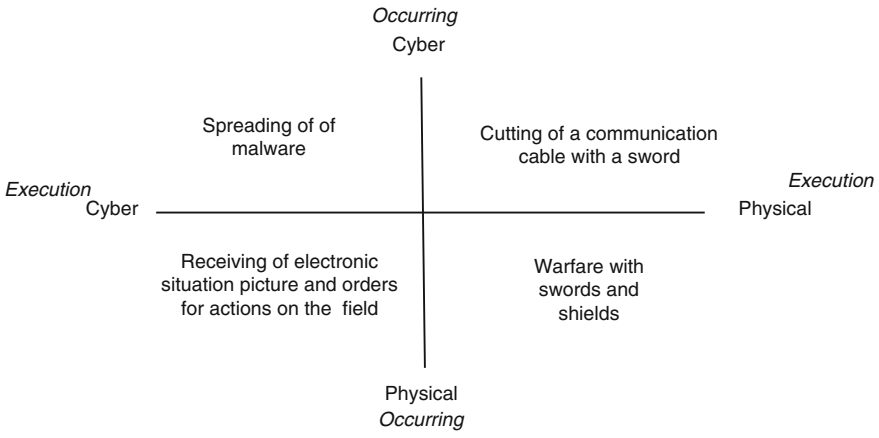

Fig. 2 The physical and the cyber world framework

requirement analysis documents. The framework can be used for assuring that both the cyber and the physical world aspects are concerned.

The framework consists of four fields: physical-physical, physical-cyber, cyber-cyber and cyber-physical fields as presented in Fig. 2. The placing of an activity to one of the fields of the framework depends on the world where the activity is executed and occurs (Kuusisto and Kuusisto 2013). An activity that is executed and occurs in the physical world is placed in the physical-physical field but an activity that is executed in the physical world and occurs in the cyber world is placed in the physical-cyber field. The number of activities executed and occurring only in the physical world is decreasing compared to the activities that are executed or occurring in the cyber world.

Recently, the interest in the kinetic cyber has increased. The kinetic cyber is about the effects occurring in the physical world caused by the activities executed in the cyber world. It is the questions of effects from the technology network to the real world. When this phenomenon is studied further, the following examples can be formed.

Traditional warfare with swords and shields is an activity that is executed and occurs in the physical world. If the warrior disconnects a communication cable with his sword he would perform an activity that is executed in the physical world but occurs in the cyber world. If the warrior has a mobile device and uses it to spread malware on the remaining part of the network he would perform an activity that is executed and occurs in the cyber world. If the warrior receives an electronic situation picture and the orders for the further actions on his mobile device and he would implement the asked actions in the physical battle space there would happen activities executed in the cyber world and occurring in the physical world.

Most crucial in the kinetic cyber are the second and higher order consequential effects caused by damage in cyber related individual entity. The damaging of the power plant stops many other functions almost immediately. Recovery time may be intolerable long causing fundamental or even permanent changes to the existence of the local community or even the society. This, as it, is nothing new. If vital functions and services of the society will fall the society suffers greatly. The new phenomenon is that this damage can be executed distantly and instantly without the physical touch to the target.

3 System Modeling Approaches on the Cyber World

3.1 The Cyber World as a Complex Adaptive System

The cyber world is continuously evolving over time. All of the details of the cyber world cannot be known by any groups of actors or by a certain moment of time. These are some of the characteristics of complexity. As Kauffman (1995) and Ball (2004) describe, complexity means that the comprehensiveness of the interacting

entities and processes is not completely or thoroughly known nor under precise, or even any control. Complex systems will emerge outputs that are not necessarily predictable in the content or in time. Complex is not completely known by any group of actors, but it may be understood at least to some degree by some actors. That is different from complicated that is a known but such a big entity that it cannot be described by any individual actor. It can be argued in a simple way that the meaning of complicatedness is relatively degrading while the meaning of complexity is increasing in the cyber world. Complicated phenomenon is deterministically defined, but complex cannot be defined deterministically. However, if the main features and phenomena of a complex system or a system of complex systems are figured out, enough understanding of the behavior principles of this system can be gained.

The complex adaptive system (CAS) theory (Holland 1996) is a widely accepted approach to increase understanding about complexity. It has an actor's point of view to the chaotic nature of a multi-actor interactive system. The CAS theory aims at to explain the adaptive behavior of an entity in its acting environment by categorizing the basic features of the entity. The CAS theory divides the basic elements of an entity in four properties and three mechanisms. A CAS entity with these elements is depicted in Fig. 3. A short description of the elements are:

(1) An aggregation is a property of an entity. It defines that an entity seeks to categorize similar parts into a class. For example, the parts of vehicles belong to a vehicle class and all the municipalities belong to local government. After the classification the members of those classes are treated as equivalent.
(2) Tagging is a mechanism that gives a descriptive symbol (name) for an aggregate.
(3) Nonlinearity is property that expresses that the outcome of the whole is not the sum of its parts.
(4) A flow is property that tells what transfers between building blocks. A flow can be information, material, radiation or symbol.

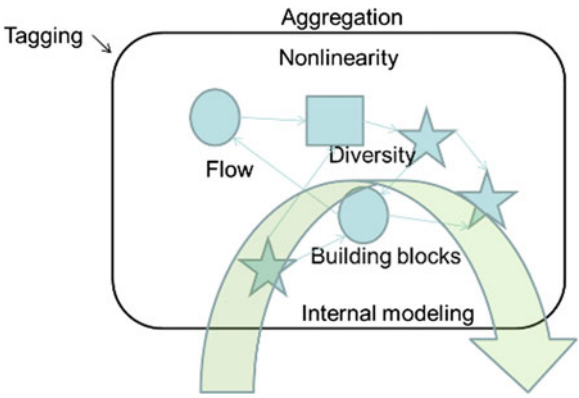

Fig. 3 A CAS entity and its basic elements (Kuusisto 2012)

(5) Diversity is property that tells that a wholeness contains certain (various) amount certain (different) kinds of nodes that have suitable role in that wholeness.
(6) An internal modeling or a schema is a mechanism that causes certain behavior of an entity, when certain stimulus occurs.
(7) Building blocks are the mechanism that enables to construct models in a simple way (Holland 1996).

The cyber world can be considered as a complex adaptive system of complex adaptive socio-technological systems. It is neither random nor accidental. It is a collection of systems' elements with certain kinds of universal features and the continuum of their interrelations. This makes the cyber world act in a non-deterministic way that becomes understandable if we perceive the system at the structural level that suits our viewpoint, see Ball (2004), Kauffmann (1995), Moffat (2003). The cyber world cannot be described thoroughly nor defined exactly by the applying of the CAS theories. However, the cyber world can be approached from a purposefully chosen worldview and viewpoint for increasing the high level understanding of its characteristics.

3.2 The Content Analysis of the Cyber World

A system modeling approach for forming of simple rules from the complex systems is depicted in Fig. 4. The modeling process proceeds from the lower left corner to the upper right corner in the figure. The forming of simple rules is based on both the increasing of understanding and the analyzing of the emergently revealing phenomena of the complex systems as well as on the categorizing and analyzing of information collected by utilizing complicated models. This approach was first presented and applied in Kuusisto (2008) in the context of the collaboration support systems.

Complex systems have a tendency to produce emergent phenomena, see Ball (2004), Kauffmann (1995), Moffat (2003). The first step in the modeling process is

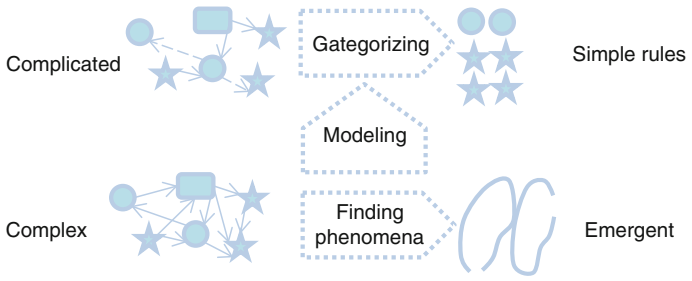

Fig. 4 The content analysis of the cyber world

to figure out these phenomena by approaching the contents of the complex system from a chosen worldview and viewpoint at the structural level that suits the viewpoint. This analysis is performed according to Krippendorff's (2013) content analysis method. The use of the social system model that will be introduced next in this chapter is one method for analyzing the contents of a complex system.

When the emergent phenomena are recognized, they can be utilized for focusing the further modeling of the complex system. This modeling often has to be based on incomplete or fragmentary information and the results of modeling sometimes show themselves as complicated models. However, simple is one component of complex. The complicated models can be simplified by abstracting and categorizing their contents further with the support of the identified phenomena. This produces information for forming of simple rules about the complex system.

The cyber world is a complex adaptive system that inevitably produces some kind of logically understandable phenomena over time. These phenomena can be identified by approaching the cyber world from a chosen worldview and viewpoint. The aim is to find out certain principles of the contents of the cyber world for focusing the more detailed analysis on the most significant issues. This chapter gives an example of finding out the emergent phenomena while the further modeling of the cyber world will remain subject of the further research. The worldview chosen in this chapter is the social system. The primary actors in the cyber world are people so the use of a social system model was chosen as a method to model the contents of the cyber world. The viewpoint of this chapter is the change of the public media focus concerning cyber-related news. Next a social system model is briefly outlined and then the system modeling approach is applied in a media survey.

3.3 A Social System Model as a Worldview to the Cyber World

Social systems have the human perspective on the society. In this chapter it is the chosen worldview to the cyber world. Kuusisto (2004) presents a general social system model based on the thinking of Habermas (1984, 1989) and Parsons (1951). This model is described in Fig. 5, which is derived from a model of organization dynamics presented in Kuusisto and Kuusisto (2009), Habermas (1984, 1989) states that a social system has an initial and a goal state and the communication orientation of the system is internal as well as external as depicted in Fig. 5.

Habermas (1989) refers to Talcott Parsons' (1951) thinking and argues that the information that directs activities of an actor consists of four basic classes: values, norms, goals, and external facts. These are described in the bottom of Fig. 5. The activities using information are on the top of the figure. These are adaptation, goal attainment, integration, and pattern maintenance. The structural phenomena of social systems that consist of culture, community, polity, and institutions are in the

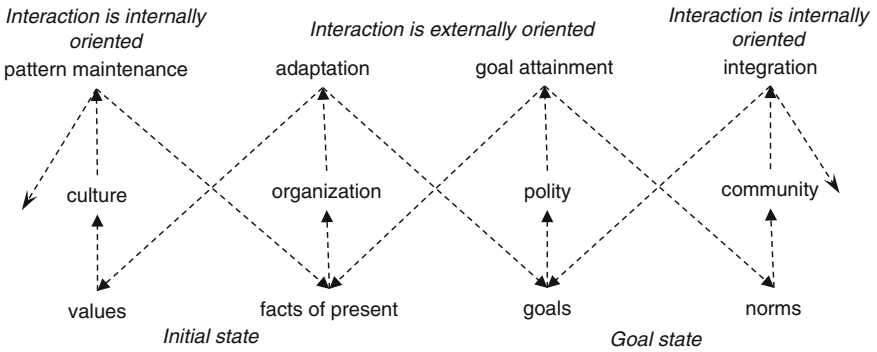

Fig. 5 A system model of a society

middle of the figure. As Habermas (1989) says, cultural systems are more solid than communities, which are again more solid than polity structures and institutions.

Interactivity relationships exist between an item in the Fig. 5 and the items above or below the item. In addition, interaction exists across neighboring action and information items. 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 neighbor parts of the system and external input of each subsystem of the comprehensive system. It can be easily recognized that this kind of a system is complex thus being emergent.

The cyber world can be considered as a system of social systems. People are acting in the structures of the cyber world guided by the structures of social systems and obeying more or less the internal norms. People acting in the cyber world are producing information both inside a social system and between other, neighboring social 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 social system model can help to create understanding about the complex nature of the ever interacting and dynamically evolving system of various subparts and phenomena of the comprehensive cyber world thus helping us to figure out relevant enough acts to make it more convenient to live in this kind of new surroundings.

4 The Content Analysis of the Cyber World

4.1 Media Surveys

The system modeling approach was applied in two media surveys in 2011 and 2012. The worldview of these empirical studies was the social system model and

the viewpoint was the change of the public media focus concerning cyber-related news. The structural level the empirical data was collected and analyzed was the state or central government level. The results of the surveys are reported in Kuusisto and Kuusisto (2013). They are one interpretation of the complex cyber world emergent phenomena.

The media surveys show that the cyber world issues are discussed on public but the focus of the discussions is changing over time. The change in the reported cyber-related news (%) according to the categories of the social system model from the year of 2011 to the year of 2012 is presented in Table 1.

The key findings of the first media survey in 2011 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 in 2012 show 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. These results mean that people want to know what kind of norms will guide the world and how the cyber world will be perceived in futures. A strong need to organize the 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?

The need to organize the cyber world is likely to be an emergent phenomenon produced from the cyber world. From the security planning and implementation point of view the information about this need can be utilized when implementing the more detailed analysis of the cyber world. This analysis can be implemented by using strategy planning methods such as the analysis of strategic level metrics or by using some of the basic strategy tools like SWOT analysis.

Table 1 The change in the reported cyber-related news (%) in Helsingin Sanomat (2011, 2012) from the autumn 2011 to the autumn 2012

	Interaction is internally oriented	Interaction is externally oriented		Interaction is internally oriented
Action	Pattern maintenance	Adaptation	Goal attainment	Integration
	−3	−4	−11	−1
Structure	Culture	Organization	Polity	Community
	0	5	0	0
Information	Values	Facts of present	Goals	Norms
	1	7	−2	7
	Initial state		Goal state	

4.2 Information Assurance

One aspect on the security planning and implementing is the securing of information. The modern society is dependent on the security aspects of the information guiding the functions and services of the society in the cyber world. Especially, the integrity of information has raised into an important position among information security attributes mutual spectrum. One consequence derived from this is that the information security practices are not anymore sufficient to meet the increasing requirements. Information security can be considered as a phenomenon of a complicated world. Information security strategies, policies, its management principles, priorities and especially toolsets can be determined in a precise way. This is expressed in the various kinds of information security standardization systems too. It may be argued that information security is a mature concept and practice, and it can be practiced in orderly manner quite well in a typical organization.

Alongside information security the concept of information assurance shall be taken under serious consideration. Assurance is typically defined as 'the state of being sure or certain about something' (Merriam-Webster 2013). So, information assurance is the state of being sure or certain about information. While information security in its traditional meaning focuses on the information itself in organizations like e.g. enterprises and states and communicates, e.g., societies, information assurance focuses on also security culture, decision-making apparatus, as well as possible activities of human and technological networks and systems. Without a well performing and context relevant information assurance societies, states, coalitions, companies or any other actors cannot fulfill their existence in the spirit of information integrity.

A comprehensive information assurance consisting of strategy, policy and execution, including information security, is needed. This does not mean only the securing of data or making it available, but also the creating of functioning and well enough together understood information assurance culture, decision-making structures, shared critical information as well as structures and functions to put relevant actions in practice. Information is the basic agent to produce relevant action, but it requires relevant structures to become realism. Well planned and proven information assurance is a key factor of that.

5 Conclusions

This chapter presented a system modeling approach for increasing understanding of those phenomena that may be important, when studying the cyber world for developing security in the society. The complex adaptive system theory was applied for understanding the nature of the cyber world and a social system model was introduced as a method to approach the cyber world. The chapter showed how some of the emergent phenomena of the complex cyber world can be identified by using

the social system model. The model was applied for studying the change in the public media focus concerning the cyber-related news.

The system modeling approach seems to support the recognition of the significant issues of the cyber world. However, this approach outlines rough pictures of the cyber world. More exact modeling approaches and methods are needed for producing information that is sufficient for forming of simple rules of the cyber world. These approaches and methods include strategy planning and implementation methods such as metrics analysis. In addition, there is a need for continuing the research work with empirical data collected from several sources and representing longer time periods for verification and validation of the presented system modeling approach.

The results of the media surveys show that the organizing of the cyber world is likely to be in the front of us. It seems that in the current cyber-era the structures of the changing world are unclear. Decision-making apparatus is changing and goals reveal themselves in a different way than before. The global field is becoming more complex requiring new skills for those acting there. Cyber advanced countries, enterprises and social networks have been interconnected together in a non-cancellable way. That interconnected world forms the modern sphere where relevant or even vital activities take place. In a positive future scenario, collectively acting communities reach towards their good future connected together within all areas of exchanging information that guarantees relevant activities.

The changes in the structures and decision-making systems of a society need to be addressed by the security planning and implementation efforts. Cultural systems are more solid than the other structures of a society. Values interact with culture and influence on norms of the society. So, the security planning and implementation activities should be focused on the security culture development and regulation preparation for assuring security in the rapidly changing cyber world. Discussions in the society are needed to be raised for increasing mutual understanding about the interpreting of values concerning security culture and norms covering the cyber world. People will adopt the security culture and accept the norms if the culture and norms are based on the mutual understanding of values. In the long run it seems that the security planning and implementing activities will move on from the regulation preparation to the cyber world community building.

References

Ball P (2004) Critical mass: how one thing leads to another. Heinemann, London

Bergson H (1911) Creative evolution. Henry Holt and Company, New York

Damasio A (1999) The feeling of what happens: body and emotion in the making of consciousness. Harcourt Brace, New York

Habermas J (1984) The theory of communicative action. Volume 1: Reason and the rationalization of society. Beacon Press, Boston

Habermas J (1989) The theory of communicative action. Volume 2: Lifeworld and system: a critique of functionalist reason. Beacon Press, Boston

Hathaway M, Klimburg A (2012) Preliminary considerations: on national cyber security. Chapter In: Klimburg A (ed), National Cyber Security Framework Manual. NATO Cooperative Cyber Defence Centre of Excellence, Tallinn

Helsingin Sanomat (2011, 2012) The main daily published newspaper printed in Finland

Hofstede G (1984) Culture's consequences: International differences in work-related values. Sage Publications, Beverly Hills

Holland JH (1996) Hidden order: how adaptation builds complexity. Perseus Books, New York

ITU-T (2008) X.1205: Overview of cybersecurity. ITU-T Recommendations, X Series: Data Networks, Open System Communications and Security. International Telecommunication Union (ITU). http://www.itu.int/rec/T-REC-X.1205-200804-I. Accessed 5 Nov 2013

ITU (2011) ITU national cybersecurity strategy guide. International Telecommunication Union (ITU). http://www.itu.int/ITU-D/cyb/cybersecurity/docs/itu-national-cybersecurity-guide.pdf. Accessed 5 Nov 2013

Kauffman SA (1995) At home in the universe: the search for the laws of self-organization and complexity. Oxford University Press, Oxford

Krippendorff K (2013) Content analysis: an introduction to its methodology, 3rd edn. SAGE Publications, Newbury Park

Kuusisto R (2004) Aspects on availability: a teleological adventure of information in the lifeworld. PhD thesis, Edita Prima Oy, Helsinki

Kuusisto R (2008) Analyzing the command and control maturity levels of collaborating organizations. In: Proceedings of the 13th international command and control research and technology symposium (13th ICCRTS) (Bellevue, Washington, 2008), Paper 028

Kuusisto R, Kuusisto T (2009) Information security culture as a social system: Some notes of information availability and sharing. In: Gupta M, Sharman R (eds) Social and human elements of information security: emerging trends and countermeasures. IGI Global, Hershey, Pennsylvania, pp 77–97

Kuusisto R (2012) Information sharing framework for agile command and control in complex inter-domain collaboration environment. In: Proceedings of the 17th international command and control research and technology symposium (Fairfax, Virginia, 2012), p 19

Kuusisto R, Kuusisto T (2013) Strategic communication for cyber-security leadership. J Inf Warfare 12(3):41–48

Luhmann N (1999) Ökologishe kommunikation, 3rd edn. Westdeutcher Verlag, Opladen/ Wiesbaden

Merleau-Ponty M (1969) The visible and invisible. Northwestern University Press, Evanston, Illinois

Merriam-webster online dictionary (2013). http://www.merriam-webster.com/. Accessed 9 Sept 2013

Moffat J (2003) Complexity theory and network centric warfare, DOD command and control research

Parsons T (1951) The social system. Free Press, London

Polanyi M (1966) The tacit dimension. Routledge, London

Schein EH (1992) Organizational culture and leadership, 2nd edn. Wiley, New York (4th edition, Jossey-Bass, 2010)

Citizens in Cyber World—Despatches from the Virtual "Clinic"

Torsti Sirén and Aki-Mauri Huhtinen

Abstract People aren't "good" or "bad". People are people, and they respond to incentives. They can nearly always be manipulated—for good or ill—if only you find the right levers. Influence is all about learning what the right levers are and how to apply them (Mackay and Tataham 2011, p. 64). Cyber as a concept has usually, implicitly at least, been understood in technological terms, as a synonym for computer- and internet-based networks. This chapter, however, approaches cyber from psychological perspective and argues that, in addition with it's technological dimension, cyber should also be considered as kind of a virtual Agora, mental battle space or global mind space where ideas can be mediated, challenged and psychological influencing rehearsed in many ways and on many platforms, ones of which are Internet and Social Medias (SOME). The chapter leans theoretically on Freudian–Lacanian psychoanalytical identity theories, both of which deal with "eternal" struggle between individual and social human past, present and future. The authors have used participant observation as their method when analyzing various discussion threads they have participated in Facebook. The main argument of the chapter is three-folded as follows: (1) Much of today´s clashes between human societies are waged in medias (including SOME) far from kinetic battlefields; (2) The speed of SOME discussions (Facebook especially) approaches the speed of face-to-face discussions, which easily may lead into intolerant comments by a participant discussant of the wider world view towards discussants of the more narrow world views; (3) Mental battles in Facebook against intolerant Freudian–Lacanian ego-fortresses may only be won by the most credible arguments and even then, not by one emancipatory capable ego alone, but with many of a kind.

T. Sirén (✉) · A.-M. Huhtinen
Department of Leadership and Military Pedagogy,
Finnish National Defence University, Helsinki, Finland
e-mail: torsti.siren@mil.fi

A.-M. Huhtinen
e-mail: aki.huhtinen@mil.fi

M. Lehto and P. Neittaanmäki (eds.), *Cyber Security: Analytics, Technology and Automation*, Intelligent Systems, Control and Automation: Science and Engineering 78, DOI 10.1007/978-3-319-18302-2_3

1 Introduction

During the pre-modern time, people gathered themselves from their caves or homes around the bonfires, the common market-place or *Polis/Agora* for sharing ideas and trading. In today's post-modern age people tend to gather around in SOME, based on Internet, as a virtual *Agora* for sharing and challenging their ideas. The invention of Internet in 1983 turned to be revolutionary platform for gathering people into virtual discussions. In 1990 Internet had existed 7 years, and only three million people had access to it worldwide; 73 % of these were living in the USA and 15 % in the Western Europe. Since then, the Internet has become a technological platform, which has changed the way we communicate. The Internet as a platform, together with Internet-based SOMEs (there is no SOMEs without discursive human beings), does form a technico-psychological c*yberspace or cyber world,* of which psychological dimension is called here as a *virtual Agora and global mind space*. In December 1995, the Internet users were 16 millions. In December 2003, internet users were already 719 millions and by September 2010, 1,971 millions (Miniwatts Marketing Group 2010; Barbosa 2006; D'Amico and Nigro 2003). Today, there are over 2,000 million Internet users and the amount is growing. People share information and meet each others in virtual meeting places of SOME like Facebook, Twitter, LinkedIn and so on. In other words, people are globally connected, which opens new possibilities to affect them and challenge their identities. The most interesting issue, concerning SOME, is that when people become active in one or several SOME applications they use those networks in spreading their individual ideas or ideas of some particular group—It is important then to have as many "friends" with various opinions as possible in Facebook as well, if wishing to influence on local, regional or global mind space (Rainie et al. 2011; See also Sirén 2013a, p. 90).

The Internet and SOME have emancipated people to express themselves even in societies, which do not support freedom of speech. Consequently, repressive governments do set Internet and SOME restrictions and people try to bypass those. This could be prescribed as technological cyberspace arms race, but it is also about universal need for individual freedom, and the Internet and SOMEs offer a combination of means and ways for satisfying that need. The current, virtually globalized world challenges all the repressive governments, unless they realize that all the governments are for their citizens with freedoms, not for their subjets with restrictions. In January 2011, e.g., so-called *Jasmin revolution* in Tunisia was, at least partially, ignited by Wikileaks, which confirmed the view of the Tunisian people about their government's corruptedness (Payne 2011). Tunisia's revolution inspired protests at least in Egypt, Libya, Yemen, Jordan and Bahrain. Even China went to set firewalls into SOME, fearing her people's claims for political freedom (Sirén 2013a, pp. 90–91).

The cyber, as a concept, became largely known in the beginning of 2000s and was "doctrinated", e.g., in the USA in 2011 by the U.S. Department of Defence. Today, cyber has become popular concept, the content of which usually has been given technological meaning only, consisting of hacking, espionage, sabotage,

denial-of-attacks and attacks on electric power grids through internet. In other words, cyber has been understood, more or less, as a synonym for such a traditional military concept as computer network operations (CNO) or electric operations security (OPSEC)—the same content, albeit different "umbrella concept". Cyber then refers to operations/actions conducted in Internet via computers—it is the fifth operational domain (in addition with land, maritime, air and space), according to the militaries of the world (Department of Defense 2011). In March 2013 [technological] cyber attacks were considered as a bigger threat than Al Qaeda in the USA (Dilanian 2013).

Cyber is still usually considered as a synonym for CNO or electric OPSEC by the states and security organizations (see, e.g., Valtioneuvosto [Council of State of Finland] 2013, p. 12). According to the Council of State of Finland, cyber is rarely used as an independent word, but usually as defining, e.g., the word security (cyber security). Even then cyber security does not consist of psychological dimension, but technological merely, at least according to the Council of State of Finland. The same actors have usually ignored cyber's psychological dimension, even if such figures as General Stanley McCrystal has argued that much of today's kinetic battles are waged in medias (i.e. in human minds) or in capitals far from the kinetic action theatre itself (Mackay and Tatham 2011). This is basically what Carl von Clausewitz argued already in 19th Century when stating that (kinetic) conflict is a clash of wills. The old saying "live and learn" must be reversed in cyber conflict, for there we "learn and live"; otherwise we die. (See Mackay and Tatham 2011, p. 155). The new metaphor of our age is the network, the interconnectedness of all things. According to Mackay and Tatham (2011, p. 161), the Net is messy and complex, a web of cause and effect, difficult to understand and impossible to contain. It requires a new set of analytical tools, a new way of seeing and working. Social network sites (SNSs) is a gate to new kind of communication, persuasion and perception management.

In this chapter, the authors do contribute in the above-presented discussion by trying to introduce novel ideas of cyber by approaching it theoretically from the perspective of *Freudian–Lacanian* psychoanalytical identity theories. By its nature, the chapter may be called as "combat report" or observations (despaches) of conducted "influence operations" in Facebook by the authors, albeit written and structured in academic way by leaning on traditional IMRD-scheme (Introduction, Method, Results and Discussion). The chapter asks: *How the influencing on habituated norms, narratives and myths of participating discussants of Facebook could be effectively rehearsed.*

2 Theory and Method

Behaviours and social connectivity of SOME users have been widely researched, but not explicitly from the perspective of psychoanalytical identity theories, which the authors of the chapter lean on. One of the few identity-related SOME research are

Oliver Mannion's article "Reading Facebook Through Lacan" (2011) and Catalina Toma's article "Feeling Better But Doing Worse: Effects of Facebook Self-Presentation on Implicit Self-Esteem and Cognitive Task Performance" (2013).

In his article, Oliver Mannion analyzes the motives of the people to join Facebook and commit themselves into virtual discussions with basically unknown "friends" by leaning implicitly on Abraham Maslow's hierarchy of needs in general and the need for self-actualization in particular (Maslow 2004 [1943]). In other words, Facebook offers one way to fulfil the human need for self-actualization, the idea of which Mannion supplements by using such Jacques Lacan's terms as the "the mirror stage", "the big Other", "the unconscious", "desire" and "jouissance" (Lacan 1977 [1966]). In other words people tend to construct their egos through Others (the mirror stage)—It is not important then to ask "Who am I", but "Who am I to Others", according to Mannion. Mannion continues by arguing that people try to signify themselves by presenting their ideal ego in their Facebook profiles with pictures, the amount of friends, interests and activities and so on. In order to present this ideal ego to Others, the Facebook users are not updating their Facebook walls to any particular individual, but to a virtual space, or a frame (the big Other), which consequently defines their position as subjects among other subjects. The Facebook users become acquainted with the proper ways to discuss with and present their ideal egos to their fellow virtual subjects (the unconscious), which may lead into narcissistic desire to earn positive recognition (*Theory of positive recognition*; See, for example, Wendt 2003, p. 496), or some sort of recognition at least, from the Others. Positive recognition always brings pleasure, but it may also lead into an addiction toward this pleasure (jouissance) (Mannion 2011).

Catalina Toma is much in the same lines with Oliver Mannion by analyzing human virtual desire. Toma leans here on so called *self-affirmation theory*. According to Toma, self-affirmation theory is predicated on the premise that people have a fundamental need for self-worth and self-integrity, or for seeing themselves as good, appropriate, worthy, and valuable. This need is satisfied by pursuing activities that boost and protect the Self. One such activity is self-affirmation, or unconscious attendance to information in the environment that captures essential aspects of the Self in a positive and accurate manner. In Facebook, e.g., nobody tells about thinking a suicide (Toma 2013).

In her above-mentioned research, Catalina Toma discusses over the question: How does exposure to media content that centers around the users themselves and their identity, interests, and social connectivity affect subsequent perceptions and behaviors. The study approaches this task by investigating perceptual and behavioral effects that stem from exposure to one's own profile on Facebook. On the perceptual front Toma asks: How Facebook profile's self-presentation affects user's evaluation of the Self, operationalized as state of self-esteem. On the behavioral front Toma tries to find out the effects of Facebook's self-presentation on cognitive task performance. In an experimental procedure of Toma's research, participants (N = 159) were randomly assigned to examine either their own Facebook profile's self-presentation, which tend to highlight social connectedness and treasured aspects of the Self, or a stranger's profile. Toma argues that Facebook users experienced a significant boost

in state of self-esteem, if an average stranger pays attention on the user's own Facebook profile. Toma also argues, e.g., that motivation to perform in Facebook decreased when participants did not experience any prior ego threat (Toma 2013). In other words, it should be necessary to "shake" the Facebook users' safe mental havens by challenging them with provocative arguments in order to increase their level of participation, if one wishes to affect participant discussants Selves.

While Oliver Mannion's and Catalina Toma's articles and arguments do offer a good starting point when planning "influence operations" in Facebook, one has to go theoretically deeper into psychological identity theories prior to launching his/her operation for being better equipped in verbal and mental struggle during the operations. Even though Mannion and Toma have theoretically leaned on psychological identity theories of Sigmund Freud and Jacques Lacan, explicitly or implicitly, these theories should be elaborated a bit further.

In addition with Mannion's and Toma's articles, psychoanalytical identity theories of Sigmund Freud and Jacques Lacan have previously used as theories in social sciences for analyzing, e.g., strategic culture and foreign policy of the Russian Federation (see, e.g., Heikka 1999, pp. 57–108). Freud's psychoanalytical identity theory is based on the relationship between mother, son and father. In *pre-oedipal phase* the connection between mother and son is unbreakable until the son realizes his difference from the mother (considered here as an analogy of getting stucked into the past narratives and myths). In *oedipal phase* the son becomes conscious of himself as well as difference between himself as well as his mother and father (considered here as an analogy of equal adoption of the past narratives and new experiences). In this phase the son may consider his father as equal or threatening competitor for gaining attention from the mother (i.e. "the new" is recognized and accepted, or "the new" is recognized, but not dared or wanted to recognize). In other words, the son may stay in the pre-oedipal phase (*Oedipus Complex*), or he may struggle for "ownership" of his mother with his father. The son also may turn into a mature individual and "decenter" the original agency (i.e. the pre-oedipal son) into a new agency that accepts the original connection to the mother, but at the same time accepts the father as equal competitor for gaining attention from the mother (Freud 1959 [1922], pp. 37–38).

Jacques Lacan supplemented Freud's theory by arguing that the defence of the ego is genetic (see more about ego-/self-discussion in, e.g., Mead 1992, p. 182, Berger and Luckmann 2005, p. 147, and Paden 2003, p. 29).[1] The formation of the ego has been compared to a fort, which an individual is desperately trying to defend. Human beings and nations analogically, are trying to find their own identity by signifying their ego with ever new qualities, which differs the ego from other egos (e.g. liberal-democratic, progressive, female, male, rational etc.). This may lead into hysterical, even paranoic defence of the "ego-fortress" (Lacan 1977 [1966], p. 5, 17; See also Miller 1999, p. 119). In this way, according to the Freud's

[1]According to William Paden, people do born into societies that provide them lenses, through which they perceive and interpret the world symbolically.

psychoanalytical identity theory, Russia, e.g., may be interpreted as being stucked into pre-oedipalian phase; it has not left the mythical idea of "The Mother-Russia", the purpose of which is to struggle against all the new thoughts introduced by liberal individuals, the West overall, or the "foreign agents" supported by the West, which may endanger the original and "safe" mythical narrative basis the Russian regime at least has habituated into. In addition, according to the Lacan's identity theory, the Russian Federation may be interpreted as having narrated even new nationalistic identity layers, which do support the defence of her "ego fortress" (the destiny of the Russian Federation, e.g., is to unite all the great civilizations of the world), which as a matter of fact further delineates Russian Federation from other egos of the world (especially from the West) (Heikka 1999, pp. 57–108).

Individuals and human societies do possess very strong myths over their individual Self and social Selves, which are symbolically represented, e.g., in Facebook profiles. Generally speaking, it may be argued that rationality would be kind of a basic element of human Self, which determines our ordinary life. In reality, however, we are often irrational beings, since we may commit ourselves in direct action and communicate each others in pre-hermeneutic way as well. In this context the Internet and SOMEs can be considered as not only facilitating our emotional "virtual desire" to "safely" express our individual and social norm-, symbol- and myth-structures, but challenging them as well, since the Internet is not systematic and fully controllable network—it is the "rhizome" as Gilles Deleuze and Félix Guattari (2004, p. 7) have argued as follows:

> The rhizome itself assumes very diverse forms, from ramified surface extension in all directions to concretion into bulbs and tubers. … any point of rhizome can be connected to anything other, and must be. … A rhizome ceaselessly establishes connections between semiotic chains, organizations of power, and circumstances relative to the arts, sciences, and social struggles.

From the perspective of psychoanalysis, our observations are methodologically speaking never innocent, but references to trauma(s) we are unable to symbolise (according to Kuronen and Huhtinen 2013). In other words, we do not juxtapose in a static field, but make sense of the field in different ways and from different viewpoints. Moreover, as psychoanalysis is a clinical art aimed at helping individuals with their problems, structural understanding can be seen as an epistemic starting point for making a beneficial intervention (analysis and therapy) rather than an epistemic end in itself. Thus, psychoanalytic practice is not restricted to mere understanding of humanity in the socio-political field, but aims at furthering the well-being of the individual attending the clinic.

Internet, SOMEs and the concept of cyber can be considered as technological and virtual rhizome, which exist everywhere affecting multifaceted social rhizome (including identity structures) of human individuals and societies (Kuronen and Huhtinen 2013). Technological and social rhizome together facilitates human emancipation from past myths, narratives and symbols by offering a way to draw new national and global grand narratives for managing the transformation from modern to post-modern cyber age. To become conscious of this possibility,

however, free citizens and suppressed subjects should be courageous enough to set their world views and identities open to public criticism. This is the challenge built inside the cyber age identity. For being efficiently able to affect peoples' world views and identities in Facebook, for example, one should set his/her profile and comments open to all participants of the Facebook. This changes our classical idea of privacy and publicity. It also may be challenging to seduce Others to participate discussions in Facebook at the certain pre-planned time, since participation is not pre-determined by an e-mail list or any other explicit relationship mechanism (See, for example, Bradley and McDonald 2011, pp. 14–15).

In this chapter the Internet and SOMEs have been considered as a virtual cyber house (see Fig. 1), which consists of the building itself, including various everyday technologies, such as electricity, heating as well as cabling, based on fibre-optics and traditional cables, etc. The mentioned building is a metaphor for the content of cyber the states and militaries speak in their cyber strategies about—for states and militaries cyber is about this metaphoric building with technological infrastructure, more or less. The cyber house, however, consists of various human aspects of life inside the building itself. Inside the metaphoric building one can also find virtual clinics (SOME clinics), where citizens are "psychoanalyzing" each others. Each clinic (Facebook, Twitter, etc.) may contain various sofas for congenial analyzers discussing with each others and each others only—"inter-sofa" discussions are rare in these clinics. This is metaphorically the psychological dimension of cyber the states

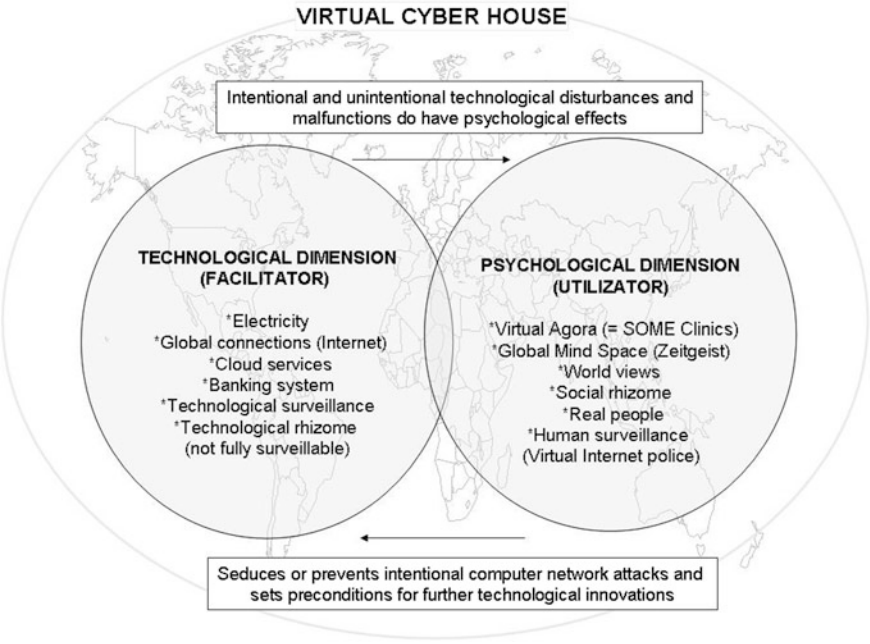

Fig. 1 Cyber space as two dimensional virtual cyber house

are basically ignored. Thus, there are no "state psychiatrists" in the mentioned clinics, psychiatrists who would direct discussions into some particular direction, or psychiatrists who would try to expand the mutual understanding between the discussants as well as widen their world views in order to meet the challenges of the future world. There can be found, however, a doorman at the lobby of the house (Virtual Internet Police), who may ask the discussants in the clinic, from time to time, to keep a quieter voice or blaming them for racist or intolerant comments, but still the clinics are lacking psychiatrists. When states are not interested in widening their citizens' mind space (despite of its benefits for the states in order to cope with ever toughening global competition between the sates) some individuals will take the initiative from the states and try to widen other participant discussants' world views between the sofas of the clinic. Seldom, however, these ad hoc "citizen psychiatrists" do manage to penetrate through the narrow-minded, racists or nationalistic ego-fortresses of "everyman discussants", but collective "citizen psychiatry" could have a chance to make deep penetrations into even most oppressive minds of the clinic by sharp arguments. But who has the right then to say for being wide-minded ad hoc "citizen psychiatrist" in SOME? The authors of this chapter do consider themselves as being such "citizen psychiatrists" with wide and open world views anyhow. However, the "citizen psychiatric sessions" in Facebook, conducted by the authors, have revealed to the authors that even how tolerant the "citizen psychiatrist" may be s/he easily slips into same oppressive or dismissive tone of discussion as her/his "patients", because of occasional frustration and the discursive speed in Facebook. If a person calls her-/himself as a liberal and tolerant, s/he has to understand even the most agonizing opinions and world-views. This does not mean, however that we should accept all the intolerant opinions and world-views, but to say that one should try to understand those first if wishing to fight the intolerance of the world.

As the authors found some theoretical reference points in Oliver Mannion's ideas and Catalina Toma's *self-affirmation theory* vis-à-vis the applied Freudian–Lacanian psychoanalytic identity theories, they also found some methodological reference points in so called *netnography*. Netnography has been used to study global ethics and perceptions of illegal peer-to-peer file-sharing, to investigate consumer activism, and to show how knowledge creation and learning occur through a reflective "virtual re-experiencing" discourse among the members of innovative online communities (Kozinets 2012, pp. 1–2). Consequently, it may be argued then that technology does not determine culture, but facilitates the reconstruction of it merely. We may adopt various roles in SOMEs, such as "interactor", "marker", "lurker" or "networker", which all are facilitated by the Internet, but are necessarily determined by the human subjectivity above all. Virtuality, however, would not exist without technology (i.e. the Internet in this context), but it would not exist without the participating actuality either (i.e. human participation) (See, for example, Kozinets 2012, pp. 33–35, and Zizek 2004, p. 26). Even if natural science and technology started to define and manipulate the human genome, they would not dominate and manipulate human subjectivity. What makes the human being "unique" is neither our genetic formula nor the way our dispositions were developed due to the influence of the environment but the unique self-relationship emerging out of the interaction between the two

(Zizek 2004, p. 118). If we then accept, as a baseline definition, that cultures are learned, consisting of symbolic systems of meaning, of which language is primary, we consequently have to accept that citizen members of online communities do modify and reconstruct human cultures through computer-mediated communication. In other words, the people we virtually meet online are not virtual—They do comprise real communities populated with real people, even though they meet each others virtually (Kozinets 2012, p. 15).

By combining all the above-mentioned theoretical and methodological notions, the authors decided to lean theoretically on themes of Freudian–Lacanian identity theories and methodologically on so called *participant observation*. The participant observation as a method, used in this chapter, needs a situation or platform the researcher is participating. Facebook, as one of many SOMEs, was chosen to be used as a clinic (for authors themselves as well as for all the other discussants) or platform for analyzing the contents of three empirical discussion threads initiated by the authors. The chosen discussion threads do focus on such ongoing and value-sensitive issues in the Finnish society as religion, immigration and gender-neutral marriage law, which are all present in the Finnish "home, religion and fatherland" -based identity construction. All of the mentioned issues are interconnected, but they have been analyzed and thematized here by using the three main themes of the Freudian–Lacanian psychoanalytic identity theories, namely (1) the defence of the ego (ego's habituation into the past narratives and myths), (2) protecting the original ego by inventing new defensive narrative layers on it (impenetrable and paranoic defence of the ego), and (3) decentering the original ego (penetrable and protean ego).

3 Preparing for Mental Battles in the Virtual Clinic

It was in 2009 when one of the authors (Sirén) joined Twitter, LinkedIn and Facebook communities for satisfying his emancipatory need to spread a message of tolerance, pluralism and reason in the Finnish mind space, which he overall considered as habituated into past narratives and borders of otherness between "Us" and "Them". There was not and still is not any clear official program or plan, shared by all the political parties of Finland, to widen the citizens' mind-sets in order to meet the challenges of ever globalizing world. One could even argue that the Finnish state and her politicians are handicapped in front of raising nationalism and racism of the Finnish society by letting the people to anchor themselves mentally into past narratives of *home* (home for the "original" and heteronormative Finns only), *religion* (Christianity) and *fatherland* (idea that is mostly based on remembrance of victories of 1939 Winter War and 1941–1944 Continuation War) against the "true Other" of the Finnish society—Soviet Union and Russian Federation as her successor-state. While it is natural that individuals try to seek mental asylum in front of new and scaring (i.e. ever changing world), it could, or should even, be one of the state's tasks to act as a psychiatric and rehearse constructive psychoanalysis with the people in all or some of the clinics of the SOME. This is to say that usually

people tend to enhance and reify the tenets of their narrow world views in SOME with the people ("friends" in Facebook) of the same world view. This is not to say that there cannot be found emancipatory capable individuals in the Facebook clinic, who could not challenge the most oppressive comments and world views, but to say that there should be more of a kind, since the future-oriented Finnish state has left her citizens on her own in virtual therapy sessions, committed by themselves.

Facebook, from all the mentioned SOMEs, was considered to be the best platform for affecting others, since Facebook is a truly democratic and interactive clinic, facilitating long and in-depth discussions. The first phase prior to launching "influence operation" ahead was to create seducing Facebook profile, which would, by design, capture and visualize the author's (Sirén) social connectivity with friends and family. In this way the profile was constructed as somehow flattering, yet honest, version of the authors's Self. Consequently, the second phase was to ask as many people to start to be "friend" of the author as possible, because it would be waste of time to "enforce" fellow-citizens to participate in "therapy sessions" unless one have enough "friends" to discuss with. Eventually this phase was terminated when the author had 1,000 "friends" in his catalogue of friends. This amount was considered as a sufficient critical mass, after which there would be no need to ask anyone to become a friend of the author in the "clinic", vice versa, the others would ask the author to become their friend. The assessment proved to be correct—Today, the author has about 1,300 "friends" in his catalogue of friends. All of the "friends" of the author are not truly "friends", but "foes" even, since it makes no sense to have discussions with friends only, if one's aim was to widen people's world views as was the aim of the author. The third phase before starting "adventures" in the virtual clinic was to define the method of seducing "friends and foes" into "intra- and inter-sofa therapy discussions" with the author. Multifaceted argument or a question turned out to be the best method for this.

4 Ego's Habituation into the Past Narratives and Myths

The first example of conducted mental battles or "therapy sessions" in Facebook is an example of a "home"-related discussion, the meaning of which was to seduce fellow citizens to discuss about the content of "Finnishness" and to test its borders of otherness. The discussion thread was launched by an argument as follows:

> Suomi on maailman paras maa monilla mittareilla, mutta ollaanko me kuitenkaan niin suvaitsevainen ja moniarvoinen "kansakunta" kuin haluaisimme olla, jotta selviytyisimme "voittajina" tulevaisuuden moniarvoistuvassa maailmassa? (Finland is one of the best countries in the world, according to many indicators, but are we really so tolerant and pluralistic "nation" as we wished to be for surviving as "winners" in the future world of ever increasing pluralism?) (Sirén 2013b)

The argument raised a long discussion with 114 comments by 14 people, in addition with 9 "Likers" who did not take part the discussion. Only one of the discussants took part the discussion clearly from the different sofa of the clinic than

all the other discussants. The argument, however, raised inter-sofa discussion, which is important to notice. The discussion soon turned to be an open inter-sofa quarrel between one "foe" and 13 "friends" of the author as well. At the same time the content of the discussion enlarged and eventually touched about all the three mentioned narratives/myths of the Finnish society. The Finnish narratives of home, religion and fatherland were touched, e.g., upon with four comments by the "foe" of the author as follows:

> Jumalauta jos mieheltä puuttuu isänmaallisuus se on pelkkä nolla. Siinä ei perkele muuta tarvita, se sisältää kodin, uskonnon ja isänmaan – Maansa myyneet ovat pelkkiä luopioita, joille kuolemakin on liian helppo rangaistus! (God damn, if a man lacks patriotism he's a pure zero. For hell's sake, religion and fatherland are the only ingredients of patriotism that are needed – Those who have sold their country are pure turncoats, for whom even the death is too easy of a punishment!)
> …
> Unohtaakko [ne sodat pitäisi] – Miksi sinulla on yleensä sananvapaus? Ilman heitä [suomalaiset sotilaat talvi- ja jatkosodassa] olisitte orjia! (Should we forget then [the past wars] – Why do you have freedom of speech overall? Without them [Finnish soldiers of Winter- and Continuation Wars] you would be slaves!)
> …
> Isänmaallisuus on sisäsyntyistä.. ainoat, joilla sitä on ovat perussuomalaiset, kipitäpä äidin helmoihin. (Patriotism is inherent.. the only ones, who have it are the True Finns [one of the many Parties of Finland], just skitter now under your mother's skirt.)
> […]
> Suomi on kohta takapajula hyysäreiden mukaan asutettu, raunioitunut maa, jossa asuu kaikki muut, mutta ei alkuperäiset suomalaiset. (Soon, as the 'tolerantists' wish, Finland will turn to be backward country, populated by all the others but the original Finns.) (Sirén 2013b)

The "friends" of the author (and the author himself) tried to penetrate collectively, more or less, through the "foe's" oppressive ego-fortress by initially ignoring her comments. It was amazing to notice that the "friends" did not attack verbally against the oppressor, but continued their "civilized" intra-sofa argumentation with each others. The author, however, and one or two his "friends" soon got frustrated on the "foe's" comments and kind of stripped their "hoods of education", forgetting the rules of civilized discussion when starting to shout from the one sofa to the other by using the same tone and language against the foe as the foe used as follows:

> Maryyyy! [keksitty nimi], nyt menee taas lujaa ja yli ymmärryksen. Ainoastaan pluralistinen Suomi voi pelastaa Suomen Suomena tulevaisuuden haasteissa (Maryyy! [the name of the "foe" changed], now you are speeding up again with the arguments that go beyond comprehension. Only the pluralist Finland may save Finland as being Finland in front of the future challenges.) (Sirén 2013b)

As the discussion went on, one of the friends suddenly launched a counter-argument to the author, supporting the foe's statements, as follows:

> Millä perusteella "moniarvoisuus", jonka käytännön toteutus on täyttää lähiöt sosiaalituella elävillä kouluttamattomilla ja kielitaidottomilla maahanmuuttajilla, tekee Suomesta "voittajan". Mikä antaa olettaa, että osaisimme ja voisimme ns. kotouttaa humanitaariset

> maahanmuuttajat paremmin kuin rikas Ruotsi? Nähdäkseni "moniarvoisuus" tarkoittaa käytännössä palavaa Kontulaa, Suvelaa ja Varissuota viimeistään vuoden 2025 kieppeillä. (On what basis "pluralism", of which practical implication is to habitate suburbs with non-educated and illiterate immigrants living on social support, will make Finland as a "winner"? What let us suggest that we would and could integrate humanitarian immigrants better into the Finnish society than the rich Sweden has managed to integrate their immigrants into her society? I believe that "pluralism", in practise, means burning Kontula, Suvela and Varissuo [quarters of Helsinki] by 2025 latest.) (Sirén 2013b)

This therapy session soon tamed down after the foe, encouraged by the above presented argument passed her last statement as follows: *Kiva on Kanadasta [tulevaisuudessa] muuttaa Suomeen, joka on muuttunut Somaliaksi.* (*It would be nice to immigrate* [*in the future*] *from Canada to Finland, when Finland has turned to be Somalia*) (Sirén 2013b).

5 Impenetrable and Paranoid Defence of the Ego

Religion turned to be an issue, which seemingly divides the identity construction of the participant citizens of the virtual clinic most. Discussants, participating in religion related discussions in virtual clinic, may be categorized as believers and non-believers. This is not any novelty, but was decided to test anew and challenge the two different identity fronts into a public discussion with each others in virtual clinic. The discussion, based on the news release according to the religious circles are practicing rituals of expelling demons and devils in one of the Lutheran chapels in Helsinki, started by the author as follows:

> Aika ufoa ja demonimaista menoa uskonnollisella kerhorintamalla. Hyvä ja ajatteluttava juttu MOT:lta. Poika kuuli leikkiessään noita demonijuttuja TV:stä ja tajusi tulla varmistamaan epäilynsä isältään kysyen, että "eihän mikään tauti voi tulla demonilta". "Ei tietenkään", sanoin. Sitten naurettiinkin yhdessä koko aikuisten ihmisten hullutuksille. Siis poika on 6-vuotias. (There can be found rather UFO- and demonic-kind of a going in religious club front. A good and thoughtful news from the MOT [a Finnish broadcasting company]. My son was playing with his toys, heard the news, but was well-informed enough to come to me for asking, "whether any disease may be caused by the demon". "Of course not", I replied. Then, both of us laughed to the follies of adults. I mean the son is six years old.) (Sirén 2013c)

The argument raised a moderate discussion with 19 comments by 6 people, in addition with 16 "Likers" altogether. It was assumed that the argument would raise a fierce inter-sofa discussion, or shouting even, in the clinic. Despite of some resentful comments from one discussant of the "religious sofa" the discursive bait was not effective enough. The first comment after the main argument, however, came from the religious-sofa, the content of which nicely prescribes the impenetrable nature of religious people's ego-fortress, according to which all in the nature and human world come from the God and his holy books, as follows:

> Kyllä sairaudet tulevat saatanan demoneilta, sehän lukee Raamatussakin, joka on Jumalan meille antama elämänopas. Kaikki eivät vain ole saaneet kykyä tajuta näitä asioita. Rajan

> takana kaikki on sitten paljastettuna, mutta epäuskoisia se ei enää auta… Raamattu on Pyhän Hengen vaikutuksesta kirjoitettu Jumalan Sana, vaikka ihminen olisi ollut siinä "kirjurina". (Diseases are truly caused by the demons, which are stated even in the Bible as a God-given guide of life for human beings. Not all of us have had a gift to realize these things, however. In the Afterlife all will be revealed, but it won´t not help infidels anymore…Bible is written under the influence of the Holy Spirit, even if human being had operated in this process as a "scribe".) (Sirén 2013c)

On the basis of this therapy session, one could presume that religious ego-fortresses are impenetrable for alternative world views. This assumption was further strengthening when the author participated (as a visitor) in another religious-related discussion thread initiated by the above-mentioned discussant from the religious sofa of the clinic. However, on the basis of inter-sofa discussions between non-religious and religious world views, it cannot be argued that discussants with religious world view would create any further defensive layers on their ego-fortresses either. What can be deduced from these discussions is that the discussant who initiates a discussion thread in the clinic, easily get frustrated by him-/herself, when noticing that inter-sofa discussion changes into shouting between sofas without any constructive outcome. Self-critically, it has to be admitted that the author perpetrated into this frustration and started to shout by himself again when ending the discussion thread as follows:

> Moniarvoisuutta ja ihan omaa maalaisjärkeä maailmaan. Raamattu on vain yksi satukirja muiden joukossa, eikä edes paras sellainen, vaikka sellaiseen asemaan sitä yhä edelleen pyritään asettamaan. (The world needs pluralism and common sense. Bible is only one storybook among others and not even the best one, even though it is still been tried to set into that position.) (Sirén 2013c)

6 Penetrable and Protean Ego

It would be easy to assume that setting the concepts of Fatherland and the Armed Forces of Finland, as "crown jewelries" of the Finnish identity construction, under any sort of discursive critique at the clinic, would end up in heavy inter-sofa shouting. This was tested by the author by initiating a discussion thread by the words as follows:

> Suomen asevoimat (puolustusvoimat Suomessa) on heteronormatiivinen ja luterilainen yhdenmukaistamislaitos? Asevoimiemme lippujuhlapäivänä puhuu luterilainen pappi luterilaisen jumalan sanomaa, vaikka useat suomalaiset eivät kuulu edes ko. kirkkokuntaan tai kuuluvat johonkin toiseen vastaavaan. Miksi ko. tilaisuudessa edes puhuu pappi? Miksi Ruotsin asevoimissa on homoseksuaaleilla oma etujärjestönsä, mutta Suomen asevoimissa ei? Miksi paljon parjatussa USA: ssakin on asevoimissa palvelevien nykyään mahdollista tunnustaa suhteensa/asennoitumisensa julkisesti, mutta meillä se on oikeastaan tabu? Miksi Suomen asevoimat ei ole mukana Pride-paraateissa, mutta näyttäytyvät kyllä esim. maatalousmessuilla? Kenen asevoimat Suomessa on? (The Armed Forces of Finland [officially the Finnish Defence Forces] is heteronormative and Lutheran standardization facility? Lutheran priest conveys God's message at the Flag Day ceremony of our Armed Forces,

> even though many Finns do not belong to that Church, or do belong to the other of a kind. Why the Lutheran priest even speaks at the mentioned ceremony? Why do homosexuals have their own union in the Swedish Armed Forces, but not in the Finnish one? Why is it so that even the maligned USA has allowed homosexuals openly confessing their relationship/stance in her Armed Forces, but homosexuality in the Finnish Armed forces is almost a taboo? Why the Finnish Armed Forces are not openly involved in Pride parades, but are involved in agricultural fairs? Who's Armed Forces we do have in Finland?) (Sirén 2012; Aljazeera 2012; Niskanen 2012)

Against all the author's expectations, the provocative opening of this thread caused rather temperate and long discussion with 50 comments and 16 Likes. Two discussants even shared the provocation in their own Facebook front pages. This was an intra-sofa discussion thread, which did not seduce any foes to participate the discussion, even though the author precisely so hoped. It was interesting to notice that the discussants were generally ready to accept the provocative questioning of the deepest content of the Finnish identity construction. Simultaneously, participant discussants warned, anyhow, that active officers (the author) should be careful in questioning traditions too actively, since it would "quickly create an image among the reservists [of the Finnish Armed Forces] that 'they [active officers] have gone totally nuts there now'" (Sirén 2012).

The first comments on the author's provocation were short and joking in their nature as follows: "Why is it not possible to wank during the breaks of the military drills [in the Finnish Armed Forces]?" The author applied the same tone when answering the question: "of course it's possible—just wank" (Sirén 2012). Soon the comments started to get more serious tone in their nature as follows:

> Onpa paljon kysymyksiä Torstilla ja varmaankin osa kyseenalaistavia. Mikään asia ei saa olla itsestään selvä. Koti, uskonto ja isänmaa – siinä varmaankin lippupäivän yhteys. Se kuka on maanpuolustushenkinen ja miten hän on sukupuolisesti suuntautunut, on aivan sama minulle… Ikävä kyllä osa suhtautuu uskontoon ja seksuaaliseen suuntautumiseen varauksella. (That's a lot of questions, Torsti, and surely part of them speculative ones. No matter may not be obvious. Home, religion, fatherland – that probably is the connection with the Flag Day. National defence spirit and sexual orientation of anyone are the same to me… Unfortunately, some of us treat religion and sexual orientation with caution.)
> …
> Olen samaa mieltä kanssasi, Torsti, siitä, että minkä takia kirkon asema on niin vahva valtiossa ja puolustusvoimissa. Itse hyödynnän SA-joukon kanssa kirkollista työtä (vaikka en kirkkoon kuulukaan), mutta kysymyksesi on oikeutettu. Sen sijaan homousasiaa pidän moniulotteisempana. (I agree with you, Torsti, about the question related to the strong status of the Church in the state and Armed Forces. I do make use of ecclesiastical work when drilling my wartime unit, but your question is justified. The gay-related issue I, however, consider for being more multi-dimensional.) (Sirén 2012)

The discussion went on much in the same tolerant and truly discursive tone as presented above, but just because of that the author tried to tempt the citizens of the clinic to make more biting comments by another follow-up provocative argument as follows:

> Kotimaa on neutraali käsite. Kenttäpiispan sanoissa ihmettelin seuraavaa osuutta: "Perimmältään vapautemme on Jumalan lahja, hänen, joka Raamatun mukaan on määrännyt kansoille niiden asumisen rajat. Pojassaan hän on vapauttanut meidät synnin ja

> kuoleman vallasta. Tässä hetkessä osoitamme kiitollisuuttamme hänelle, jolta kaikki hyvät lahjat tulevat." Kenen Jumala, miksi Raamattu, miten niin se Jumala on asettanut asumisemme rajat jne? (Homeland is a neutral concept. The Speech of the Army Bishop included following elements, which I wondered: "Fundamentally, our freedom is a gift from God, the God who has determined the borders for the nations to be live within, according to the Bible. In His Son He has released us from the power of sin and death. At this moment we show gratitude to Him, whom all the good gifts come from." Who's God, why Bible and how on earth God has determined borders for our living and so on?) (Sirén 2012; Niskanen 2012)

The argument fulfilled its purpose, since at least one discussant started to doubt the author's rationality. This episode ended the third discussion thread and nicely grasped the whole idea of the need for emancipatory capable citizens for emancipating ourselves from our past narratives, traditions and narrow world views in order to construct our human world more pluralist one as follows:

> Älä nyt Torsti, oletko ihan tosissasi? (Come on, Torsti, are you kidding me?)
> …
> Ihan tosissaan se on "vedetty". Mikä kohta "tökkäsi?" (Totally seriously stated. Which part you do not agree on?)
> …
> Ainoastaan tuo armeija, kirkko, homoseksuaalisuus, ei muu. Olen aika konservatiivinen puolustusvoimia koskevissa näkökannoissani ja en henkilökohtaisesti tykkää juuri tuollaisesta ravistelusta, vaikka pointtisi noin järjen tasolla ymmärränkin. (Only that part, which is related to the Armed Forces. Church and homosexuality [i.e. basically all that the author argued], nothing else: I am rather conservative in my views concerning the Armed Forces and personally I do not like such shaking, even though I understand your point of view at the level of common sense.)
> …
> Mikään ei etene, jos ei ravistele – omaa päätämme erityisesti meidän itse kunkin. (There is no progress, unless one shakes – especially our own heads individually by each of us.) (Sirén 2012)

7 Discussion

Cyber is generally understood as technological dimension of the Internet-based networks. In this chapter the authors, however, have argued that SOMEs, structured on technological basis of the Internet, are to be understood as a social and psychological dimension of cyber—Without SOMEs there would be no need for such a concept as cyber. States have only recently woken up to social and psychological dimensions of cyber, but are still ignoring the possibilities of making full use of those dimensions.

Social media do offer a possibility to deal with social injustices both for citizens of liberal democracies as well as subjects of authoritarian states. In this way citizens and subjects are able to take a progressive stand and construct human world a better place to live in. In mental "battles" conducted in SOMEs, citizens of liberal democracies may criticize grievances of their societies more or less freely. Citizens

of liberal democracies may also promote "wonderfulness" of their societies, if they have considered their societies as such. Authoritarian governments, on the other hand, are handicapped for allowing their subjects to criticize and challenge the "given" symbols, myths and narratives of their societies in SOMEs. Then, the only way for authoritarian regimes is to increase monitoring of SOMEs or cut down their subjects' internet connections timely or totally.

This chapter has treated SOME as Facebook-based virtual clinic, where participant discussants conduct "therapy sessions" with each others. The discussants of the clinic occupy homogeneous or heterogeneous kind of world views. Mentioned world views have been treated here as "therapy sofas" of the clinic. These therapy sofas may conduct intra-sofa discussions with each others or inter-sofa discussions between sofas of the clinic. The main point, however, is that this virtual clinic is missing psychiatrist. Lack of "state-psychiatrist" (i.e. states do not participate virtual therapy sessions actively) has lead into situation where each participant discussant of the clinic act as psychiatrist for each others.

If citizen wishes to spread some message in the mentioned clinic, s/he has to have as many followers, likers, friends and foes in the clinic as possible, because it is in vain to participate in intra-sofa therapy sessions with discussants of homogeneous world views merely. Followers, likers and friends with wide world view may support the argument or psychological engineering initiated by one of them, but it is not possible to enforce them to do it. Thus, the initiating argument of a discussion thread always have to be provocative for raising as wide of an interest among the sofas of the clinic as possible.

Results of the participant observation, conducted by the authors in the Facebook clinic have been reported as despatches from the "front" (i.e. first impressions, based on one's own experiences). These first-impression "combat reports" have been collected from three chosen discussion threads, which one of the authors (Sirén) has initiated by provocative arguments, the contents of which have focused on the assumingly sensitive elements of the Finnish identity construction (i.e. home, religion and fatherland). It was realized that the speed of the Facebook discussions sometimes approaches the speed of face-to-face discussions, which easily may lead into intolerant comments by a discussant of the wider world view towards discussants of the narrower world views. It was also realized that the mental battles in Facebook against intolerant Freudian–Lacanian ego-fortresses may only be won by the most credible arguments and even then, not by one emancipatory capable ego alone, but with many of a kind. The most effective way to break the neighbouring sofa's intolerant ego-fortress in the clinic may, however, be that some emancipatory capable discussant of the mentioned intolerant neighbouring sofa would start intra-sofa emancipatory infiltration into the minds of the discussants of the same sofa. The authors, however, left this to be researched for the future researches.

All in all this research revealed that we all are habituated into our national (or some other) narratives and myths, even if we may think otherwise. This counts to the authors as well, even though we may think otherwise. This is not to say that we cannot emancipate ourselves from these, but to say that our "curtain" of education and life experience is very thin, when our deepest values have been challenged

during the "mental operations" in SOME clinics. One has to understand all the possible world views, if ever starting any kind of mental influence operation in Facebook clinic, in order to have any chance for gaining success in these operations. We do not have to accept all kinds of world views, however, but to infiltrate and widen the borders of narrow and intolerant ego fortresses is truly an art, the content of which largely still remains some sort of mystery for the authors even after this academic rehearsal.

References

Aljazeera (2012) US army allows uniforms in gay pride parade. http://www.aljazeera.com/news/americas/2012/07/20127205506531551.html. Accessed 18 June 2013

Barbosa GT (2006) Internet use 1990. www.worldmapper.org. http://www.worldmapper.org/posters/worldmapper_map335_ver5.pdf. Accessed 10 March 2011

Berger P, Luckmann T (2005) Todellisuuden sosiaalinen rakentuminen [The social construction of reality]. Gaudeamus, Helsinki

Bradley AJ, McDonald MP (2011) The social organization. How to use social media to tap the collective genius of your customers and employees. Harvard Business Review Press, Boston

D'Amico G, Nigro S (2003) Gulf War and Iraqi freedom war: a comparison between the media's role. http://cyber.law.harvard.edu/wsis/DAmico.html. Accessed 12 March 2011

Deleuze G, Guattari F (2004) A thousand plateaus: capitalism and schizophrenia. Continuum, London

Department of Defense (2011) Strategy for operating in cyberspace. http://www.defense.gov/news/d20110714cyber.pdf. Accessed 5 August 2013

Dilanian K (2013) Cyber-attacks a bigger threat than al qaeda, officials say. Los Angeles Times, March 12 2013. http://articles.latimes.com/2013/mar/12/world/la-fg-worldwide-threats-20130313. Accessed 5 August 2013

Freud S (1959 [1922]). Group psychology and the analysis of the ego. W.W. Norton & Company, New York

Heikka H (1999) Beyond neorealism and constructivism: desire, identity, and Russian foreign policy. In: Hopf T (ed) Understandings of Russian foreign policy (introduction). Pennsylvania State University Press, Pennsylvania, pp 57–107

Kozinets RV (2012) Netnography: doing ethnographic research online. Sage, London

Kuronen T, Huhtinen A-M (2013) Leadership as zero-institution. Conference paper, "Re-Working lacan at work" conference, Paris, 14–15 June 2013

Lacan J (1977 [1966]) Écrits: a selection (translational by Alan Sheridan). W.W. Norton & Company, New York

Mackay A, Tatham S (2011) Behavioural conflict: why understanding people and their motivations will prove decisive in future conflict. Military Studies Press, United States

Mannion O (2011) Reading facebook through lacan. NZ Sociol 26(1):143–154

Maslow AH (2004 [1943]) A theory of human motivation. Psychol Rev 50:370–396. http://www.altruists.org/f62. Accessed 2 January 2014

Mead GH (1992) Mind, self, and society: from the standpoint of a social behaviorist. In: Morris CW (ed) Works of George Herbert Mead, vol 1. University of Chicago Press, United States

Miller J-A (ed) (1999) The ethics of psychoanalysis 1959–1960: the seminar of Jacques Lacan, book VII. Routledge, East Sussex

Miniwatts Marketing Group (2010) Internet growth statistics: today's road to e-commerce and global trade internet technology reports. http://www.internetworldstats.com/emarketing.htm. Accessed 8 March 2011

Niskanen H (Army Bishop, Finnish Armed Forces) (2012) Kenttähartaus Lippujuhlan päivän paraatissa 4.6.2012 [Field Devotional at the Flag Day Parade 4th June 2012]. http://www.puolustusvoimat.fi/wcm/Erikoissivustot/Lippujuhla2012/Suomeksi/Paraatipuheet+ja+tervehdykset/Kenttapiispan+puhe/. Accessed 18 June 2013

Paden W (2003) Interpreting the sacred: ways of viewing religion. Review updated edition. Beacon Press, New York

Payne M (2011) Bradley manning and the jasmine revolution. http://www.thepaltrysapien.com/2011/01/bradley-manning-and-the-jasmine-revolution. Accessed 14 March 2011

Rainie L, Purcell K, Smith A (2011) The social side of the internet. Pew research center. http://www.pewinternet.org/Reports/2011/The-Social-Side-of-the-Internet.aspx. Accessed 14 March 2011

Sirén T (2012) Facebook discussion thread, 10 August 2012. https://www.facebook.com/torsti.siren/timeline/2012. Accessed 7 June 2013. Material at the possession of author (Sirén)

Sirén T (2013a) Winning wars before they emerge: from kinetic warfare to strategic communications as a proactive and mind-centric paradigm of the art of war. Universal Publishers, Boca Raton

Sirén T (2013b) Facebook discussion thread, 25 May 2013. https://www.facebook.com/torsti.siren/timeline/2013. Accessed 7 June 2013. Material at the possession of author (Sirén)

Sirén, T (2013c) Facebook discussion thread, 25 February 2013. https://www.facebook.com/torsti.siren/timeline/2013. Accessed 7 June 2013. Material at the possession of author (Sirén)

Toma CL (2013) Feeling better but doing worse: effects of facebook self-presentation on implicit self-esteem and cognitive task performance. Media Psychol 16(2):199–220. http://dx.doi.org/10.1080/15213269.2012.762189. Accessed 4 June 2013

Valtioneuvosto [Council of State of Finland] (2013) Suomen kyberturvallisuusstrategia (Valtioneuvoston periaatepäätos) [Kyberstrategy of Finland (Decision in Principle by the Council of State)]. Turvallisuuskomitean sihteeristö [Secretariat of the Security Committee], Helsinki

Wendt A (2003) Why a world state is inevitable. Eur J Int Relat 9(4):491–542

Zizek S (2004) Organs without bodies: deleuze and consequences. Routledge, New York

Powers and Fundamental Rights in Cyber Security

Riitta Ollila

Abstract Protection of privacy and confidential communications are crucial fundamental rights in cyber security. The protection of privacy and confidential communications are twofold in the meaning that active security steps in communications may require interference with confidential communications. The detection and profiling of potential threats may raise suspects on innocent participants of communications. The NCSC-FI inside the Communications Authority has the initial task and powers to monitor the cyber security. The bill for the Code of Information Society introduces new obligations for information security and preparation for emergency situations. If new powers will be granted to authorities they must narrowly tailored and limited to the necessary measures. The interference with confidential communications in information retrieval requires legal remedies against misuse of powers and constitutional accountability of security authorities.

1 Introduction

The threats in cyber space have developed from hacking activism to crimes and espionage. Cyber terrorism and cyber warfare are escalations of threats. The powers of authorities to observe the threats in cyber space are still based on powers of crime investigation and powers of communications authority in information security supervision. These powers rest on detecting specific crimes and security threats in communications networks. The cyber security is not a legal concept as such and it does not mean extra powers to authorities. The powers of authorities must be based on existing legislation. If the threats escalate to massive attacks that risk the vital functions in society the powers to individual crime investigation are not perhaps sufficient.

R. Ollila (✉)
University of Jyväskylä, Jyväskylä, Finland
e-mail: riitta.h.ollila@jyu.fi

M. Lehto and P. Neittaanmäki (eds.), *Cyber Security: Analytics, Technology and Automation*, Intelligent Systems, Control and Automation: Science and Engineering 78, DOI 10.1007/978-3-319-18302-2_4

The National Cyber Security Centre Finland (NCSC-FI) has been established within Finnish Communications Regulatory Authority (FICORA) since 1st of January 2014. The NCSC-FI continues previous activities of CERT-FI within the FICORA. The NCSC-FI prepares guidance and agreements concerning national security activities and the handling of international classified information. The NCSC-FI will monitor cyber security threats of national interest and produce advanced situation awareness services to its constituents. To facilitate that, the NCSC-FI collects and correlates information from a variety of sources.

The establishing of NCSC-FI has been based on government resolution 24.1.2013 concerning cyber security strategy. The activities of the centre have been manifested on the programme of the centre. According to the Article 119 of the Constitution Act the powers of public authorities must regulated by an act of Parliament if they use public power in the meaning that they decide the rights of individuals. The question follows that does NCSC-FI use public power and does it interfere with constitutional rights of individuals? The activities of NCSC-FI rest on the powers prescribed in the Act on the Protection of Privacy in Electronic Communications. However, the powers rest on the previous CERT-FI activities and the act has not been amended to introduce new powers to NCSC-FI activities.

The powers of police aim on crime investigation and prevention of disorder and crime. In cyberspace and communications networks those powers are exercised by intercepting phones and messages and monitoring subscriber data. The general observation of cyber space is approved for the purpose of threats and prevention of criminal activities. The use of the observation data is possible for the protection of national security and for the prevention of crimes and immediate dangers to public security.

The Finnish security intelligence service and military service have no right to general and unlimited surveillance in cyber security. The Department of Defence has set up a working group for preparing legislation concerning powers in cyber security and interference with fundamental rights of individuals. I examine the existing powers of authorities in cyber security and needs for amendments in legislation. The prevention of infringements and damages are primary purposes for cyber security surveillance and not the breaches of fundamental rights of ordinary people. However, the temptations to use data for other possible purposes are alternatives if surveillance authorities get masses of personal data. I examine the obvious risks and safeguards against misuse of personal data.

2 Constitutional Protection of Personal Data and Confidential Communications

The privacy and protection of personal data must be considered in circumstances when data refers to natural persons. Article 10 § of the Constitution Act specifies the protection of personal data: “More specific provisions on the protection of

personal data shall be prescribed and specified by an Act of Parliament". The bases for the processing of personal data must prescribed by law. The Constitutional Committee is the committee in Finnish Parliament that interprets the constitution. The Constitutional Committee has regarded that the purpose of processing and the content of data, the further processing of data including deliveries of data and the legal safeguards of the data subject must be prescribed by law.

The Constitutional Committee has developed the following principles of interpretation in its practice:

1. The basis of the filing systems must be prescribed by law in details. The period of storing and the removal of personal data cannot be prescribed on the level of administrative decrees inferior to acts of Parliament (PeVL 7/1997 and 11/1997)
2. The data subjects must have legal safeguards. The safeguards might vary from the data subject's right of access to the Data Ombudsman's right of access whether the data relating to him are processed and whether the processing fulfils the conditions prescribed by law (PeVL 7/1997).
3. The processing of personal data of public authorities must be based on justified reasons (like crime investigation, PeVL 7/1997)
4. The prescribed-by-law requirement is followed by the requirement of accuracy. "The pressing reasons" are not adequately accurate if they have not been concretized at the level act (PeVL 27a/1998).

The Constitutional Committee has considered in its practice that wide and non-specified access to personal data justified on the necessity reasons for the activities of authority are problematic. According to the Data Protection Act, purpose specification and duty based on law are general conditions for the processing of personal data in activities of authorities. Is purpose of the authority enough for the processing if the authority does not deal with personal data on the consent of the person concerned? If the authority collects information related to individuals it interferes with private life or confidential communications of individuals protected in the Article 10 of the Constitution. The interference with those fundamental rights should be based on justified reasons and should be prescribed by law.

The purpose specification, data quality principles and general legitimacy conditions of processing are the crucial conditions for the disclosure of personal data. All these conditions must prevail and the disclosure must be compatible with the purpose. The special categories of processing have their own special conditions for processing. I have scheduled all the conditions of disclosure in a chain in which the purpose specification and data quality principles must always be fulfilled, but the legitimacy conditions and special categories of processing have their own alternative paths (Table 1).

The general legitimacy conditions and special categories of processing form each of them special paths and categories of processing (Ollila 2004). If there is no consent or contract to the processing of personal data it must be based on legal obligation of the authority. Purpose specification is the prevailing condition for processing of personal data but it must be regulated by law or by consent. If the

Table 1 The purpose specification, data quality principles and general legitimacy conditions of processing

Purpose specification	Data quality principles	Legitimacy conditions of processing
• Adequate • Relevant • Not excessive • Accurate • Up to date	• Consent • Contract • Legal obligation • Vital interest • Client, employee, member • Merger • Payment service • Public service • Permission of the data board	• Sensitive information • Scientific • Historical • Statistic • Credit information • Who is who records • Genealogy • Direct marketing

authority does not have direct relationship with the customer in administrative matters, the conditions for the processing of personal data must be prescribed by law.

The conditions for interference with confidential communications have been prescribed in Article 10 § 3 of the Constitution Act. In criminal investigation necessary restrictions may be provided by an Act of Parliament in the investigation of offences that endanger the security of an individual, society or home. In legal proceedings and in security checks and during deprivation of personal freedom the restrictions must be prescribed by an Act of Parliament but there are no conditions for the quality of the restrictions and offences.

The protection of confidential communications can be divided in the vertical relationship between the state and the citizen and in the horizontal relations between private participants. In the vertical relationship between the state and the citizen the main issue is the restraint from the interference by the state with private communications and the justifications for interference. The restraint from interference conforms to respect for the confidential communications. The crime investigations and legal proceedings are the main reasons for interferences with the secrecy of communications.

In horizontal relations between the participants and service providers of the communications, the main issue is the protection of confidentiality against violations of others and the security obligations of the service providers to guarantee the confidentiality of the communications. The protection of confidentiality in horizontal relations requires active safeguards and legislation from the state to prevent the violations and to impose on security standards for the confidentiality. The telecom operators are obliged to take care of the security in their networks.

The protection of privacy and confidential communications are twofold in the meaning that active security steps in communications may require interference with communications (Ollila 2005). The detection of threats requires identification of those participants who threaten communications. The profiling of potential threats may raise suspects on innocent participants of communications.

3 The Powers of Communications Authority

The Finnish Communications Authority carries out general supervision and powers based on the Act on Protection of Privacy in Electronic Communications. It collects information and investigates violations and threats to information security in communications networks and services. The NCSC-FI agency carries out the data security tasks inside the communications authority. NCSA-FI specialises in information assurance matters in cases concerning technical information security and the security of telecommunications. The NCSA-FI is one of national security authorities besides the Ministry of Defence, the Defence Command and the Finnish Security Intelligence Service. The Ministry of Foreign Affairs acts as leading National Security Authority in implementing international security obligations.

The NCSC-FI has access to the identification information of subscribers in the context of vulnerabilities of communications and data offences. This right covers both the metadata and the content of telecommunications. The NCSC-FI does not need any consent of court for these surveillance activities. The absence of legal remedies has been justified on the reason that NCSC-FI receives only IP-addresses of data and they do not deliver the data to other authorities. They do not process personal data.

4 The Activities of NCSC-FI

- Reports from telecom operators and other enterprises obliged by law
- Surveillance of public sources in traditional and social media and rumours
- Networks of data exchange in autoreporter and Havaro programs
- Analysing data

According to the Article 20 of Act on Protection of Privacy in Electronic Communications, the measures for information security include powers:

Measures for maintaining information security may include:

(1) Automatic analysis of message content;
(2) Automatic prevention or limitation of message conveyance or reception;
(3) Automatic removal from messages of malicious software that pose a threat to information security;
(4) Any other comparable technical measures.

The automatic analysis of message content means that no natural person reads the content. If the contents of a message have been manually processed, the sender and recipient shall be informed of the processing.

The powers prescribed in the Article 20 are proposed to be granted to all telecom operators and service providers according to the Information Society Code bill. The telecom operators and service providers will have similar powers than the Communications Authority for necessary measures to maintain information

security. The telecom operators and service providers shall have powers to process identification information for information security purposes.

The telecom operators and service providers are obliged to inform the communications authority about security violations. The NCSC-FI deals with those reports and it has entitled to access any identification data, location data or messages for the carrying out of its duties if the data is necessary for clarifying significant violations or threats to information security. The NCSC-FI shall destroy any information and data when this information is no longer necessary for carrying out its duties or for any criminal case concerning the information.

The subscribers and telecom operators are entitled to fight against malicious messages and programs in order to prevent damages. Those malware programs and malicious messages as such do not belong to the protected confidential communications and freedom of expression. Subscribers have the right to process their own data in order to prevent malware programs and viruses.

The analysis of public websites and social media is open to everyone and even for authorities. Do the authorities process data in the meaning of data protection law if they copy and store the public data for future purposes? This retrieved content contains personal data if any of the source web pages do. The authority can observe the changes of the web page by retrieving successive source web pages and analysing changes of the content. According to the program of NCSC-FI, enriching public data with confidential data and combining and processing data from different public sources means processing personal data if the data can be connected to individuals. The operation of search engines has been analysed in the opinion of Advocate General Jääskinen delivered to the European Court of Justice in Google Spain judgement.

Second, the search results displayed by an internet search engine are not based on an instant search of the whole World Wide Web, but they are gathered from content that the internet search engine has previously processed. This means that the internet search engine has retrieved contents from existing websites and copied, analysed and indexed that content on its own devices. This retrieved content contains personal data if any of the source web pages do.

If the Google search engine can retrieve contents from existing websites and index that content on its own devices, it is hard to deny similar activities for domestic authorities. However, any further processing of personal data remains under domestic data processing law and European data processing directive. The European Court of Justice has considered in case C-73/07 that the clipping archives of media including already published material remain under data processing directive. Compared to public websites this could mean that public websites are like clipping archives in the further processing of personal data.

5 The Powers of Police in Cyber Space

The powers of police rest on coercive means legislation for detecting crimes that already have happened. Typically the police inspect computer breaks-in, cyber fraud, communications disturbing and espionage through communications systems. The powers for surveillance to detect threats and prevent crimes rest on the general powers of police legislation. The powers for monitoring and processing data from public sources and web pages rest on general powers of police legislation. The conditions and permissions regulated in the Coercive Means Act shall not be applied to surveillance of public sources. The interception of phones and messages means interference with confidential parts of communications. The surveillance of openly available parts of messages does not mean interception in the meaning of Coercive Means Act. According to Article 23 in Chap. 10 of the Coercive Means Act, the surveillance of computer and its contents means interference with secret parts of it. The permission for surveillance is not needed for physical surveillance and processing of data obtained from public sources. However, the act on processing personal data in police activities shall be applied to all processing of personal data concerning natural person.

6 The Powers in Escalated Threats

The NCSC-FI inside the Communications Authority has the initial task to monitor the cyber security. The NCSC-FI will monitor cyber security threats of national interest and produce advanced situation awareness services to its constituents. However, the powers in situations where the cyber threats escalate to massive attacks that threaten the communications infrastructures and critical physical infrastructures have not been regulated by law. The bill for the Code of Information Society includes obligations for information security and preparation for emergency situations. The bill is under consideration in Parliament in 2014.

The bill for the Code of Information Society requires every telecom operator and service provider to take measures for information security of their services. The telecom operators and service operators have similar powers than the Communications Authority to automatic analysis, prevention and removal of messages that may include malicious software. They must report the threats to information security they have identified to Communications Authority. They must report the vulnerabilities of their services to their subscribers and other clients if the vulnerabilities cause serious threats to their services.

The telecom operators and the possessors of radio frequencies are obliged to make plans for preparation to emergency situations. The Communications Authority has powers to impose more precise regulations on those preparation plans. The telecom operators must take care that the control and maintenance of their critical communications infrastructures can be restored to Finland in

emergency situations. The proposals in the Code of Information Society bill increase requirements for preparation for crisis and emergency situations.

The Emergency Act requires military attack or threat of a military or corresponding attack that endangers the vital functions of society in order that the government can take granted those emergency powers. In cyber attacks the distinction between military attack and escalation of cybercrimes is not visible and clearly observable as in traditional military attacks. The Emergency Act requires that authorities use those emergency powers in proportionate manner for only necessary purposes.

The Ministry of Defence has ordered a working group to consider the powers of security authorities for the implementation of Cyber Strategy and obvious needs for new legislation. The working group considers if it is necessary to grant new powers for security authorities in information retrieval. Those powers must be balanced in relation to obligations running from constitutional and human rights. The interference with confidential communications in information retrieval requires legal remedies against misuse of powers and constitutional accountability of security authorities. The constitutional and human rights must be guaranteed in cyber space and information society.

References

Committee for Constitutional Law, Finnish Parliament (2012) Perustuslakivaliokunnan lausunto PeVL 18/2012 vp–HE 66/2012 vp

European Court of Justice (2013) Opinion of Advocate General Jääskinen delivered on 25 June 2013: Case C–131/12 Google Spain SL and Google Inc. v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González, section 34

Judgement of the Court (Grand Chamber) 16 December 2008 Tietosuojavaltuutettu (Data Ombudsman) v. Satakunnan Markkinapörssi Oy and Satamedia Oy (Directive 95/46/EC–Scope–Processing and flow of tax data of a personal nature–Protection of natural person Freedom of expression), Case C-73/07

Ollila R (2004) Freedom of speech and protection of privacy. PhD thesis, Edita Publishing Oy, Helsinki. http://www.edilex.fi

Ollila R (2005) Turvallisuus ja urkinta. Lakimies 5(2005):781–791

Part II
Cyber Security Threats, Legality and Strategy

Coder, Hacker, Soldier, Spy

Kenneth Geers

Abstract A cyber attack is best understood not as an end in itself, but as a means to a wide variety of ends, some of which have serious legal, political, military, or economic ramifications. Cyber attacks may be employed for any purpose: espionage, crime, activism, terrorism, or war. They are used for competitive advantage in any and every form of human conflict. This chapter seeks to help cyber defenders classify attacks appropriately so that they can most efficiently allocate finite resources to combat this rising threat.

1 Introduction

A cyber attack is best understood not as an end in itself, but as a means to a wide variety of other ends, some of which have tangible political, military, criminal, and social consequences. A cyber attack is not a strategy, but a tactic that may be employed as one of many other, cyber and non-cyber tactics toward the attainment of a broader strategy. A cyber attacker's ultimate goal could be anything—personal amusement, intellectual property theft, political revolution, terrorism, or even international war.

In any competition, rivals strive for competitive advantage. Information technology, computer networks, and hacking are simply new ways to obtain it. Computer hardware and software are vulnerable to many different types of attack, including data theft, data denial, and data modification. Hackers are good at exploiting these vulnerabilities, but a cyber attack can have much more impact if the hacker is part of a larger team of criminals, soldiers, or spies.

This chapter explores five types of cyber conflict—espionage, crime, activism, terrorism, and war. Each type is hard to define precisely, and there is gray area between them. However, it is already clear, even at the very dawn of the Internet era,

K. Geers (✉)
FireEye, Reston, VA, USA
e-mail: KGeers@ATLANTICCOUNCIL.ORG

M. Lehto and P. Neittaanmäki (eds.), *Cyber Security: Analytics, Technology and Automation*, Intelligent Systems, Control and Automation: Science and Engineering 78, DOI 10.1007/978-3-319-18302-2_5

that each type is not only theoretically possible, but that there are numerous real-world examples to prove it. Finally, the author hopes that an improved ability to classify cyber attacks will help cyber defense personnel to better understand the threats they face, and as a result have the ability to prepare their defenses more effectively.

2 Cyber Espionage

Let's take espionage first, because in this case a hacker does not actually do anything to your data—except take it and read it. However, the amount of data that hackers are able to steal has already made this generation the Golden Era of Espionage. And as more and more of our lives are played out online, and as once-isolated computers are connected to the Internet, the level of sensitivity of the stolen data continues to rise.

However tenuous, there are wires and/or radio waves that connect every computer on planet Earth. It is a hacker's job to figure out how to transfer data between any two of them, and professional hackers are quite good at making this happen. Today, digitized information is stolen on an industrial scale, and data analysts possess powerful tools to mine that data very quickly, in order to sort the wheat from the chaff.

Nation-states have engaged in cyber espionage for at least the past 30 years. From the Cuckoo's Egg, to Moonlight Maze and Titan Rain, U.S. national laboratories and government agencies have been prime targets for foreign hackers. And the dynamic is accelerating, not slowing down. In 2008, Russia was the prime suspect in what U.S. Deputy Secretary of Defense William Lynn called the "most significant breach of U.S. military computers ever"—a USB-vector attack on Central Command (CENTCOM) (Lynn 2010). In 2012, Duqu, Flame, and Gauss targeted Iran (Bencsáth 2012) and the "Mahdi" malware (Simonite 2012) compromised engineering firms, government agencies, financial houses, and academia throughout the Middle East (Zetter 2012). Iran, for its part, is believed to have compromised a Dutch digital certificate authority, after which it issued more than 500 fraudulent certificates for major companies and government agencies (Charette 2011a, b), and possibly placed a Trojan horse program within the Simurgh "proxy" software, that is used by human rights activists (Marquis-Boire 2012). Back in 2009, the plans for a new U.S. Presidential Marine Corps 1 helicopter were found on a file-sharing network in Iran (Borak 2009). And in 2013, India was thought to be responsible for "Operation Hangover," a large-scale cyber espionage campaign in which Pakistani IT, mining, automotive, legal, engineering, food service, military, and financial networks were affected ("Operation …" 2013).

However, in 2013, the elephant in the room is China, which is currently the biggest single player in worldwide cyber espionage. Its U.S. target list alone seems exhaustive: in government, the Department of Energy, which manages nuclear technology (Gerth and Risen 1999); in the military, the plans for the most advanced

U.S. fighter jet, the F-35 (Gorman et al. 2009a); in technology, Google, Intel, and Adobe (Gross 2011); in business, Morgan Stanley and the U.S. Chamber of Commerce (Gorman 2011); in media, *The New York Times*, *Wall Street Journal*, and *Washington Post* (Perlroth 2013a, b).

And China is not merely interested in the U.S. In Europe, Chinese hackers have targeted the UK House of Commons (Warren 2006), British businessmen (Leppard 2010), and the German Chancellery (Espionage 2007). In India, one highly advanced attack on a Navy headquarters reportedly used a USB vector to bridge the "air-gap" between a compartmentalized, standalone network and the Internet (Pubby 2012). In South Korea, many personal computers and PDAs belonging to South Korean government leadership were compromised (Ungerleider 2010), as well as an Internet portal that held personal information for 35 million Korean citizens (Mick 2011). In Japan, many economic sectors have been compromised, and Chinese hackers are believed to have even stolen classified documents (McCurry 2011; "China-based …" 2011). In Australia, the blueprints for the Australian Security Intelligence Organization's new $631 million building were stolen ("Report: Plans …" 2013). And worldwide, researchers discovered that China-based hackers controlled a cyber espionage network in over 100 countries ("Tracking …" 2009). Last but not least, in 2010, a Chinese telecommunications firm transmitted erroneous routing information for 37,000 computer networks, misrouting some Internet traffic through China for 20 min, and exposing data from 8,000 U.S. networks, 1,100 Australian networks, and 230 French networks (Vijayan 2010).

Since at least 2010, many, likely China-based APTs have targeted the governments, militaries, and businesses of ASEAN, the Southeast Asian geopolitical and economic group composed of Brunei, Burma (Myanmar), Cambodia, Indonesia, Laos, Malaysia, Philippines, Singapore, Thailand, and Vietnam. The targeted industries include telecommunications, transportation, oil and gas, banks, and think tanks. The usual motivation is to gain tactical or strategic advantage within the political, military, and/or economic domains (Finkle 2011).

We believe that many of these regional economic organizations are attractive targets to APT campaigns because the information they possess is valuable—while their level of cyber security awareness is low. Often, these organizations have inconsistent system administration, infrequent software patch management, and/or poor policy control. Thus, many of these networks are "low-hanging fruit" for attackers. And to make matters worse, once compromised, they are used as staging grounds for further attacks on regional targets, through the installation of illicit command-and-control servers, the abuse of legitimate email accounts, and the dissemination of stolen office documents as "bait."

FireEye researchers are following numerous APT actors in this region, including BeeBus, Mirage, Check Command, Taidoor, Seinup and Naikon. Their most common tactic is spear-phishing, often using legitimate "decoy" documents that are related to the victim's national economy, politics, or regional events such as ASEAN summits, Asia-Pacific Economic Cooperation (APEC) summits, energy exploration, or military affairs.

For example, one group behind many noteworthy attacks, called the "Comment Crew" (Sanger et al. 2013) appears to be a large contractor to the Chinese government. One of its campaigns can be considered "strategic" in every sense of the word, focusing on U.S. aerospace and defense industries (Pidathala et al. 2013). The Comment Crew is so large, in fact, that when the Federal Bureau of Investigation (FBI) decoded only one of the group's stolen caches of information, if printed out, it would have created a stack of paper taller than a set of encyclopedias (Riley and Lawrence 2012).

Of course, it is highly unlikely that China alone is conducting cyber espionage. In 2006, the China Aerospace Science and Industry Corporation (CASIC) found spyware on its classified network ("Significant …" 2013). In 2007, the Chinese Ministry of State Security stated that foreign hackers were stealing Chinese information, "42 %" from Taiwan and "25 %" from the United Sates (*Ibid*). In 2009, Chinese Prime Minister Wen Jiabao announced that a hacker from Taiwan had stolen his upcoming report to the National People's Congress (*Ibid*). In 2013, Edward Snowden, a former system administrator at the National Security Agency (NSA), published documents suggesting that the U.S. conducted cyber espionage against China (Rapoza 2013); and the Chinese computer emergency response team (CERT) stated that it possessed "mountains of data" on cyber attacks by the U.S. (Hille 2013).

But what exactly are they looking for? In fact, they are looking for the same types of information that spies have always sought—leadership communications, research and development data, economic negotiating positions, political secrets, military plans, etc. The list of priorities for a national intelligence service is endless. In short, anything that might give a government a competitive advantage vis-à-vis its rivals (both internal and external).

The high return on investment has turned cyber espionage into an industrial-scale enterprise, which the lack of effective mitigation strategies on the part of cyber defenders suggest that this is not a problem that will go away soon. For now, the Russian government has recently taken the extreme measure of buying a shipment of old-fashioned typewriters (Ingersoll 2013).

3 Cyber Crime

Criminals are no strangers to technology. Counterfeiting, for example, is as old as money itself—which means ancient Greece. Today, counterfeiters likely have it much easier, as both money and intellectual property exist in electronic bits that can be transmitted around the world at light speed. In 2011, the European Union's carbon trading market was hacked, resulting in the theft of more than $7 million in credits, forcing the market to shut down temporarily (Krukowska and Carr 2011). In 2012, the European Aeronautic Defence and Space Company (EADS) and German steelmaker ThyssenKrupp fell victim to major attacks by Chinese hackers, who almost certainly used cyber attacks for strategic economic gain (Rochford 2013).

For businesses, one important trend has emerged. They should always be on the lookout for advanced persistent threat (APT) cyber attacks just before and during international negotiations. In 2011 alone, the European Commission complained of widespread hacking before an EU summit ("Serious …" 2011), the French government was compromised prior to a G-20 meeting (Charette 2011b), and at least 10 Norwegian defense and energy companies were hacked during large-scale contract negotiations, via phishing that was specifically tailored to each company (Albanesius 2011). In the economic sphere, the U.S.-based International Monetary Fund (IMF) fell victim to a phishing attack in 2011 that was described as a "very major breach" (Sanger and Markoff 2011).

Cyber crime can simply be the use of IT to facilitate traditional crime such as theft, fraud, or a tactical strike such as hacking an alarm system before unauthorized entry. All of this is possible at the individual, organizational, or nation-state level. In the case of a nation-state attack, there will be no law enforcement agency on the planet capable of completing a successful prosecution; the only recourse would be diplomatic, economic, or military coercion.

Cyber crime can also be "cyber-only" crimes such as sharing illegal images, files, or sensitive information. And professional hacker groups also fall into this category; they offer cyber-specific goods and services to anyone, from unknown individuals to governments, that include denial-of-service attacks and ownership of previously compromised networks. In the future, as more of our lives transpire online, the cyber-only form of crime will only grow. Perforce, this evolution will continue to push law enforcement ever further into cyberspace.

The incredible, asymmetric power of computers, networks, and computer hacking is as available to criminals as it is to anyone else. Today more than ever, non-state actors possess the tools they need to challenge the authority and indeed the supremacy of a nation-state. In 2010, the "Pakistani Cyber Army" defaced and subsequently shut down the website of the Central Bureau of Investigation, India's top police agency ("India …" 2010). In 2011, German Police found that servers used to locate serious criminals and terrorism suspects had been penetrated following a successful phishing attack ("Hackers …" 2011).

Unfortunately, there is often a gray area between state and non-state cyber attacks, some of which include criminal activity. FireEye researchers have even seen one nation-state develop and use a sophisticated Trojan, and later (after its own counter-Trojan defenses were in place) sell it to cyber criminals on the black market (Geers 2013).

Russia Internet space, for example, has created some complications for cyber defenders as they attempt to classify cyber attacks properly. In 2009, Russian hackers were blamed in "Climategate," a breach of university research in England that appeared intended to undermine international negotiations on climate change mitigation (Stewart and Delgado 2009; "Global …" 2011). And at FireEye, analysis has shown that some Russian "back doors" into compromised systems are hard to distinguish from advanced cyber criminal break-ins.

4 Cyber Activism

You say you want a revolution? Hacktivism is the art of mixing hacking and activism, thanks to groups with names like Cult of the Dead Cow and even individuals like Edward Snowden. Today, anyone who owns an Internet-connected computer possesses the historical equivalent of a printing press and a radio transmitter. Cyberspace is the perfect venue to promote any political cause, provided that people are willing to eventually migrate to the street or to the ballot box. Websites can be used to sell widgets or to bring down governments—just ask Hosni Mubarak.

Historically, nation-states have enjoyed a near-monopoly on the use of violence, and the ability to calibrate international tension. But in the Internet era, anyone can participate in international affairs, and this creates a new challenge for national decision makers. Ordinary citizens can participate in international conflict—either through the dissemination of propaganda or by computer hacking—without reporting to any traditional chain of command. Thus, one fascinating aspect of conflict in the Internet era is the phenomenon of "patriotic hackers" who ostensibly wage cyber war on behalf of their nations, if not their governments.

At the very dawn of the World Wide Web, in the mid-1990s, Russia was engaged in a protracted struggle over the fate of Chechnya; in the course of events there, the Chechens became pioneers in cyber propaganda, and the Russians became pioneers in shutting down their websites. In 1998, when Russian ally Serbia was under attack from NATO over the fate of Kosovo, pro-Serbian hackers jumped into the fray, targeting NATO with denial-of-service (DoS) attacks and at least twenty-five strains of virus-infected email. In 2001, following the downing of a U.S. Navy plane and the prolonged detention of its crew in China, pro-U.S. and pro-China hackers began a patriotic hacker war with uncertain consequences for political leadership. In 2007, Russia was the prime suspect in the most famous international cyber activist attack to date—the punitive distributed DoS on Estonia for having moved a Soviet-era statue. And in 2008, during the Russo-Georgian war, analysts noted that there will be a close relationship between cyber and conventional operations in all future military campaigns (Geers 2008). In 2010, the "Iranian Cyber Army" disrupted Twitter and the Chinese search engine Baidu, redirecting users to Iranian political messages (Wai-yin Kwok 2010).

The Arab-Israeli conflict, which evokes interest, argument, and passion worldwide, is a good case study in the use of cyber attacks to promote and defend a cause. Since at least the year 2000, pro-Israeli hackers have targeted websites of political and military significance in the Middle East (Geers 2008). But as an advanced industrial nation, Israel is also dependent on information technology, and in fact has proven to be quite vulnerable to cyber attacks, which often target the Israeli economy. In 2009, during Israel's military operation in Gaza, hackers briefly paralyzed many government sites with a DDoS attack from at least 500,000 computers. The 2009 attack consisted of four independent waves, each stronger

than the last, peaking at 15 million junk mail deliveries per second. The Israeli "Home Front Command" website, which plays a key role in national defense communications with the public, was down for 3 h. Due to technical similarities with the 2008 cyber attack on Georgia during its war with Russia, Israeli officials surmised that the attack itself may have been carried out by a criminal organization in the former Soviet Union, and paid for by Hamas or Hezbollah (Pfeffer 2009).

Today, Syria is in the midst of a civil war, so there is a lot of cyber activity to analyze. The most prominent hacker group by far is the Syrian Electronic Army (SEA), which is loyal to Syrian President Bashar al-Assad. SEA has conducted DDoS attacks, phishing, pro-Assad defacements, and spamming campaigns against governments, online services, and media that are perceived hostile to the Syrian government. SEA has hacked *Al-Jazeera*, Anonymous, Associated Press (AP), BBC, *Daily Telegraph*, *Financial Times*, *Guardian*, Human Rights Watch, National Public Radio, *New York Times*, Twitter, and more (Fisher and Keller 2011; "Syrian …" 2013). Its most famous exploit was an announcement via AP's Twitter account that the White House was bombed and President Obama injured—after which stock markets briefly dipped more than $200 billion (Manzoor 2013).

In the month of July 2013 alone, SEA compromised three widely used online communications websites: Truecaller (the world's largest telephone directory) (Khare 2013), Tango (a video and text messaging service) (Kastrenakes 2013; Albanesius 2013), and Viber (a free online calling and messaging application) (Ashford 2013). Successful compromises such as these are significant because they could give Syrian intelligence access to the communications of millions of people, including political activists within Syria who might then be targeted for espionage, intimidation, and/or arrest.

To compromise its victims, SEA often sends socially-engineered, spear-phishing emails to lure opposition activists into opening fraudulent, weaponized, and malicious documents. If the recipient falls for the scam, Trojan Horse, Remote Access Tool (RAT) software is installed on the victim's computer that can give the attacker keystrokes, screenshots, microphone/webcam recordings, stolen documents, and passwords. And of course, SEA likely sends all of this information to a computer address lying within Syrian government-controlled Internet Protocol (IP) space for intelligence collection and review (Tsukayama 2013).

As a category of hacker, activists will normally not possess the ability to employ the advanced cyber attacks that are characteristic of a nation-state. But what they do have in their favor is dynamism. Individual actors can operate on a whim, and respond quickly to ongoing events in the international arena. By contrast, government bureaucracies are notoriously bad at spontaneity and flexibility. So with cyber activism, the relative sophistication of an attack may be calculated less in the technological sophistication of the attack, and more in the clever and unpredictable ways in which the cyber attacks unfold.

5 Cyber Terrorism

To some degree, all advanced industrial economies have become dependent on the Internet, and would appear to be vulnerable to cyber terrorism. But in 2013, there is still no clear-cut case of cyber terrorism.

There are a few borderline reports. In 1998, the L0pht hacker group warned that it could shut down the Internet within 30 min (Schneier 1998). In 2000, a disgruntled employee wirelessly dumped 800,000 L of untreated sewage into Australian waterways (Smith 2001). In 2008, a CIA official informed a conference of critical infrastructure providers that unknown hackers, on multiple occasions, had been able to disrupt the power supply in various foreign cities (Nakashima and Mufson 2008). In 2013, the Iranian media reported that the Syrian army had carried out a cyber attack against the water supply of the Israeli city of Haifa. Prof. Isaac Ben-Israel, a cyber security adviser to Prime Minister Benjamin Netanyahu, said that the report was false but that cyber attacks on critical infrastructures do pose a "real and present threat" to Israel (Yagna 2013).

For the moment, most non-state terrorist organizations likely still fear nation-state surveillance more than they trust the uncertain prospects for cyber terrorism. There are still no cases of cyber terrorists causing serious physical damage or human casualties via the Internet. However, as with any other organization or subculture, terrorist groups are likely satisfied with the Internet's ability to help them organize, recruit, raise money, and disseminate propaganda.

Worldwide, the overall level of professionalism on the side of cyber defense has risen considerably. The ability of nation-states to police their own cyber jurisdiction is steadily improving. For cyber terrorism, this means that to an increasing degree, would-be cyber terrorists may need some level of state sponsorship.

So let's consider this dynamic more closely, in possibly the most likely candidate to be a sponsor of state-sanctioned cyber terrorism—North Korea. Pyongyang, an ally of China, would seem to be stuck in something of a cyber Stone Age, especially relative to prosperous South Korea, which is allied to the U.S. South Korea currently has the fastest download speeds in the world (McDonald 2011), and is home to schools where the students will be issued not books but computer tablets by 2015 (Gobry 2011).

However, from North Korea's perspective, the Internet offers a way not only to project national power but also to do it in a highly asymmetric way. North Korean defectors have described a burgeoning cyber warfare department of 3,000 personnel, largely trained in China or Russia. They have stressed that North Korea has a growing "fascination" with cyber attacks as a cost-effective way to compete against conventionally superior foes. They believed that North Korea is growing increasingly comfortable and confident in this new warfare domain, assessing that the Internet is not only vulnerable to attack but that a cyber attack strategy can create psychological pressure on the West. Toward this end, North Korea has focused on disconnecting its important servers from the Internet, while building a dedicated "attack network" (Fisher 2013).

What cyber attacks have we seen from North Korea thus far? In 2009, North Korea launched its first major assault on U.S. and South Korean government websites. There was little damage done, but the incident still gained wide media exposure (Choe and Markoff 2009). By 2013, however, the threat actors had matured. A group dubbed the "DarkSeoul Gang" was responsible for at least 4 years' worth of high-profile attacks on South Korea, including DDoS campaigns and malicious code that wiped computer hard drives at banks, media, ISPs, telcos, and financial companies—overwriting legitimate data with political messages ("Four …" 2013). Suspected North Korean attacks on U.S.-specific institutions include U.S. military elements in South Korea, the U.S.-based Committee for Human Rights in North Korea, and the White House.

6 Cyber War

Last but not least … is cyber war a myth, or is it reality? Skeptics focus on the fact that computer hacking is a technical discipline, not a missile or a bomb. And they like to point out that the number of known human casualties caused by cyber attacks is zero. However, the skeptics likely underappreciate the broader point that hacking is not an end in itself, but a means to a wide variety of ends—some (not all) of which could have serious political, military, or economic ramifications.

Because computers are used to manage everything today, including a nation's critical infrastructure and military weapons systems, a successful hack can lead to the sabotage of the larger system it helps to manage. In a national security context, think air defense, power grid, nuclear missiles. Any national leader would consider such an attack an act of war, and be forced to respond in some way. And these examples represent "hard" military targets; in other words, they are likely well-defended, but still perhaps vulnerable computer networks. But there are many other types of targets for hostile intelligence or military organizations that are "soft" targets and almost certainly poorly defended, including public media, university research, and government contractors such as logistics providers. For the practitioners of psychological operations and information warfare, such targets represent a fertile battlefield.

We are still only at the dawn of the Internet era, but there are numerous cyber security incidents that suggest cyber war may be a part of our common future. In 2007, Israel reportedly disrupted Syrian air defense networks via cyber attack (with some collateral damage to its own domestic networks) in order to facilitate the Israeli Air Force's destruction of an alleged Syrian nuclear facility (Carroll 2007). In 2008, there was clear evidence that computer network operations played a supporting role in Russian military advances during its invasion of Georgia ("Overview …" 2009). In 2009, French Navy planes were grounded following an infection by the Conficker worm (Willsher 2009). Also in 2009, Iraqi insurgents used $26 off-the-shelf software to intercept live video feeds from U.S. Predator drones, likely giving them the ability to monitor and evade U.S. military operations

(Gorman et al. 2009b). In 2012, the UK admitted that hackers had penetrated its classified Ministry of Defense networks (Hopkins 2012). In 2013, DHS reported that 23 gas pipeline companies were hacked in a potential sabotage operation (Clayton 2013), Chinese hackers were seen at the U.S. Army Corps of Engineers' National Inventory of Dams (Gertz 2013), and Russia (following the example of the U.S., China, and Israel) announced that it is creating cyber warfare-specific units (Gorshenin 2013).

But the most clear-cut example analysts have today of what we might call "cyber war" is Stuxnet (Sanger 2012), which was equally shocking on both the tactical and the strategic levels. At the tactical level, Stuxnet stands in stark contrast to computer worms such as Slammer and Code Red, which tried to compromise as many computers as possible. Stuxnet did the opposite, attempting to compromise *as few as possible*. Next, its malicious behavior was elegantly concealed under a veneer of apparently "legitimate" operational data; ultimately, however, the malware would cause centrifuge failure and destruction. In terms of a tactical military operation, Stuxnet (remember, this is just a piece of computer code!) to some degree replaced a squadron of fighter aircraft that would have violated foreign airspace, dropped laser-guided bombs, and left a smoking crater in the Earth's surface (*Ibid*). From a strategic perspective, Stuxnet was revolutionary in that it offered the prospect of tangible political and military gains in the international arena. The operation has been characterized as a "cyber missile" that was designed with a singular goal in mind—to disrupt the Iranian nuclear enrichment program, and help the international community prevent the arrival of a new member into the world's "nuclear club."

The Stuxnet "family" of malware, which likely includes Duqu, Flame, and Gauss (Bencsáth 2012), introduced the world to a new class of software—"military grade" software. Such programs—which in the first place are never even supposed to see the light of day—arrive at their destination in an encrypted form (often with stolen or forged digital certificates), and can only be decrypted and installed on a specific target device. Such tactics help the malware to evade the prying eyes of cyber defenders, and considerably increase the obstacles to discovery and reverse engineering. Beyond all this, consider that Stuxnet employed multiple zero-day exploits, and employed a world-first computational achievement in a forced cryptographic "hash collision" (Goodin 2012).

So if offensive military strikes are possible in the cyber domain, what options are available to the victim by way of retaliation? Does the victim keep the fight in the cyber domain, or can the response come in the form of a traditional military (or terrorist) assault? In 2012, Iran appears to have chosen the first option. A hacker group called the "Cutting Sword of Justice" used the "Shamoon" virus to attack the Saudi Arabian national oil company Aramco, deleting data on three-quarters of Aramco's corporate PCs (including documents, spreadsheets, e-mails, and files) and replacing them with an image of a burning American flag (Perlroth 2012). And over the past year, another group called *Izz ad-Din al-Qassam* launched "Operation Ababil," a series of DoS attacks against many U.S. financial institutions including the New York Stock Exchange (Walker 2013). Finally, in 2013 the *Wall Street*

Journal reported that Iranian actors had increased their efforts to compromise U.S. critical infrastructure (Gorman and Yadron 2013).

In 2013, President Obama signed a directive that the U.S. should aid allies who come under foreign cyber attack (Shanker and Sanger 2013), and as a step toward cyber détente, the U.S. and Russia signed an agreement to build a cyber "hotline"—similar to that used for nuclear scares during the Cold War—to help defuse any computer-related crises in the future (Gallagher 2013).

7 Conclusion

The future is unknown, but this chapter has shown that the power of computers and computer networks is being exploited by everyone—including spies, criminals, activists, terrorists, and soldiers. There are known, clear examples of cyber attacks that have been used to facilitate espionage, crime, activism, terrorism, and war.

In the future, as the Internet expands its reach into every aspect of our national and civil societies, and as we organically deepen our dependence upon it, the number of cyber attack examples, as well as their potential impact on our lives, is likely to increase. At the individual level, we must worry about freedom and privacy. At the national level, decision makers will remain focused on crime, terrorism, war, and revolution.

Could cyber attacks down a power grid or a financial market? In theory, yes … in practice, we still do not know. But we do know that the Internet provides governments with an incredibly powerful tool to manipulate our lives, and to conduct surveillance against not only their adversaries, but against their own citizens. Therefore, in the Internet era, it is vital that governments place an increasing emphasis on respect for both the rule of law and the Laws of War.

References

Albanesius C (2011) Norway cyber attack targets country's oil gas systems. PCMag. http://www.pcmag.com/article2/0,2817,2396611,00.asp. Accessed 17 Nov 2013

Albanesius C (2013) Tango messaging app targeted by Syrian Electronic Army. PCMag. http://www.pcmag.com/article2/0,2817,2422129,00.asp. Accessed 17 Nov 2013

Ashford W (2013) Syrian hacktvists hit second mobile app in a week. Computer Weekly. http://www.computerweekly.com/news/2240201656/Syrian-hacktvists-hit-second-mobile-app-in-a-week. Accessed 17 Nov 2013

Bencsáth B (2012) Duqu, flame, gauss: followers of stuxnet. BME CrySyS Lab RSA. http://www.rsaconference.com/writable/presentations/file_upload/br-208_bencsath.pdf. Accessed 17 Nov 2013

Borak D (2009) Source in Iran views Marine One blueprints. Marine Corps Times. http://www.marinecorpstimes.com/article/20090303/NEWS/903030307/Source-in-Iran-views-Marine-One-blueprints. Accessed 17 Nov 2013

Carroll W (2007) Israel's cyber shot at Syria. Defense Tech. http://defensetech.org/2007/11/26/israels-cyber-shot-at-syria/. Accessed 17 Nov 2013

Charette R (2011a) 'Spectacular' cyber attack gains access to France's G20 files. IEEE Spectrum. http://spectrum.ieee.org/riskfactor/telecom/internet/spectacular-cyber-attack-gains-access-to-frances-g20-files. Accessed 17 Nov 2013

Charette R (2011b) DigiNotar certificate authority breach crashes e-government in the Netherlands. IEEE Spectrum. http://spectrum.ieee.org/riskfactor/telecom/security/diginotar-certificate-authority-breach-crashes-egovernment-in-the-netherlands. Accessed 17 Nov 2013

China-based servers in Japan cyber attacks (28 Oct 2011) The Indian Express. http://www.indianexpress.com/news/chinabased-servers-in-japan-cyber-attacks/866665/. Accessed 17 Nov 2013

Choe S-H, Markoff J (2009) Cyberattacks jam government and commercial web sites in US and South Korea. The New York Times. http://www.nytimes.com/2009/07/09/technology/09cyber.html?_r=0. Accessed 17 Nov 2013

Clayton M (2013) Exclusive: cyberattack leaves natural gas pipelines vulnerable to sabotage. The Christian Science Monitor. http://www.csmonitor.com/Environment/2013/0227/Exclusive-Cyberattack-leaves-natural-gas-pipelines-vulnerable-to-sabotage. Accessed 17 Nov 2013

Espionage report: Merkel's China visit marred by hacking allegations (2007) Spiegel. http://www.spiegel.de/international/world/espionage-report-merkel-s-china-visit-marred-by-hacking-allegations-a-502169.html. Accessed 17 Nov 2013

Finkle J (2011) 'State actor' behind slew of cyber attacks. Reuters. http://www.reuters.com/article/2011/08/03/us-cyberattacks-idUSTRE7720HU20110803. Accessed 17 Nov 2013

Fisher M (2013) South Korea under cyber attack: is north Korea secretly awesome at hacking? The Washington Post. http://www.washingtonpost.com/blogs/worldviews/wp/2013/03/20/south-korea-under-cyber-attack-is-north-korea-secretly-awesome-at-hacking/. Accessed 17 Nov 2013

Fisher M, Keller J (2011) Syria's digital counter-revolutionaries. The Atlantic. http://www.theatlantic.com/international/archive/2011/08/syrias-digital-counter-revolutionaries/244382/. Accessed 17 Nov 2013

Four years of DarkSeoul cyberattacks against South Korea continue on anniversary of Korean war (2013) Symantec. http://www.symantec.com/connect/blogs/four-years-darkseoul-cyberattacks-against-south-korea-continue-anniversary-korean-war. Accessed 17 Nov 2013

Gallagher S (2013) US, Russia to install 'cyber-hotline' to prevent accidental cyberwar. Ars Technica. http://arstechnica.com/information-technology/2013/06/us-russia-to-install-cyber-hotline-to-prevent-accidental-cyberwar/. Accessed 17 Nov 2013

Geers K (2008) Cyberspace and the changing nature of warfare. SC Magazine. http://www.scmagazine.com/cyberspace-and-the-changing-nature-of-warfare/article/115929/. Accessed 17 Nov 2013

Geers K, Kindlund D, Moran N, Rachwald R (2013) World war C: understanding nation-state motives behind today's advanced cyber attacks. FireEye Labs. http://www.fireeye.com/resources/pdfs/fireeye-wwc-report.pdf. Accessed 17 Nov 2013

Gerth J, Risen J (1999) 1998 report told of lab breaches and China threat. The New York Times. http://www.nytimes.com/1999/05/02/world/1998-report-told-of-lab-breaches-and-china-threat.html?pagewanted=all&src=pm. Accessed 17 Nov 2013

Gertz B (2013) Dam! sensitive army database of US dams compromised; Chinese hackers suspected. The Washington Times. http://www.washingtontimes.com/news/2013/may/1/sensitive-army-database-us-dams-compromised-chines/?page=all. Accessed 17 Nov 2013

Global warning: New Climategate leaks (2011) RT. http://rt.com/news/global-warming-climate-controversy-009/. Accessed 17 Nov 2013

Gobry P-E (2011) South Korea will replace all paper with tablets in schools by 2015. Business Insider. http://www.businessinsider.com/south-korea-tablets-in-schools-by-2015-2011-7. Accessed 17 Nov 2013

Goodin D (2012) Crypto breakthrough shows flame was designed by world-class scientists. Ars Technica. http://arstechnica.com/security/2012/06/flame-crypto-breakthrough/. Accessed 17 Nov 2013

Gorman S (2011) China hackers hit US chamber. Wall Street Journal. http://online.wsj.com/news/articles/SB10001424052970204058404577110541568535300. Accessed 17 Nov 2013
Gorman S, Yadron D (2013). Iran hacks energy firms, US says. Wall Street Journal. http://online.wsj.com/news/articles/SB10001424127887323336104578501601108021968. Accessed 17 Nov 2013
Gorman S, Cole A, Dreazen Y (2009a) Computer spies breach fighter-Jet project. Wall Street Journal. http://online.wsj.com/news/articles/SB124027491029837401. Accessed 17 Nov 2013
Gorman S, Dreazen Y, Cole A (2009b) Insurgents hack US drones. Wall Street Journal. http://online.wsj.com/news/articles/SB126102247889095011. Accessed 17 Nov 2013
Gorshenin V (2013) Russia to create cyber-warfare units. Pravda. http://english.pravda.ru/russia/politics/29-08-2013/125531-cyber_warfare-0/. Accessed 17 Nov 2013
Gross MJ (2011) Enter the cyber-dragon. Vanity Fair. http://www.vanityfair.com/culture/features/2011/09/chinese-hacking-201109. Accessed 17 Nov 2013
Hackers infiltrate German police and customs service computers (2011) Infosecurity Magazine. http://www.infosecurity-magazine.com/view/19494/hackers-infiltrate-german-police-and-customs-service-computers/. Accessed 17 Nov 2013
Hille K (2013) China claims 'mountains of data' on cyber attacks by US. Financial Times. http://www.ft.com/intl/cms/s/0/921f47cc-cdce-11e2-a13e-00144feab7de.html#axzz2kwXMwlAz. Accessed 17 Nov 2013
Hopkins N (2012) Hackers have breached top secret MoD systems, cyber-security chief admits. The Guardian. http://www.theguardian.com/technology/2012/may/03/hackers-breached-secret-mod-systems. Accessed 17 Nov 2013
India and Pakistan in cyber war (2010) Al-Jazeera. http://www.aljazeera.com/news/asia/2010/12/20101241373583977.html. Accessed 17 Nov 2013
Ingersoll G (2013) Russia turns to typewriters to protect against cyber espionage. Business Insider. http://www.businessinsider.com/russia-turns-to-typewriters-for-secrets-2013-7. Accessed 17 Nov 2013
Kastrenakes J (2013) Syrian Electronic Army alleges stealing 'millions' of phone numbers from chat app Tango. The Verge. http://www.theverge.com/2013/7/22/4545838/sea-giving-hacked-tango-database-government. Accessed 17 Nov 2013
Khare A (2013) Syrian Electronic Army hacks truecaller database, gains access codes to social media accounts. iDigital Times. http://www.idigitaltimes.co.uk/articles/492337/20130719/syrian-electronic-army-hacks-truecaller-database-gains.htm. Accessed 17 Nov 2013
Krukowska E, Carr M (2011) EU carbon trading declines after alleged hacking suspends spot market. Bloomberg. http://www.bloomberg.com/news/2011-01-20/carbon-trading-declines-as-eu-regulator-halts-spot-market-on-hacking-probe.html. Accessed 17 Nov 2013
Leppard D (2010) China bugs and burgles Britain. The Sunday Times. http://www.thesundaytimes.co.uk/sto/news/uk_news/article196465.ece. Accessed 17 Nov 2013
Lynn WJ (2010) Defending a new domain: the Pentagon's cyberstrategy. Foreign Aff 89(5):97–108. http://www.foreignaffairs.com/articles/66552/william-j-lynn-iii/defending-a-new-domain. Accessed 17 Nov 2013
Manzoor S (2013) Slaves to the algorithm: are stock market math geniuses, or quants, a force for good? Ottawa Citizen. http://www.ottawacitizen.com/business/Slaves+algorithm+stock+market+math+geniuses+quants+force+good/8707964/story.html. Accessed 17 Nov 2013
Marquis-Boire M (2012) Iranian anti-censorship software 'Simurgh' circulated with malicious backdoor. Citizenlab. https://citizenlab.org/2012/05/iranian-anti-censorship-software-simurgh-circulated-with-malicious-backdoor-2/. Accessed 17 Nov 2013
McCurry J (2011) Japan anxious over defence data as China denies hacking weapons maker. The Guardian. http://www.theguardian.com/world/2011/sep/20/china-denies-hacking-attack-japan. Accessed 17 Nov 2013
McDonald M (2011) Home internet may get even faster in South Korea. The New York Times. http://www.nytimes.com/2011/02/22/technology/22iht-broadband22.html. Accessed 17 Nov 2013

Mick J (2011) Chinese hackers score heist of 35 million South Koreans' personal info. Daily Tech. http://www.dailytech.com/Chinese+Hackers+Score+Heist+of+35+Million+South+Koreans+Personal+Info/article22284.htm. Accessed 17 Nov 2013

Nakashima E, Mufson S (2008) Hackers have attacked foreign utilities, CIA analyst says. Washington Post. http://www.washingtonpost.com/wp-dyn/content/article/2008/01/18/AR2008011803277.html. Accessed 17 Nov 2013

Operation hangover: Q&A on attacks (2013) Symantec. http://www.symantec.com/connect/blogs/operation-hangover-qa-attacks. Accessed 17 Nov 2013

Overview by the US-CCU of the cyber campaign against Georgia in August of 2008 (2009) A US-CCU special report, U.S. Cyber Consequences Unit. http://www.registan.net/wp-content/uploads/2009/08/US-CCU-Georgia-Cyber-Campaign-Overview.pdf. Accessed 17 Nov 2013

Perlroth N (2012) In cyberattack on Saudi firm U.S. sees Iran firing back. The New York Times. http://www.nytimes.com/2012/10/24/business/global/cyberattack-on-saudi-oil-firm-disquiets-us.html. Accessed 17 Nov 2013

Perlroth N (2013a) Washington post joins list of news media hacked by the Chinese. The New York Times. http://www.nytimes.com/2013/02/02/technology/washington-posts-joins-list-of-media-hacked-by-the-chinese.html. Accessed 17 Nov 2013

Perlroth N (2013b) Wall Street Journal announces that it too was hacked by the Chinese. The New York Times. http://www.nytimes.com/2013/02/01/technology/wall-street-journal-reports-attack-by-china-hackers.html. Accessed 17 Nov 2013

Pfeffer A (2009) Israel suffered massive cyber attack during Gaza offensive. Haaretz. http://www.haaretz.com/news/israel-suffered-massive-cyber-attack-during-gaza-offensive-1.278094. Accessed 17 Nov 2013

Pidathala V, Kindlund D, Haq T (2013) "Operation Beebus," FireEye advanced threat report—2H 2012. FireEye Labs. http://www2.fireeye.com/rs/fireye/images/fireeye-advanced-threat-report-2h2012.pdf. Accessed 17 Nov 2013

Pubby M (2012) China hackers enter Navy computers plant bug to extract sensitive data. The Indian Express. http://www.indianexpress.com/news/china-hackers-enter-navy-computers-plant-bug-to-extract-sensitive-data/968897/. Accessed 17 Nov 2013

Rapoza K (2013) U.S. hacked China universities, mobile phones, Snowden tells China Press. Forbes. http://www.forbes.com/sites/kenrapoza/2013/06/22/u-s-hacked-china-universities-mobile-phones-snowden-tells-china-press/. Accessed 17 Nov 2013

Report: Plans for Australia spy HQ hacked by China (2013) USA Today. http://www.usatoday.com/story/news/world/2013/05/28/china-hackers-australia-spy-agency-headquarters/2364863/. Accessed 17 Nov 2013

Riley M, Lawrence D (2012) Hackers linked to China's Army seen from EU to D.C. Bloomberg. http://www.bloomberg.com/news/2012-07-26/china-hackers-hit-eu-point-man-and-d-c-with-byzantine-candor.html. Accessed 17 Nov 2013

Rochford O (2013) European space, industrial firms breached in cyber attacks: report. Security Week. http://www.securityweek.com/european-space-industrial-firms-breached-cyber-attacks-report. Accessed 17 Nov 2013

Sanger D (2012) Confront and conceal: Obama's secret wars and surprising use of American power. Broadway Books, New York, pp 188–225

Sanger D, Markoff J (2011) I.M.F. reports cyberattack led to 'very major breach'. New York Times. http://www.nytimes.com/2011/06/12/world/12imf.html. Accessed 17 Nov 2013

Sanger D, Barboza D, Perlroth N (2013) Chinese army unit is seen as tied to hacking against U.S. The New York Times. http://www.nytimes.com/2013/02/19/technology/chinas-army-is-seen-as-tied-to-hacking-against-us. Accessed 17 Nov 2013

Schneier B (1998) Click here to bring down the internet. Schneier on Security. https://www.schneier.com/essay-003.html. Accessed 17 Nov 2013

'Serious' cyber attack on EU bodies before summit (2011) BBC. http://www.bbc.co.uk/news/world-europe-12840941. Accessed 17 Nov 2013

Shanker T, Sanger D (2013) U.S. helps allies trying to battle Iranian hackers. New York Times. http://www.nytimes.com/2013/06/09/world/middleeast/us-helps-allies-trying-to-battle-iranian-hackers.html. Accessed 17 Nov 2013

Significant cyber incidents since 2006 (2013) Center for strategic and international studies. http://csis.org/publication/cyber-events-2006. Accessed 17 Nov 2013

Simonite T (2012) Bungling cyber spy stalks Iran. MIT Technology Review. http://www.technologyreview.com/news/429046/bungling-cyber-spy-stalks-iran/. Accessed 17 Nov 2013

Smith T (2001) Hacker jailed for revenge sewage attacks. The Register. http://www.theregister.co.uk/2001/10/31/hacker_jailed_for_revenge_sewage/. Accessed 17 Nov 2013

Stewart W, Delgado M (2009) Were Russian security services behind the leak of 'Climategate' emails? Daily Mail. http://www.dailymail.co.uk/news/article-1233562/Emails-rocked-climate-change-campaign-leaked-Siberian-closed-city-university-built-KGB.html. Accessed 17 Nov 2013

Syrian Electronic Army (2013) Wikipedia. http://en.wikipedia.org/wiki/Syrian_Electronic_Army. Accessed 17 Nov 2013

Tracking GhostNet: Investigating a cyber espionage network (2009) Information warfare monitor. http://www.nartv.org/mirror/ghostnet.pdf. Accessed 17 Nov 2013

Tsukayama H (2013) Attacks like the one against the New York Times should put consumers on alert. The Washington Post. http://articles.washingtonpost.com/2013-08-28/business/41530263_1_hackers-security-researchers-fireeye. Accessed 17 Nov 2013

Ungerleider N (2010) South Korea's power structure hacked, digital trail leads to China. Fast Company. http://www.fastcompany.com/1696014/south-koreas-power-structure-hacked-digital-trail-leads-china. Accessed 17 Nov 2013

Vijayan J (2010) Update: report sounds alarm on China's rerouting of U.S. Internet traffic. Computerworld. http://www.computerworld.com/s/article/9197019/Update_Report_sounds_alarm_on_China_s_rerouting_of_U.S._Internet_traffic. Accessed 17 Nov 2013

Wai-yin Kwok V (2010) Baidu hijacked by cyber army. Forbes. http://www.forbes.com/2010/01/13/baidu-cyber-attack-markets-technology-china.html. Accessed 17 Nov 2013

Walker D (2013) Hacktivists plan to resume DDoS campaign against U.S. banks. SC Magazine. http://www.scmagazine.com/hacktivists-plan-to-resume-ddos-campaign-against-us-banks/article/283474/. Accessed 17 Nov 2013

Warren P (2006) Smash and grab, the hi-tech way. The Guardian. http://www.theguardian.com/politics/2006/jan/19/technology.security. Accessed 17 Nov 2013

Willsher K (2009) French fighter planes grounded by computer virus. The Telegraph. http://www.telegraph.co.uk/news/worldnews/europe/france/4547649/French-fighter-planes-grounded-by-computer-virus.html. Accessed 17 Nov 2013

Yagna Y (2013) Ex-General denies statements regarding Syrian cyber attack. Haaretz. http://www.haaretz.com/news/diplomacy-defense/ex-general-denies-statements-regarding-syrian-cyber-attack.premium-1.525941. Accessed 17 Nov 2013

Zetter K (2012) Mahdi, the Messiah, found infecting systems in Iran, Israel. WIRED. http://www.wired.com/threatlevel/2012/07/mahdi/. Accessed 17 Nov 2013

Cyber Warfare

Rain Ottis

Abstract This chapter explores the concept of cyber warfare from two different angles. First, from the perspective of public international law on armed conflict (henceforth—international law). This is important to understand, since it addresses the role of cyber warfare in the context of using (armed) force in international conflicts. Second, the chapter explores cyber warfare as a developing military capability, which is finding its place among other (often more mature) capabilities, such as electronic warfare or missile defense. Instead of covering the breadth of each topic, the chapter identifies the key points that help understand the nature of cyber warfare.

1 Introduction

Cyber warfare and (military) cyber operations have become popular terms in recent years. However, in many cases the terms are used with little understanding of the underlying concept. For example, referring to the events in Estonia in April and May of 2007 as cyber warfare, or even "Web War one" (Davis 2007) understandably creates confusion about the existence and the nature of cyber warfare. This chapter tries to clarify these issues.

The chapter is based on the point of view that cyber warfare is primarily a military endeavor. In addition to military cyber operations, the concept covers (covert) offensive cyber operations by intelligence agencies, but excludes espionage. The term *cyber warfare* is used to refer to the discipline as a whole (including doctrine, tactics, technologies, etc.), while the term *cyber operation* is used to refer to a specific act utilizing the methods and/or means of cyber warfare.

R. Ottis (✉)
Tallinn University of Technology, Tallinn, Estonia
e-mail: rain.ottis@ttu.ee

R. Ottis
Jyväskylä University, Jyväskylä, Finland

M. Lehto and P. Neittaanmäki (eds.), *Cyber Security: Analytics, Technology and Automation*, Intelligent Systems, Control and Automation: Science and Engineering 78, DOI 10.1007/978-3-319-18302-2_6

Although cyber warfare has a defensive and an offensive side, only the latter is discussed, since there are far fewer conceptual and legal problems with defense, and the defensive actions and technologies are well known and not unique to the military entities. Note that for the sake of clarity the concepts like active defense or proactive defense are considered to be offensive in nature. This does not mean that all forms of offensive cyber operations should be considered threatening. For example, Red Teaming or probing one's own systems for vulnerabilities is a commonly accepted and relatively uncontroversial technique, where offensive tools and methods are used.

The chapter does not cover cyber espionage, cyber crime or hacktivism, even though modern (cyber) conflicts often include, overlap with, or have direct links between some or all of those elements and military operations. Consider, for example, the role of so-called 'patriotic hackers' in the 2008 Russia-Georgia war (Carr 2009).

The cyber warfare capability discussion is not limited to a strict binary view of war and peace, but exists throughout the spectrum of violence (and corresponding law). To illustrate the vagueness of this spectrum, consider the case of Stuxnet, where a (allegedly) US cyber operation was used to sabotage an Iranian uranium enrichment facility, even though those states were not engaged in an armed conflict at the time. Another example from the same conflict is from 2011, when Iran was able to remotely manipulate a US military drone aircraft to land in Iran so that it could be captured. While the exact details of how this was accomplished are not publicly available, it is speculated that some form of cyber warfare capability was involved. While US has been worried for years about potential attacks against its infrastructure, recently exposed documents indicate that the US has already launched numerous cyber operations against other states. Therefore, cyber warfare capability has clearly evolved from a theoretical possibility to reality (regardless of how the above examples are categorized in legal terms) (See, for example, Falliere et al. 2011; Gellman and Nakashima 2013; Peterson 2011; Rawnsley 2011).

The chapter is divided into four parts. After the introduction, the second part looks at cyber warfare from the perspective of international law. It introduces some key terms of art and illustrates the significance of those by answering two questions. This part is necessarily focused on the more violent end of the spectrum. The third part looks at cyber warfare as a (emerging) military capability that is still developing. The chapter closes with some concluding remarks.

2 Cyber Warfare from the Perspective of International Law

When considering cyber warfare from the perspective of international law, it is important to clarify some legal terms of art, namely, *armed conflict*, *armed attack* and *threat or use of force*. Instead of *war* or *warfare*, the international law generally uses the term *armed conflict* (there are exceptions, such as the term *prisoner of war*). Hence the phrase Law of Armed Conflict (LOAC), which refers to the body

of international law that addresses the legal aspects of the conduct of hostilities in international relations.

It should be noted that international law as such cannot and does not prevent armed conflict. The de facto decision to engage in armed conflict rests with the states and is often more concerned with the political and military reality than with legal frameworks, especially if the latter have some room for interpretation. It is likely that similar flexible analysis is seen in future conflicts involving cyber operations.

From the perspective of cyber warfare, one of the key instruments of international law is the United Nations Charter which, among other things, sets the standard for the use of (or restraint thereof) force between states. While the "threat or use of force" is generally prohibited [Article 2(4)], the Charter does offer a few exceptions. Specifically, Article 42 allows the UN to sanction operations to "maintain or restore international peace and security" and Article 51 makes the exception for states that engage in (collective) self-defence in response to an "armed attack" (United Nations 1945).

The UN Charter (United Nations 1945) does not provide a clear definition for *use of force* or *armed attack*. While other sources of international law (such as treaties and state practice) do provide some guidance, there is still substantial ground for interpretation, especially for emerging topics like cyber warfare or operations. On one hand this is a problem, because states can (and often do) interpret these thresholds differently. On the other hand this provides longevity to the law. Consider, for example, if the Charter (or some other document) were to specifically list the types of things that would qualify as an armed attack. That document would have to be updated whenever a new technology or tactic of sufficient influence is developed, such as cyber operations or nano-technology. This update mechanism would likely be too slow to matter in real conflicts.

When new technologies and tactics do become available, they will need to be analyzed from the perspective of international law. Specifically, whether they fall under it at all and if so, how does international law apply in that case. In terms of cyber warfare, cyber weapons, cyber operations etc. this debate is still ongoing, but there is already plenty of research and policy available. Some states have developed internal manuals and policies on the subject and many scholars have tackled various aspects of it over the years (for example, Schmitt 1999, 2002; Ziolkowski 2012). Perhaps one of the most comprehensive international treatments of the subject is the Tallinn Manual on the International Law Applicable to Cyber Warfare (Schmitt 2013), which explores relevant questions on sovereignty, state responsibility, use of force, conduct of hostilities, etc. While the Tallinn Manual as such is not binding to states, it does provide a thorough analysis of existing law and state practice (in terms of cyber warfare) at the time of writing.

Let us explore two questions in order to illustrate why international law matters in the context of cyber warfare. First, can an offensive cyber operation or a series of such operations alone reach the threshold of armed attack? If the answer is no, then there is little to stop states from using offensive cyber operations against each other. If the answer is yes, however, then the risk level changes significantly for the

attacking side, because the target may legally use force in response (self-defense under Article 51 of the UN Charter), if the operation in question does meet the armed attack threshold. It is important to realize that there is no requirement of symmetricity in self-defense. If one side conducts a cyber operation that rises to the level of armed attack, then the other side can counter with cruise missiles, naval action, cyber operation, etc. There are other factors that must be considered, such as the proportionality of the response, but those do not necessarily limit the type of force used. While the threshold of "scale and effects" remains under debate, it seems clear that the international law community is generally in agreement that a cyber operation can reach the status of armed attack (Schmitt 2013).

Second interesting question, if theoretical, is whether cyber war can exist on its own. That is, can there be a *proper* war that only takes place in or via cyberspace, without troop movements, conventional weapons, etc.? Since it seems possible that a cyber operation can reach the level of armed attack (triggering an armed conflict) and the following actions by opposing actors could also be limited to cyber operations, then the answer is yes. However, this type of scenario does not seem likely, as it would needlessly restrict the playing field of the opposing actors, who are likely to also have other options at their command.

While the armed attack threshold is high, it does not mean that *lesser* cyber operations do not matter. Once an armed conflict has begun (regardless of the type or legality of the event that started it), cyber operations will be subject to LOAC. They will have to follow the principles of distinction (targeting only military objectives), proportionality, etc. just as all other military operations (Schmitt 2013).

3 Cyber Warfare as a Military Capability

The field of military cyber operations or cyber warfare is still rapidly developing and therefore has not yet settled within the framework of more mature warfighting disciplines and technologies, such as air combat or surface-to-surface missiles. This process of experimentation and accumulation of practical experience occurs quite naturally whenever a new technological or tactical development makes its way to the battlefield. The overview of the first decades of air combat (featuring aircraft that are heavier than air) in the first half of the twentieth century by Rattray (2001) provides many parallels to the arguments that we have seen in recent years about the nature of cyber warfare. For example, the discussion about the need for a new type of soldier or a new career model for 'warfighters' that have no direct contact with the enemy (a similar argument is currently ongoing with the military drone pilots) (See, for example, Rattray 2001; Conti and Surdu 2009; Hoagland 2013).

Perhaps the best analogy within the widely accepted and more mature military capabilities is electronic warfare. As with cyber warfare, the direct effects of electronic warfare are aimed at technology and are difficult to detect with human senses. Both are very technical in nature. Both are considered part of Information Operations by JP 3-13, even though JP 3-13 uses the term cyberspace operations

(the 2006 version of JP 3-13 used the term computer network operations (CNO), which was further categorized into computer network defense (CND), computer network attack (CNA) and computer network exploitation (CNE)) (Joint Chiefs of Staff 2006, 2012). The main difference seems to be the time they have had to mature, since electronic warfare has been around for much longer.

Cyber warfare capability can be viewed from two very different contexts. First, cyber warfare can be viewed as part of information operations, where the emphasis is on information. That is—managing or manipulating the content and availability of information to affect the relevant interest groups (friendly, neutral or hostile). This is a soft approach, where cyber is one of the ways to affect the decision process (ability) of the adversary (Joint Chiefs of Staff 2012). Possible examples include defacements (replacing the content of the target system), using adversary communications infrastructure to disseminate information (such as sending text messages containing propaganda or warnings of an impending attack), degrading or disabling the communication systems for the purposes of information blockade, etc.

Second, cyber warfare can be viewed in the context of information technology enabled (and dependent) military capabilities. Over the recent decades, warfare through superior technology and networking capability (as illustrated by Operation Desert Storm in 1991) has been discussed under different names, such as network centric warfare, network enabled capability, or revolution in military affairs. This second approach to cyber warfare is targeting technology, specifically the information technology that enhances the effectiveness of 'conventional' weapons and tactics. The role of offensive cyber operations is to manipulate (destroy, degrade, deny, etc.) the enemy systems so that their fighting capacity is (significantly) lowered, if not completely neutralized. For example, corrupting logistics databases, disabling GPS terminals, degrading the quality of sensor feeds, etc.

A common argument against (the usefulness of) cyber warfare is that it is difficult to produce a lasting effect, unlike with most conventional capabilities where injuries, death and destruction are fairly easy to achieve. The target of an offensive cyber operation may have to restart the server, install a patch, re-install the operating system, etc., but generally the system can be returned to operational status, given enough time. In case of distributed denial of service attacks, the target may just have to wait until the attack ends and the service is restored. Therefore, it should be clear that most cyber operations will indeed have a temporary effect on the target system. However, it is much more important to consider the effects that cyber operations have on the larger operations that they support.

Consider, for example, the Israeli air strike on the alleged nuclear facility in Syria in September of 2007. The Syrian air defenses did not work as intended and the Israeli planes made it back safely after destroying the facility. While details of the operation are not published, there has been a lot of speculation about the use of electronic or cyber warfare capabilities to temporarily neutralize the Syrian air defense system. Regardless of the actual method used to achieve the result, this example clearly demonstrates the (potential) value of temporary effects (by cyber operations), as long as they are well integrated with a larger operation (See, for example, Adee 2008; Fulghum 2007).

Another common misconception about offensive (military) cyber operations is that they are just a push of a button away, requiring little or no prior planning. While there are examples like this, they are typically not significant in the context of military operations. For example, it is easy to assemble a botnet before a conflict and use it to launch a distributed denial of service attack against any target that is connected to the same network (generally speaking—the Internet). Similarly, it is possible to launch (relatively) untargeted attacks with little or no advance preparation against targets of opportunity by scanning for and exploiting known vulnerabilities.

However, what about mission critical targets that are not connected to outside networks (air gapped), use custom hardware or software, or need to be affected in a very specific time or way? What if the attack must remain stealthy for months or years after it is executed? Consider, for example, the case of Stuxnet, where a piece of software was used to sabotage a uranium enrichment plant in Iran. A cyber operation of this type would require a lengthy preparation period, which includes collecting information about the target system (hardware and software versions, network diagram, the physical process to be affected, control systems, security measures, etc.), developing and testing the attack software (which may require custom or simulated hardware and software), and delivering it to the target (jumping the air gap). Therefore, such attacks can only be ordered, if the preparation has been done beforehand (as a contingency plan) or if there is enough time to complete the preparation for the attack. This was likely the case with Stuxnet, where the preparation period seems to have been several years (See, for example, Falliere et al. 2011, Sanger 2012).

In practical terms, this means that there are two types of offensive cyber operations—against targets of opportunity and against (previously) specified targets. Arguably, the latter are more dangerous, but rare. Such operations can be prepared as contingencies, if the likely opponent and targets are determined years or months in advance. However, this preparation also increases the risk of detection (collecting information on a specific target, building similar test sites, intelligence leaks, etc.) and may escalate tensions between the two actors.

4 Conclusion

The military capability of cyber warfare is undergoing a period of rapid development in the understanding of the applicability and application of established norms, both in terms of international law and of military tradition. States experiment with the boundaries of policy, law and technology in order to find the right balance, reminiscent of the discussion about air warfare a century ago.

Unfortunately, the term *cyber warfare* is often used very liberally, sometimes encompassing hacktivism, cyber crime or even espionage. This chapter explored cyber warfare from the point of view that it is a primarily military endeavor.

Specifically, the chapter explored cyber warfare from the perspective of international law and as a developing military capability.

It is clear that cyber operations conducted in the context of (international) armed conflict fall under international law. Perhaps most importantly, offensive cyber operations with sufficient scale and effect can rise to the level of armed attack, which enables the victim state to use force in self-defense. Once an armed conflict is under way, however, all military cyber operations will have to follow LOAC. Notably, cyber operations must comply with the principles of proportionality, distinction, etc.

As a military capability, cyber warfare is still finding its place among the other (often more mature) disciplines. However, the doctrinal questions about cyber warfare are not new—the same questions were asked about airplanes and radars, when these technologies entered the battlefield. Cyber warfare can be viewed from two main positions. In one the aim of cyber warfare is to (assist) manipulate the content of systems (information, data), in order to affect the decision making process of the target audience. In the other, the aim of cyber warfare is to manipulate systems in order to weaken or neutralize the adversary's technical (military) capabilities.

Cyber warfare will continue to change with the development and adoption of new technologies. Advances in artificial intelligence, the Internet of Things and quantum computing are just a few possible sources for innovation in cyber warfare. However, it will still be subject to international law and will likely have to pass through a turbulent settling period, as has been the case with other military developments in the past.

References

Adee S (2008) The hunt for the kill switch. IEEE Spectrum, May. http://spectrum.ieee.org/semiconductors/design/the-hunt-for-the-kill-switch/0. Accessed 02 Oct 2013

Carr J (2009) Inside cyber warfare: mapping the cyber underworld. O'Reilly Media, Sebastopol

Conti G, Surdu J (2009) Army, navy, air force, and cyber—is it time for a cyberwarfare branch of military? IAnewsletter 12(1):14–18

Davis J (2007) Hackers take down the most wired country in Europe. Wired Magazine, 15.09. http://www.wired.com/politics/security/magazine/15-09/ff_estonia. Accessed 02 Oct 2013

Falliere N, Murchu LO, Chien E (2011) W32.Stuxnet dossier. Symantec. http://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/w32_stuxnet_dossier.pdf. Accessed 02 Oct 2013

Fulghum D (2007) Why Syria's air defenses failed to detect Israelis. Aviation week, Oct 3. http://www.aviationweek.com. Accessed 02 Oct 2013

Gellman B, Nakashima E (2013) U.S. spy agencies mounted 231 offensive cyber-operations in 2011, documents show. Washington Post. http://www.washingtonpost.com/world/national-security/us-spy-agencies-mounted-231-offensive-cyber-operations-in-2011-documents-show/2013/08/30/d090a6ae-119e-11e3-b4cb-fd7ce041d814_story.html. Accessed 02 Oct 2013

Hoagland B (2013) Manning the next unmanned air force: developing RPA pilots of the future. Policy paper, center for 21st century security and intelligence at brookings

Joint Chiefs of Staff (2006) Information operations. Joint Publication 3-13

Joint Chiefs of Staff (2012) Information operations. Joint Publication 3-13
Peterson S (2011) Exclusive: Iran hijacked US drone, says Iranian engineer (Video), Christian Science Monitor, Dec 15. http://www.csmonitor.com/World/Middle-East/2011/1215/Exclusive-Iran-hijacked-US-drone-says-Iranian-engineer-Video. Accessed 02 Oct 2013
Rawnsley A (2011) Iran's alleged drone hack: tough, but possible. Wired, Dec 16. http://www.wired.com/dangerroom/2011/12/iran-drone-hack-gps/. Accessed 02 Oct 2013
Rattray G (2001) Strategic warfare in cyberspace. MIT Press, Cambridge
Sanger D (2012) Obama order sped up wave of cyberattacks against Iran. NYTimes.com, June 1. 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. Accessed 02 Oct 2013
Schmitt M (1999) Computer network attack and use of force in international law: thoughts on a normative framework. Columbia J Transnatl Law 37:885–937
Schmitt M (2002) Wired warfare: computer network attack and jus in bello. Int Rev Red Cross 84 (846):365–399
Schmitt M (ed) (2013) Tallinn manual on the international law applicable to cyber warfare. Cambridge University Press, Cambridge
United Nations (1945) Charter of the United Nations
Ziolkowski K (2012) Ius ad bellum in cyberspace—some thoughts on the "Schmitt-criteria" for use of force. In: Czosseck C, Ottis R, Ziolkowski K (eds) Proceedings of the 4th international conference on cyber conflict. NATO CCD COE Publications, Tallinn

Deception in the Cyber-World

William Hutchinson

Abstract Like any other communication medium, cyber-space has been used for deception since its inception. Originally a medium that gave immediate global range to deceptive messages, it also provided a medium to contradict any deceptive message sent. Of course, messages are not necessarily true or false but convey an opinion about reality that the recipient accepts or does not. The main concern of managers of this information has been that the messages have not been corrupted by those with malevolent intent. Hence, at its simplest level the integrity of the message (in the information security sense) is the primary objective. With more complex messages the use of propaganda techniques that attempt to influence opinions are of concern. A medium such as the public Internet with its low cost of entry and ubiquitous access is ideal for this and, because of its multi-media and interactive format, gives a much better success rate that 'conventional' media. Cyber-space over the last few years has rapidly entered a new phase with almost universal use of mobile online devices that many individuals and organisations are becoming increasingly dependent on. In this environment two other developments have significant implications for the practice of deception which changes the degree to which it changes the relationship of machines, deception and humans. These new factors are: the development of neuroscience and its associated technologies, and networked robotics. These are examined in this chapter and the consequences for deception at the level of individuals and large groups are examined.

1 Introduction: Setting the Scene

Deception is an intrinsic part of all life; it aids in survival. In society, humans practice deception in subtle and complex ways, and as they live in an information bubble where the data that comes into their system are imperfect. Thus, decisions

W. Hutchinson (✉)
Security Research Institute, Edith Cowan University, Joondalup, Australia
e-mail: w.hutchinson@ecu.edu.au

M. Lehto and P. Neittaanmäki (eds.), *Cyber Security: Analytics, Technology and Automation*, Intelligent Systems, Control and Automation: Science and Engineering 78, DOI 10.1007/978-3-319-18302-2_7

made are also imperfect. Humans, like all living things, have constrained sensors as well as processors that are programmed to 'expect the expected' and therefore can be deceived and can be fooled by illusions that they consciously know are incorrect but cannot reverse that illusion. Animals and plants use deception as a part of their survival (Forbes 2009). Even in higher animals such as humans, examples such as the 'hollow mask' illusion (Hill and Johnston 2007) illustrate this. This phenomenon is even more surprising as the person observing the images of the masks know it is an illusion but the brain does not allow the 'correct' image to be observed. Another surprising occurrence is the lack of seeing something in plain sight because of inattention and/or distraction—the so called 'Gorilla in Our Midst' phenomenon (Simons and Chabris 1999)—a physical entity is obviously there when inattention is removed. This illustrates that deception can occur with physical phenomena as well as purely conceptual. In the cyber-world, these can still take place but another dimension has been added—an artificially created virtual world. It is especially relevant to those tasked with observing screens such as those used with Closed Circuit Television systems.

This chapter will examine deception in this environment as well as the developing networked, robotic world.

2 Why Deception?

All of our input stimuli and data are limited by capability and opportunity. The data we receive are filtered by ourselves, our society and our limited knowledge. To influence anything, an influencer needs to make sure that society and the individuals within it have well understood information boundaries. Whilst this is anathema to libertarians, it really cannot be any other way. Humans are limited by themselves and their social setting. Hence, even if the information given is believed to be the 'truth' it can only be a well-meaning version of it. However, if an organisation or individual wants to 'sell' and idea for their own benefit it must involve deception of some type. Cynically, Machiavelli (1532/1983) noted this when he urged 'princes' to set aside 'normal' moralistic behaviour but at same time must appear to be merciful and humane. This principle has permeated many contemporary leaders. Deception only makes sense if there is a perceived truth—if this idea becomes flexible, as it often does in this fast moving relativist world, then there is no such thing as deception: it just becomes another version of reality. Deception in the modern world has gone passed just not telling the truth but has developed into an art where the information environment is shaped to conform to a desired perception.

Information and communication have always been fundamental to obtaining and holding on to power. Influence and deception are intrinsic factors in these processes. There is a link between power and influence; both have the ability to change behaviour. Influence can be political, authoritative, economic, financial or social. However, all need relationships between the social actors to develop with the population affected (Castells 2007).

Deception has a desired end product and the efforts can be designed for a limited operational objective or a long term strategic one. The theory of influence and also strategy can assist in understanding of deception. In influence theory Cragin and Gerwehr (2005, pp. 14–21) developed a model with three stages needed to enable influence and subsequently change behaviour. Starting from the simple, short term stage of 'compliance'; this requires a short term effort and uses tactics such as coercion and enticement. The second stage 'conformity' requires a medium term effort and uses tactics such as social and environmental manipulation. Lastly, comes 'conversion' which requires a long term effort and attempts to shape the worldview of people and limits the scope of their perception.

A similar three stage model regarding power was developed by Lukes (2005) where three distinct levels of power are introduced. The first view is where one party can make another party do something they would not have otherwise done for example, when an enemy surrenders. The second view of power includes the previous view but where a party creates or reinforces values and practices within a situation that limit the scope of options for another. Here the conflict may be overt or covert. The third view of power adds the idea of shaping perceptions and cognitions so various options are not even thought about; conflict can be latent in these situations. It is this latter situation that deception operations try to emulate. Each of these models of influence and power have a continuum from short term goals to much longer term strategic goals. As far as the target of deception goes a target population that is 'converted' and a level of power that matches with the third level of power would be the ultimate goal—if the effort was worth it.

3 The Cyber-World: Another Dimension

This chapter speculates about the ***consequences*** of social behaviour and the implementation of digital technology within the information bubble or perhaps it should be named an 'information shroud'. It is something that optimistic, technologically driven people seldom seem to do. The attitude is often: 'We can build these gadgets it is up to fate to see where they go'—a sort of *laissez faire* philosophy for technology and its consequences. However, technology does have implications for society and these should be critically examined.

The cyber world can be viewed from two perspectives:

- The cyber-world is a part of the normal world with its constraints. Here normal, natural laws apply, or
- The cyber-world is an entity in its own right with its own rules and reality.

It is the same dichotomy as that portrayed in the fictional movie the *Matrix* (1999) or in Gibson's (1984) novel *Neuromancer*: "The Matrix: a consensual hallucination experienced daily by billions of users: A world within a world…" In fact, the dichotomy between the real world and the perceived world is what the deceiver needs to exploit; there never is a completely 'true' world as it is

constructed dynamically within the brain, and in the contemporary world, by intelligent machines, as well. Whilst this might seem fanciful, the change from a basically computer phobic general population into one that has grasped the networked, intelligent, mobile devices with a speed and enthusiasm that is remarkable. The number of applications has grown exponentially and with this a dependency on these machines for communication, social life and business. The uptake and use of the associated software such as *Facebook* and *Twitter* has enabled a true matrix of people and organisations. With this has come another opportunity—to deceive mass or targeted audiences with equal ease. People's beliefs and perceptions are governed in large part by the data they can receive—when this data is received over an electronic network it can be manipulated. However, this is more than the 'normal' censorship of limiting information; it is more dynamic where pleasing and easy to understand information can be targeted to individuals or groups simply by choosing a subset of the population. Deception based on the standard principles of 'Hiding the Real' (dissimulation) and/or 'Showing the False' (simulation) can be sent by skilled multimedia savvy influencers. Of course, this applies to others with opposing messages and so a competition arises to mould peoples' perception of the truth. A fuller classification and examples of the types of deceptive tactics cab be found in Bell and Whaley (1991), and Bennett and Waltz (2007).

Deception on the Matrix (that is, the combination of networks) is relatively easy with the multimedia, psychological, and technical cyber-skills available. The joint development of networks, attractive easy to use computer devices, artificial intelligence, robotics and neuroscience have come together to produce a matrix that is fundamentally different from the rather 'quaint' Internet of a few years ago. In the past, deception on the Internet or just in cyberspace was much like the physical world that preceded it. For example, propaganda could be presented as it was before in newspapers or television—this new medium allowed deceptive practices to be performed to targeted audiences in a convincing manner. Because the audiences could now produce their own input through such avenues as social media, there was a general perception that the users were controlling the agenda.

The world is developing in such a way that all major information networks, data storage and presentation of information will be digital. Many organisations have no contingency plans for sending or storing their data that are not digitally based. The positive side of digital data is that it is so flexible and easy to change; however, this attribute makes it vulnerable unauthorized change. The flexibility of digital data makes the copying or changing of files whether images, videos, text, sound or instructions easy to perform. The implication of this is that the acceptance of any digital file should be taken with care. Deception relies on subtly changed perceptions and the digital environment is full of opportunities for the deceiver.

4 The Strategic Use of Deception in the New World

The strategic use of information and digital technology needs to be examined for it is only at this level that we make sense of the tactical consequences in the world as a whole. As such there is no intention to concentrate here on the use of Social Media which is really just a tool much like any other mass media tool but rather the more fundamental point of how influence works. But let us first look at what goes on at the moment.

Deliberate deception occurs within such activities as Influence and Psychological Operations, Public Affairs and Strategic Communications (by governments), and marketing and public relations (in industry), each of these have a slightly different meaning but really are tantamount to the same thing: getting the targeted population to think and behave in a specific way. The end goal is that these techniques should change behaviour not just opinion.

Until about a decade ago most of the strategies were executed via kinetic means or the mass media but tended to be controlled by a central 'power'. It was a 'one to many' process where the influence or deception was sold to a population and followed a specific plan to produce the desired effect. However, recently these tactics have not worked as well as the world is now truly electronically interconnected. However, the old techniques still work—early indications occurred in the 1991 Iraq War when Western perceptions were controlled perfectly by the Allied media machine. Despite relatively open and accessible communication networks the general public were given and acceptable version with little dissidence. This was achieved by jingoistic and incessant propaganda as well as restricted access to a variety of viewpoints. The major media outlets were 'taken over' much the same happened in the 2003 invasion of Iraq—it was 'deception in plain sight'.

The contemporary cyber-world consists of many people and machines connected wirelessly. If mobile digital devices were viruses we would think that we were living through a pandemic. After about 50 years of the general public being resistant, by and large, to the charms of computers, there is now a situation of an epidemic of people who are now 'in love' with their personal machines and their associated software. A person cannot get on a train, a bus or aircraft without most people lovingly and almost desperately clutching the intelligent device—"if I lost you—my life would be over". The development and expansion of mobile technologies have had a profound effect on the communication of information. This has enabled the ability to influence individuals directly and groups indirectly (via a virally spread story from targeted individuals to their peer network). The nature of mobile device usage is that it can be assumed that a target can be reached 24 hours a day. The features of this technology mean that users can be targeted at an individual user basis, or be chosen by their geographic position, in fact, when coupled with databases and search engines any criteria stored in those databases can be used if they can be linked to user identifications. At a service level this can mean that all mobile devices in a geographic area can be alerted to some impending peril such as

a bush fire, or at a more devious level, a group with an age range can be targeted to sell a product or an idea wherever they are.

Mobile devices have become ubiquitous in the last decade and are ideal for influence campaigns. Fogg (2003) explains the attractiveness of mobile digital devices and offer three metaphors to explain their attraction:

- The heart metaphor: we love them
- The wristwatch metaphor: they are always with us
- The magic wand metaphor: they have so much capability.

This attraction influences the owners of these networked devices—almost all of the information used by the owners come through these devices. A perfect target for deceivers: a trusted source of information and a real time target for whatever the deceiver needs to promote. This technical euphoria is relevant when neuro-systems are discussed later in the chapter.

The effect this has had on influence and deception strategists is that they now have populations with a definitive and self-selected series of target-audiences that are eager and willing to give up all their personnel information that can be sorted, sifted and allocated. Influence has never been easier. True there are some dissidents but they can be managed as the vast majority are happy taking in the given social norms. 'Give the people what they think they want'.

Frederick the Great said "The army is Prussia and Prussia is the army" such was the intrusion the army had made to Prussian society. The army was everywhere. In modern society could we say "The Internet is society and society is the Internet?" or perhaps "The intelligent mobile phone is society and society is the iPhone".

5 Old Style Cyber-Deception: New Style Cyber-Deception

Deception is most often associated with criminal or politico-military areas and cyber-deception is no different. In terms of security: attacks are either based on brute force or deception (or both). In network terms security can be physical but it also virtual. Attackers use deception either by the means of software or manipulating people directly. So the attack would be aimed at the software, hardware or wetware. Network attack methods include sniffer attacks, identity spoofing attacks and so on: all of these involve deception. Defending systems can also involve deception, for instance, by the use of honeynets (bogus networks that can be attacked as if they are 'real') or honeyfiles (files with attractive name but are really 'non-existent' to the live system so no-one should access them, however, if an attempt is made a flag is sent to security).

Deceptions on cyber-systems are basically targeted towards either people, software, data stores or system components. Much has been written about all of them and some are very technical and change at a rapid speed—just note the rate of amendments to virus checkers. The number of virus, spyware and firewall (attempted) attacks is enormous.

Humans have two systems of thought—one that is fast automatic and emotional, and a slower logical one that requires more cerebral effort. The first system does not have time to stop and think, it is a 'doer' not a thinker. The second is more ponderous but comes to more resilient conclusions. The speed of the contemporary world and cyber-systems seems to be pushing us to rely on the faster system. In fact, it actively encourages us to do that—computer simulation games encourage our children to make decisions immediately. Society expects managers to come up with immediate solutions. As a sideline this is why it is imperative to have enforceable contingency plans that are up to date and relevant. The Fukushima nuclear power station disaster was a supreme example of a predictable disaster that was overlooked because of time lines, financial constraints and a 'can-do' attitude when the impacts were enormous and brushed under the carpet because of expediency and self-interest (Willacy 2013).

"In the beginning…there was the Internet" which was a subset of the real world, is it now the world itself or at least its pervasive nervous system? Has it created a whole new world with its own life forms and environments within which mere humans, and the Information they get, can be controlled?

The concept of cyber-space being a subset of the real world is one that could be readily understood but the phenomenon of a semi-autonomous cyber-realm is not. The cyber-world does not have the same time frame as the 'real' world; nor does it have the same constraints of distance, location, boundaries, and jurisdictions. It is fluid, reduces barriers of entry and expression, obscures identity and avoids responsibility. However, this environment brings with it changes and challenges (a post-modern word for 'problems') that have not been faced or researched. But the power dimension has not really changed—the corporate/government world still dominates and has the potential to expand as the cultures in the world tend to amalgamate into one. Societies are virtually becoming consistent and thus prone to the same deceptions and influences. Political efforts in the cyber-domain support 'real' objectives in the physical domain—at least, this is the conventional model although in the single cyber-world domain efforts are self-referencing and reinforce themselves with no reference to the real world. Thus, deception becomes almost meaningless as reality is flexible and can be invented almost at will.

All this is developing in a world which is networked so power is exercised in different ways from the past. Traditionally power was the exercise of physical power where coercion ensured compliance, now it could be argued that power can be exercised more in the information realm where power ensures conformity. However, this still requires effort and is quite reversible. The battle for the human mind is largely executed in the processes of communication. Digital networks extend the ability to communicate into all domains of social life and can increasingly become customised. As a result the power relationships that control society are constantly changing—or at least appear to be. Power is the ability of a social actor to impose, in one way or another, their will over other social actors. Political legitimacy is often governed by the ability to effectively use communication media to impose their will. These media are not the holders of power in themselves but constitute the space where power is decided. The rise of self-communication with

(apparently) self-generated content has changed the variety of material on the networks—it is the job of the strategic influencer to ensure that the agenda in which these supposedly independently viewpoints are made is a suitable one.

The extent and reach of digital networks ensure the ability to extend into everyone's lives and although power relationships are conflictive, technology might make the conflict between rulers and the ruled disappear. This brings the next stage indoctrination, where cognitive power requires not only control of the information produced and received (traditionally through censorship and propaganda) but the thought processes themselves are manipulated. The end product of this should be 'conversion'.

Enter modern neuroscience as another player in the game of influence, deception and power. A significant event here was the successful use of an implant called *Braingate* which enabled a woman, who had lost the use of her limbs after a stroke, to control a robot arms by thought processes alone. Another innovation is the use of brain scans that have the potential to do such things as vet people for security clearance or assess their ability to be selected as a clearance diver. Already low cost helmets that enable the wearers to remotely control things such things as weapons and Unmanned Aerial Vehicles have been produced. How long a queue outside the new suppliers would there if these were made available to the public? The passing of weak electrical signals through the skull using trans-cranial direct current stimulation can change people's performance. Rao (2013) outlines the Brain Computer Interfaces (BCIs) being investigated. This technique has already been used to improve people's ability to recognise Improvised Explosive Devices (IEDs) and snipers. BCIs have been considered to connect brains directly to military equipment—who then controls the technology the person or the system? All this could produce a new world where everyone was connected; we are almost there with iPhones. With brain activity networked, the ability for manipulation is manifold.

Throughout history, powers of physical coercion, plus communication and information have been fundamental to the exercise of control; now direct control of thought can be added those who seek power over citizens' minds. Shaping minds is much longer lasting than physical coercion. Those that can control the physical, informational, and cognitive space will control all those wired to the system; and if iPhones are anything to go by the populace will be willing and eager to join. These non-coercive techniques are already being used to secure compliance, obedience, as well as behavioural and attitudinal change.

A good example of shaping minds is the use of military oriented electronic games that can realistically simulate battle scenes. The use of unreal scenarios can train people to gain the skills needed in a real world situation where the situation is no longer a simulation but look remarkably like it. This very sophisticated software can give the user the illusion of being in a real situation; so how does the user know when it is real or not (Stahl 2010). As the principles of neuroscience 'invade' the cyber-world the integration of these technologies can create a new world with new rules far in excess of the cyber/physical world divide mentioned in Sect. 2. As software and artificial intelligence advance the ability to actually change cognitive

processes online with the use of 'software agents' that start of as a problem solving tool for an individual but a co-evolution could take place where it would be difficult to separate who/what made a decision about a topic—the person or the software agent (Clark 2014). In other words, the cognitive part of the human is indistinguishable from the cyber. The fighting of a war as a game becomes the same for the controller in cyberspace whether it is real of not. Nevertheless, this illusion has a profound effect on the receivers of the outcomes—if it is purely in the cyber-sphere then it remains a game, if not it is war in the physical war with all the explicit suffering and psychological trauma involved (Singer 2009). This is deception having a real world effect.

6 Networked Robots

With the march of electronic networks has developed a series of computer driven devices that use modern technology that integrates artificial intelligence, computer technology, and neuroscience. A good example of this integration is the development of drones: robots that can independently work either within human control or outside it. Whilst these machines are not exclusively used for military purposes their rapid evolution and development has certainly been financed by militaries around the world. Singer (2009) gives a detailed account of the profound effect this will have on the military and international power balances as well civilian implications. The idea of mobile networks is almost at its ultimate when a person situated in America or Britain can control an aerial drone in Afghanistan for surveillance or attack—killing from around the globe in real time (Benjamin 2013).

Classical deception tactics can be used against drones as they can be against a conventional attack. Wesson and Humphries (2013) show how 'conventional' practices such as 'spoofing' or 'jamming' allow a deceiver to take control of an unmanned aerial vehicle (UAV); whilst Peacock and Johnstone (2013) show how a specific drone can be disabled. Both these tactics involve deceiving the machine: whilst more sophisticated approached such as falsifying the signals sent back to the 'pilot' require much more expertise. However, camouflage of the target can also be used. However, these are limited because of the use of various sensors on drones such as infra-red (thermal) and radar. The idea of hiding from a networked robot is becoming increasingly difficult because of the vast number of sensors that can be used. Intelligence gathering has moved on from purely sight to registering the full electro-magnetic spectrum, sonic and seismic emissions to behavioural observations such as the gait and walking patterns of individuals (Clark 2010). Of course kinetic means can also be used or more indirect deceptive tactics such as propaganda. The game changes slightly if the drone is autonomous as then the deception must be aimed at the vehicle itself that is, its sensors, its logic processes, or other functioning parts such as its navigation system or other internal systems. As Libicki (2007) outline there are three basic modes of attack that on the *physical layer* (that is, the manipulation of bits and the physical components that enables this), the

syntactic layer (those instructions and services that enable the system(s) to use information), and the *semantic layer* that enables the interpretation of the system's information—each one of these can be deceived.

7 Summary

The Internet and other networks are being driven by such developments as nanotechnology, software techniques, storage and communications technology, neuroscience, and robotics to develop a new cyber-world. The new mix of sensors, processors and a myriad of add-on peripheral are creating a 'new world'. There are some (for example, Barrat 2013) who believe that the development of neuroscience will put an end to the human era as machines take over and displace humans, or at least our cognitive functions. Others (such as Satel and Lilienfeld 2013) think this is overblown and, although caution should be use, it is just a part of the evolution of science and technology. Whatever, the outcome the use of deception will change significantly even if the basic principles and reason for doing it will not be. Deception can be altruistic but often it is just a desire to inflict a worldview on people for some sort of advantage. Initially, such things as the Internet promised an ability to choose which worldview you could believe. Of course, it was a good vehicle for deception as well. The possible future of thought transfer across the wireless network has profound implications with a new level of deception one that goes straight to a human cognition rather than via the senses.

Deception is manipulation, and the developing cyber-world and its ubiquitous networks could be an opportunity for deceivers to work at a global level. Whilst there are checks and balances in the system, it is also possible to withhold information in this world by steering populations to specific and popular sites. Deception as an art form is not dead just more sophisticated.

References

Barrat J (2013) Our final invention: artificial intelligence and the end of the human era. Thomas Dunne Books, New York

Bell JB, Whaley B (1991) Cheating and deception. Transaction Publishers, New Brunswick

Benjamin M (2013) Drone warfare: killing by remote control. Verso, London

Bennett M, Waltz E (2007) Counterdeception principles and applications for national security. Artech House, Norwood

Castells M (2007) Communication, power and counter-power in the network society. Int J Commun 1:238–266

Clark A (2014) Mindware: an introduction to the philosophy of cognitive science, 2nd edn. Oxford University Press, Oxford

Clark RM (2010) The technical collection of intelligence. CQ Press, Washington, DC

Cragin K, Gerwehr S (2005) Dissuading terror: strategic influence and the struggle against terrorism. RAND, Santa Monica

Fogg BJ (2003) Persuasive technology: using computers to change what we think and do. Morgan Kaufmann Publishers, San Francisco
Forbes P (2009) Dazzled and deceived: mimicry and camouflage. Yale University Press, New Haven
Gibson W (1984) Neuromancer. Harper Collins, London
Hill H, Johnston A (2007) The hollow-face illusion: object-specific knowledge, general assumptions or properties of the stimulus? Perception 36(2):199–223. doi:10.1068/p5523
Libicki MC (2007) Conquest in cyberspace: national security and information warfare. Cambridge University Press, New York
Lukes S (2005) Power: a radical view, 2nd edn. Palgrave MacMillan, Basingstoke
Machiavelli N (1532/1983) The prince (tran: Bull G). Penguin Books, London
Matrix (1999) Matrix, DVD, 131 minutes. Village Roadshow Films
Peacock M, Johnstone M (2013) Towards detection and control of civilian unmanned aerial vehicles. In: Valli C (ed) Proceedings of the 14th Australian information warfare conference (Perth, 2013). Edith Cowan University, Perth, pp 9–15
Rao RPN (2013) Brain-computer interfacing: an introduction. Cambridge University Press, New York
Satel S, Lilienfeld SO (2013) Brainwashed: the seductive appeal of mindless neuroscience. Basic Books, New York
Simons DJ, Chabris CF (1999) Gorillas in our midst: sustained inattentional blindness for dynamic events. Perception 28(9):1059–1074
Singer PW (2009) Wired for war: the robotics revolution and conflict in the 21st century. Penguin Books, New York
Stahl R (2010) Militainment Inc.: war, media and popular culture. Routledge, New York
Wesson K, Humphreys T (2013) Hacking drones. Sci Am 309(5):54–59
Willacy M (2013) Fukushima. MacMillan, Sydney

Legal Framework of Cyber Security

Eneken Tikk-Ringas

Abstract The subject of cyber security has come to blend the social, economic, political and military implications of uses of ICTs by different actors for diverse purposes. Despite the absence of a single dedicated legal framework to address the cyber domain, cyberspace and actions in it are addressed by numerous legal disciplines and normative instruments that, unfortunately, do not always provide immediate and convincing remedies to current cyber security issues. This chapter will outline the scope and core areas of cyber security from a legal perspective; introduce selected legal instruments and authorities addressing cyber security in its multiple facets; and some recent conclusions about the applicability and sufficiency of cyber security law to deal with emerging cyber security concerns. It will conclude with a discussion of some of the reasons behind the diminishing legal certainty in this field and the potential implications of declining authority of law in the context of cyber security.

1 Introduction

In 1996, Judge Frank Easterbrook shocked the legal community and the attendees of a conference held at the University of Chicago, by proposing that there was no more a (need for) "law of cyberspace" than there was a "Law of the Horse" (Easterbrook 1996). The law applicable to specialized endeavours, such as, in his conception, cyberspace, was to be found in general rules and any legal disciplines to be were to illuminate the entire law (Easterbrook 1996). A heated debate followed in legal communities (most notably Lessig 1999). Less than two decades later the dilemma seems to have taken care of itself—cyberspace is a subject that cannot be addressed without illuminating the entire law and although it has not

E. Tikk-Ringas (✉)
Baltic Defence College, Tartu, Estonia
e-mail: Eneken.Tikk-Ringas@baltdefcol.org

M. Lehto and P. Neittaanmäki (eds.), *Cyber Security: Analytics, Technology and Automation*, Intelligent Systems, Control and Automation: Science and Engineering 78, DOI 10.1007/978-3-319-18302-2_8

developed into a core legal area, addressing it as a distinct legal framework or discipline is useful to outline countless rules applicable to behaviour in cyberspace.

The subject of cyber security has come to not just include, but blend the social, economic, political and military implications of uses of ICTs by different actors for diverse purposes. After about 40 years of evolution, legal norms addressing behaviour in cyberspace and, consequently, the legal framework of cyber security, constitute an amalgam of regulatory authority and normative mechanisms.

The entirety of legal regimes and instruments applicable to cyber security is challenging to comprehend by any one legal expert and impossible to fit under any distinct area of law. There is no single legal instrument or authority to principally address this set of issues. Instead, cyber security related legal problems fall under several legal subjects, disciplines, and instruments. The absence of a single dedicated legal framework, the wide range and multiplicity of cyber security issues, and perhaps the legacy of early Internet law discourse, has led many to conclude that existing law is irrelevant or insufficient for resolving contemporary cyber security issues (e.g., Barlow 1996, in their respective fields also Hollis 2007). Recent cyber operations and modi operandi of both governments and non-state actors in cyberspace justify such scepticism.

On the other hand, there are authoritative conclusions and evidence about the applicability of existing law in and to cyberspace. In the past few years, several revisions of the existing legal frameworks have taken place, indicating flaws and inconsistencies in existing legal disciplines, yet concluding that several legal frameworks and instruments are applicable in cyberspace (e.g., Comprehensive study on cybercrime 2013; Schmitt 2013).[1]

With divided views as to the existence and sufficiency of legal tools to shape behaviour in cyberspace, at the core of the cyber security and law conversation is the divide between commitments and enactment, theory and practice, words and action. Indeed it is the same group of countries that, calling for cyber peace, are announcing and deploying military cyber capabilities. Despite verbal reaffirmations of human rights online and calls for wider ratification of the Council of Europe's Cyber Crime Convention, privacy and law enforcement against cyber crime are of questionable efficacy.

This chapter will explain why cyber security cannot be categorically viewed as a lawless domain, yet why the existing normative instruments do not always provide immediate and convincing remedies to current cyber security issues. After outlining the scope and core areas of cyber security from a legal perspective, the chapter will introduce the legal instruments and authorities that currently address cyber security in its multiple facets, accompanied by selected positions on the applicability of

[1]See also conclusion of the UN Human Rights Committee that human rights apply online as they do offline (Human rights document 2012) and the note of the UN Group of Governmental Experts that international law, in particular the UN Charter, is applicable and is essential to maintaining peace and stability and promoting an open, secure, peaceful and accessible ICT environment (Statement on consensus achieved by the UN Group of Governmental Experts on cyber issues, US State Press Release, June 7 2013).

existing legal instruments by relevant authorities. It will introduce some recent authoritative conclusions about the applicability and sufficiency of cyber security law and conclude with a discussion of some of the reasons behind the diminishing legal certainty in this field and the potential implications of declining authority of law in the context of cyber security.

2 Cyber Security from a Legal Perspective

The scope and focal point of the word pair 'cyber security' can be established by emphasizing the various applications of both terms. Each of the words allows wide interpretation.

'Cyber' as a combining form (cyber-) or as an adjective refers to 'relating to (use of) computers and networks' (Merriam-Webster Dictionary: Cyber 2013). Relevant technologies have undergone a mind-bending evolution since their emergence in the middle of the 20th century and in the current discourse form part of information and communication technologies (ICT).[2] Attacks against computers and the data they held emerged already in late 1950s and early 1960s, i.e. long before the Internet was in place. Cyber incidents, which subsequently raise questions regarding the appropriate balance between individual freedom and collective security, have been subject to legal debates for more than forty years. Teenage hacking arrived on the scene in the 1970s, and then grew in 1980s. Signals intelligence communities got accustomed to the Internet as it expanded across the world in 1980s. Early covert government operations exploiting ICTs date back to the same time. In 1990s hacking for political and social objectives emerged, blossoming in the 2000s. Economic gain factor came with the boom of .com and e-commerce in early 1990s and developed into a huge criminal industry by the beginning of the millennium. By the end of the last decade both cyber crime and politically motivated cyber incidents accounted for a substantial share of all cyber attacks. Now, as Farwell and Rohozinski conclude after an analysis of Stuxnet and operation Olympic Games, the international community has entered the era of state-on-state cyber conflict, potentially war, both with state and non-state actors involved (Farwell and Rohozinski 2011; Farwell and Rohozinski 2012).

Security as the state of being protected or safe from harm or things done to make people or places safe (Merriam-Webster Dictionary: Security 2013) has also expanded, both as a concern and condition, with the emergence of new threats and threat actors in cyberspace. The latest expansion of the concept that too many associates with technical safety and routine information assurance, has occurred in the field of national security. Denial-of-service of government functions, sabotage

[2]According to (Protection of mission critical functions to achieve trusted systems and networks, DoD instruction 5200.44. Department of defense of United States of America 2012) ICT includes all categories of ubiquitous technology used for the gathering, storing, transmitting, retrieving, or processing of information.

of industrial processes for political reasons and launching cyber militias against competitors all carry a strong national security undertone.

Added to rather broad scope of bot terms is the ambiguity they carry in the political context. What constitutes a 'national security' issue is far from agreed among the international community. In fact, the degree of imposing national jurisdiction on persons, objects and events is often subject to tension and disagreement between governments. It is essential to note that even with the event of computer security rising to the interest threshold of national security not all concerns related to uses of ICTs become a strategic issue per se. Wolfers emphasizes that when specific policy formulas gain popularity (as 'cyber security' is today) one must carefully scrutinize such concepts to avoid permitting everyone to label whatever policy he favors with an attractive name (Wolfers 1952). At the same time purely technical and tactical approaches to cyber security may not suffice in the context of cross-border and cross-context threats.

The above connotations and their variations coexist, interrelate and feed into the general discussion of 'cyber security'. Curiously, most cyber threat trends have survived over time and created unique symbiosis. These days, crime syndicates are developing and deploying tools that prove useful for governments to use as part of their cyber arsenal. Script kiddies with patriotic attitude form militias that successfully fly below the international legal framework radar. Business information system guardians, however, turn out to be a useful link in the chain of cyber defense.

Rating cyber risk among top three business risks in 2013, Lloyds notes that the perception of this risk is evolving from financial crime to political and ideological attacks (Lloyd's risk index 2013). In other terms, boundaries between personally, economically, ideologically, politically motivated and toned cyber incidents are blurring. Similar conclusions can be made when looking at incidents like taking offline the business operations of Saudi Aramco (see more Bronk and Tikk-Ringas 2013) or denial-of-service against large US banks *earlier this year*.

Takedown of the Interpol, CIA and Boeing websites, the suspension of alternative currency Bitcoin's trading floor, the mass theft of passwords from professional networking site LinkedIn, large-scale exfiltration of defense secrets from governments and defense contractors, shutdown of the largest oil company's business operations and secretive online surveillance programs are not predicated, executed or remedied by the same set of tools, methods and actors.

The words 'cyber' and 'security' combined serve as an umbrella to a number of disciplines from technical data and information assurance to resilience and reliability of business and industrial processes to (military) operations security. In sum, therefore, cyber security constitutes a comprehensive and complex area of political, industrial, technical, and operational concerns involving and combining public and private, civil and military ambitions and engagement and therefore being covered by very different legal ramifications, both national and international.

3 Instruments and Areas of Law Addressing Cyber Security

Attempts to create a structured thinking and understanding of legal aspects of cyber security require patience and creativity. Also, there are many ways to structure one's thinking about cyber security and law.

One can determine the legal areas and concepts, which in sum comprise the legal framework of cyber security, from cyber concept updates. Following a recent conceptualization of cyberspace, this global domain consists of interdependent network of information technology infrastructures, including the Internet, telecommunications networks, computer systems and embedded processors and controllers (Cyberspace operations 2010). All these components are subject to numerous legal ramifications. Internet law alone comprises legal issues like e-commerce, torts, consumer protection, privacy, cybercrime, content regulation and intellectual property protection (Rustad 2009). ICT infrastructure and telecommunications networks are subject to several legal regimes such as the law pertaining to international telecommunications, satellites and undersea cables. Different areas of content regulation have emerged over time (personal data protection, xenophobic content, child pornography). Somewhat loosely, ICT components are subject to product liability and Intellectual Property protection regimes.

From the perspective of public international law, cyber security issues fall primarily under seven legal areas that correspond to the main fields of cyberspace—infrastructure, access and content regulation, economic utility of ICTs and various malicious acts (crime, national security threat and use of force) in and against cyberspace. Individual categorizations of applicable international legal areas vary (see, e.g., Ziolkowski 2013).

Issues like freedom of information and the right to privacy are incorporated under the human rights law. The UN Human Rights Council has confirmed that human rights apply in the same way both online and offline (Human rights document 2012). Universal Declaration of Human Rights (Universal Declaration of Human Rights 1948), the European Convention on Human Rights (European convention on human rights 1950) and many other regional instruments provide for the right to privacy and access to information. Data exchange regulation is relatively well established in the EU countries. Directive 95/46/EC (Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995) (Data Protection Directive) serves as the basis for personal data protection legal acts in nearly 30 advanced information societies. Under Article 10 of the European Convention of Human Rights a general right exists to receive and impart information. According to Article 4(2) of the Directive on Privacy and Electronic Communications (Directive 2002), in case of a particular risk of a breach of the security of the network, the provider of a publicly available electronic communications service must inform the subscribers concerning such risk.

Criminal law delivers mechanisms for cross-border investigations and prosecution, extradition and enforcement of court decisions. Such arrangements are

based on both multilateral and bilateral agreements. Despite low ratification rates, the Council of Europe Cyber Crime Convention (Convention on cybercrime 2001) provides a legal basis to investigate and prosecute cyber crime. Under Article 23 of the Cyber Crime Convention, the Parties are requested to co-operate with each other through the application of relevant international instruments on international co-operation in criminal matters, arrangements agreed on the basis of uniform or reciprocal legislation, and domestic laws, to the widest extent possible for the purposes of investigations or proceedings concerning criminal offences related to computer systems and data, or for the collection of evidence in electronic form of a criminal offence (Branscomb 1986). There are several other instruments of the Council of Europe addressing the issue of cross-border legal/judicial cooperation in criminal matters. The most relevant in the context of pre-trial investigation of cybercrime is the European Convention on Mutual Assistance in Criminal Matters with its additional protocols. The Estonian request for assistance to the Russian Federation in 2007 was based on the Mutual Legal Assistance Agreement. Experts of the Council of Europe have concluded that existing international conventions and other instruments promoting the harmonization of both national substantive and procedural law and also international cooperation are applicable to these misuses of the Internet for terrorist purposes.

The law of telecommunications outlines the legal regime for operating and maintaining the international telecommunications infrastructure. Special provisions related to undersea cables derive from the law of the sea and certain aspects of satellite communications fall under space law. Scholarly works explain how the provisions of International Telecommunication Union (ITU) treaties apply to cyber security (Rutkowski 2011). In 1986, one of the first legal studies on global communications networks was released by the American Bar Association (Branscomb 1986). The authors concluded that the right of the public to correspond by means of international services of public correspondence is established both under the ITU Convention and in treaties of friendship, commerce and navigation. It was also noted that communications facilities within a country would not be covered by such agreements and that the right to transmit data to a particular country would depend on the approach to data transmission by that country, (Branscomb 1986, p. 19). Rutkowski concludes that general network security obligations derive from the existing ITU Convention (Rutkowski 2011).

Several instruments in the field of international trade and commerce deal with cyberspace as an economic environment, addressing issues such as intellectual property protection, taxes, tariffs and consumer protection.

The Law of Armed Conflict (LOAC) comprises jus ad bellum and jus in bello and thereby dictates the legal premises and limitations for waging war. The right to self-defence in response to cyber incidents under the law of armed conflict needs to be assessed in the light of the broader concept of the use of cyber force extensively discussed in legal theory before and after the rise of cyberspace (see, e.g., Brownlie 1963, also Benatar 2009). With reference to the Nicaragua case, Schmitt concludes that the use of force line must lie somewhere between economic coercion and the use of armed force, (Schmitt 1998–1999, p. 914). Schmitt further offers criteria for

delimiting economic and political coercion from the use of armed force and accentuates that it is not any use of force, but an "armed attack" which gives a state the right to respond in self-defence, (Schmitt 1998–1999, p. 920). The Tallinn Manual puts the Law of Armed Conflict into the context of contemporary cyber operations, offering a 'cyber' restatement of hard security law.

All these areas of law have got a decent amount of coverage in both academic and practitioner communities, potentially offering solutions to acute cyber security concerns. Herrera observes that the "placeless-ness" of the Internet is not absolute and trusted computing elements, like certificates, firewalls, digital rights management or static IP addresses, turn segments of the Internet into a tightly controlled and surveyed space (Herrera 2005). Cox explains that many governments have successfully regulated cyberspace and that nations are increasingly acting in concert to deal with the seemingly borderless nature of cyberspace (Cox 2002). Codified in the International Law Commission's Articles on State Responsibility (hereinafter also "the Articles"), the legal foundations of state responsibility define a wrongful act or omission to a state to be when (a) attributable to the state under international law; and (b) constituting a breach of an international obligation of that state (International Law Commission 2001).

Less attention has been paid to national security law in the context of international cyber security. This field is made up from the outside margins of legal provisions. For instance, article 27 (4) of the Budapest Convention on Cyber Crime entitles a Party to refuse assistance under the Convention if the request concerns an offence which the requested Party considers a political offence or an offence connected with a political offence, or if the requested Party considers that execution of the request is likely to prejudice its sovereignty, security, *ordre public* or other essential interests. This means that at least from a purely legal perspective a state can choose not to apply the Convention in case such a decision supports its national interests. Article 34 of the ITU Convention allows Member States to cut off, in accordance with their national law, telecommunications which may appear dangerous to the security of the State or contrary to its laws, to public order or to decency. The Council of Europe Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (Article 9(2)) allows derogation from the provisions of this convention as provided for by the law of the Party when it constitutes a necessary measure in a democratic society in the interests of protecting State security, public safety, the monetary interests of the State or the suppression of criminal offences or is necessary for protecting the data subject or the rights and freedoms of others. Indeed, in the light of recent cyber operations several countries seem to be intensively exploiting these legal premises.

Another field yet to be authoritatively restated is that of state responsibility. Responsibility constitutes the core element of the legal order and reaches through all areas and levels of law. In international law, responsibility is considered the 'necessary corollary' of the equality of states, one of fundamental principles of international law. With reference to SS Wimbledon, (S.S. Wimbledon (U.K. v. Japan) 1923, pp. 4, 25), Pellet notes that the possibility for a State to incur responsibility 'is an attribute of State sovereignty' just as the right of entering into

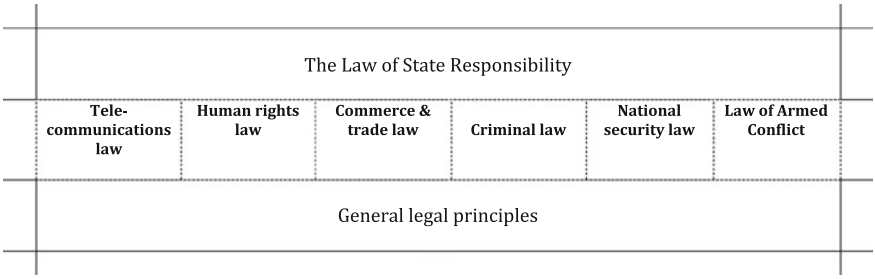

Fig. 1 Scheme of public international law applicable to cyber security

international engagements, (Pellet 2010, p. 4). A situation where a State would enjoy immunity from responsibility would be inconceivable under international law due to the equality of States as this prerogative of sovereignty extends to all sovereigns due to their equality (Declaration on principles of international law concerning friendly relations and cooperation among States in accordance with the Charter of the United Nations. Resolution 2625), proscribing the duty to respect the personality, territorial integrity and political independence of other states and the right to freely choose and develop its political, social, economic and cultural systems. The law of State Responsibility comprises the principles governing when and how a state is held responsible for a breach of an international obligation. Unlike other, substantive areas of international law, the law of state responsibility determines, in general, when an obligation has been breached and what the legal consequences of that violation are. Because of this generality, the rules of state responsibility can be applied in the context of any primary international obligations. ILC Draft Articles of State Responsibility establish the conditions for certain state behavior to qualify as internationally wrongful act; the circumstances under which acts of officials, private individuals and other entities are attributable to a state, and the consequences of liability (Fig. 1).

While many of the above areas of law have been developed with cyber security issues in mind or recently updated in this context, there has not been much discussion about the applicability of some non-cyber-specific legal concepts in and to cyberspace, with the notable exception of the Law of Armed Conflict. Especially, the applicability of general principles of law, such as for instance good faith and humanity; and selected legal concepts (neutrality, friendly relations, non-intervention, and sovereignty) deserve more attention by legal community in the light of upcoming international law talks at the UN.

Table 1 Legally binding instruments of the European Union and the Council of Europe addressing aspects of cyber security

1981	COE	Convention nr 108 The Protection of Individuals with Regard to Automatic Processing of Personal Data
1991	EU	Dir 91/250/EEC The Legal Protection of Computer Programs
1995	EU	Dir 95/46/EC Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data
	EU	Res 96/C 329/01 on the Lawful Interception of Telecommunication
	EU	Dir 95/144/EC Common Information Technology Security Evaluation Criteria
1996	EU	Dir 96/9/EC The Legal Protection of Databases
1997	EU	Dir 97/7/EC The Protection of Consumers in Respect of Distance Contracts
1998	EU	Dir 98/84/EC The Legal Protection of Services Based on, or Consisting of, Conditional Access
1999	EU	Dir 1999/93/EC Framework for Electronic Signatures
2000	EU	Dir 2000/31/EC Electronic Commerce
2001	COE	Con No 185 Cybercrime
2002	EU	Dir 2002/58/EC Privacy and Electronic Communications
	EU	Dir 2002/19/EC Access Directive
	EU	Dir 2002/20/EC Authorisation Directive
	EU	Dir 2002/21/EC Framework Directive
	EU	Dir 2002/22/EC Universal Service Directive
2003	EU	Dir 2003/98/EC Re-Use of Public Sector Information
	COE	Add Prot No 189 (C 185) Acts of a Racist and Xenophobic Nature Committed Through Computer System
2005	EU	Dec 2005/222/JHA Attacks Against Information Systems
	EU	Dir 2006/24/EC Data Retention
2009	EU	Dir 2009/140/EC Electronic Communications, Networks and Service

A critical look at legal instruments reveals that the majority of legally binding obligations derive from the European Union and Council of Europe instruments.[3] This turns many existing legal provisions into pointers for a wider international norms dialogue. One can also observe that in most other regions the focal point of cyber security regulation has been cyber crime and instead of binding instruments a wide-spread practice of model laws and regional framework conventions has emerged.[4] In the light of cyber incidents widely covered in other chapters it is evident that legal action is rarely invoked against actors when incidents carry a political undertone and that law does not prove a sufficient tool in the hands of law enforcement in the absence of air-gapped national legal frameworks and improved

[3]See Table 1 for a list of EU regulation in the field of cyber security.

[4]See Table 2 for selected regional approaches to cyber security.

Table 2 Selected Regional and International Initiatives to Reform Cyber Security Law

1996	UNCITRAL	Model Law on Electronic Commerce
2000	ASEAN	e-ASEAN Framework Agreement
2001	CIS	Agreement on Cooperation in Combating Offences related to Computer Information
2002	Commonwealth	Commonwealth Model Law on Electronic Evidence
	Commonwealth	Commonwealth Model Law on Computer and Computer Related Crimes
2004	OAS	Strategy to Combat Threats to Cybersecurity
	League of Arab States	Model Law on Combating Information Technology Offences
2005	APEC	Cybersecurity Strategy to ensure Trusted, Secure and Sustainable Online Environment
2007	ECOWAS	Harmonization of the Legal Framework Governing ICTs in West African States
2008		East African Community Draft Legal Framework for Cyberlaws
2009	SCO	Agreement on Cooperation in the Field of International Information Security
	ECOWAS	Draft Directive on Fighting Cybercrime
2010	League of Arab States	Convention on Combating Information Technology Offences
2011	COMESA	Cybersecurity Draft Model Bill
		East African Community Draft Legal Framework for Cyberlaws
	ECOWAS	Directive on Fighting Cyber Crime within ECOWAS
2012	African Union	Convention on the Establishment of a Legal Framework Conducive to Cybersecurity in Africa (Draft)
	SADC	Model Law on Computer Crime and Cybercrime

capability of attribution, currently still the prerogative of the most capable cyber powers.

On national level, cyber security is broken down into even more legal areas. Sieber lists the first six waves of law reforms under western legal systems: privacy protection (mid-1970s); computer-related economic crime repression (early 1980s); protection of intellectual property (mid-1980s); illegal and harmful contents (1980s); reforms in the field of criminal procedural law (late 1980s); and security law[5] (1990s) (Sieber 1998, pp. 26–31). Each of these waves has created new legal concepts that are difficult to bring under uniform common nominators due to differences in legal systems and structures. The US Congressional Research Service has concluded that more than 50 national statutes address various aspects of cyber security either directly or indirectly, including areas such as information security

[5]Sieber refers to security as minimum obligations for security measures in the general public interest, including prohibitions and export controls for cryptography, (Sieber 1986, p. 5).

management, protection of critical infrastructure, information sharing, protection of privacy, cybercrime and the cyber security workforce (Fisher 2013).

Computer and network security is subject to numerous standards that combined with legal principles such as duty of care, due diligence and legal frameworks for business and government operations constitute established and enforceable thresholds. Under the EU Directive 95/45/EC the controller[6] must implement appropriate technical and organizational measures to protect personal data against accidental or unlawful destruction or accidental loss, alteration, unauthorized disclosure or access, in particular where the processing involves the transmission of data over a network, and against all other unlawful forms of processing (Article 17 (1)). Under the Directive 2000/31/EC, the standard for hosting provider's liability is provided. In case of actual knowledge of illegal activity and awareness of facts and circumstances from which the illegal activity or information is apparent, the information service provider is obliged to act to remove or to disable access to such information (Article 14(1)). The Council of Europe's Cyber Crime Convention sets an international standard for cyber crime obliging Parties to adopt such legislative and other measures as may be necessary to establish as criminal offences under domestic law, when committed intentionally, the access to the whole or any part of a computer system without right (Article 2).

Besides numerous legal instruments, norms pertaining to cyber security are abundant in international policy instruments. As Warner notes, during the half-century that it has taken the cyber issue to mature, numerous policies, standards, and doctrines have emerged, which consequently shaped U.S. responses to the broader emergence of the cyber security issue in the late 1990s (Warner 2012). The same is true for Europe and several other countries. Over the past two to three decades, international organizations have adopted more than 150 policy statements dealing with aspects of uses of ICTs ranging from privacy and consumer protection to duties of telecommunication infrastructure providers, to the criminalization of computer and network related offenses. State practices regarding the implementation of these policies has evolved and matured over time, and they have established and deployed numerous national and collective strategies and doctrines (see, e.g., Strategies and policies. NATO Cooperative Cyber Defense Centre of Excellence 2015).

In addition to treaties and other legal instruments, case law of the European Court of Human Rights (ECHR) and the European Court of Justice (ECJ) addresses issues like privacy, freedom of information, and criminal proceedings in the context of information technology (see Tikk and Talihärm 2010). Studies on state practice and customary law have emerged in the cyber field. As Brown observes, even if one disregards the Siberian pipeline incident and considers Moonlight Maze (1998–2001) the first major state-on-state cyber incident, one can assess the past dozen

[6]Article 2(d) in (Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995) defines controller as the natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data.

years of general practice to determine what constitutes customary law in cyberspace. Incidents which occurred during this period have set precedents for what states consider acceptable cyber behavior. What is remarkable is the lack of protest from nations whose systems have been degraded in some way by malicious cyber activity (Brown and Poellet 2012, p. 132). Polanski has discussed international regulation of the Internet from the perspective of customary law (Polanski 2007). Demidov also suggests that a closer look on existing legal principles and instruments and their interpretation by States to assess which State practice is sufficient to qualify and be accepted as customary law (Demidov 2012).

Finally, one can approach the legal framework of cyber security by looking at the legal ramifications related to numerous open questions. In the light of increasing amount and sophistication of cyber attacks, a general obligation to secure networks, services, data and content is imminent. Elements of such obligation can be found in the law of telecommunications, regulation of ISP activities, principles of duty of care and due diligence. Some national legal developments indicate a separate standard for an obligation to secure and protect for critical infrastructure providers. It is yet to be discussed what role the law of product liability and consumer protection will play in the context of cyber security. It is thus fair to ask whether existing agreements and practices in the field of telecommunication services and data protection establish a duty to secure information systems and services under their jurisdiction. One could argue that a duty to cooperate and provide mutual assistance in the event of cyber attacks exists under collective security treaties; mutual criminal assistance agreements; and general principles of good faith and friendly relations.

With the cross-border nature of most incidents a general obligation to assist is essential. It has been suggested that collective defence provisions of the North Atlantic Treaty would apply in case of a cyber attack rising to the threshold of an armed attack. Yet there are other 24/7 networks of assistance, such as under the Budapest Convention. An upcoming set of obligations is related with reporting an incident and possibly sharing information about threats and vulnerabilities. Obligations to prosecute and assist with investigations are reflected in several treaties yet may require revisiting in the context of national security exceptions and the Russian refusal to cooperate with Estonia after the 2007 attacks. In other words, practice does not offer good answers to what satisfies an obligation to cooperate (time, quality—how they apply in other areas—what is reasonable?) and fails to furnish the criteria of 'unable and unwilling' standard in context.

In the light of the concerns tabled by Russia and China it is essential to understand to what extent and under which circumstances different entities (state, ISP, victim) have a right to restrict access to and from their networks, data and services? Cases of Syria and Egypt shutting down Internet on their territory indicate state practice in this field, as does the case of Estonia in 2007.

A well-established area of law in need of restatement and better enforcement is that of privacy. Under the EU regulation, individuals have access to and oversight

of their data as processed by gov't and private sector. The same right does not apply internationally and data collected for one purpose can be used for other purposes? Privacy law feeds directly into the requirement and justification for surveillance and needs to be balanced with the legal concepts of data retention and requirements of national security. Apparently, Intellectual Property law needs to undergo a similar analysis in the absence of effective mechanisms to curb economic and industrial espionage.

Involved jurisdictions, implemented tools and methods, underlying motivations and perpetrating actors all define the legal landscape for each unique incident. It is necessary to analyse the essence and context of an incident to conclude the (availability of) an appropriate response and action. While from a purely technical perspective, cyber incidents may involve similar targets (government websites), methods (defacement, denial of service attacks) and even context (political tensions between two nations), the legal remedies to such incidents may drastically vary.

4 Implications of Diminishing Legal Certainty

Scholarly work and current practices do not say much about how international law will resolve a specific problem or support a specific goal; in fact, law is often reactionary rather than an anticipatory in its development. In a complex and multi-faceted area like cyber security, where the respective national capabilities, concerns, and priorities for various states differ dramatically, one must first examine the concrete problem that a given law is expected to resolve for the international community, for an alliance, or a single actor. This challenge highlights two additional points. First, it is difficult to provide a legal assessment of a case without facts and [by failing to follow?] appropriate procedures. From an analytical perspective, even high-profile cyber incidents leave legal scholars with more questions than answers. Second, problem-solving is further complicated by the fact that any one legal area or concept faces difficulties accommodating all types of cyber issues, such as accommodating the multi-stakeholder model of Internet governance with exclusive state control over the Internet, or applying cyber warfare to the law of armed conflict while effectively ruling out cyber warfare in the first place. Multiple areas of legal expertise and hardly aligning levels of authority involved complicate legal action. Finally, clever approaches by malicious actors exploit flaws in existing legal frameworks, by combining jurisdictions with weak legal structures with problems of attribution and the lack of law enforcement resources.

Scepticism about the relevance of existing legal framework to resolve cyber security challenges has grown over the years both in the absence of an overarching legal framework but also due to rather low rate of successful remediation of the constantly increasing amount of cyber incidents. Lack of decisive legal action tends to extend over several subsets of cyber security: despite wide-ranging privacy guarantees in existing legal instruments individuals complain loss of privacy online; against millions of 'cyber attacks' daily the statistics of complaints and claimed

losses only operates with hundreds of thousands; and most notably, high profile cyber-attacks with political undertone result in no legal action or are blown over by political rhetoric.

Some of the differences between statistical and real numbers of cyber incidents can be blamed on the inconsistency of definitions, non-uniform legal categorization and inconsistent statistics. Yet the comprehensive study on cybercrime conducted by the UNODC in 2013 observes that the proportion of actual cybercrime victimization reported to the police ranges upwards from just 1 %, (Comprehensive study on cybercrime 2013 p. 117, 119). Symantec has concluded that 80 % of individual victims of core cybercrime acts do not report the crime to the police (Norton cybercrime report 2012). Between 50 and 100 % of cybercrime acts encountered by the police involve a transnational element, (Comprehensive study on cybercrime 2013, p. 118). Although cyber incidents occur globally and are reported locally, the low ratio of reported cyber incidents talks of challenges in law enforcement and legal cooperation.

There are alarming signs of emerging *laissez-faire* attitude towards cyber security law. Taking justice into their own hands, victimized entities and individuals often turn to self-help or self-remediation, thereby further undermining the legal certainty in cyber domain. The issue of 'hack-back' illustrates the choices often made by the private sector to trust their own operational expertise over formal, lengthy and revealing legal procedures. Similarly, calculating losses of identity theft and credit card fraud into business risk and remedying customers at the expense of profits has resulted in the trend of non-legal mitigation of essential segments of cyber crime and low requirement of attention to and remediation of the incident by customers.

Dealing with most cyber incidents today cannot be regarded as one-off engagement. An intrusion into a telecommunication provider's network can be a single incident, but also part of a campaign or operation that only becomes evident when comparing notes with other providers, CERTs or victims at a later stage. Before an incident reaches the threshold of a national security relevant or even armed attack, it could be handled and monitored by a number of institutions. The facts gathered about the origin, duration, technicalities, etc. of the incident may play a significant role in determining a proportionate response in case the attack eventually involves national security authorities and the military. It is therefore essential that cases are not rested at the borders of one's business operations or previous experience.

Operations in cyberspace offer no better observations. Instead, they highlight growing tensions between executive and parliamentary authorities around the extent and essence of cyberspace operations or the expectation of privacy are illustrative of

a wider trend of diminishing value of the rule of law. Analyzing operation Olympic Games (see Farwell and Rohozinski 2012)[7], an alleged cyber operation of the US and Israel to retard Iran's progress in developing nuclear weapons, Farwell and Rohozinski suggested that by devising the operation, White House exceeded its jurisdiction either under the US Constitution, which reserves to Congress the right to declare war, or under the War Powers Resolution (Farwell and Rohozinski 2012).

Given the divided and diffuse focus of international organisations, the primary responsibility for addressing cyber security threats and responses currently lies with national governments. Without a thorough understanding of threats and remedies available on the national level any debate on the international level would lack focus. Only after national homework the areas for common concern can be referred to in detail sufficient for constructively debating additional remedies needed.

After a thorough review of existing instruments, it may appear that law already contains enforceable obligations for infrastructure and service providers to secure their services. This revelation may change the calculus of those who previously concluded that only changes of national law would clarify the scope of obligations for some stakeholders.

5 Conclusion

A consensus is emerging regarding the applicability of international law in and to cyberspace. There is plenty of evidence of existing international treaty law to affect cyber affairs between state and non-state actors. The United States, Russia, and China have disagreed strongly on a number of key issues, including how to define and frame cyber issues and the appropriate model of Internet governance, and the admissibility of cyber warfare. The UN GGE report in 2010 concluded that 'existing agreements include norms relevant to the use of ICTs by States' (Report of the group of governmental experts on developments in the field of information and telecommunications in the context of international security. A/65/201, General Assembly, United Nations 2010) and reiterated in the 2013 report that international law applies to cyber conflict (Statement on consensus achieved by the UN Group of Governmental Experts on cyber issues, US State Press Release, June 7 2013). The newly published Tallinn Manual on International Law Applicable to Cyber Warfare

[7]Stuxnet, the best known component of the operation Olympic Games, was the first alleged identified instance of weaponised computer code or malware targeting the Iranian uranium enrichment program and widely regarded as a 'use of force'. Two other targeted computer viruses for espionage surfaced shortly after: Duqu in September 2011 and Flame in May 2012. Allegedly 20 times more complex than Stuxnet, Flame affected computers in Lebanon, the United Arab Emirates, the West Bank and Iran. It is said to have gathered intelligence by logging keyboard strokes, recording conversations by activating microphones, and taking screen shots. At Iran's oil ministry and oil-export terminal, the virus also erased information on hard discs while gathering information. Many attribute it to the United States and Israel. These allegations remained unconfirmed by either government.

explains how one can apply norms of International Law of Armed Conflict and the International humanitarian Law to cyber operations (Schmitt 2013).

Areas of potentially applicable law have mostly crystallized by now. In light of both increasingly sophisticated and widespread cyber attacks and intensifying interstate confrontation, many non-hypothetical questions and interest statements relevant to international law have been raised on an international level.

The developments and the amount cyber security related norms adopted under a variety of regional and international mandates leave little doubt that negotiations in the field of international law and cyberspace are on the rise. With the question now being framed 'how' instead of 'if', states will need to discuss and agree which existing norms are relevant to cyber security and what is the extent of their applicability to current cyber affairs.

This is not an easy task to accomplish. Legally binding norms relevant to cyber security are found in different legal areas reaching from human rights and telecommunications to e-commerce, cyber crime, national security law and the Law of Armed Conflict. The geopolitical reach of potentially applicable international norms is often less than global and national interpretations of distinct norms vary considerably as do scholarly opinions. Political guidance about cyber security is scattered around dozens and dozens of instruments dating back to 1980s. Yet many issues and solutions raised over time are still directly relevant to those with which we are struggling today. Among other caveats, legal scholars have noted early on the basic political dichotomy between the United States and the rest of the world over the ownership and control of communications systems; the dangers resulting from the concepts of territoriality and statehood from the cross-border nature of global communications; and the difficulty that existing laws have in effectively responding to communication models that do not discriminate between types of communication services (Branscomb 1986).

One of the challenges comes from the fact that most existing norms require a thorough review and possibly restatement in the context of contemporary cyber security in order to 'qualify' as solutions to widespread cyber contestation. The Tallinn Manual undertakes to put the Law of Armed Conflict into the context of contemporary cyber operations, thus offering a 'cyber' restatement of hard security law. Internationally recognized duties of non-intervention and prohibition of use of force, although globally applicable, require a more coherent understanding of what they mean in cyber context. In turn, 'cyber law' comprising everything from privacy and copyright to computer crime, needs to undergo a review from (hard) security perspective. Only after re-weighing individual rights and national security considerations we will be able to apply privacy norms within the paradigm of national security and exercise national security with business and civil interests in mind.

Numerous cyber security norms are regional and therefore lack uniform implementation and enforcement, potentially calling for codification and extension of their applicability. In some fields, especially telecommunications, competition of regional and international norms may occur due to the lack of common understanding of 'cyber' and difference of national interests and strategies. Also, different

schools of legal thought and practice lead to diverse national norms, occasionally even between like-minded countries.

Reclaiming legal remediation initiative in the field of cyber security will be challenging for international organizations and governments alike. Yet a failure to do so is almost not an option.

References

Additional protocol to the European convention on mutual assistance in criminal matters. Council of Europe, Strasbourg, signed on 17 March 1978. http://www.conventions.coe.int/Treaty/EN/Treaties/HTML/099.htm

Agreement of the Republic of Estonia and the Russian Federation on legal assistance and legal co-operation in civil, family and criminal matters. Signed on 26 January 1993, entered into force on 19 March 1995. Riigi Teataja (State Gazette) II 1993, 16, 27; RT II 2002, 14, 58

Barlow JP (1996) A declaration of the independence in cyberspace. http://homes.eff.org/~barlow/Declaration-Final.html Accessed 5 May 2011

Benatar M (2009) The use of cyber force: Need for legal justification? Goettingen J Int Law 1 (3):375–396

Branscomb AW (ed) (1986) Toward a law of global communications networks. Longman, New York

Bronk C, Tikk-Ringas E (2013) The cyber attack on saudi aramco. Survival: Global Politics and Strategy, 55(2):81–96

Brown G, Poellet K (2012) The customary international law of cyberspace. Strateg Stud Quar 6 (3):126–145

Brownlie I (1963) International law and the use of force by states. Clarendon Press, Oxford

Committee of Experts on Terrorism (CODEXTER) 12th meeting in Strasbourg, 23–24 April 2007, Meeting report CODEXTER (2007) 16. http://www.coe.int/t/dlapil/codexter/Source/meetings/codexter_m12_report_en.pdf

Comprehensive study on cybercrime (2013). United Nations, 2013 http://www.unodc.org/documents/organized-crime/UNODC_CCPCJ_EG.4_2013/CYBERCRIME_STUDY_210213.pdf

Convention on cybercrime (2001). ETS 185, Council of Europe, signed on 23 November 2001, entry into force on 1 July 2004

Cox N (2002) The regulation of cyberspace and the loss of national sovereignty. Inf Commun Technol Law 11(3):241–253

Cyberspace operations (2010). Air Force Doctrine Document 3–12, United States Air Force, 2010

Declaration on principles of international law concerning friendly relations and cooperation among States in accordance with the Charter of the United Nations. Resolution 2625 (XXXV) of 24 October 1970, General Assembly, United Nations

Demidov O (2012) International regulation of information security and Russia's national interests. Security Index: A Russ J Int Secur 18(4):15–32

Dinstein Y (2005) War, aggression and self-defence, 4th edn. Cambridge University Press, Cambridge

Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market (Directive on electronic commerce). Official Journal, (L 178), 17/7/2000, pp 0001—0016

Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications

sector (Directive on privacy and electronic communications). Official Journal, (L 201), 31/07/2002

Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. Official Journal, (L 281), 23/11/1995, pp 0031–0050

Easterbrook FH (1996) Cyberspace and the Law of the Horse. University of Chicago Legal Forum

European convention on human rights (1950). European Court on Human Rights, Strasbourg, 1950

European convention on mutual assistance in criminal matters (1959). Council of Europe, Strasbourg, signed on 20 April 1959. http://www.conventions.coe.int/Treaty/en/Treaties/Html/030.htm

Farwell JP, Rohozinski R (2011) Stuxnet and the future of cyber war. Survival: Global Politics and Strategy, 53(1):23–40

Farwell JP, Rohozinski R (2012) The new reality of cyber war. Survival: Globa Politics Strateg, 54 (4):107–120

Farwell J, Rohozinski R (2012) The new reality of cyber war. Defence iQ. http://www.defenceiq.com/cyber-defence/articles/the-new-reality-of-cyber-wa/

Fisher EA (2013) Federal laws relating to cybersecurity: overview and discussion of proposed revisions. CRS Report for Congress, Congressional Research Service, Washington, DC

Herrera GL (2005) Cyberspace and sovereignty: thoughts on physical space and digital space. Prepared for the 1st international CISS/ETH conference on the information revolution and the changing face of international relations and security, Lucerne, Switzerland, May 2005, pp 23–25

Hollis DB (2007) Why states need an international law for information operation. Lewis Clark Law Rev 11(4):1023–1061

Human rights document A/HRC/20/L.13 (2012). United Nations, 2012

International Law Commission (2001) Draft articles on responsibility of states for internationally wrongful acts. Supplement No. 10 (A/56/10), chp.IV.E.1, available at: http://www.refworld.org/docid/3ddb8f804.html

Lessig L (1999) The Law of the Horse: What cyberlaw might teach. Harvard Law Rev 113:501–546

List of the treaties coming from the subject-matter: Legal co-operation in criminal matters. Council of Europe (2008) http://www.conventions.coe.int/Treaty/Commun/ListeTraites.asp?MA=20&CM=7&CL=ENG. Accessed 11 Apr 2008

Lloyd's risk index (2013). http://www.lloyds.com Accessed 5 May 2011

Merriam-Webster Dictionary: Cyber (2013). http://www.merriam-webster.com/dictionary/cyber Accessed 5 May 2011

Merriam-Webster Dictionary: Security (2013). http://www.merriam-webster.com/dictionary/security Accessed 5 May 2011

Norton cybercrime report (2012) symantec, 2012

Pellet A (2010) The definition of responsibility in international law. In: Crawford J, Pellet A, Olleson S, Parlett K (eds) The law of international responsibility. Oxford University Press, Oxford, pp 3–16

Polanski PP (2007) Customary law of the internet. In: the search for a supranational cyberspace law. T.M.C Asser Press, The Hague

Protection of mission critical functions to achieve trusted systems and networks, DoD instruction 5200.44. Department of defense of United States of America, 2012

Report of the group of governmental experts on developments in the field of information and telecommunications in the context of international security. A/65/201, General Assembly, United Nations, 2010

Rustad ML (2009) Internet law in a nutshell. Thomson Reuters, Saint Paul, Minnesota

Rutkowski A (2011) Public international law of the international telecommunication instruments: cyber security treaty provisions since 1850. Info 13(1):13–31

Schmitt MN (1998–1999) Computer network attack and the use of force in international law: thoughts on a normative framework. Columbia Journal of Transnational Law 37:885—937
Schmitt MN (ed) (2013) Tallinn manual on the international law applicable to cyber warfare. Cambridge University Press, New York
Second additional protocol to the European convention on mutual assistance in criminal matters. Council of Europe, Strasbourg, signed on 8 November 2001. http://www.conventions.coe.int/Treaty/EN/Treaties/HTML/182.htm
Sieber U (1986) The international handbook on computer crime: computer-related economic crime and the infringements of privacy. Wiley, New York
Sieber U (1998) Legal aspects of computer-related crime in the information society: COMCRIME study. European Commission
S.S. Wimbledon (U.K. v. Japan) (1923) Publications of the permanent court of international justice, Series A—No. 1; Collection of Judgments and Orders, A.W. Sijthoff's Publishing Company, Leyden
Statement on consensus achieved by the UN Group of Governmental Experts on cyber issues, US State Press Release, June 7, 2013. http://www.state.gov/r/pa/prs/ps/2013/06/210418.htm
Strategies and policies. NATO Cooperative Cyber Defense Centre of Excellence 2015. www.ccdcoe.org/strategies-policies.html
Tikk E, Talihärm A-M (eds) (2010) International cyber security legal and policy proceedings. CCD COE, Tallinn
Universal Declaration of Human Rights (1948). United Nations, 1948
Warner M (2012) Cybersecurity: A pre-history. Intell National Secur 27(5):781–799
Wingfield T, Tikk E (2010) Frameworks for international cyber security: the cube, the pyramid, and the screen. Tikk E, Talihärm A-M. International Cyber Security Legal and Policy Proceedings. CCD COE, Tallinn, pp 16–22
Wolfers A (1952) National security as an ambiguous symbol. Political Sci Quart 67(4):481–502
Ziolkowski K, ed. (2013) Peacetime regime for state activities in cyberspace: international relations and diplomacy. NATO, Tallinn, http://www.ccdcoe.org/publications/books/Peacetime-Regime.pdf

Finnish Cyber Security Strategy and Implementation

Antti Sillanpää, Harri Roivainen and Martti Lehto

Abstract Technical and automated solutions and information networks, which make planning, guidance and implementation possible fast and in a cost-efficient way, are widely used in Finnish information society. The flipside of this development is increased dependency on extensive and complicated technical systems and information networks. Failures in these systems or, for example, in their power supply may rapidly affect comprehensive security in society. Threats against security in society have become more multifaceted and, as a consequence, more complicated. Threats can no longer be divided clearly into military and non-military threats or internal and external threats; they are often interconnected with each other, difficult to predict and likely to occur at short notice. Internal security and external security are more and more intertwined and often it is not appropriate or even possible to separate them from each other. Cyber-related espionage, crimes and operations between states have been growing and the trend seems to be on the increase (Turvallinen Suomi: Tietoja Suomen kokonaisturvallisuudesta 2013).

1 Introduction

Finland's Security Strategy for Society which was updated in 2010 takes into account cyber threats with which are meant threats against information environments and information networks in the information society. A number of states and

A. Sillanpää (✉) · H. Roivainen
Security Committee, Ministry of Defence, Helsinki, Finland
e-mail: antti.sillanpaa@turvallisuuskomitea.fi

H. Roivainen
e-mail: harri.roivainen@turvallisuuskomitea.fi

M. Lehto
Department of Mathematical Information Technology of the University of Jyväskylä, Jyväskylä, Finland
e-mail: martti.lehto@jyu.fi

M. Lehto and P. Neittaanmäki (eds.), *Cyber Security: Analytics, Technology and Automation*, Intelligent Systems, Control and Automation: Science and Engineering 78, DOI 10.1007/978-3-319-18302-2_9

international organisations have also raised the question of preparing for cyber threats in the planning and development of their activities. Cyber threats have emerged as a central element when considering the overall security in society, both in Finland and in other countries of high technology. Because of the very nature of broad-based security threats, preparing for and addressing them requires strong national and international cooperation and agreed cooperation arrangements.

Finland's Cyber Security Strategy which was published as a Government Resolution on 24.1.2013 is part of the implementation of the comprehensive security strategy of society. In the strategy the key objectives and policies are defined with which the challenges to the cyber domain are addressed and its functioning is secured. Through the policy-making of the Cyber Security Strategy and the measures needed to implement them Finland will be able to manage on the national level the intentional and unintentional adverse effects of the cyber domain, to address them and to recover from them. The implementation of the strategy has been started and it will be implemented in the Finnish way by observing the principles of comprehensive security where each administrative branch and actor is responsible for its part for achieving the objectives presented in the strategy.

2 Comprehensive Security

Comprehensive security as an operating model in Finland is the result of a development that spans decades. The work started in 1958 in the form of regional contingency planning which developed into the concept of total national defence. During the past decade the concept of *comprehensive security* emerged. In the Foreign and Security Policy Report of 2012 *preparedness* is defined as follows:

> The preparedness of Finnish society materialises under the comprehensive concept of security, which means that society's vital functions are secured through collaboration between the authorities, the business community, organisations and citizens.

Comprehensive security is not only about cooperation between administrative branches but refers to the diverse ways of benefitting from society's resources and the continuity of Finland's preparedness model. The functioning of society is secured so that each actor for its part is responsible for preparedness and taking action in times of crisis in order to achieve the jointly set national objective.

Finland's Security Strategy for Society which was published as Government Resolution on 16.12.2010 defines the grounds for comprehensive security in society. Based on the comprehensive security concept, the functions vital to society and the threats against them, the responsibilities and strategic tasks of the different administrative branches and the grounds for crisis management are defined in the Strategy. It also contains the principles of the follow-up and development of its implementation as well as the principles of training for preparedness and crisis management. The Security Strategy for Society underlines preparedness and the securing of society's vital functions in all circumstances. The functions vital to

society include the management of Government affairs, international activity, Finland's defence capability, internal security, functioning of the economy and infrastructure, the population's income security and capability to function, and psychological resilience to crisis. The strategy's threat scenarios concern the electric power supply, telecommunication and electronic ICT systems, transport systems, public utilities, food supply, financial and payment systems, availability of public funding, health and welfare of the population, major accidents, extreme natural phenomena and environmental threats, terrorism and other criminality that endanger social order, border security, a political, economic and military pressure and the use of military force. Supported by the meeting of the heads of preparedness of the ministries, the Security Committee is responsible for the joint monitoring and development of the Strategy's implementation.

The Government resolution on comprehensive security was published on 5.12.2012 with the aim to clarify the organisation of and responsibilities for comprehensive security, particularly at government level. The head of preparedness in each ministry assists the permanent secretary in the implementation of preparedness and security-related tasks. The meeting of permanent secretaries and the meeting of the heads of preparedness are permanent cooperation bodies. The meeting of preparedness secretaries assists the heads of preparedness. In accordance with the resolution the Security Committee was established to serve as a permanent cooperation body, in pursuit of proactive preparedness for comprehensive security. Its members include a state secretary, who is appointed for the duration of the Prime Minister's office, permanent secretaries of the ministries, Secretary General of the President's office, Chief of Defence Command Finland, Chief of the Finnish Border Guard, CEO of the National Emergency Supply Agency, Director-General of the Department for Rescue Services, National Commissioner of Police, and Director-General of the Finnish Customs. The Security Committee is tasked to:

- assist the Government and its ministries in preparations for comprehensive security and in coordinating such preparedness
- do follow-up on and assess how the changes in Finland's security and defence policy environment and in society impact comprehensive security arrangements
- do follow-up on the measures taken by different branches and levels of administration to maintain and develop comprehensive security related preparedness
- when necessary, integrate extensive and important preparedness related matters such as the coordination of national preparedness and the development of cooperation forms, operating models, research and exercises (Government Resolution: Comprehensive Security 2012).

One of the tasks of the Security Committee is to coordinate the preparation of the key security related strategies. It is not tasked to lead or steer when a disturbance or another type of crisis situation or emergency occurs; competent authorities are responsible for preparedness and steering measures in their administrative branches. The Security Committee coordinates preparedness for cyber security, monitors the implementation of the Cyber Security Strategy and makes proposals for further

development. The Security Committee acts in close cooperation with other cooperation bodies which coordinate cyber security related matters in connection with their own tasks.

The Government steers and monitors proactive preparedness and each ministry does the same in their respective areas of responsibility. Examples of proactive preparedness include contingency plans and plans to manage incidents, technical and structural advance preparations in view of emergency conditions or disturbances, cooperation models and operating models, training and exercises, and allocated facilities and critical resources. Preparedness also covers planning for communication and information processing during emergency conditions and disturbances. The Government, administrative authorities, agencies, establishments and municipalities are responsible for implementing proactive preparedness measures.

In disturbances, the existing legislation and the principles introduced in the Security Strategy for Society are followed while the competent authority steers operations. The responsible authority starts the incident management related measures and informs of the situation in real time, accurately and in accordance with the agreed practices. Other authorities will participate in the management of the situation by providing executive assistance according to regulations as required. The premise is to take action through the regular organisation of authorities and in accordance with regular operating models.

3 Cyber Security as a Part of Comprehensive Security

Society's growing information intensity, the increase of foreign ownership and outsourcing, integration between information and communication technologies, the use of open networks as well as the growing reliance on electricity have set totally new requirements for securing society's vital functions in normal conditions, during serious disturbances in normal conditions and in emergency conditions.

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. By exploiting system vulnerabilities, the openness of the cyber domain makes it possible to carry out attacks from all over the world. Such vulnerabilities exist in human action, organisational processes and the ICT technology being used. It is very difficult to protect oneself against complex and sophisticated malware, and to identify or locate the perpetrators. The increasing proliferation of information technology in industrial production and control systems has created new vulnerabilities and possible targets for cyber attacks. 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.

In the Finnish model the cyber domain is also seen as a possibility and a resource. A safe cyber domain makes it easier for both individuals and businesses to

plan their activities, which in turn boosts economic activity. A properly working environment also improves Finland's appeal for international investors. In addition to these, cyber security itself is a new and strengthening business area. National cyber security is interconnected with the success of Finnish companies.

As Finland's Cyber Security Strategy will be implemented in line with the principles of the Security Strategy for Society the competent authorities are responsible for incident management and the related preparedness. The rule of law and the existing division of duties are followed in cyber incident management. The same principles are followed in normal conditions and in emergency conditions. The authorities' division of duties and the modi operandi of the cooperation bodies will remain as they are in normal conditions. The Ministry of Transport and Communications is responsible for securing the functions of ICT systems while the Ministry of Finance is responsible for securing the IT functions and information security as well as the common service systems in central government. The Cyber Security Strategy and its implementation are a part of the implementation of the Security Strategy for Society. The goal is to maintain the uninterrupted and safe flow of different functions in everyday life and during disturbances. The following principles are followed in developing cyber security (Finland's Cyber Security Strategy 2013):

1. As cyber security is an essential part of the comprehensive security of society the approach for its implementation follows the principles and procedures established in the Security Strategy for Society.
2. Cyber security relies on the information security arrangements of the whole society. Cyber security depends on appropriate and sufficient ITC and telecommunications networks security solutions established by every actor operating in the cyber world. Various collaborative arrangements and exercises advance and support their implementation.
3. The approach for the implementation of cyber security is based on efficient and wide-ranging information-collection, an analysis and gathering system as well as common and shared situation awareness, national and international cooperation in preparedness. This requires the establishment of a Cyber Security Centre as well as the development of 24/7 information security arrangements for the entire society.
4. Cyber security arrangements follow the division of duties between the authorities, businesses and organisations, in accordance with statutes and agreed cooperation. Rapid adaptability as well as the ability to seize new opportunities and react to unexpected situations demands strategic agility awareness and compliance from the actors as they keep developing and managing the measures which are aimed at achieving cyber security.
5. Cyber security is being constructed to meet its functional and technical requirements. In addition to national action, inputs are being made into international cooperation as well as participation in international R&D and exercises. The implementation of cyber security R&D and education at different levels

does not only strengthen national expertise, it also bolsters Finland as an information society.
6. Cyber security development will heavily invest in cyber research and development as well as in education, employment and product development so that Finland can become one of the leading countries in cyber security.
7. In order to secure cyber security development, Finland will see to it that appropriate legislation and incentives exist to support the business activities and their development in this field. Basic know-how in the field is gained through business activity.

4 Cyber Security Strategy

On political level the development of cyber security is one of the key projects of the Government. According to the Government programme, Finland will prepare a cyber strategy on national information security and will actively participate in international cooperation in the field. As part of this work the Government resolution on Finland's Cyber Security Strategy was published on 24 January 2013.

While the objective set by the political leadership provides a clear mandate to look for new operational models in the cyber domain the work relies, however, on the existing strengths. As a country of high technology, Finland has excellent preconditions for becoming a leading country in the field of cyber security. The private and public sectors as well as different administrative branches have a long history of close and trustable cooperation. The vision of Finland's cyber security is that:

- Finland is capable of securing the functions vital to society against cyber threats in all situations
- Citizens, authorities and businesses can benefit from the cyber domain and the competence generated to protect it in efficient ways both nationally and internationally
- In 2016, Finland is internationally recognised as a forerunner in preparing for cyber threats and managing the incidents caused by them.

The cyber security strategy and its background dossier suggest what kind of investments society should make for cyber security. Here the focus is on strategic policy-setting through which preconditions can be created to fulfil a national vision for cyber security as described above. Below are the strategic guidelines which are followed by evaluations by the writers.

1. Create an efficient collaborative model between the authorities and other actors for the purpose of advancing national cyber security and cyber defence

The first guideline gives the clearest outline of the strategy as part of the concept of comprehensive security. Common modi operandi have been used and trained in Finland, but when cyber incidents that advance fast and cross easily state

boundaries and areas of responsibilities occur the role of common modi operandi is highlighted. Comprehensive situation awareness, training and exercises and all development of capabilities require cooperation in the new situation. The exchange of information between authorities and the business life also emerges as a key element.

2. Improve comprehensive cyber security situation awareness among the key actors that participate in securing vital functions of society

The second guideline underlines the importance of situation awareness. The goal is to produce up-to-date, complete and analysed information on vulnerabilities and incidents and their impact on society with the view to generate a comprehensive picture about the possible escalation of problems. A broad-based situation picture thus incorporates political, military, social, cultural, technical & technological and economic perspectives. In practice it is very hard to create such a comprehensive view in a fast moving situation.

The transformational ability of cyber threats poses great challenges to the proactive approach. Researchers are seldom capable of anticipating what is going to happen and when, especially if nothing similar has happened before. If, however, a phenomenon was in some way foreseeable the credibility of an assessment would be low. A radically new thing cannot be taken in by an audience or a readership.

The role of the proactive approach can therefore be divided between a more operational situation picture and research. The task of research is to look for completely new forms of manifestation while a situation picture starts with combining separate observations and comparing them with similar cases as quickly as possible. Because of its multidimensional nature the creating of a cyber security situation picture is based on a networked approach. The future Cyber Security Centre will be in the hub of these activities. Its services include collecting information and creating, compiling, maintaining and disseminating the situation picture. As the Centre will not be able to do this on its own it will cooperate closely with national and international partners. However, the establishment of the Centre does not change the fact that cyber incident damage control is the responsibility of the authorities and businesses that the incident concerns. The cooperation between the Cyber Security Centre and the Government Situation Centre (GOVSITCEN) ensures that the situation awareness reaction capability is improved also on political and operational levels.

3. Maintain and improve the abilities of businesses and organisations critical to the vital functions of society as regards detecting and repelling cyber threats and disturbances that jeopardise any vital function and their recovery capabilities as part of the continuity management of the business community

As the main part of critical infrastructure is privately owned and commercially run the role of businesses is central in producing comprehensive security. For a long time, a number of businesses have been involved in significant work in the security of supply organisation. The organisation is network-based, maintaining and developing security of supply under the principle of private-public partnership. About two thousand businesses and organisations which are critical to society's vital functions contribute to the National Emergency Supply Agency.

There is a long history of security and contingency planning in businesses where cyber incidents are a new phenomenon. By improving their own preparedness the businesses also support preparedness in entire society. Risk assessment, detection and identification of incidents and minimising detrimental effects were taken up as objectives in the setting of the guidelines. When detrimental effects are minimised and recovery capability is shortened, tolerance is improved. The security of supply organisation supports this activity through reports, instructions and training.

4. Make certain that the police have sufficient capabilities to prevent, expose and solve cybercrime

As a rule, when offences are related to the cyber domain the police are the competent authority in preventing, solving and reporting offences for the bringing of charges. An information network is a profitable environment with an attractive risk-reward ratio for perpetrators to commit crimes with an economic or terrorist purpose. Also the more traditional organised crime exploits the vulnerabilities of information networks and information systems. Cyber attacks can be used to endanger society's critical infrastructure and carry out terrorist strikes. In addition to terrorist acts, also frauds, the sexual abuse of children and industrial espionage are more and more perpetrated in the cyber domain.

The Strategy and its implementation require that the police have sufficient powers and resources in the new operating environment. In matters related to powers it is useful to compare Finnish practices with those of foreign counterparts. International cooperation and exchange of information are extremely important to counter cross-border crime. The strongest partners are to be found in the law-enforcement authorities and organisations, such as Europol, of the EU countries.

5. The Finnish Defence Forces will create a comprehensive cyber defence capability for their statutory tasks

In Finland, 'cyber' has not become a separate military dimension; its effects show in all arms of service and widely in the Defence Forces. According to the guidelines, a military cyber defence capability encompasses intelligence as well as cyber attack and cyber defence capabilities. To guarantee capabilities, intelligence and proactive measures in the cyber world will be developed as elements of other military force. From the military point of view, the cyber situation picture must be compiled to generate an early warning, ensure preparation time and make impact. In the same way as the police, defence administration must reflect on the matters related to powers and possible legislative needs.

The Defence Forces have traditionally exercised with other authorities and actors; the cyber dimension may have an effect on the nature of joint exercises and increase the demand for them.

6. Strengthen national cyber security through active and efficient participation in the activities of international organisations and collaborative forums that are critical to cyber security

In international cooperation the authorities are responsible for their own branches. The aim is that each actor practises cooperation with international forerunners so that the exchange of information and learning speeds up own learning process.

In addition to individual partners, the guidelines highlight the importance of the Nordic countries, the European Union and many international organisations in the development of cyber security. These include the UN, the OSCE, Nato and the OECD. The role of the EU which cooperates with both states and organisations is becoming more important also in this area.

7. Improve the cyber expertise and awareness of all societal actors

As the Finnish educational system has often been ranked as the best in the world it should be included in the clearly strong areas of the cyber security strategy. As the importance of cyber security grows there will be an increasing demand for a professional approach and increased general awareness. Improving competence has been seen as the most cost-efficient way to improve security. The available education and training as well as research contributions should meet these needs. Research, product development, and education and training on different levels support society and its members by providing functional competencies in a new situation.

After the publication of the Cyber Security Strategy businesses and educational establishments have started to develop new forms of cooperation. The Finnish Funding Agency for Technology and Innovation (TEKES) which offers public financing supports also the development of cyber networks. Financed mainly by TEKES, a top strategic centre for science, technology and innovation in Finland DIGILE is tasked to bring together and systematise research and development and make it more effective. Another example is the Cyber Security Laboratory, which is a project by VTT Technical Research Centre of Finland and the Finnish Information Security Cluster, FISC. Its new services will improve preparedness and the degree of domestic origin of information security. An example of cooperation between educational establishments is the Innovative Cities programme (INKA) which is led by the city of Jyväskylä and involves universities and polytechnics.

8. Secure the preconditions for the implementation of effective cyber security measures through national legislation

When looked at from the legislative point of view, both the events in the cyber world and the resulting incidents in the real world should be taken into account. This is why, in addition to the interests of the police and the Defence Forces as stated earlier cyber related matters easily lead to identifying the needs for change.

Legislation should provide competent authorities and other actors in different fields with the sufficient means and powers to secure society's vital functions and, especially, the security of the state against cyber threats. When assessing specific administrative branches, obligations based on international treaties may be identified, resulting in the need to review provisions. This may concern the exchange of information, for example.

The possibility of private individuals and businesses to benefit effectively from the secure cyber domain in highlighted in the strategic vision. The concretisation of possibilities also means that legislation must provide a favourable environment for the development of business activities.

9. Assign cyber security related tasks, service models and common cyber security management standards to the authorities and actors in the business community

The first guideline which is about creating a collaborative model is closely connected with the guideline discussed here. The allocation of responsibilities and tasks is highly important but it also seems to be in some respects difficult. When working on a guideline the risk assessments and maturity analyses of administrative branches emerge; on the basis of these significant problem areas and management tools can be identified. A national and sector-specific implementation programmes highlight tools with which it is possible to address challenges and make progress towards the desired vision. The public sector supports the drafting of action plans for the business community in cooperation with the security of supply organisation. The models of the security of supply organisation provide support also in this work.

10. The implementation of the Strategy and its completion will be monitored

The implementation of the Cyber Security Strategy follows the regular order: ministries and agencies are responsible for the work allocated to their administrative branches. The competent authority makes the decisions on the basis of what has been provided by law. Ministries, agencies and establishments are to include the resources for the implementation of the Cyber Security Strategy in their operating and financial plans. The Government Information Security Management Board (VAHTI) which is located in the Ministry of Finance will process and coordinate the central government's key information security and cyber security guidelines.

It is the task of the Security Committee to monitor and coordinate the implementation of the strategy. The objectives of coordination are to avoid overlapping measures, identify possible shortcomings and establish responsibilities. Implementation and steering will be discussed at a later point (Fig. 1).

Fig. 1 Strategic guidelines of the Finland's Cyber Security Strategy

5 Generic Cyber Strategy Process

5.1 Character of the Process

A strategy normally implies organisational change. It involves being future-oriented with the point being to do things in the future better than they had previously been done. The strategy process entails realising an organisation's potential to achieve something that is more developed or better compared to the present situation. The process encompasses the current situation and its problems, and creates the desire to think outside of the present framework and proceed towards the future. There are many ways to define the strategy process and to describe its different phases. The prevailing view embraces a continuing strategy process in which the phases of the process are continually repeated and create continuous growth and development.

What does *strategic* mean? Strategy, as a concept, is normally associated with the action of top management and has great importance attached to it. While the strategic decision of an organisation or an institution can be construed as being "strategic", can its implementation at lower levels of the organisation be regarded as strategic action?

The goal of cyber security is to determine and confirm different actors' roles and responsibilities, identify top priorities and create the development schedules and methods of verification for the implementation. Rather than being the ultimate objective, a cyber security strategy is continuous process. No country has started this process from scratch—nor has any country managed to finalise the process.

The cyber strategy process can be divided into three sub-processes, which are: strategic analysis, strategic priority and the implementation of the strategy.

Within the central government, the total management of cyber security is normally done at three levels. From the perspective of cyber security the highest level encompasses the Government and the cabinet committee, which are responsible for providing political guidance for cyber security as well as preparing for the important issues in internal security and comprehensive defence. The political level bears the prime responsibility for the strategic priority.

As part of their comprehensive security duties, the cyber security committee or the security committee prepares and implements cross-sectoral cyber security coordination. Its tasks and responsibilities can greatly vary from country to country. Normally, it is at this level where the cyber security strategy is prepared and its implementation managed. Strategic analysis and strategic priority, elements of the cyber strategy process, materialise through close interaction between top management and the security committee.

The functional level of cyber security comprises different administrative branches which carry out the measures included in the cyber security strategy, such as contingency and recovery planning, management of disruptions as well as cyber security R&D (Fig. 2).

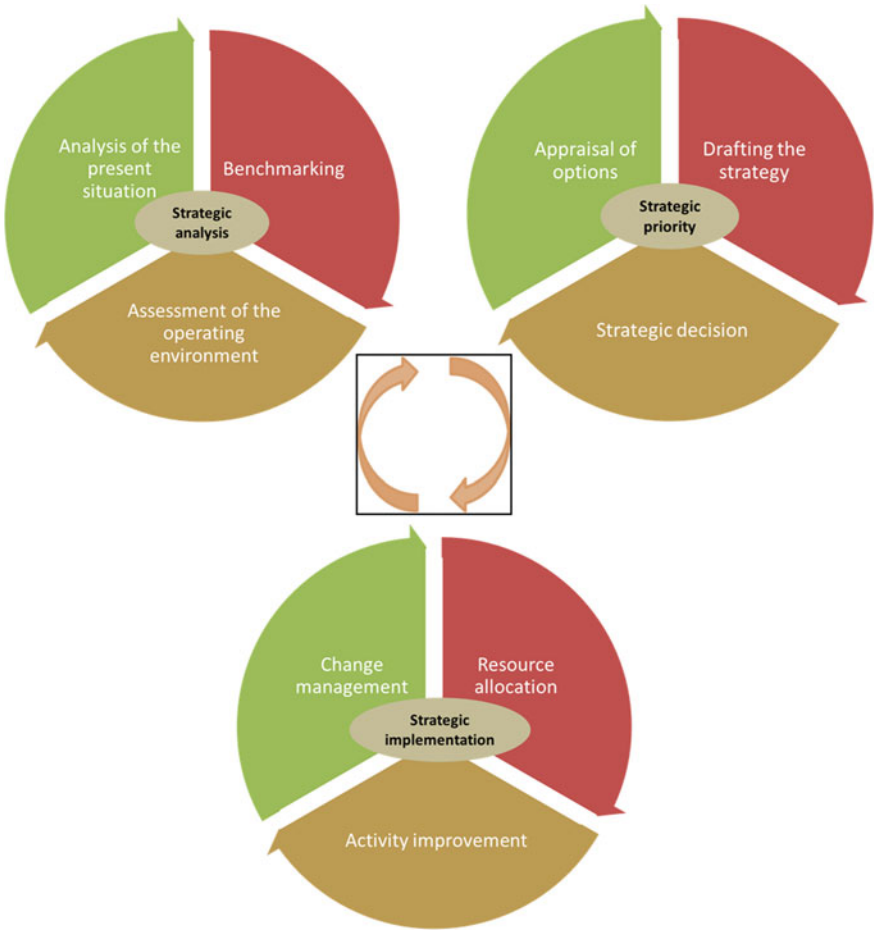

Fig. 2 The cyber strategy process

5.2 Strategic Analysis

Strategic analysis can be divided into three sub-processes, which are: analysis of the present situation, assessment of the operating environment and benchmarking. The strategic analysis produces a preliminary study which lays the foundation for the strategic priority.

The analysis of the present situation determines our own status, i.e. how we are situated in relation to the operating environment and its different elements. When it comes to cyber strategy this entails an analysis of the cyber threat environment, identification of the vulnerabilities in society's vital functions, and an assessment of the consequent risks. The analysis generates risk models and threat scenarios from

the perspectives of citizens, the business community and society's vital functions. Furthermore, cyber security performance, including its shortcomings, is analysed.

The process of assessing the operating environment identifies the phenomena within the cyber security environment, drafts the required definitions for the strategy, and identifies the already ongoing national cyber security projects, including ancillary plans and projects. Cyber security is not an insular element within the security environment. Rather, it is a part of the comprehensive national security.

The benchmarking process acquires information from other countries' cyber security strategies and identifies such best practices which can be used locally. Benchmarking also provides the opportunity to foster international cooperation as part of the national cyber security strategy.

The strategic analysis phase produces a preliminary study which places us in the national and international cyber environment, laying the groundwork for continued action in the form of political guidance and further analyses.

5.3 Strategic Priority

Strategic priority can be divided into three sub-processes, which are: appraisal of options, strategic decision and drafting the strategy. The strategic analysis produces a cyber security strategy which lays the foundation for implementing cyber security measures and the strategy itself.

Through the process of appraising the options the options for the strategy's different sectors are defined, such as cyber security management structure models or options for the cyber security centre's configuration and practices. The appraisal generates different threat scenarios, including recommended solutions. The process puts emphasis on realistic economic resources and competence. This phase defines the desired end state and the required resources (competence, personnel, financing, premises and equipment). The road to the desired end result may include alternative courses.

Organisations must be more efficient in setting and achieving their goals. In an increasingly volatile operating environment organisations must be able to ask the right questions in the midst of constant change, as well as learn and evaluate different future scenarios. (Flood and Jackson 1991, pp. 143–157).

Within the process of appraising options the task of futurology is to provide grounded opinions from the future's point of view, including the many alternative courses of development, to act as a foundation for planning, decision-making and action. Futurology explores the main courses of development and intentions which have or are providing glimpses into the future reality. Through research possible alternative options for events and decisions in human organisations—both causal and intentional—can be formed, as well as an assessment of changes on value and factual bases. Scientifically valid futurology circumvents reactive action, making proactive strategic planning and decision-making possible (Malaska 2000).

Through the strategic decision-making process the desired end state is selected, including the measures required to achieve it, from the different options. Furthermore, the desired cyber capabilities and the necessary courses of action are defined. The strategic decision provides the groundwork for the final draft of the strategy. Political guidance is embodied in the strategic decision.

The draft cyber security strategy determines the structure of the strategy and its format. The drafting phase includes several iterations during which the written text is read, re-read and polished into its final form. The different versions of the draft strategy are promulgated as extensively as possible so as to receive a sufficient number of comments for consideration. Interim presentations guarantee that the strategic decision is correctly interpreted in the draft. The drafting phase culminates in the presentation of the strategy and its adoption by the political leadership.

5.4 Implementing the Strategy

The two first phases of the strategy process have resulted in the adoption of the strategy document. The strategy includes a plan for its implementation and another plan on how to maintain the strategy process as a "living document". The implementation plan can also be prepared as a separate document after the strategy has been published. In this context the implementation of the strategy can be divided into three sub-processes, which are: change management, activity improvement and resource allocation.

In the change management phase the strategy will be implemented by applying the strategy's recommendations in practice at different administrative and organisational levels. A measuring and monitoring system will be created for change management, making it possible to gauge the success of implementation. The state of the implementation process will be regularly reported on to the political leadership.

Improvement of activity is a process which monitors developments in the cyber domain and, when necessary, steers administrative units in implementing the principles of the strategy. The goal is to constantly maintain the "big picture" of the cyber world and, if necessary, provide guidance on drafting sub-strategies. This process also triggers strategy reviews, at which time the work commences in the strategic analysis phase. The organisation responsible for the work shall indicate the need to completely review the cyber security strategy to the political leadership.

Resource allocation is an important element in implementing the strategy. At the practical level, efficiency and effectiveness are directly proportional to the economic and mental resources at hand. The Government's budget steering creates the framework for cyber security resourcing. Within their budgetary authority, accountable administrative units allocate resources for the practical implementation of cyber security, such as the creation of a situation picture, R&D and training.

The cyber security process must be dynamic and capable of reacting to changes in the operating environment. Proper definitions of the strategy process and

management levels assist in the implementation of the correct action at the correct levels as well as in preserving the content of the strategy at the appropriate level. The top level of central government must set political objectives upon which actors within the public administration and the private sector can act and achieve the goals as effectively as possible. A national security committee or a dedicated cyber security committee is a proven solution when it comes to implementing the cyber security strategy, or any continuous development process.

6 Finnish Implementation and Steering Process

Finland's national Cyber Security Strategy has attracted a lot of attention. The English version has been downloaded more often than the Finnish one which reflects interest shown also on the international level. Estimated empirically, the attention given to the strategy in Finland has been more critical than the attention abroad. It has received praise especially for its cross administrative and across society approach. Meagre resources so far allocated to the Cyber Security Centre and few concrete measures have been criticised.

The action plan to implement the strategy process addresses for its part the above mentioned problems and relies on said the positive elements. The purpose of the action plans is also to concretise the guidelines. The Security Committee drafts an action plan for national implementation but it is dependent on the views of the administrative branches. Each ministry and the National Emergency Supply Agency are entitled to suggest measures and practices or to propose amendments to the legislation to be included in national implementation. Measures should be based on identified shortcomings in the current situation. The purpose of the coordinating role of the Security Committee is to avoid duplication, identify possible shortcomings and make certain who the responsible actor is.

Administrative branches bring important and concrete measures to the national programme to further cyber security. Administrative branches work in parallel or start their own implementation programme. Measures which are below the national level or which remain clearly within the administrative branch stay in the administrative branch's own programme. The capability of administrative branches to prepare implementation programmes varies greatly: some ministries are very advanced in their work while others have only just started. The national programme is drafted first so that administrative branches can benefit from it in their own work.

The Security Committee will follow the materialisation of the measures and principles referred to in the national implementation programme as well as the progress of the sector-specific implementation programmes (Finland's Cyber Security Strategy 2013). The participation of the business community will also be assessed. To implement cyber security it is necessary to carry out coherently the strategy's principles also on regional and local levels. This requires that there is enough cooperation between various actors and the best practices are used.

The follow-up of the strategy's implementation should enable timely and well-targeted maintenance and development measures. This will also produce real-time information on whether resources have been correctly allocated, in line with the strategy's objectives. The Security Committee will report annually to the Government on the implementation of the cyber security strategy.

The two-way communication of the Cyber Strategy is important for development. There will be observations on regional or local levels which can be utilised in the follow-up work. The involvement of actors taking measures ensures that strategy papers are not overlooked but will provide the basis for planning own activities.

The annual review by the Security Committee ensures that the strategy is up-to-date and measures make progress. The Security Committee will also make proposals as to how the strategy can be further developed. Sections of the Cyber Security Strategy are likely to be incorporated into the Security Strategy for Society when it will be next updated.

Digital reality and our understanding of the problems relating to it change rapidly. This presents great challenges to keep legislation, instructions, practices and budgeting up-to-date. It is necessary to ensure a continuous development process and the up-to-date preparedness of all key actors. Cyber security concerns all of us.

References

Comprehensive Security (2012) Government Resolution, 5 Dec 2012

Finland's Cyber Security Strategy (2013) Government Resolution, 24 Dec 2013

Finnish Security and Defence Policy (2012) Prime Minister's Office Publications 1/2013. http://vnk.fi/julkaisukansio/2012/j05-suomen-turvallisuus-j06-finlands-sakerhet/PDF/VNKJ0113_LR_En.pdf

Flood RL, Jackson MC (1991) Creative problem solving: total systems intervention. Wiley, Chichester

Malaska P (2000) Knowledge and information in futurology. Foresight 2(2):237–244

Turvallinen Suomi: Tietoja Suomen kokonaisturvallisuudesta (2013) http://www.puolustusvoimat.fi/wcm/erikoissivustot/kokonaisturvallisuus/suomeksi

Part III
Cyber Security Technology

Clustering-Based Protocol Classification via Dimensionality Reduction

Gil David

Abstract We propose a unique framework that is based upon diffusion processes and other methodologies for finding meaningful geometric descriptions in high-dimensional datasets. We will show that the eigenfunctions of the generated underlying Markov matrices can be used to construct diffusion processes that generate efficient representations of complex geometric structures for high-dimensional data analysis. This is done by non-linear transformations that identify geometric patterns in these huge datasets that find the connections among them while projecting them onto low dimensional spaces. Our methods automatically classify and recognize network protocols. The main core of the proposed methodology is based upon training the system to extract heterogeneous features that automatically (unsupervised) classify network protocols. Then, the algorithms are capable to classify and recognize in real-time incoming network data. The algorithms are capable to cluster the data into manifolds that are embedded in low-dimensional space, analyzed and visualized. In addition, the methodology parameterized the data in the low-dimensional space.

1 Introduction

Network traffic classification and recognition is a critical component in many Internet applications such as traffic control (prioritize browsing over other traffic), lawful interception (requires recognition of all voice traffic), general data mining to track end user behavior, identification of specific applications, guarantee of QoS, eliminating garbage transmission and thus ensures that the bandwidth is utilized efficiently, etc.

G. David (✉)
Department of Mathematical Information Technology, University of Jyväskylä,
P.O. Box 35 (Agora), 40014 Jyväskylä, Finland
e-mail: gil.david@jyu.fi

M. Lehto and P. Neittaanmäki (eds.), *Cyber Security: Analytics, Technology and Automation*, Intelligent Systems, Control and Automation: Science and Engineering 78, DOI 10.1007/978-3-319-18302-2_10

Until recently, classification was mainly done by payload inspection of traffic packets, to look for an application signature (string or regular expressions, message structure, etc.). While payload inspection techniques have worked well in the past, they suffer from the following major limitations:

1. Applications such as Skype use techniques which provide a random-looking header and (optionally) payload to avoid passive protocol identification;
2. Signature development scalability, which affects the response time of traffic analysis equipment vendors (from the time a new method/protocol is introduced until a signature is found, distributed and absorbed), becomes a major issue;
3. Payload becomes encrypted;
4. Many new protocols per year are introduced;
5. Many protocols become stealthy;
6. Payload inspection may violate the privacy of the sender.

Currently, there are many protocols in use and new protocols are frequently being introduced, thus making the reactive development of a signature for each new protocol a challenging task on one hand and impractical on the other hand.

Non-payload based protocol classification methods are required for two broad tasks: Classification of network traffic into application families and recognition of specific applications, even if they try to hide (stealthy) or masquerade. For example, some families of interest include:

- Voice (e.g., Skype, GoogleTalk, Yahoo messenger);
- Video over P2P (e.g., Joost, Zatoo) or video over HTTP (e.g., YouTube);
- Chat (e.g., Yahoo, MSN messengers);
- P2P (e.g., BitTorrent, Gnutella, Emule);
- Gaming.

The challenge is to develop robust and performance-efficient methods, which can classify and recognize the network traffic accurately even if the application behavior is dynamic (e.g., changes between versions) that can be used for identifying applications from the same family. The focus is on classification that is based on the application/host behavior and not on the payload internal information.

Following are several desired properties that a solution should satisfy. The algorithm should work in real time and classify network traffic soon after a session starts (as opposed to classification of sessions upon termination). The algorithm should be suitable for operation on network equipment assigned at the edge of the network, on links that connect a number of subscribers (e.g., 10,000) to the Internet cloud. In addition, it may be assumed that as a classification result, the network traffic will be controlled (bandwidth limited, etc.), which may cause network traffic patterns of an adaptive application to change. In addition, payloads become encrypted. These make payload inspection useless. Therefore, in contrast to the common deep packet inspection (DPI), the proposed solution does not inspect the payload.

In this paper, we propose a unique framework that is based upon diffusion processes and other methodologies for finding meaningful geometric descriptions in high-dimensional datasets. We will show that the eigenfunctions of the generated

underlying Markov matrices can be used to construct diffusion processes that generate efficient representations of complex geometric structures for high-dimensional data analysis. This is done by non-linear transformations that identify geometric patterns in these huge datasets that find the connections among them while projecting them onto low dimensional spaces. Our methods automatically classify and recognize network protocols.

The main core of the proposed methodology is based upon training the system to extract heterogeneous features that automatically (unsupervised) classify network protocols. Then, the algorithms are capable to classify and recognize in real-time incoming network data. The algorithms are capable to cluster the data into manifolds that are embedded in low-dimensional space, analyzed and visualized. In addition, the methodology parameterized the data in the low-dimensional space.

2 Related Work and Mathematical Background

Our proposed methods are based on dimensionality reduction methods [diffusion maps (DM) (Coifman and Lafon 2006a), geometric harmonics (Coifman and Lafon 2006b)] and clustering algorithms. Therefore, we present here short descriptions of these methods.

2.1 Dimensionality Reduction

2.1.1 Diffusion Maps (DM)

This section describes the diffusion framework that was introduced in Coifman and Lafon (2006a, b) and Nadler et al. (2006). Diffusion maps and diffusion distances provide a method for finding meaningful geometric structures in datasets. In most cases, the dataset contains high dimensional data points in $\mathbb{R}^n$. The diffusion maps construct coordinates that parameterize the dataset and the diffusion distance provides a local preserving metric for this data. A non-linear dimensionality reduction, which reveals global geometric information, is constructed by local overlapping structures. Let $\Gamma = \{x_1, \ldots, x_m\}$ be a set of points in $\mathbb{R}^n$ and μ is the distribution of the points on Γ. We construct the graph $G(V, E)$, $|V| = m$, $|E| \ll m^2$, on Γ in order to study the intrinsic geometry of this set. A weight function $W_\varepsilon \triangleq w_\varepsilon(x_i, x_j)$ is introduced that measures the pairwise similarity between the points. For all $x_i, x_j \in \Gamma$, the weight function has the following properties:

symmetry: $w_\varepsilon(x_i, x_j) = w_\varepsilon(x_j, x_i)$
non-symmetry: $w_\varepsilon(x_i, x_j) \geq 0$

positive semi-definite: For all real-valued bounded function f defined on Γ,

$$\int_\Gamma \int_\Gamma w_\varepsilon(x_i, x_j) f(x_i) f(x_j)\, d\mu(x_i)\, d\mu(x_j) \geq 0. \tag{1}$$

One of the common choices for W_ε is

$$w_\varepsilon(x_i, x_j) = e^{-\frac{\|x_i - x_j\|^2}{\varepsilon}}. \tag{2}$$

The non-negativity property of W_ε allows to normalize it into a Markov transition matrix P where the states of the corresponding Markov process are the data points. This enables to analyse Γ as a random walk. The construction of P is known as the *normalized graph Laplacian* (Chung 1997).

Formally, $P = \{p(x_i, x_j)\}_{i,j=1,\ldots,m}$ is constructed as

$$p(x_i, x_j) = \frac{w_\varepsilon(x_i, x_j)}{d(x_i)}, \tag{3}$$

where

$$d(x_i) = \int_\Gamma w_\varepsilon(x_i, x_j)\, d\mu(x_j) \tag{4}$$

is the degree of x_i. P is a Markov matrix since the sum of each row in P is 1 and $p(x_i, x_j) \geq 0$. Thus, $p(x_i, x_j)$ can be viewed as the probability to move from x_i to x_j in *one* time step. By raising this quantity to a power t (advance in time), this influence is propagated to nodes in the neighborhood of x_i and x_j and the result is the probability for this move in t time steps. We denote this probability by $p_t(x_i, x_j)$. These probabilities measure the connectivity among the points within the graph. The parameter t controls the scale of the neighborhood in addition to the scale control provided by ε.

Construction of

$$\tilde{p}(x_i, x_j) = \frac{\sqrt{d(x_i)}}{\sqrt{d(x_j)}} p(x_i, x_j). \tag{5}$$

which is a symmetric and positive semi-definite kernel leads to the following eigen-decomposition:

$$\tilde{p}(x_i, x_j) = \sum_{k \geq 0}^{m} \lambda_k v_k(x_i) v_k(x_j). \tag{6}$$

A similar eigen-decomposition

$$\tilde{p}_t(x_i, x_j) = \sum_{k \geq 0}^{m} \lambda_k^t v_k(x_i) v_k(x_j) \tag{7}$$

is obtained after advancing t times on the graph. Here p_t and $\tilde{p}_t$ are the probability of transition from x_i to x_j in t time steps.

A fast decay of $\{\lambda_k\}$ is achieved by an appropriate choice of ε. Thus, only a few terms are required in the sum in (7) to achieve a relative cover $\eta(\delta)$ for $\delta > 0$.

A family of diffusion maps was introduced in Coifman and Lafon (2006a). They are $\Phi_t(x)_{m \in \mathbb{N}}$ given by

$$\Phi_t(x) = \begin{pmatrix} \lambda_0^m v_0(x) \\ \lambda_1^m v_1(x) \\ \vdots \end{pmatrix}.$$

The map $\Phi_m : \Gamma \to l^{\mathbb{N}}$ embeds the dataset into a Euclidean space. They also introduced the *diffusion distance*

$$D_t^2(x_i, x_j) = \sum_{k \geq 0} (\tilde{p}_t(x_i, x_k) - \tilde{p}_t(x_k, x_j))^2.$$

This formulation is derived from the known random walk distance in Potential Theory. It is shown (see Coifman et al. 2005 for a proof) that the diffusion distance can be expressed in terms of the right eigenvectors of P:

$$D_t^2(x_i, x_j) = \sum_{k \geq 0} \lambda_k^{2t} (v_k(x_i) - v_k(x_j))^2.$$

It follows that in order to compute the diffusion distance, one can simply use the eigenvectors of $\tilde{P}$. Moreover, this facilitates the embedding of the original points in a Euclidean space $\mathbb{R}^{\eta(\delta)-1}$ by:

$$\Xi_t : x_i \to (\lambda_0^t v_0(x_i), \lambda_1^t v_1(x_i), \lambda_2^t v_2(x_i), \ldots, \lambda_{\eta(\delta)}^t v_{\eta(\delta)}(x_i)).$$

This also provides coordinates on the set Γ. Essentially, $\eta(\delta) \ll n$ due to the fast spectral decay of the spectrum of P. Furthermore, $\eta(\delta)$ depends only on the primary intrinsic variability of the data as captured by the random walk and not on the original dimensionality of the data. This data-driven method enables the parametrization of any set of points—abstract or not—provided the similarity matrix of the points W_ε is available.

Nadler et al. (2006) extended the diffusion map theory for data that is sampled from some unknown probability distribution and for dynamic systems that evolve in time according to a stochastic partial differential equation.

2.1.2 Geometric Harmonics

Geometric Harmonics (GH) is a method for extending the low-dimensional embedding space, which was computed on the input dataset, for the arrival of new data points to determine whether or not they belong to the precomputed clusters. Assume that $\Gamma = \{x_1, \ldots, x_m\}$ is a set of points in $\mathbb{R}^n$ and Φ_t is its diffusion embedding map. The geometric harmonic scheme extends Φ_t into a new dataset $\bar{\Gamma}$ in $\mathbb{R}^n$. We first overview the Nyström extension which is a common method for extension of functions from a training set to new data points. Let $\varepsilon \geq 0$ be a scale extension and consider the eigenvectors and eigenvalues of a Gaussian kernel of width ε on the training set Γ:

$$\lambda_l \varphi_l(x) = \sum_{y \in \Gamma} e^{\frac{-\|x-y\|}{\varepsilon^2}} \varphi_l(y), \quad x \in \Gamma. \tag{8}$$

Since the kernel can be evaluated for any $x \in \mathbb{R}^n$, (8) can be written as

$$\bar{\varphi}_l(x) = \frac{1}{\lambda_l} \sum_{y \in \Gamma} e^{\frac{-\|x-y\|}{\varepsilon^2}} \varphi_l(y), \quad x \in \mathbb{R}^n. \tag{9}$$

This is known as the Nyström extension. The extension distance of $\bar{\varphi}_l$ from the training set is proportional to ε. Beyond this distance, the numeric extension vanishes. Let f be a function on the training set Γ. Since the eigenfunctions of the kernel form a basis of the set of functions on the training set, f can be decomposed as

$$f(x) = \sum_l \langle \varphi_l, f \rangle \varphi_l(x), \quad x \in \Gamma. \tag{10}$$

Using the Nyström extension, we can define f for points in $\mathbb{R}^n$ to be

$$\bar{f}(x) = \sum_l \langle \varphi_l, f \rangle \bar{\varphi}_l(x), \quad x \in \mathbb{R}^n. \tag{11}$$

When choosing a Gaussian of width ε, the extension distance is proportional to ε. The geometric harmonics scheme allows this extension range to grow by considering two observations:

1. The scale ε of the kernel used for extending should be as large as possible;
2. This scale should not be the same for all functions that we try to extend. It should depend on the complexity of the function.

An iterative process for finding an appropriate extension ε for a given function is described in Coifman and Lafon (2006b) and Lafon et al. (2006).

2.2 Clustering Techniques

Clustering in data-mining methodologies identifies related patterns in the underlying data. A common formulation of the problem is: Given a dataset S of n data-points in $\mathbb{R}^d$ and an integer k, the goal is to partition the data-points into k clusters that optimize a certain criterion function with respect to a distance measure. Each data-point is an array of d components (features). It is a common practice to assume that the dataset was numerically processed and thus a data-point is a point in $\mathbb{R}^d$.

The partition includes k pairwise-disjoint subsets of S, denoted as $\{S_i\}_{i=1}^k$, and a set of k representative points $\{C_i\}_{i=1}^k$ which may or may not be part of the original dataset S. In some cases, we also allow to classify a data-point in S as an *outlier*, which means that the data-point is an anomaly. In some applications, outliers can be discarded, while other applications may look for these outliers.

The most commonly used criterion function for data clustering is the square-error:

$$\phi = \sum_{i=1}^{k} \sum_{x \in S_i} \|x - C_i\|^2 \tag{12}$$

where in many cases C_i is the *centroid* of S_i, i.e. the mean of the cluster S_i is

$$\text{mean}(S_i) = \frac{\sum_{x \in S_i} x}{|S_i|}. \tag{13}$$

Finding an exact solution to this problem is NP-hard even for a fixed small k. Therefore, clustering algorithms try to approximate the optimal solution, denoted ϕ_{opt}, to a given instance of the problem. In this chapter, we use the k-means algorithm for the clustering of the data, however, it can be replaced with other centroid based techniques [for example, SOM (Kohonen 1990)].

2.2.1 k-Means and Its Derivatives

The k-means algorithm (Lloyd 1982) is probably the most popular clustering algorithm used today. Although it only guarantees convergence towards a local optimum, its simplicity and speed makes it attractive for various applications.

The k-means algorithm works as follows:

1. Initialize an arbitrary set $\{C_i^0\}_{i=1}^k$ of k centers. Typically, these centers are chosen at random with uniform probability from a dataset S.
2. Assign each point $x \in S$ to its closest center. Let S_i^j be the set of all points in S assigned to the center C_i^j in phase j.

3. Recompute the centers $C_i^{j+1} = mean(S_i^j)$.
4. Repeat the last two steps until convergence. It is guaranteed that ϕ in (12) is monotonically decreasing.

The k-means algorithm may yield poor results since $\frac{\phi}{\phi_{opt}}$ is not guaranteed to be bounded. However, if a bounded error is wanted, it is possible to use other initialization strategies instead of Step 1. In other words, we can *seed* the k-means algorithm with an initial set $\{C_i^0\}_{i=1}^k$ that was chosen more carefully than just sampling S at random.

k-means++ (Arthur and Vassilvitskii 2007) is such a method. Instead of choosing $\{C_i^0\}_{i=1}^k$ at random, in k-means++ the following strategy is used:

1. Select C_1^0 uniformly at random from S.
2. For $i \in \{2, \ldots, k\}$, select $C_i^0 = x'$ at random from S with probability $D_i(x')^2 / \sum_{x \in S} D_i(x)^2$, where $D_i(x) = \min_{i \in \{1,\ldots,i-1\}} \|x - C_i\|_2$.

The use of this approach to seed the k-means algorithm guarantees that the expectation of $\frac{\phi}{\phi_{opt}}$ is bounded (see Arthur and Vassilvitskii 2007).

Kernel k-means (Zhang and Rudnicky 2002) is a kernel clustering scheme, which is based on k-means for large datasets. It introduces a clustering scheme which changes the clustering order from a sequence of samples to a sequence of kernels. It employs a disk-based strategy to control the data.

3 Evaluation Datasets

Our protocol classification and recognition results are based on datasets that contain real network traffic that was captured in real-time from a big enterprise network. The datasets contain network traffic between hundreds of users from the internal network to other users in the internal network and to other users from the Internet. This data is not available to the public.

The datasets consists of 41,000 network flows that were generated by a commercial product that is integrated into the enterprise network. Each network flow contains session statistics of packets from different ranges. For example, a single network flow can contain statistics of packets 30–90 or packets 90–190 or packets 190–240 etc., of a specific session. Each network flow corresponds to a single network session. As a result, a session that contains 240 packets is represented by three different corresponding network flows: The first network flow represents packets 30–90, the second network flow represents packets 90–190 and the third network flow represents packets 190–240.

Each network flow is labeled (classified) by a commercial payload-based product that uses signatures (patterns) for protocol classification. This product uses deep

packet inspection techniques for payload analysis and labeling. Here are some possible labels for each network flow:

1. HTTP (e.g., audio, video, binary, flash and browsing);
2. P2P (e.g., Gnutella, BitTorrent, Emule and Warez);
3. Mail (e.g., pop3 and smtp);
4. Chat (e.g., MSN Messenger and ICQ);
5. Voice (e.g., sip and Skype).

Each session flow contains three groups of parameters:

1. The protocol labeling of a network flow as it was labeled by the commercial product:
   - Protocol name—the exact name of this protocol (e.g., BitTorrent file transfer, BitTorrent networking, video over HTTP, binary over HTTP, Skype VOIP);
   - Protocol family—the general family name of this protocol (e.g., P2P, HTTP, mail).
2. Non-payload statistics that were gathered for the network flow by the commercial product:
   - Packets range—the sampled packet range (the range of packets in the whole session);
   - Start packet—the index of the first data packet (the first packet that contains payload) that was sent during the packets range;
   - Total packets—the total number of packets, including the start packet, that were sent from both directions (client and server) during the packets range;
   - Total client packets—the total number of packets that were sent by the client during the packets range;
   - Total server packets—the total number of packets that were sent by the server during the packets range;
   - Client average packet—the average size of the packets that were sent by the client during the packets range;
   - Server average packet—the average size of the packets that were sent by the server during the packets range;
   - Client packet rate—the number of packets per second that were sent by the client during the packets range;
   - Server packet rate—the number of packets per second that were sent by the server during the packets range;
   - Client bit rate—the number of Kbits per second that were sent by the client during the packets range;
   - Server bit rate—the number of Kbits per second that were sent by the server during the packets range;
   - Client average packet variance—the variance of the average size of the packets that were sent by the client during the packets range;
   - Server average packet variance—the variance of the average size of the packets that were sent by the server during the packets range;

- Average time interval between packets—the average inter-arrival time between packets that were sent in both directions during the packets range;
- Time interval variance—the variance of the average inter-arrival time between packets that were sent in both directions during the packets range;
- Client total volume—the total of bytes (data) that were sent by the client during the packets range;
- Server total volume—the total of bytes (data) that were sent by the server during the packets range;
- Total time—the number of seconds that were elapsed from the arrival of the first packet to the arrival of the last packet during the packets range;
- Volume ratio—the ratio between the client total volume and the server total volume.

3. The identification of hosts that participate in the session:

- Server IP—the IP address of the server;
- Client IP—the IP address of the client;
- Server port—the port number of the server;
- Client port—the port number of the client;
- Protocol—the transport layer protocol (udp/tcp).

The corpus set for our evaluation process was generated from the full network flows by selecting the data points according to the following guidelines:

1. Select only network flows that were associated to protocols that have at least 1 % of presence in all the protocols (at least 1 % of the network flows belong to the same protocol as the network flow);
2. Select only network flows that their protocol was well-labeled by the commercial product. In other words, these are the network flows that the commercial product knew how to label them.

The resulting corpus set was processed and analyzed by the flow-oriented traffic analyzer that is presented in Sect. 4.

4 Traffic Analyzer

The protocol classification datasets contain three groups of parameters (Sect. 3):

1. The protocol labeling of the network flow;
2. The non-payload statistics that was gathered for the network flow;
3. The identification of hosts that participate in the session.

In order to prove the validity of the proposed protocol classification and recognition (PCR) and its capabilities according to the challenges that were described in Sect. 1, we choose to use only some of the parameters from each network flow without using any payload-based methods and parameters and without relying on

the well-known ports of specific applications. We also choose to use only the first range of packets in each session in order to prove that the protocol classification can be determined at the very beginning of each session. As a result, we use only parameters from the second group of parameters: Non-payload statistics of each flow.

Figure 1 presents a flow diagram of the traffic analyzer.

5 Sequential Application of the Flow-Oriented Traffic Analyzer

1. Get a new arrival of a network flow;
2. For each network flow, the analyzer checks the range of this packet. The analyzer handles only network flows for the first packets range (for example, packets 30–90 in each session). All the other network flows are ignored;
3. The analyzer selects the following statistics from each network flow (these selected statistics represent this flow):

   - Start of data packet;
   - Number of packets (total from client and server);
   - Average packet size (client, server) and their standard deviation;
   - Packet rate (client, server);
   - Packet bit rate (client, server);
   - Total volume (client, server) and their ratio;
   - Time interval between packets and its standard deviation;
   - Total time (duration) of the analyzed flow.

The output from this process is a statistics matrix that contains a single row for every network flow. Each row contains the above 18 values (features) that have different scales. For example, the total number of packets is at most 60 (the packet range is always 30–90 in this example) whereas the total size can be tens of thousands (of bytes). The resulting corpus set contains 5,500 network flows. Therefore, the statistics matrix from the network flows dataset, which was produced by the traffic analyzer, is of size $5,500 \times 18$. This matrix is the input for the protocol classification processes (training and classification). These 18 values are the features we decided to choose to demonstrate the performance of the algorithm. This represents one option to choose from. Of course, fewer, more and other features with other relations among the features can be chosen. The choice of features vector is critical and should be done with the understanding of the "physical" nature of the underlying communication network.

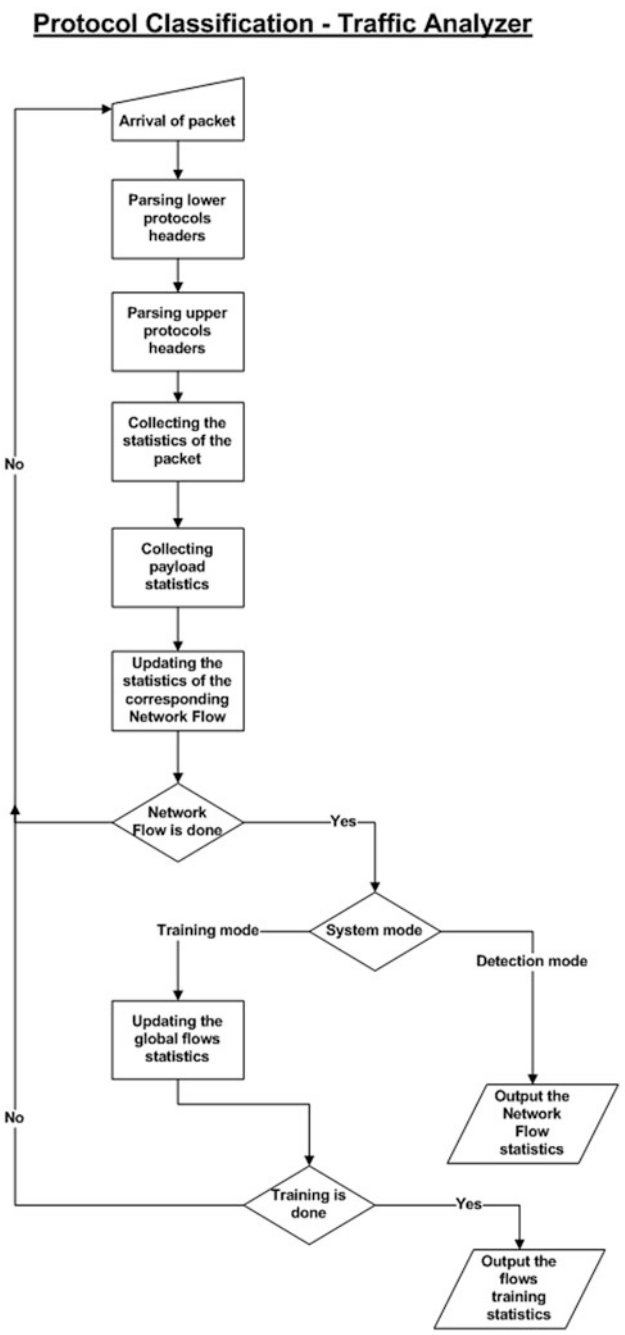

Fig. 1 Protocol classification traffic analyzer

6 Clustering-Based Protocol Classification via Dimensionality Reduction

This section describes how to classify protocols via dimensionality reduction. The algorithms are explained via their application to commercial network that was described in Sect. 3. But the proposed methodology is general and generic and fits other types of datasets and networks.

The input to this process is the matrix that was generated by the protocol classification traffic analyzer (see Sect. 4). This matrix contains statistical information on the meta-data (non-payload statistics). This matrix contains the statistics on various meta-data values that were collected by the traffic analyzer.

The protocol classification and recognition is performed in real-time. The protocol classification and recognition contains an initial training step for a predetermined training period and only then the classification is applied in real-time. The training step generated the infrastructure for the classification step.

The next section provides a full description of the proposed PCR algorithm.

6.1 Outline of the Real-Time Protocol Classification Process

This algorithm has two sequential steps:

1. *Training*: Study and analysis of the training datasets and projecting them onto lower-dimension space. Then, clustering, classification and recognition of the projected data points are applied. This is done in a offline mode for a predetermined period of time interval for every network. Its frequent application depends on the behavior profile of the system that this PCR classifies. The output from this step enables to classify protocols in Step 2. Figure 2 shows a high level diagram of the training step.

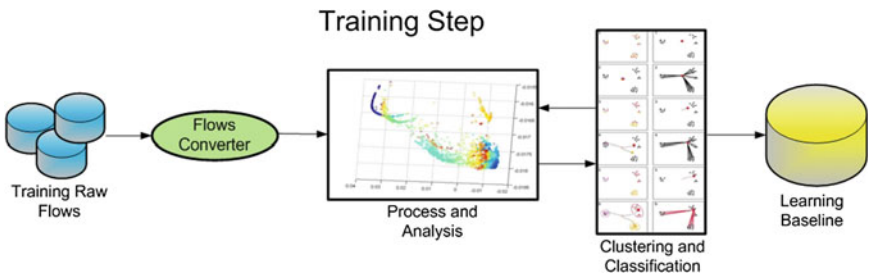

Fig. 2 Protocol classification and recognition: training step diagram

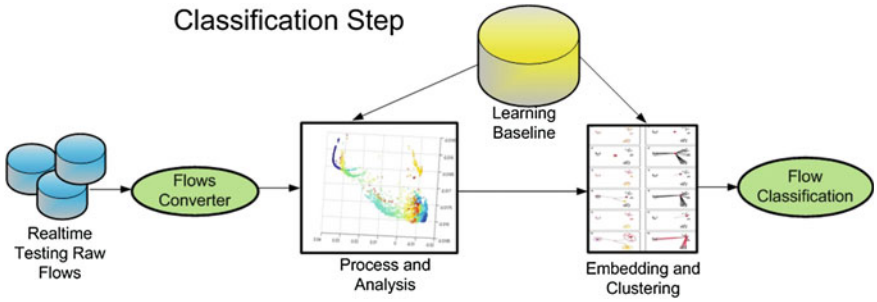

Fig. 3 Protocol classification and recognition: real-time classification step diagram

2. *Classification and recognition*: Application of automatic unsupervised tools that enable real-time classification and recognition of network traffic and detection of problems (anomalies, trojans, data leakage and intrusions). It classifies each newly arrived data point to be either one of the already known protocols (normal, according to the training phase) or abnormal (does not belong to any of the clusters from the training set). Figure 3 shows a high-level diagram of the real-time classification step.

The training step is based on the offline intrusion detection algorithm that was described in David (2009). However, there are several modifications and extensions that are done to this algorithm. The normalization step is replaced by a logarithm normalization. This normalization is very fast and efficient. However, the rest of the steps in the offline algorithm remain the same. At the end of this process we apply the K-Means (MacQueen 1967) clustering to the reduced data. Each cluster is classified as one of the application families that were described in Sect. 3. The output of this training step is the baseline profile for a real-time classification process.

Figure 4 presents a flow diagram of the training mode process.

The first step in the real-time classification step involves taking the logarithm of newly arrived data point. After this newly arrived data point was normalized, we apply the geometric harmonics to extend the (reduced) embedding baseline matrix with the newly arrived data point. As a result, we get a new embedding for the newly arrived data point. Finally, we use the baseline profile in order to classify the newly arrived data point as belonging to one of the training clusters (normal) or as abnormal. This way we classify all the data points and detect all the anomalies in real-time.

Figure 5 presents a flow diagram of the classification mode process.

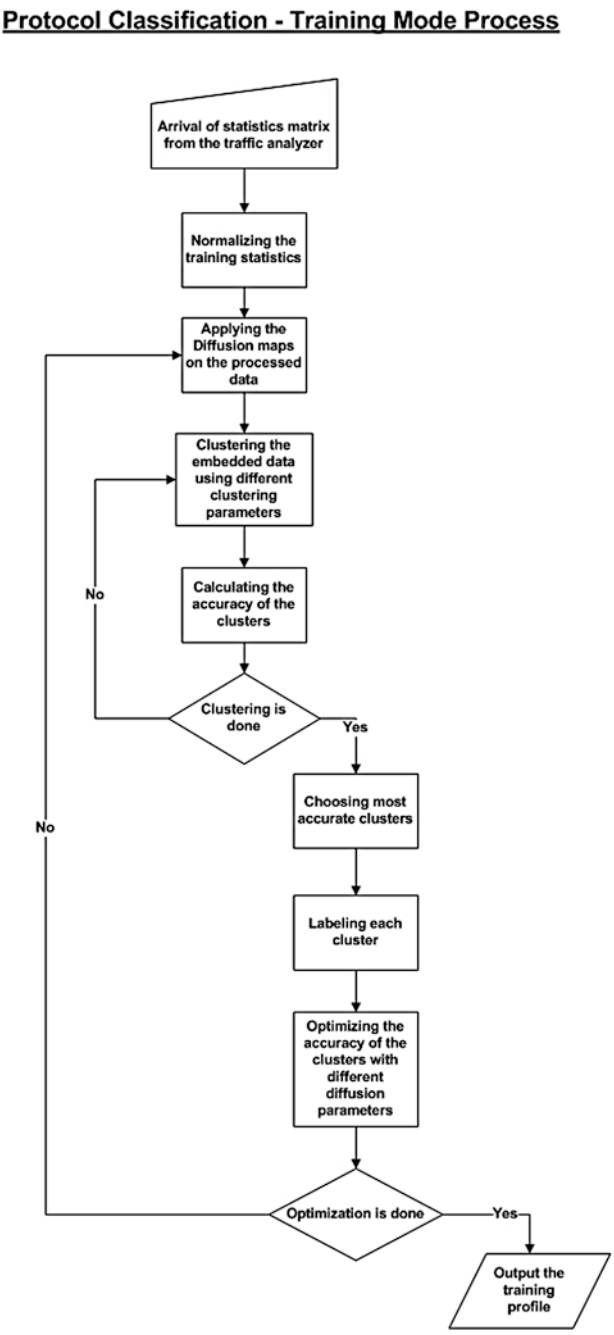

Fig. 4 Protocol classification: training mode process

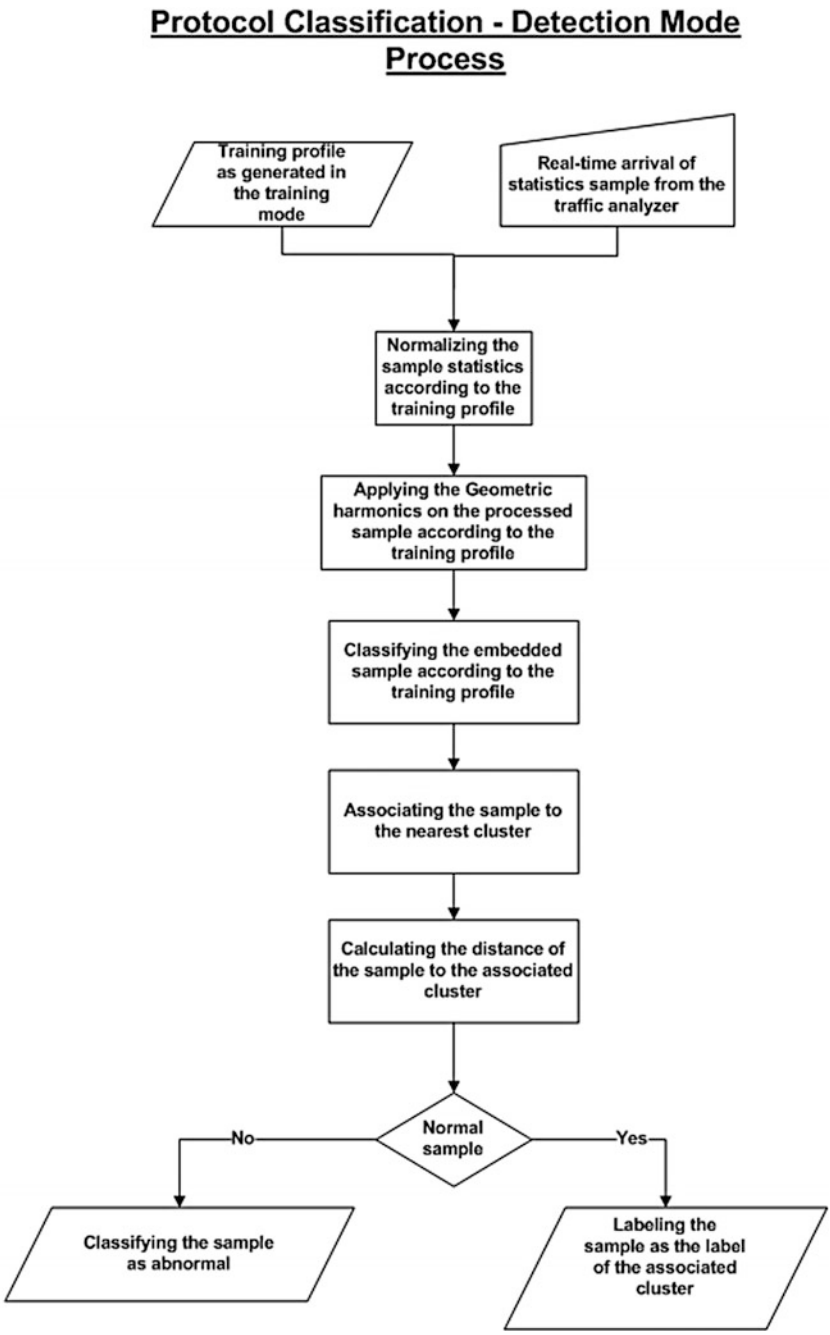

Fig. 5 Protocol classification: classification mode process

6.1.1 High Level Description of the PCR Algorithm

Here are the main steps of the PCR algorithm:

1. Training step:
   (a) Application of the traffic analyzer to produce a matrix with statistical data. The selection process of the features vector was described in Sect. 4;
   (b) Each column (feature vector) in the matrix is normalized by using the logarithm of its value;
   (c) The normalized matrix is processed by the diffusion maps to derive its embedding matrix:
      - The pairwise distances in the normalized matrix are computed according to the weighted Euclidean distance method where the weight vector is randomly selected.
      - The distances matrix is analyzed by the application of diffusion maps (see Sect. 2.1.1). It returns the first 10 eigenvectors. These eigenvectors are the embedding matrix.
      - The weight vector and the computed embedding matrix are saved as a baseline for the real-time classification Step 2.
   (d) Clustering and classification using the embedding:
      - All points in the embedding matrix are clustered using the K-Means clustering:
        – The value of K is initialized.
        – K-Means clustering is applied.
        – The accuracy of each of the K clusters is determined.
        – This process is repeated with different values for K. K that achieves the best accuracy is chosen.
      - Each of the K clusters is classified (labeled) as one of the application families.
      - Each cluster is represented by its centroid.
      - The radius of each cluster is determined by finding the maximum distance between the centroid of the cluster and each of the points that associated with the cluster.
      - The K centroids and their corresponding radiuses and labels (the corresponding classified cluster) are saved as a baseline for the real-time classification in Step 2.
   (e) Optimizing the accuracy of the clustering:
      - Steps (c) and (d) are repeated with different weight vectors.
      - The weight vector the achieves the best accuracy is chosen. It is saved as a baseline for the real-time classification in Step 2.

2. Classification step:

   (a) On-line (can be done in real-time) application of the traffic analyzer to produce the statistics of the newly arrived data point (row vector);
   (b) Each value (feature) in this new row vector is normalized by using the logarithm of its value;
   (c) This normalized row vector is processed by the application of the geometric harmonics to derive its embedding vector:

      - This row vector is analyzed by the application of the geometric harmonics using the baseline embedding matrix (was computed and saved in the training step) and the baseline weight vector. It returns the matrix extension. This extension is the new embedding vector of the newly arrived data point.

   (d) Classification of the newly arrived data point as either belonging to a known application protocol of one of the training clusters (normal) or not belonging which means that it is an abnormal:

      - The distances between the embedding vector of the newly arrived data point and the K centroids (were computed and saved in the training step) are computed according to the weighted Euclidean distance (the weight vector was computed and saved in the training step).
      - The newly arrived data point is associated with the cluster of its nearest centroid (one of the K centroids).
      - A data point, which is mapped to a cluster that its distance from the centroid of the cluster is larger than the baseline radius of this cluster, is classified as an abnormal point. Otherwise, it is classified as a normal point and its classification is determined according to the baseline label of the corresponding cluster (the radiuses and the labels were computed and saved in the training step).

6.1.2 Detailed and Formal Description of the PCR Algorithm

1. Training Step:

Processing the training dataset: Let H be a dataset such as network flows. Let C be a matrix of size $m \times n$ that was produced from H by the traffic analyzer. m is the number of network flows and n is the number of features that were gathered for every network flow.

Normalization of the matrix C: The matrix C is normalized by taking the logarithmic value of each feature. Assume $r, 1 \leq r \leq m$, is a row in C denoted by $c^r \stackrel{\Delta}{=} \{c_{ri} : 1 \leq i \leq n\}$. The normalized vector a^r is constructed by $a^r = \log(c^r)$, $r = 1, \ldots, m$. The normalized training matrix A is saved as baseline for the real-time detection step.

Processing the normalized matrix A—derivation of its embedding matrix Ψ: We reduce the dimensionality of the data from n (number of features) to 10 (it can be any number). This process applies the diffusion maps. We denote the row vector $i, 1 \le i \le m$, in the normalized matrix A to be $\overrightarrow{a_i} \stackrel{\Delta}{=} \{a_{ik} : 1 \le k \le n\}$. We compute for A its pairwise distances matrix $\tilde{A}$ whose entries are $\tilde{a}_{ij}$ using one of the following distance metrics:

Euclidean distance metric:

$$\tilde{a}_{ij} \stackrel{\Delta}{=} \left\{ \sqrt{(\overrightarrow{a_i} - \overrightarrow{a_j}) \cdot (\overrightarrow{a_i} - \overrightarrow{a_j})^T} : i,j = 1, \ldots, m \right\}.$$

Weighted Euclidean distance metric:

$$.\tilde{a}_{ij} \stackrel{\Delta}{=} \left\{ \sqrt{(\frac{\overrightarrow{a_i} - \overrightarrow{a_j}}{\overrightarrow{w}}) \cdot (\frac{\overrightarrow{a_i} - \overrightarrow{a_j}}{\overrightarrow{w}})^T} : i,j = 1, \ldots, m \right\}.$$

where $\overrightarrow{w} \stackrel{\Delta}{=} \{w_k : 1 \le k \le n\}$ is a weighting factor vector. As w_k becomes larger, the influence of the kth feature on the distance between $\overrightarrow{a_i}$ and $\overrightarrow{a_j}$ becomes smaller.

Cosine distance metric:

$$\tilde{a}_{ij} \stackrel{\Delta}{=} \left\{ \left(1 - \frac{\overrightarrow{a_i} \cdot \overrightarrow{a_j}^T}{\sqrt{\overrightarrow{a_i}^T \cdot \overrightarrow{a_i}} \cdot \sqrt{\overrightarrow{a_j}^T \cdot \overrightarrow{a_j}}} \right) : i,j = 1, \ldots, m \right\}.$$

Mahalanobis distance metric:

$$\tilde{a}_{ij} \stackrel{\Delta}{=} \{ \sqrt{(\overrightarrow{a_i} - \overrightarrow{a_j}) \cdot \Sigma^{-1} \cdot (\overrightarrow{a_i} - \overrightarrow{a_j})^T} : i,j = 1, \ldots, m \text{ and } \Sigma \text{ is the sample covariance matrix} \}.$$

Σ can also be the features matrix.

We build a Gaussian kernel

$$K_{ij} = e^{-\frac{\tilde{a}_{ij}}{\varepsilon}}, \quad i,j = 1, \ldots, m.$$

Since ε is fixed for all entries in $\tilde{A}$, it gives a coarse scaling control. A finer scaling control can be achieved as follows: First, we build the initial Gaussian kernel $\tilde{K}_{ij}$ with the fixed scale control ε,

$$\tilde{K}_{ij} = e^{-\frac{\tilde{a}_{ij}}{\varepsilon}}, \quad i,j = 1, \ldots, m.$$

Then, we build a Gaussian kernel with a finer scale control

$$K_{ij} = e^{-\frac{\tilde{a}_{ij}}{\sum_{q=1}^{m} \tilde{K}_{iq}}}, \quad i,j = 1,\ldots,m.$$

This finer scale control provides better and compact description of the local geometric properties of the pairwise distances matrix $\tilde{A}$. This process is repeated until the scale factor is sufficiently fine and K_{ij} represents optimally the nature of the local geometry of $\tilde{A}$. K_{ij} is normalized into a matrix P_{ij} by one of the following methods:

Graph Laplacian matrix:

$$P_{ij} = \frac{K_{ij}}{\sqrt{\sum_{q=1}^{m} K_{iq}} \cdot \sqrt{\sum_{q=1}^{m} K_{jq}}}.$$

Laplace-Beltrami matrix: First, we compute the graph Laplacian matrix

$$\tilde{P}_{ij} = \frac{K_{ij}}{\sqrt{\sum_{q=1}^{m} K_{iq}} \cdot \sqrt{\sum_{q=1}^{m} K_{jq}}}.$$

We repeat this process and get the Laplace-Beltrami matrix

$$P_{ij} = \frac{\tilde{P}_{ij}}{\sqrt{\sum_{q=1}^{m} \tilde{P}_{iq}} \cdot \sqrt{\sum_{q=1}^{m} \tilde{P}_{jq}}}.$$

Since P_{ij} is a symmetric positive semi-definite kernel, it enables the following eigen-decomposition:

$$..P_{ij} = \sum_{w \geq 1}^{m} \lambda_w v_w\left(\overrightarrow{a_i}\right) v_w\left(\overrightarrow{a_j}\right)\vec{x}..$$

where λ_w are the eigenvalues and v_w are the eigenvectors. Finally, we build the embedding matrix Ψ. We denote the ith column of Ψ by Ψ^i.

We take the first 10 eigenvectors of the eigen-decomposition of that process. The outputs of this process are the weight vector w and the embedding matrix Ψ. w and Ψ are saved for the real-time (online) detection step.

Clustering and classification of the points in the embedding Ψ: We use the embedding matrix Ψ to cluster the points in the dataset using K-Means clustering

(MacQueen 1967). For every K, $5 \leq K \leq 100$, we apply the K-Means clustering to Ψ. For each K, denote the output from the K-Means clustering by

$$\text{CLUST}^K \stackrel{\Delta}{=} \{\text{CLUST}_i^K : 1 \leq i \leq K\}$$

and

$$\text{CENTR}^K \stackrel{\Delta}{=} \{\text{CENTR}_i^K : 1 \leq i \leq K\}.$$

CLUST_i^K contains the points in Ψ that were associated by the K-Means to the ith cluster in CLUST^K and CENTR_i^K is the corresponding centroid value in CLUST_i^K. We compute the radius RADIUS_i^K for every CLUST_i^K by calculating the maximum distance between all the points in Ψ that belong to CLUST_i^K and between CENTR_i^K. We denote the radiuses of the K-Means clusters CLUST^K by

$$\text{RADIUS}^K \stackrel{\Delta}{=} \{\text{RADIUS}_i^K : 1 \leq i \leq K\}.$$

For each i, the points in CLUST_i^K are grouped according to their labels (application type from the training). The classification CLASS_i^K of the cluster CLUST_i^K is determined according to the label of the most popular group. Denote this group by POP_i^K. For each i, the accuracy ACCU_i^K of the cluster CLUST_i^K is computed by

$$\text{ACCU}_i^K = \frac{|\text{POP}_i^K|}{|\text{CLUST}_i^K|},$$

where $|\text{POP}_i^K|$ is the number of points in CLUST_i^K that belong to the most popular group and $|\text{CLUST}_i^K|$ is the total number of points in CLUST_i^K. We denote the accuracy of the K-Means based clusters CLUST^K by

$$\text{ACCU}^K \stackrel{\Delta}{=} \{\text{ACCU}_i^K : 1 \leq i \leq K\}.$$

We repeat this process for every K, $5 \leq K \leq 100$. Finally, we choose the K that achieves the best accurate results. We save the K classifications CLASS^K, the corresponding K centroids CENTR^K and the corresponding K radiuses RADIUS^K as a baseline for the on-line (real-time) classification in Step 2.

Optimizing the clustering accuracy: We recall from the training step that while processing the normalized matrix A for derivation of its embedding matrix Ψ, we use a random weight vector w. We repeat the training step with different weight vectors where each time we use a different random vector w. The weight vector w, which achieves the best accuracy results ACCU^K, is chosen. This vector is saved as a baseline for the real-time classification in Step 2.

The outputs from this training step are the normalized matrix A, the weight vector w, the 10 dimensional embedding matrix Ψ and the parameters of the best (most accurate) K-Means clusters (the corresponding classifications CLASS^K, the centers CENTR^K and the radiuses RADIUS^K). These outputs are the baseline parameters for the real-time (online) classification process described in Step 2.

2. Classification Step:

Real-time processing of a newly arrived data point: Let P be a row vector of size $1 \times n$ that is produced by the traffic analyzer in real-time for every new network flow where n is the number of features that are gathered for every flow. These features were described in Sect. 4.

On-line normalization of the data point P: The vector P is normalized by taking the logarithmic value of each feature. The normalized vector A is constructed by $A = \log(P)$. At the end of this process, the original row vector P is replaced by the normalized row vector A.

Processing the normalized vector A—derivation of its embedding vector ψ: The baseline embedding matrix Ψ, which was saved in the training step, is used. The dimensionality of A is reduced from n to 10. This process uses the application of geometric harmonics, which was described in Sect. 2.1.2, to extend the embedding matrix Ψ with the newly calculated normalized vector A. We use the baseline weight vector w to compute the pairwise distances of A in this process. The output of this process is the extension of the matrix. This extension is the new embedding vector ψ of the newly arrived data point.

Real-time classification of a newly arrived data point ψ: The baseline embedding matrix Ψ, the baseline weight vector w and the baseline K-Means parameters CLASS^K, CENTR^K and RADIUS^K, which are the baseline corresponding classifications, centers and radiuses), respectively, which were saved in the training step (Step 1), are used. These baseline parameters are used to classify the newly arrived data point ψ as either belonging (normal) to the application protocol of one of the training clusters or not (abnormal). The distances between the embedding vector ψ of the newly arrived data point and the K centroids CENTR^K are computed using the weighted Euclidean distance metric and the baseline weight vector w. The newly arrived data point is associated with the cluster of its nearest centroid (out of the K centroids). We denote the distance of ψ from its nearest centroid by DIS_j^K where j is the cluster number that the new data point is associated with. The newly arrived data point is classified as abnormal if its distance from the centroid of its associated cluster is larger than the corresponding radius. Formally, ψ is abnormal if $\text{DIS}_j^K \geq \text{RADIUS}_j^K$. Otherwise, ψ is classified as a normal point and its classification label is CLASS_j^K.

The output of this real-time process is a ten-dimensional embedding vector ψ and a decision mechanism that determines whether the newly arrived data point is normal or abnormal. This classification step determines whether each newly arrived data point is either normal or abnormal.

Our real-time classification algorithm was tested on real datasets from a big enterprise network. The inputs are the statistics of these datasets (as generated by the traffic analyzers). Section 7 presents some experimental results on these datasets after the application of PCR algorithm.

7 Experimental Results

In Sect. 7.1, PCR algorithm was tested on datasets that contain real network traffic that were captured from a big enterprise network (Sect. 3). In Sect. 7.2, it was tested on 16 public datasets from UCI repository (Blake and Merz 1998), and its clustering results were compared to several known clustering algorithms. These datasets are from different domains and the purpose of this experiment is to show that PCR is generic and can fit in different domains.

7.1 Protocol Classification and Recognition

The inputs are the statistics of the sessions of these datasets [as was generated by the flow-oriented traffic analyzer (Sect. 4)]. Section 7.1.1 presents the experimental results on the training dataset that contains 5,500 data points. Section 7.1.2 presents the experimental results on the testing dataset that contains 2,000 data points.

7.1.1 Experimental Results on Training Datasets

This section presents the experimental results on the training dataset that contains 5,500 data points. We recall from Sect. 6 that the dimensionality reduction of the statistics matrix [that was generated by the flow-oriented traffic analyzer (Sect. 4)] involves a computation of the pairwise distances in this matrix. The metric that is used for computing the pairwise distance is one of the following (see Sect. 6.1.2):

- Euclidean (L_2) distance;
- Mahalanobis distance;
- Cosine distance;
- Weighted Euclidean (WL_2) distance.

In this section, we present the experimental results for each metric.

We recall from Sect. 6 that the points in the embedded matrix are clustered using the K-Means clustering for different values of K ($5 \leq K \leq 100$).

We recall from Sect. 6.1.2 that the accuracy $ACCU^i$ of the cluster $CLUST^i$ is computed by

$$ACCU^i = \frac{|POP^i|}{|CLUST^i|}$$

where $|POP^i|$ is the number of points in $CLUST^i$ that belong to the most popular group and $|CLUST^i|$ is the total number of points in $CLUST^i$.

For each distance method and for each K-Means value, we compute the number of clusters where their accuracies in the K clusters are more than a certain accuracy level. We define the ratio between this number and the total number of clusters (K) as the inter-cluster accuracy. In order to evaluate the performance of different distance metrics and different K-Means values we use the following accuracy levels:

1. Accuracy level of 95 %;
2. Accuracy level of 90 %;
3. Accuracy level of 85 %.

Furthermore, we compute the number of flows that are associated to clusters whose accuracy is more than a certain accuracy level. We define the ratio between this number and the total number of flows as the inter-cluster cover.

This section presents the experimental results in the following order:

1. The inter-cluster accuracy results of the classification algorithm according to different distance metrics and according to different K-Means values;
2. The inter-cluster cover results from the classification algorithm according to different distance metrics and according to different K-Means values.

The X-axis in each figure represents different K values and the Y-axis represents the ratio results (accuracy or cover) in percentages.

The Inter-Cluster Accuracy Results from the PCR Algorithm

Figure 6 shows the inter-cluster accuracy results from the PCR algorithm according to the four different distance metrics and according to different K where the accuracy is 95 %.

PCR with Mahalanobis distance and PCR with Weighted Euclidean (WL_2) distance achieve the best results in terms of accuracy. PCR with Cosine distance achieves less accurate results while PCR with Euclidean (L_2) distance achieves the worst results in terms of accuracy.

Figure 7 shows the inter-cluster accuracy results from the PCR algorithm according to the four different distance metrics and according to different K values where the accuracy is 90 %.

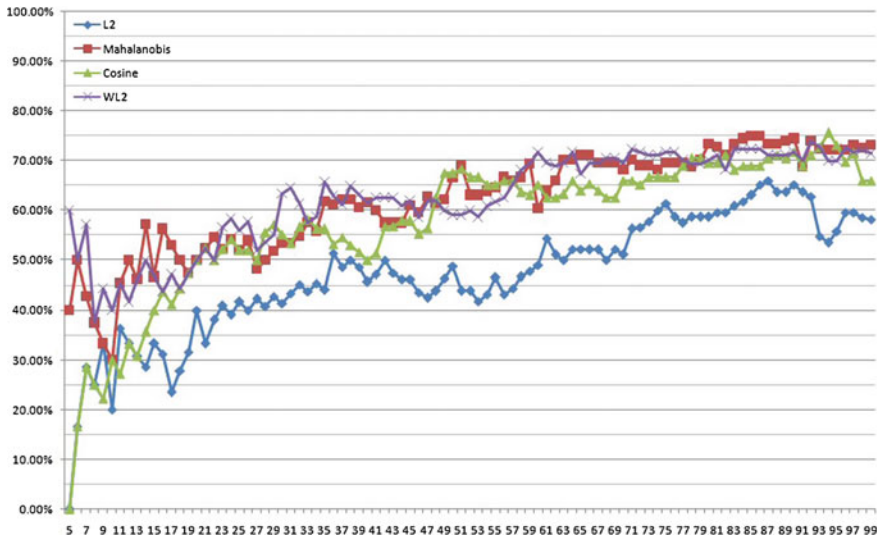

Fig. 6 Protocol classification and recognition: inter-cluster accuracy results with accuracy level of 95 %

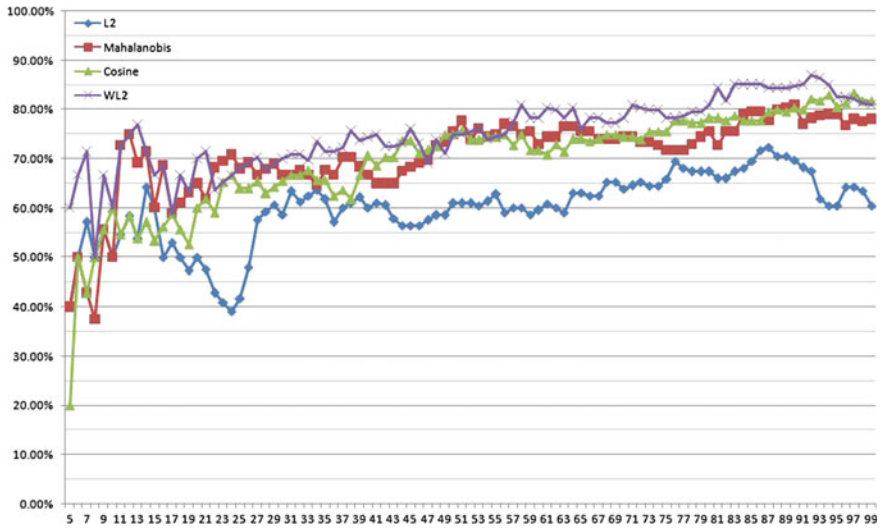

Fig. 7 PCR performance: inter-cluster accuracy results with accuracy level of 90 %

For the first 35 clusters, PCR with Mahalanobis distance and PCR with Weighted Euclidean (WL_2) distance achieve the best results in terms of accuracy. PCR with Cosine distance achieves less accurate results while PCR with Euclidean (L_2) distance achieves the worst results in terms of accuracy. From cluster 35 to

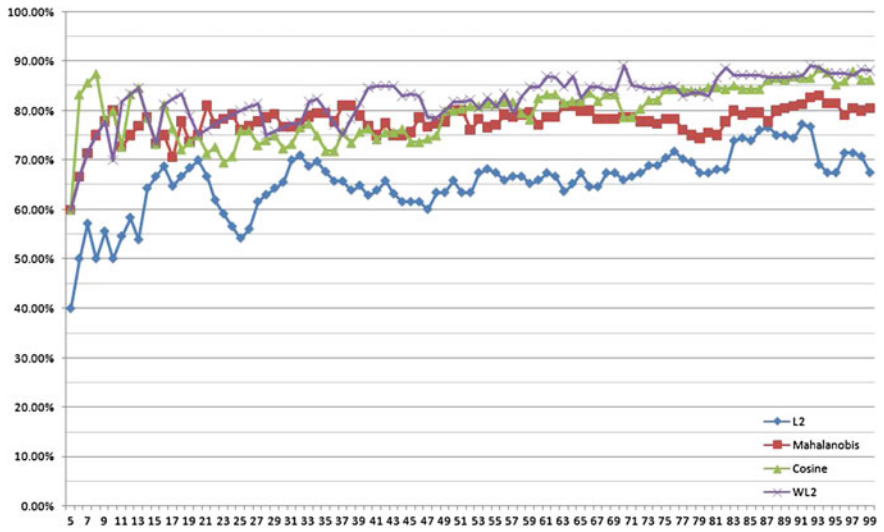

Fig. 8 Protocol classification and recognition: inter-cluster accuracy results with accuracy level of 85 %

cluster 50, PCR with Weighted Euclidean (WL_2) distance achieves the best results in terms of accuracy. From cluster 50 to cluster 60, there is no significant difference between the Mahalanobis, Cosine and Weighted Euclidean (WL_2) distances while PCR with Euclidean (L_2) distance still achieves the worst results. From the 60th cluster, PCR with Weighted Euclidean (WL_2) distance achieves the best results in terms of accuracy.

Figure 8 shows the inter-cluster accuracy results from the PCR algorithm according to the four different distance metrics and according to different K values where the accuracy is 85 %.

For the first 20 clusters, there is no significant difference between the Mahalanobis, Cosine and Weighted Euclidean (WL_2) distances while PCR with Euclidean (L_2) distance achieves the worst results in terms of cover. From cluster 20 to cluster 40, PCR with Mahalanobis distance and PCR with Weighted Euclidean (WL_2) distance achieve the best results in terms of accuracy. PCR with Cosine distance achieves less accurate results while PCR with Euclidean (L_2) distance achieves the worst results in terms of accuracy. From cluster 40 to cluster 50, PCR with Weighted Euclidean (WL_2) distance achieves the best results in terms of accuracy. From cluster 50 to cluster 60, there is no significant difference between the Mahalanobis, Cosine and Weighted Euclidean (WL_2) distances while PCR with Euclidean (L_2) distance still achieves the worst results. From the 60th cluster, PCR with Weighted Euclidean (WL_2) distance achieves the best results in terms of accuracy.

The Inter-Cluster Covers Results from the Classification Algorithm

Figure 9 shows the inter-cluster cover results from the PCR algorithm according to the four different distance metrics and according to different K values where the accuracy of the clusters is 95 %.

For the first 20 clusters, PCR with Weighted Euclidean (WL_2) distance and PCR with Mahalanobis distance achieve the best results in terms of cover. PCR with Cosine distance and PCR with Euclidean (L_2) distance achieve the worst results. From cluster 20 to cluster 50, PCR with Weighted Euclidean (WL_2) distance achieve the best results in terms of cover. From the 50th cluster, PCR with Mahalanobis distance achieves the best results.

Figure 10 shows the inter-cluster accuracy results from the PCR algorithm according to the four different distance metrics and according to different K values where the accuracy of the clusters is 90 %.

For the first 20 clusters, PCR with Weighted Euclidean (WL_2) distance and PCR with Mahalanobis distance achieve the best results in terms of cover. PCR with Cosine distance achieves the worst results. From cluster 20 to cluster 50, PCR with Weighted Euclidean (WL_2) distance achieve the best results in terms of cover while PCR with Euclidean (L_2) distance achieve the worst results. From the 50th cluster, there is no significant difference between the Mahalanobis, Cosine and Weighted Euclidean (WL_2) distances while PCR with Euclidean (L_2) distance still achieves the worst results.

Figure 11 shows the inter-cluster cover results from the PCR algorithm according to the four different distance metrics and according to different K values where the accuracy of the clusters is 85 %.

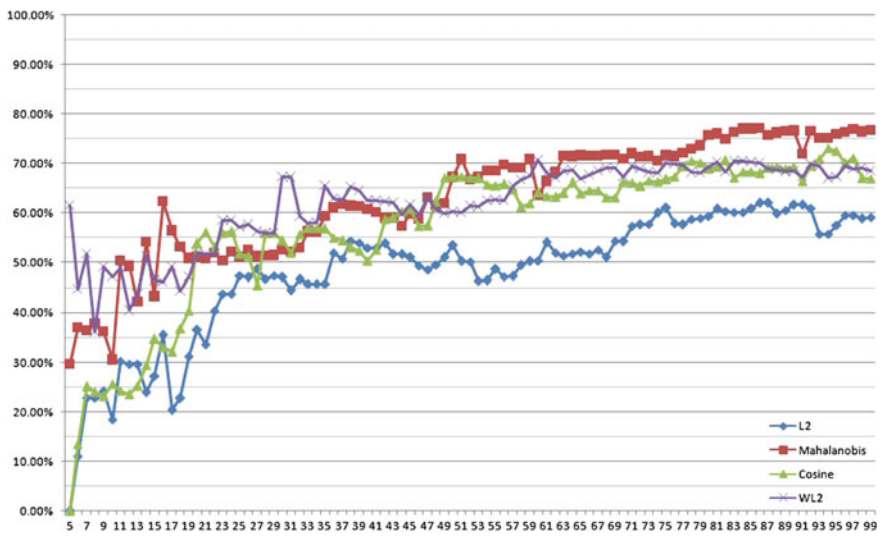

Fig. 9 PCR performance: inter-cluster cover results with accuracy level of 95 %

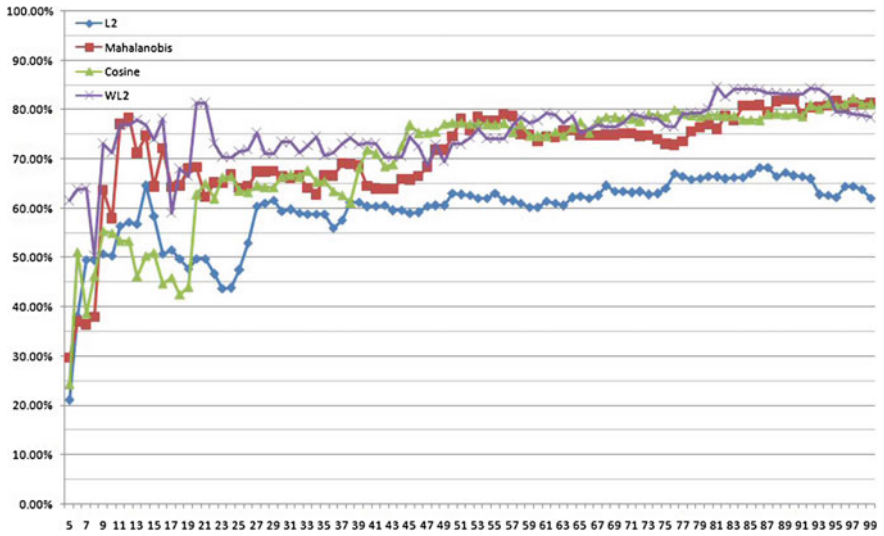

Fig. 10 PCR performance: inter-cluster cover results with accuracy level of 90 %

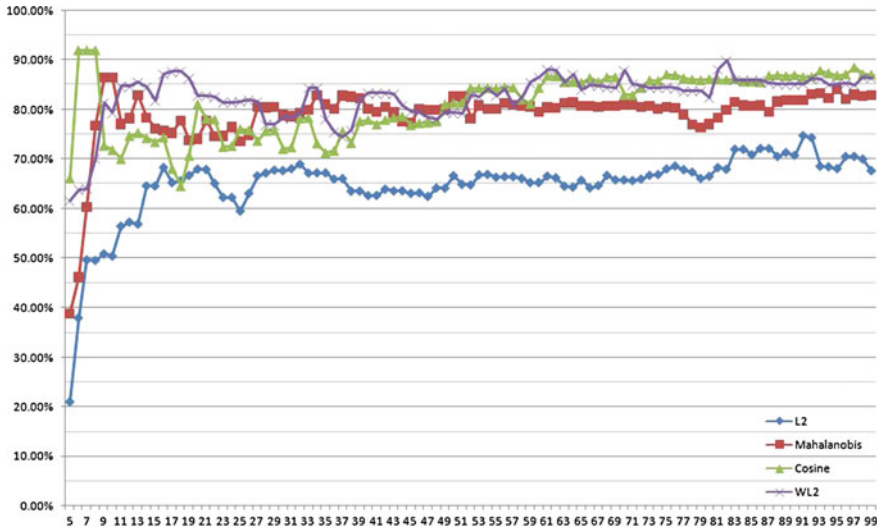

Fig. 11 PCR performance: inter-cluster cover results with accuracy level of 85 %

For the first 30 clusters, PCR with Weighted Euclidean (WL_2) distance and PCR with Mahalanobis distance achieve the best results in terms of cover. PCR with Euclidean (L_2) distance achieves the worst results. From cluster 30 to cluster 50, there is no significant difference between the Mahalanobis, Cosine and Weighted

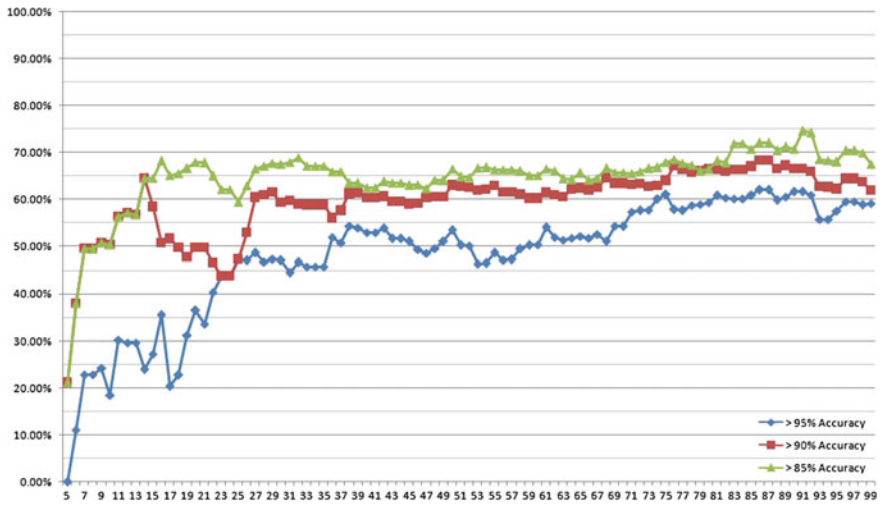

Fig. 12 PCR performance: inter-cluster cover results for the Euclidean distance metric

Euclidean (WL_2) distances while PCR with Euclidean (L_2) distance still achieves the worst results. From the 50th cluster, PCR with Weighted Euclidean (WL_2) distance and PCR with Cosine distance achieve the best results in terms of cover.

Figure 12 shows the inter-cluster cover results from the PCR algorithm according to the Euclidean distance metric and according to different K values where the accuracy of the clusters is 95, 90 and 85 %, respectively.

For the first 15 clusters, the cover results for clusters with accuracy of 90 and 85 % is identical while the results for clusters with accuracy of 95 % is the worst. For clusters 15 to 40, the cover results for clusters with accuracy of 85 % is the best while the results for clusters with accuracy of 95 % is still the worst. From the 40th cluster, there is no significant difference between the cover results for clusters with accuracy of 90 and 85 % while the results for clusters with accuracy of 95 % is the worst.

Figure 13 shows the inter-cluster cover results from the PCR algorithm according to the Mahalanobis distance metric and according to different K values where the accuracy of the clusters is 95, 90 and 85 % respectively.

For the first 80 clusters, the cover results for clusters with accuracy of 85 % is the best while the results for clusters with accuracy of 95 % is the worst. From the 80th cluster, there is no significant difference between the cover results for clusters with accuracy of 90 and 85 % while the results for clusters with accuracy of 95 % is still the worst.

Figure 14 shows the inter-cluster cover results from the PCR algorithm according to the Cosine distance metric and according to different K values where the accuracy of the clusters is 95, 90 and 85 %, respectively.

The cover results for clusters with accuracy of 85 % is the best while the results for clusters with accuracy of 95 % is the worst.

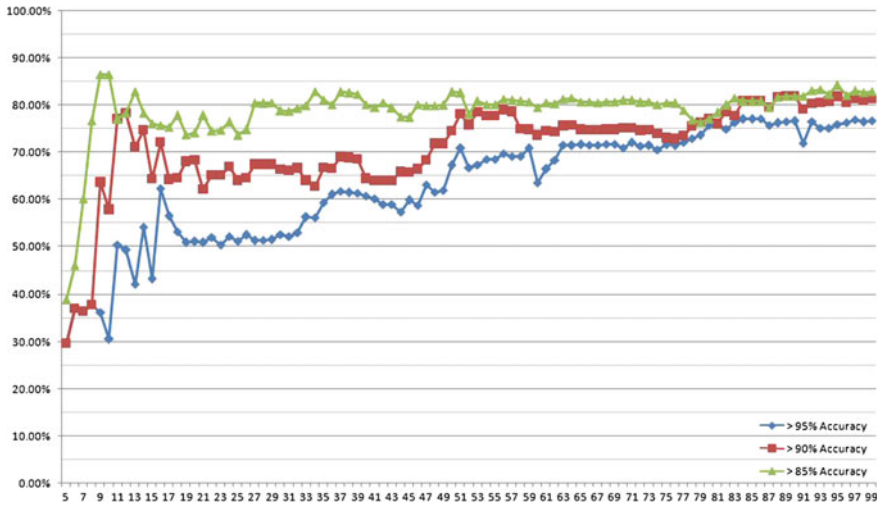

Fig. 13 PCR performance: inter-cluster cover results for the Mahalanobis distance metric

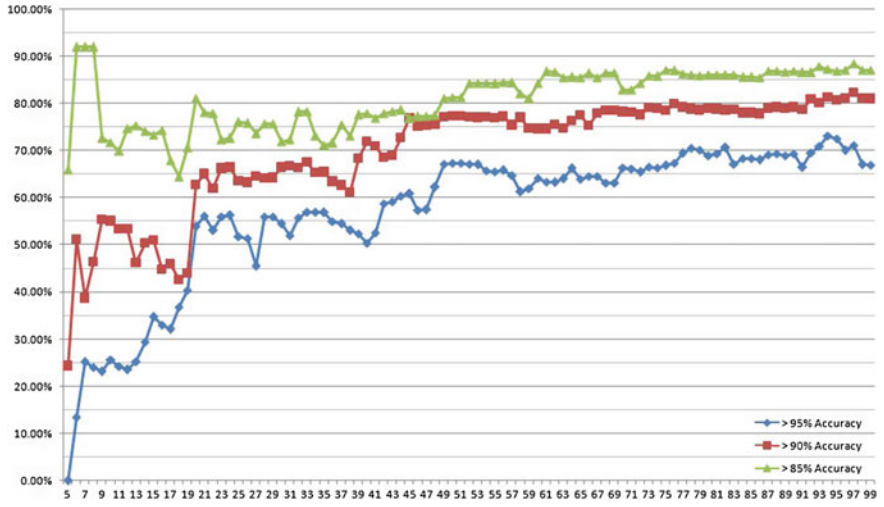

Fig. 14 PCR performance: Inter-cluster cover results for Cosine distance metric

Figure 15 shows the inter-cluster cover results from the PCR algorithm according to the Weighted Euclidean distance metric and according to different 95 K values where the accuracy of the clusters is 95, 90 and 85 % respectively.

For the first 80 clusters, the cover results for clusters with accuracy of 85 % is the best while the results for clusters with accuracy of 95 % is the worst. From the 80th cluster, there is no significant difference between the cover results for clusters with accuracy of 90 % and 85 % while the results for clusters with accuracy of 95 % is still the worst.

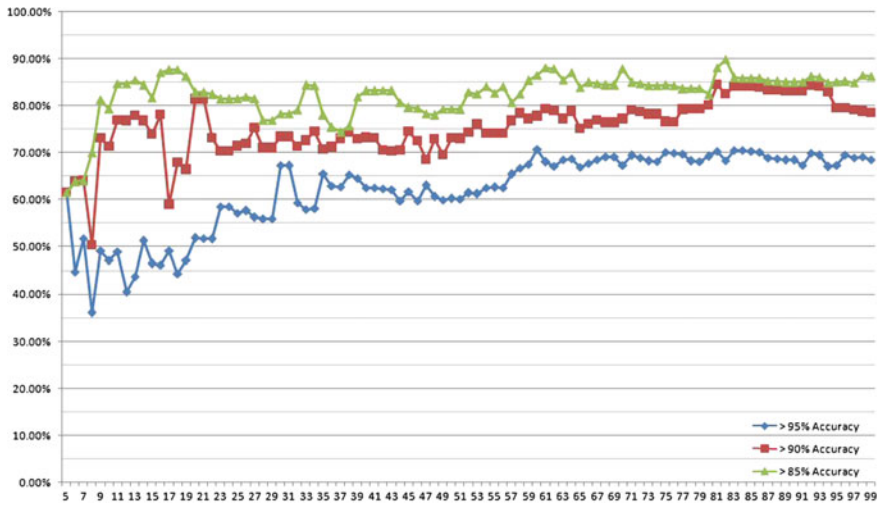

Fig. 15 PCR performance: inter-cluster cover results for Weighted Euclidean distance metric

7.1.2 Experimental Results on Testing Datasets

This section presents the experimental classification results on 2,000 data points. The baseline profile, which is used for the classification, was generated by the PCR algorithm in the training phase. The baseline profile was built by using the Weighted Euclidean distance metric where $K = 55$.

In the classification phase, each flow (data point) in the testing dataset was associated to one of the baseline clusters according to the PCR algorithm. Classification rate was computed as the ratio between the number of testing flows, which were classified successfully, and the total number of testing flows. The mis-classification rate was computed as the ratio between the number of testing flows, which were classified unsuccessfully, and the total number of testing flows. **97 %** from the testing flows were classified successfully while only 3 % from the testing flows were mis-classified. 2.8 % of the miss-classified flows were associated to clusters that include their protocol type.

7.2 UCI Datasets

We used in our experiments 16 public datasets from UCI repository (Blake and Merz 1998). These datasets belong to wide variety of domains and problems in data mining and machine learning. Table 1 shows the properties of each dataset. The first

Table 1 The properties of the datasets that were used to analyze the performance of the algorithms

Dataset name	Description	Type	Num. of dimensions	Num. of classes	Num. of samples
Iris	Sizes of Iris plants	Numerical	4	3	150
Wine	Chemical analysis of wines	Numerical	13	3	178
Glass	Elements and characteristics of glasses	Numerical	9	7	214
Segmentation	Image segmentation data	Numerical	19	7	2310
Ionosphere	Radar data	Numerical	34	2	351
Heart SPECTF	Diagnosis of cardiac Single Proton Emission Computed Tomography (SPECT) images	Numerical	44	2	80
Ecoli	Protein localization sites	Numerical	7	8	336
Yeast	Protein localization sites	Numerical	8	10	1484
Sat	Multi-spectral values of pixels in a satellite image	Numerical	36	6	4435
Pageblock	Blocks of the page layout of documents	Numerical	10	5	5473
Soybean	Soybean characteristics	Nominal and numeric	33	19	310
Dermatology	Clinical and histopathological characteristics of patients	Nominal and numeric	33	6	366
Adult	Data from the census bureau database	Nominal and numeric	14	2	30,612
Heart Cleveland	Diagnosis from Cleveland clinic foundation	Nominal and Numeric	13	5	303
Zoo	Characteristics of animals	Nominal and numeric	17	7	101
Vowel	LPC derived log area ratio coefficients	Nominal	12	11	528

column in the table presents the names of the datasets. The second column presents the description of the dataset. The third column presents the type of the dataset: some datasets contain numerical data, some contain nominal data and some contain mix of numerical and nominal data. The fourth column presents the number of dimensions (features) in each dataset: the lowest number of dimensions is 4 and the highest is 44. The fifth column presents the number of classes in each dataset: the minimal number of classes is 2 and the maximal number is 19. The sixth column presents the number of samples in each dataset: the minimal number of samples is 80 and the maximal number is 30,612.

Table 2 Descriptions of the compared algorithms where num = numerical and nom = nominal

Algorithm name	Description	Refs.	Data type
Random	Each point is randomly associated to one of the k clusters—this method is used as a reference point, in order to estimate the performance of the clustering algorithms	–	–
PCR	Our proposed method in this paper	–	Num
k-means	One of the most used clustering algorithm that was designed to cluster numerical data (see Sect. 2)	Lloyd (1982)	Num
k-means++	An algorithm for choosing the initial values for k-means clustering. It initializes the cluster centers before proceeding with the standard k-means optimization iterations (see Sect. 2)	Arthur and Vassilvitskii (2007)	Num
Kernel k-means	Kernel k-means maps the data to a higher dimension feature space using a non-linear function. Then, it partitions the points by linear separations in the new space (see Sect. 2)	Zhang and Rudnicky (2002)	Num
PCA k-means	First, we used PCA in order to reduce the dimensionality of the data by projecting the data onto the first principal components. Then, we cluster the projected data using k-means	Pearson (1901) and Lloyd (1982)	Num
Birch	An agglomerative hierarchical algorithm that is used for clustering large numerical datasets	Zhang et al. (1996)	Num
k-modes	Emerged from the k-means algorithm and it was designed to cluster categorical datasets	Huang (1998)	Nom
k-prototypes	The k-prototypes algorithm integrates the k-means and k-modes algorithms to enable clustering of mixed data	Huang (1998)	Mix
Gower k-means	The k-means algorithm applied to the Gower similarity index to enable clustering of mixed data	Lloyd (1982) and Gower (1971)	Num, nom, mix

In order to evaluate the performance of PCR, we compare its clustering results to several known clustering algorithms. Table 2 describes each of these clustering algorithms. The last column in this table presents the type of the data that each algorithm is designed to cluster: some algorithms were designed to cluster numerical data, some nominal data and some mix of numerical and nominal data.

For each algorithm we measured the accuracy (purity) index (Tan et al. 2005).

Table 3 presents the accuracy results from the application of the different algorithms to the evaluation datasets. For each clustering algorithm and each dataset, we chose k as the number of classes in the dataset (according to Table 1).

Application of the random clustering to Ionosphere and Pageblock datasets achieved the same results as all the algorithms achieved (on these two datasets). Although the main purpose of PCR was to classify and recognize network data and protocols, PCR achieved best results for 5 out of the 14 remaining datasets.

Table 3 Overall accuracy results: comparison between PCR and different algorithms

	Iris	Wine	Glass	Segmentation	Ionosphere	Heart SPECTF	Ecoli	Yeast	Sat	Pageblock	Soybean	Dermatology	Adult	Heart Cleveland	Zoo	Vowel
PCR	***0.9***	0.71	***0.65***	0.62	0.89	***0.75***	***0.87***	0.53	***0.73***	0.9	0.68	0.41	0.75	0.67	0.87	0.34
Random	0.39	0.41	0.41	0.19	***0.89***	0.56	0.45	0.35	0.26	***0.9***	0.24	0.32	0.74	0.55	0.43	0.17
k-means	0.89	0.7	0.59	0.59	0.89	0.68	0.86	0.53	***0.73***	0.9	0.74	0.42	0.75	0.56	0.86	0.29
k-means++	0.89	0.7	0.6	0.59	0.89	0.68	0.85	0.53	***0.73***	0.9	***0.73***	0.37	0.75	0.59	0.86	0.31
Kernel *k*-means	0.96	0.5	***0.65***	0.24	0.89	0.71	***0.87***	***0.56***	0.28	0.89	0.68	0.41	0.74	0.6	0.89	0.31
PCA *k*-means	0.89	0.7	***0.65***	0.62	0.89	0.68	0.79	0.49	***0.73***	0.9	0.6	0.43	0.75	0.56	0.86	0.29
BIRCH	0.89	0.7	0.58	0.63	0.89	0.66	0.86	0.53	***0.73***	0.9	0.76	0.43	0.75	0.55	0.83	0.31
k-modes	0.59	0.46	0.47	0.24	0.89	0.64	0.47	0.38	0.48	0.9	0.63	0.8	***0.81***	***0.83***	***0.9***	0.15
k-prototypes	–	–	–	–	–	–	–	–	–	–	–	0.77	0.75	0.67	0.88	***0.42***
Gower *k*-means	0.88	***0.96***	0.56	***0.64***	0.89	0.56	0.82	0.53	***0.73***	0.9	0.59	***0.85***	0.79	0.58	0.88	0.17

Bolded italic numbers represent the best achieved accuracy

8 Conclusion

We presented a unique framework that is based upon diffusion processes and other methodologies for finding meaningful geometric descriptions in high-dimensional datasets. We showed that the eigenfunctions of the generated underlying Markov matrices can be used to construct diffusion processes that generate efficient representations of complex geometric structures for high-dimensional data analysis.

Our proposed methods automatically classify and recognize network protocols with high accuracy and classification rate, while minimizing the mis-classification rate. In addition, it showed good results while testing on general purpose clustering tasks.

References

Arthur D, Vassilvitskii S (2007) k-means++: the advantages of careful seeding. In: SODA '07 Proceedings of the eighteenth annual ACM-SIAM symposium on discrete algorithms, Philadelphia, PA. SIAM, pp 1027–1035

Blake CL, Merz CJ (1998) UCI repository of machine learning databases. University of California, Department of Information and Computer Science. http://www.ics.uci.edu/mlearn/MLRepository.html

Chung FRK (1997) Spectral graph theory, volume 92 of CBMS Regional Conference Series in Mathematics. AMS

Coifman RR, Lafon S (2006a) Diffusion maps. Appl Comput Harmon Anal 21(1):5–30

Coifman RR, Lafon S (2006b) Geometric harmonics: a novel tool for multiscale out-of-sample extension of empirical functions. Appl Comput Harmon Anal 21(1):31–52

Coifman RR, Lafon S, Lee AB, Maggioni M, Nadler B, Warner F, Zucker SW (2005) Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. Proc Natl Acad Sci USA 102(21):7426–7431

David G (2009) Anomaly detection and classification via diffusion processes in hyper-networks. PhD thesis, Tel Aviv University

Gower JC (1971) A general coefficient of similarity and some of its properties. Biometrics 27 (4):857–871

Huang Z (1998) Extensions to the *k*-means algorithm for clustering large data sets with categorical values. Data Min Knowl Discov 2(3):283–304

Kohonen T (1990) The self-organizing map. Proc IEEE 78(9):1464–1480

Lafon S, Keller Y, Coifman RR (2006) Data fusion and multicue data matching by diffusion maps. IEEE Trans Pattern Anal Mach Intell 28(11):1784–1797

Lloyd S (1982) Least squares quantization in PCM. IEEE Trans Inform Theory 28(2):129–137

MacQueen JB (1967) Some methods for classification and analysis of multivariate observations. In: Proceedings of the 5th Berkeley symposium on mathematical statistics and probability, vol 1, Berkeley, CA. University of California Press, pp 281–297

Nadler B, Lafon S, Coifman RR, Kevrekidis IG (2006) Diffusion maps, spectral clustering and reaction coordinates of dynamical systems. Appl Comput Harmon Anal 21(1):113–127

Pearson K (1901) On lines and planes of closest fit to systems of points in space. Philos Mag 2(6):559–572

Tan P-N, Steinbach M, Kumar V (2005) Introduction to data mining. Addison-Wesley, Boston

Zhang R, Rudnicky AI (2002) A large scale clustering scheme for kernel *k*-means. In: Proceedings of the 16th international conference on pattern recognition (ICPR 02), vol 4, New York. IEEE, pp 289–292

Zhang T, Ramakrishnan R, Livny M (1996) BIRCH: An efficient data clustering method for very large databases. SIGMOD Rec 25(2):103–114

Timing and Side Channel Attacks

Nezer Zaidenberg and Amit Resh

Abstract How would you know the US pentagon is planning an attack on Iraq? One possible plan is to infiltrate the pentagon using spies, flipping traitors etc. But this sounds like lots of work and it is a dangerous work. That is the direct approach. Another possible plan is to ask the pizza delivery guys in the area. People planning an attack probably adds up to lots of people urgently trying to meet deadlines, staying late in the office and ordering pizza. So the pizza delivery guys know about a pending attack! The pizza delivery guys do not know the nature of the attack but they know "something is up" in the pentagon because for a few days people are staying late at the office and ordering pizza at irregular hours. The pizza approach is the side-channel attack. This attack on the pentagon is not a direct channel attack. No spies were used. No attack on the pentagon defences. It is a side channel attack. Attack on the side effects of planning something. The people who plan need to work extra time and they also need to eat.

1 Introduction

In computing security there are numerous side channel attacks on side effects of verifying efficacy. These are not attacks that are designed on the primary line that was protected but on its side effects. A trivial such example is overcoming smartphone PIN protection on stolen smartphones. Normally there are 10,000 PIN combinations. However if the attacker can decipher the PIN digits by studying the smudge left on the device by the owner's fingers the problem can become a 1:24 problem, which is significantly easier (Genkin et al. 2013).

N. Zaidenberg (✉) · A. Resh
Department of Mathematical Information Technology,
University of Jyväskylä, Jyväskylä, Finland
e-mail: nezer@trulyprotect.com

A. Resh
e-mail: amit@trulyprotect.com

M. Lehto and P. Neittaanmäki (eds.), *Cyber Security: Analytics, Technology and Automation*, Intelligent Systems, Control and Automation: Science and Engineering 78, DOI 10.1007/978-3-319-18302-2_11

This type of attack is very powerful. Schneier (2012) estimates that the NSA is able to break AES encryption using side-channel attack. Though no such method has been published.

2 Hypervisor Blue Pills and Red Pills

2.1 Subverting and Blue Pill Concept

A hypervisor is a type of software that allows running multiple operating systems on single hardware. The hypervisor treats a guest operating system in a similar fashion to the way an operating systems treats processes. The hypervisor manages the memory map for multiple operating systems in a similar fashion to the MMU of operating systems for processes. The hypervisor has a similar MMU for I/O devices, so one operating system is not effected by the I/O of another OS.

x86 allows multiple protection rings, the 2 most commonly used is Ring 0 for the Operating System (supervisor mode) and ring 3 for user code. Recently, x86 added ring-1 protection, a ring for the hypervisor, which adds instructions to create and trap OS operations. There are two different instruction sets for AMD and Intel hardware (AMD-v and VT-x instructions) but the two instruction sets are mostly polymorphic.

There are two possible hypervisors in x86 environments. Type 1(bare metal hypervisor) works directly over the hardware. The machine boots directly to the hypervisor. The hypervisor can then be used to boot other operating systems. We do not deal with type 1 hypervisor.

Type 2 hypervisor, which interests us, starts as a process (in ring 3) under an Operating system (The operating system boots over the hardware normally). After the process starts it executes instructions that allow it to become a hypervisor. Thus the hypervisor gains more permissions then the OS that started it. After starting (as a ring 3 process) the hypervisor transfers the OS that started the process to a type of guest (getting into ring-1) (Fig. 1).

Rutkowska (2006) introduced the “bluepill” concept, a hypervisor which is barely noticeable. The hypervisor starts as a user process but gains complete control of the system (getting the user to run the process in the first place is a different problem). Rutkowska describes several such methods for example by hotel maids attacking hotel guest laptops etc.

2.2 Local Hypervisor Red Pills—Direct and Sub-channel Attack

How can the user know he is not running some hypervisor rootkit (such as the blue pill) on his computer? Direct attacks on the Bluepill involve actively trying to attack

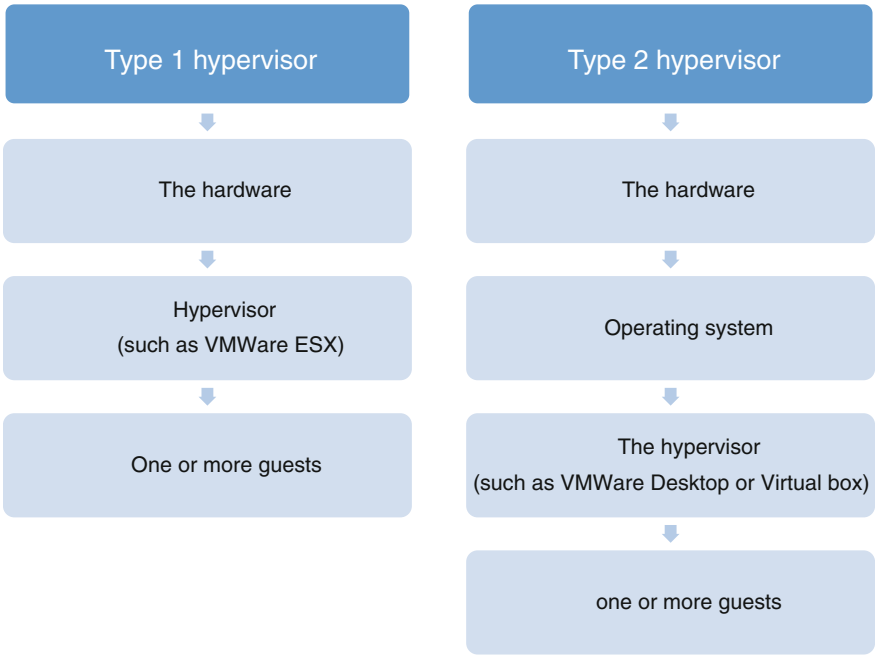

Fig. 1 Type 1 and type 2 hypervisor

the hypervisor for effects that may not be hidden well or calling instructions that should be allowed if no hypervisor is installed but should be prevented if an hypervisor is already running.

It has been shown that the Bluepill can be developed to hide masking detections attempts very well by direct means and even allowing infinitely recursive hypervisor calls. This led to sub-channel indirect attacks on the blue pill by measuring side effects of running the hypervisor (Rutkowska 2006).

Perhaps attacks in this context are not clear. As it is the hacker that tries to install an invisible root kit the attacks in this context are by the lawful user who tries to detect and remove the "invisible" rootkit.

While a malicious hypervisor can mask itself from an attacker there are certain attacks that are bound to cause side effects. For example calling the CPUID assembler instruction should normally take about 200 cycles. But if a hypervisor is involved it will consume roughly 5000 cycles as a result of CPU context switches between guest and hypervisor mode (CPUID always exits to the hypervisor).

Thus an attack on a naïve bluepill could be:

1. Measure CPU clock tick
2. Call CPUID
3. Measure CPU clock tick
4. If difference between 1 and 3 is greater the 400 cycles warn about blue pill

Off course as hypervisor based rootkits become more complex this attack can also be prevented. For example by trapping the call to get the CPU time.

There now exists a cat and mouse game in which more side effects can be measured (for example cache misses) and more attacks can be prevented by the hypervisor which cause more side effects that can also be attacked (Rutkowska and Tereskin 2007).

2.3 Remote Hypervisor Red Pills

The method described above demonstrates how the local computer can detect a malicious hypervisor rootkit as well as means that the hypervisor can use to hide itself better by catching more and more instructions.

The problem is that the hypervisor can control the guest, tricking it into believing the hypervisor does not exist. Also since red pill tools are offered to users, malicious hypervisor authors can reverse engineer any red pill and build a hypervisor that manages to mask itself against said red pill. However we can actually eliminate the cat and mouse game.

The problem lies with our check that is made locally using measuring tools that are all under the hypervisor control. However, the user can run his sanity test not locally, but using a cloud server.

The hypervisor doesn't control the cloud thus a 3rd party can efficiently detect if hypervisor is running. (It may be possible for the hypervisor to control the cloud response as it is caught by the OS but it is also possible for the cloud server to inform the user he is running a malicious rootkit using sub channel that the hypervisor does not control such as another computer).

Kennell and Jamieson (2003) have suggested a method for remote verification of the genuiness of a virtual machine.[1] Kennel et al. also include a mechanism to exchange an encrypted key with the authenticated host which was removed from our summary (Shankar et al. 2004).

The jest of the kennel method can be summarized as following:

1. The cloud generates a random test. Tests are not identical and contain multiple steps.
   In each step side effects from the previous step are entered as input to the new step.
2. The test executes in the inspected host.
3. The inspected host sends a response.
4. The cloud serve verifies the response as well as the time it took to generate it
5. If the test was successful and within an allotted time the cloud server concludes that the host is genuine.

[1]The papers uses the term "genuinity" however the correct English term is genuiness. We will use the correct English term in this chapter.

The steps in Kennell Memory Test include scans of the memory and instructions that run on the inspected OS. The side effects include TLB misses and other effects that are bound to produce different side effects if a malicious hypervisor is running.

Kennell and Jamieson (2003) argue that if the host is running some hypervisor there are bound to be different side effects. If the host is running an efficient emulator that also emulates all side effects the response will take too long to arrive.

3 Invisible Character Differences

Let's assume we have some login screen (with username and password) (Fig. 2).

The login can fail for many reasons like:

- Username does not exist
- Username is expired
- Username is locked
- Password is incorrect

Of course for an attacker the different cases call for different behaviour. If the username is wrong there is no future with said username. If the username has expired the attacker may try again later. If the username is locked it is possible that the attacker activities have been detected. It is also possible to try later. Of course if

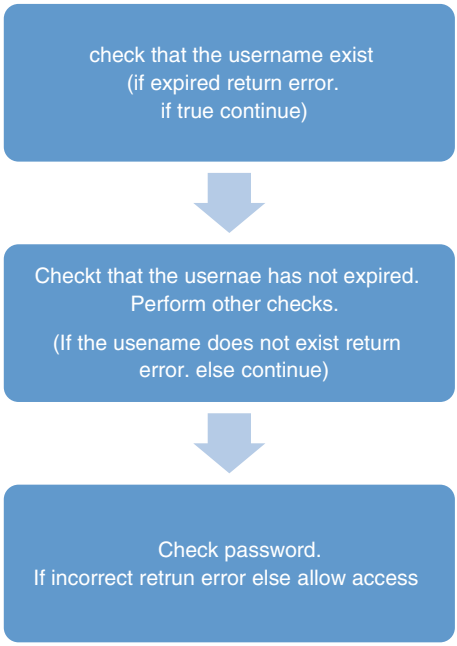

Fig. 2 Username and password failure process

the password is incorrect the attacker now has a correct username. The attacker can used the correct username for guessing the password (using dictionaries or brute force) or use the correct username for other attacks.

Assuming an attacker only wish to find out the correct username it would be critical to have all failure screens look identical. As by having a different failure screen for each case—guessing the correct username by brute force would be possible.

However, if the screens look identical, it is possible that several "invisible" differences exist. For example, if the communication is http communication transmitting a webpage, different web properties (setting cookies etc.) may exist for each of the cases.

4 Timing Attacks

Timing attacks occur when measuring the time it takes a system to respond. If the timing to receive a reply varies, not because of random variance in the delivery medium (such as the network time) but due to differences in responses and failure (for example different types of failure), an attacker can use the time variance to realize what type of failure has occurred. The information obtained using time measurements can later be used to attack the system.

4.1 GameCube DVD Password Attack

The GameCube (Nintendo game console) was not supposed to play recorded DVD (pirated copies). The original DVD was released with copy protection system that was not trivial to replicate. Pirated copies were supposed to be detected by the copy protection system and be rejected by the system.

The DVD however had a programmable override (modchip) for the protection. The override may have been used by Nintendo in their system development. Had the password remained hidden we would never have heard about it.

However Nintendo checked the password using memcmp comparing byte after byte. The verification process ended when the first incorrect password byte was detected (returning failure) or when all bytes were compared successfully. Thus if the password got the first byte correctly the password check will be just slightly longer (checking two bytes instead of one) then if we got the first byte wrong (checking only one byte). Using this method one by one, all bytes that encompass the password can be revealed.

Thus the password was indeed leaked and 2nd generation GameCube modchips appeared (Domke 2004). Even though this method was well known, Nintendo had an identical problem with the Wii, which was released 5 years later (Domke 2006).

Assume a webserver has a password protected section were username and password are required to login. When the user types his or her username and password, the algorithm from Sect. 2 occurs. The System first checks for the username. If no such User exists (or if the user has expired) the system returns with an error. If the user is OK then the system now verifies the password. If the password is not OK then an error message will appear.

Assuming the answer is immediate, by timing the response times an attacker can use this timing attack to reveal correct usernames. Furthermore, even if the response occurs over some network which adds random delay (but similar random delay to both correct username and incorrect username—an attacker may still be able to guess the password (Domke 2004).

Adding short random delays to password checking does not prevent timing attacks. As long as the delays are not significant it can be shown that an attacker can still distinguish between two classes such as incorrect username and correct username but incorrect password (Domke 2004).

5 AES Side-Channel Attacks

5.1 AES Background

AES (Advanced Encryption Standard) is the standard electronic-data symmetric-key encryption algorithm, specified by the NIST (US National Institute of Standards & Technology) since 2001 (AES 2001). It was labelled by the NIST as FIPS publication 197 and is based on the Rijndael algorithm, proposed by two cryptographers from Belgium: Daemen and Rijmen (2013). In addition to being adopted by the US Government it is also used for data transfer and communication worldwide as a successor to the Triple-DES, DES and RCx cryptographic algorithms. This finds use in many implementations, most notably the SSL3/TLS protocol, as well as disk encryption and authentication.

AES is used to encrypt and decrypt fixed blocks of 128 bits (16 bytes). The cryptographic algorithm uses three possible key sizes: 128 bits, 192 bits or 256 bits. AES encryption and decryption is performed in several iterations, called “Rounds”. During each round, 4 steps (only 3 steps in the last Round) are performed on the intermediate data block to progress the encryption from a 16-byte plaintext (*ptext*) to a ciphertext (*ctext*). During decryption, similar steps are performed to achieve the opposite: decrypting the *ctext* to restore the *ptext*. The number of rounds used depends on the key size: 10, 12 and 14 rounds are used for key sizes of: 128, 192 and 256 bytes.

The encryption process can be summarized as follows:

A. Key expansion: The original 128-bit key is expanded to 10, 12 or 14 Round-keys. The data block in each round is combined with the Round-key corresponding to that Round.

B. Initial round: Each *ptext* byte is combined with the original key
C. Rounds (all but last): activate 4 transformations on the data buffer: SubBytes; ShiftRows; MixColumns; Combine with round-key
D. Last round: activate 3 transformations: SubBytes; ShiftRows; Combine with round-key

The decryption procedure is similar, using the same original key but inverse-transformations.

5.2 AES Software Implementation

Software implementations of AES normally make use of lookup tables in favour of performance and efficiency. The lookup tables are used to quickly determine the transformation results, which are activated in the AES rounds. While theoretically it is possible to calculate the transformations without resorting to lookup tables, using only arithmetic, logic and Boolean operations, this carries a significant performance toll that is naturally avoided.

Accessing lookup-tables that have a substantial size (as is the case for AES encryption/decryption) interacts with the underlying architecture's cache mechanism. As we shall see below, this provides an opening for a side-channel attack crafted to reveal the key.

5.3 Cache Memory

Modern CPU systems have several levels of storage. They differ generally by a capacity versus access-speed trade-off.

Cache memory is used to buffer memory contents obtained during a memory cycle, so it can be provided much faster when it is needed in a following memory cycle. Cache operations are generally transparent to software and dedicated hardware is used to manage the buffering and utilization cycles of cache.

Cache memory is subdivided into cache-lines. When a memory element is stored in cache, an entire cache-line (which contains that memory element) is stored in cache. The cache circuitry kicks in at every physical memory access.

Cache operation during a memory read cycle is best explained with an example:

> When a memory location is read-in by the processor, the cache is first inspected to determine if the required value can be obtained from the cache. If it can, this event is called a "cache hit" and the value is provided to the CPU directly from the cache. Such a memory cycle is significantly faster than retrieving the value from main memory. If the required value is not in cache, it is called a "cache miss" and the value must be retrieved from main memory. In this case the value is both provided to the CPU and stored in cache. The cache-miss event causes an entire cache-line to be retrieved from memory and stored to cache,

called a cache-line fill. The next access to any memory location within that cache-line will be a cache-hit.

When a cache-miss occurs and, as explained above, a cache-line is retrieved and stored in cache—some previously stored cache-line needs to be evicted from the cache to make room for the new one. Usually cache-lines are evicted according to an LRU (Least-Recently-Used) algorithm.

As mentioned above, memory elements are stored to the cache in integral quantities of cache-lines. The address of a memory element is subdivided into 3 fields. The *index* field determines which cache-slot will be used. Note that all address locations that contain the same *index* value, share the same cache-slot. Each cache-slot also stores the *tag*, in addition to the cache-line content. The tag is used to define the exact memory location of the cache-line that is currently in the cache-slot. The LSBits are the *offset* field and define the offset of the memory element within the cache-line.

This mapping is referred to as the 'Cache Association'. See Fig. 3 that depicts the cache structure and Association. When a cache-line fill operation occurs, a cache-slot will first be evicted (written back to memory) and then filled with the new cache-line contents.

The Fig. 3 above depicts the mapping between a memory address and a specific cache-slot. This is called a 1-way association. Modern cache designs boost performance by increasing the number of available cache-slots for each *index* value. When more slots exist, a cache-line is not necessarily evicted when a new memory location with the same index needs to be stored to cache.

For example, in a 2-way association, each cache-line index has 2 separate available slots. When a new cache-line is filled, only one slot must be evicted to make room for the new one. See Fig. 4. An LRU (Least-Recently-Used) algorithm is usually used to decide which slot should be evicted. This architecture boosts performance, since recently used memory values have a better chance of remaining in cache and therefore increasing the cache-hit ratio. It is not uncommon in modern CPU architectures to have 4-way and 8-way set associative cache designs.

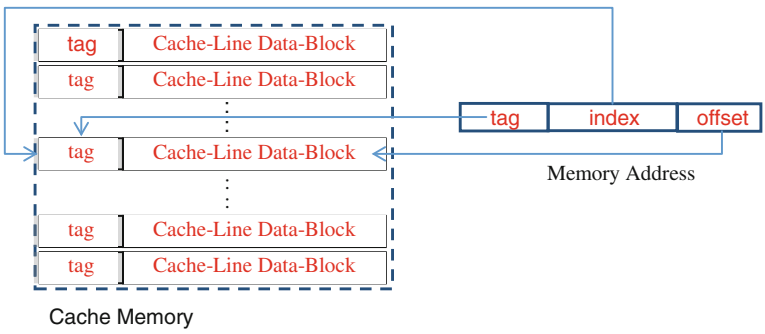

Fig. 3 Cache structure and memory association

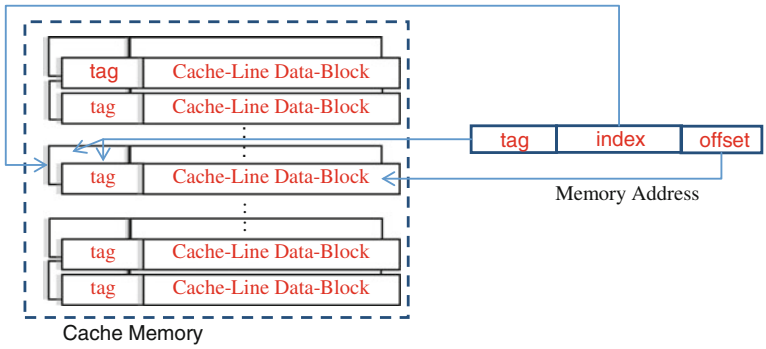

Fig. 4 2-way set associative cache

5.4 Side Channel Attacks on AES

Several attack strategies have been developed against the AES algorithm. Some attempt to acquire the memory contents of the AES application using different types of stealthy techniques, such as DMA attacks, Cold-Boot attacks or use of Firewire, PCI, etc. Once captured, memory contents can be analysed to retrieve the key used.

Another category of methods resort to an indirect strategy, which does not attempt to access the AES application directly, but takes advantage of side-effects that occur in the computer system as a result of the execution of the AES application and which can eventually lead to revealing the key. As explained above, these belong to the "Side-Channel" attacks category.

Cache-timing side-channel attacks are based on the fact that the processor accesses a cached memory element (*cache-hit*) at a significantly faster cycle time than that of a non-cached one (*cache-miss*). Different applications on the same system are protected from each other with Virtual memory; however the same underlying cache structure services all processes that run in parallel on the same CPU. Multiprocessing is supported by virtually all Operating-systems in use today. Consequently, if one process affects the cache subsystem, another parallel process can measure that affect even if it is restricted to accessing its own, private, address space. The timing differences between a cache-hit and a cache-miss are a factor of x10–x20. This leaves ample leeway for one process, running along-side another process on the same CPU to accurately measure those affects.

Recall that software driven AES applications make extensive use of lookup-tables. When a lookup-table entry is referenced it will be retrieved from the cache in the event of a cache-hit. Otherwise, in the event of a cache-miss it must be retrieved from main memory with a time penalty. An attacker routine, which runs in parallel to the AES process can measure the time of the encryption or decryption and compare the measurements when a specific lookup-table entry exists and then does not exist in the cache. One way for the attacking process to achieve this is to evict the specific lookup-entry from cache. To do that, it only needs to reference memory

elements from its own memory space, which has the same *index* as that of the lookup-table entry. Doing so for enough memory elements (at least the cache association ways) guarantees that the lookup-table entry is evicted from cache. For example, in a 4-way associative cache, 4 references to memory elements with the same *index* will evict the lookup-table entry. Following this, 2 consecutive AES decryptions are triggered and timed. If the first time measurement is longer than the second, it can be concluded that the specific lookup-table entry in question was referenced. Otherwise, the converse is true. Since lookup-table references occur as a function of the key value, repeating this process for different lookup-table entries can be used to reveal the key value. Additional details can be found in the work of Osvik et al. (2006).

Alternative methods may be employed by the attacking process to employ the same principles. For example, the entire cache may be evicted (this is usually a single instruction), followed by triggering of an AES decryption. The state of the cache would then reflect all the lookup-table entries that have been referenced. Now the attacking process can time references to memory elements in its own memory space, which have the same *index* value as a specific lookup-table entry. If the measured time corresponds to a cache-hit, it can be assumed that the lookup-table entry was referenced during the AES process. If the measured time corresponds to a cache-miss the converse can be concluded.

AES software implementation may be written to obscure the use of lookup-tables in such a way that it becomes impossible to relate its use to key values. Alternatively the implementation can avoid use of lookup-tables altogether. The down side of these methods is the performance penalty. Use of lookup-tables is the fastest (albeit vulnerable) implementation.

In 2010 Intel introduced a new instruction set in the Westmere processor family to perform AES calculations in hardware. This instruction set is dubbed AES-NI. The instruction set consists of 6 instruction op-codes: 2 for key expansion and 4 for encryption and decryption. By using these instructions the entire AES process is carried out in hardware in fixed, data-independent, timing. As a result, cache-timing attacks become completely useless (Gueron 2012).

6 Power Based Attacks

RSA is a common cryptographic method that relies on mathematical operations (mainly multiplications and divisions). However multiplication has different power requirements for bits containing 0 (hereby “0-bits”) compared to bits containing 1 (hereby “1-bits”). Due to the arithmetic nature of multiplication, multiplication involving “0-bits” is equivalent to NOP. Multiplication of “1-bits” on the other hand consumes more power for CPU operations.

Assuming the attacker has access to the platform were RSA or similar protection algorithm is running, during validation of the correct key. It has been shown that

using the power consumption of the platform the attacker can detect the "1-bits" in the key thus breaking the encryption (AES 2001).

Protection against power based attacks involves doing similar operations for "0-bits" and "1-bits" by doing random computations for "0-bits". This method increases the computation time of RSA and similar algorithms as it adds random computations. It is also still vulnerable to attacks as computation is not 100 % identical, however it has been shown that by adding random CPU work to "0-bits" the power consumption gap between "0-bits" and "1-bits" can be eliminated.

Closely related to Power-analysis side-channel attacks, are Acoustic side-channel attacks. Computer systems emit (ultrasonic) acoustic sound as a result of current surges through electronic components, such as capacitors and coils. Monitoring and analysing these sounds can reveal the underlying current consumption graph of the computer system. Therefore, a cryptographic system may be attacked in much the same way as one whose power usage is monitored. For implementation details see the work of Genkin et al. (2013). Acoustic monitoring has a distinct advantage in that a physical connection is not required, as measurements can be achieved solely by using a sensitive microphone.

Countermeasures that defeat these attacks may be to generate a random variety of sounds in the same spectrum, while computing the critical cryptographic algorithms. White-noise can also be used to acoustically drown the side-channel emissions.

References

Advanced Encryption Standard (AES) (2001) Federal Information Processing Standards Publication 197, United States National Institute of Standards and Technology (NIST)

Daemen J, Rijmen V (2013) AES Proposal: Rijndael s.l, National Institute of Standards and Technology, p 1

Domke F (2004) Console hacking 2004. In: CCC 2004

Domke F (2006) Console hacking 2006. In: CCC, 16 Nov 2006

Genkin D, Shamir A, Tromer E (2013) RSA key extraction via low-bandwidth acoustic cryptanalysis. tau.ac.il, 2013

Gueron S (2012) Intel® Advanced Encryption Standard (AES) Instructions Set–Rev 3.01. s.l

Kennell R, Jamieson LH (2003) Establishing the genuinity of remote computer systems. In: SSYM'03 Proceedings of the 12th USENIX Security Symposium, vol 12, pp 295–310, USENIX Association Berkeley, CA, 2003

Osvik DA, Shamir A, Tromer E (2006) Cache attacks and countermeasures: the case of AES. In: Pointcheval D (ed), Topics in Cryptology—CT-RSA 2006. Lecture Notes in Computer Science, vol 3860. Springer, New York, pp 1–20

Rutkowska J (2006) Introducing blue pill, the invisible things lab's blog. http://theinvisiblethings.blogspot.fi/2006/06/introducing-blue-pill.html

Rutkowska J, Tereskin A (2007) IsGameOver() Anyone? Technical presentation at Black Hat, Las Vegas, Invisible Things Lab, 2 Aug 2007

Schneier B (2012) Can the NSA break AES? Schneier on security blog. www.schneier.com/blog/archives/2012/03/can_the_nsa_bre.html

Shankar U, Chew M, Tygar JD (2004) Side effects are not sufficient to authenticate software. Report No. UCB/CSD-04-1363, Sept 2004, University of California, Berkeley, CA

Knowledge Discovery from Network Logs

Tuomo Sipola

Abstract Modern communications networks are complex systems, which facilitates malicious behavior. Dynamic web services are vulnerable to unknown intrusions, but traditional cyber security measures are based on fingerprinting. Anomaly detection differs from fingerprinting in that it finds events that differ from the baseline traffic. The anomaly detection methodology can be modelled with the knowledge discovery process. Knowledge discovery is a high-level term for the whole process of deriving actionable knowledge from databases. This article presents the theory behind this approach, and showcases research that has produced network log analysis tools and methods.

1 Network Anomaly Detection

A modern communications network is a collection of interconnected computers. The number of these computers and the detailed communications paths are usually unknown in the global view, they are just anonymous parts of a huge machinery. It is very challenging to know who are connected to the network and what kind of clients and servers there are at a certain time point. Therefore, many hiding places exist for malicious clients. There is a growing need to find these attackers and prevent their actions by cyber security methods.

Dynamic web services are vulnerable to unknown intrusions that grant access to systems that are not meant for the public audience. The trust between the network

This article is partly based on the author's dissertation (Sipola 2013). Author's current affiliation is with CAP Data Technologies.

T. Sipola (✉)
Department of Mathematical Information Technology, University of Jyväskylä,
P.O. Box 35 (Agora), 40014 Jyväskylä, Finland
e-mail: tuomo.sipola@iki.fi

M. Lehto and P. Neittaanmäki (eds.), *Cyber Security: Analytics, Technology and Automation*, Intelligent Systems, Control and Automation: Science and Engineering 78, DOI 10.1007/978-3-319-18302-2_12

participants is broken as legitimate features can be used for such unwanted access (Mukkamala and Sung 2003). However, server logs contain interesting information about network traffic. Finding attacks from these logs naturally improves security of the services. So called intrusion detection systems analyze the logs in order to find malicious traffic and the attackers (Di Pietro and Mancini 2008).

1.1 Fingerprinting

Traditional cyber security measures are based on fingerprinting, or pre-defined rules. Incoming traffic is compared to a database of known attacks, and then filtered according to the rules. All firewalls, IP blacklists and traditional intrusion detection systems are based on this approach. It is an effective way of controlling the network but with today's dynamic web services and protocols it is not possible to control everything with static rules. Moreover, the existing channels can be used for attacks, and blocking them would inhibit the normal use of the network.

1.2 Anomaly Detection

Anomalies, or outliers, are data samples that differ vastly from the baseline traffic. Anomaly detection techniques are used in many application areas, one of which is security. There are many anomaly detection methods, each having their own advantages (Chandola et al. 2009). Anomalies in networks are in a sense very common occasion. There is always happening something new and the whole network as a system is evolving as time advances. However, not all new events are the kind of anomalies that system administrators are interested in. It is a natural state that the network changes.

Therefore, there is a need to build a profile, or baseline, of clean network functioning. Anomaly detection differs from fingerprinting in that it finds events that differ from the baseline traffic even if such deviations are not previously known. Fingerprinting on the other hand reveals and prevents only known threats.

2 Network Environment

Figure 1 shows the different components of a model network. The Internet connects all the computers to each other. Normal users make queries to the servers and get responses with content. Attackers, on the other hand, try to access the servers, collect information or disrupt operations. Some servers are not protected at all while

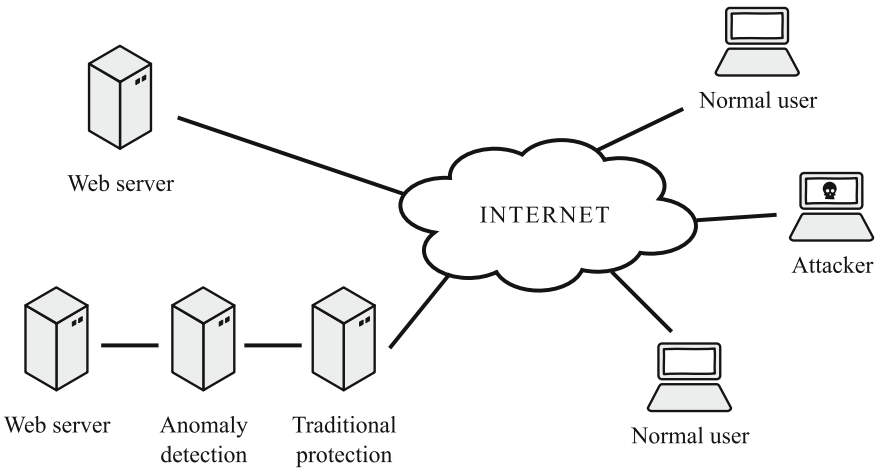

Fig. 1 Schematic network environment

others rely on traditional protection, such as firewalls. The anomaly detection system is deployed to the protected server in addition to traditional measures.

3 Knowledge Discovery Process

The anomaly detection methodology can be modelled with the knowledge discovery process. Knowledge discovery is a high-level term for the whole process of deriving actionable knowledge from databases. Presenting data mining as a part of the knowledge discovery process places the technical challenges in the broader scope. The knowledge discovery process from databases (KDD) suggests the steps that are needed to extract business knowledge from available data (Brachman and Anand 1996; Fayyad et al. 1996a, b, c).

Figure 2 shows the schematic workflow during the knowledge discovery process. The steps are described in more detail below.

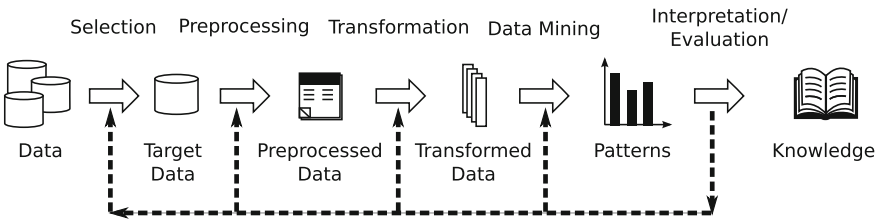

Fig. 2 Steps of the knowledge discovery process according to Fayyad et al. (1996b)

3.1 Databases

In the context of cyber security, the relevant databases include server and user logs, captured packet data and ultimately any collected information about the functioning of the system. The technical implementation of the data store is a challenge since the volumes of such data masses are huge. However, if all the data is not needed for later reference, storing is not a big problem.

3.2 Selection

In the data selection step the most relevant data sources are selected. It is usually the case that too many data sources and data features are available. Those sources that might contain important information should be selected, which requires domain knowledge. The selection process heavily depends on the knowledge discovery task. In cyber security tasks this would mean that the log files and the time frames should be selected. Some other sources might also be needed to give background information.

3.3 Preprocessing

Once the target data is defined, it needs to be preprocessed. Besides preprocessing, data cleaning is also a relevant. Converting and collecting the data to correct formats takes effort. Noise removal is a typical preprocessing step. Incomplete data entries and known large changes need to be accounted for. Combining and cleaning various databases is a rather mechanical process but sometimes poses problematic situations. The amount of work needed at this stage is usually underestimated in practical work.

3.4 Transformation

Preprocessed data is transformed to a more suitable form for clustering and classification. This involves feature extraction and selection, dimensionality reduction and other transformations. The selected features depend on the data mining case, as does the number of needed final features.

3.5 Data Mining

The data mining step itself tries to extract patterns from the transformed data. Summarization, classification, regression and clustering are some common tasks at this stage. The goal of the knowledge discovery process is matched with the relevant data mining methods. Often, this includes exploratory analysis of the data, which helps in deciding the most suitable models and parameters for the methods. The end products of this stage are rules, decision trees, regressions and clustering of the data.

3.6 Interpretation and Evaluation

Finally, the business value of the results should be understood. In this step, interpreting the patterns creates the actual final result of the knowledge discovery process. Evaluation of the obtained results is also important because some assumptions might have been wrong, the data might have been insufficient or there might have been some problem during the previous steps. Not all results have business significance even if they present some new information that has scientific or statistic novelty.

This research focuses on the transformation and data mining steps of the knowledge discovery process. The other steps are intimately connected to the utilized data, but the data mining methods are usually separate modules that can be described as independent systems. This is not to say, however, that in all cases the selection of data mining method is independent of the data and knowledge discovery problem. Correct tools should be used with differing datasets.

The actual data mining step usually uses a machine learning method. A supervised machine learning system is first trained using known data labeling (hence supervision), and then the performance of the system is tested. The testing results reveal the quality of the learned model, provided that the testing material adequately reflects the data in a real situation. This idea of training and testing is illustrated in Fig. 3.

4 Some Proposed Approaches

This section introduces some approaches to knowledge discovery from network logs. These research articles provide one point of view to the problem, but they are not the only ones.

Diffusion map-based approaches have been used for SQL injection detection and network traffic classification (David 2009; David and Averbuch 2012; David et al. 2010). These methodologies extract the relevant features from the data and create a

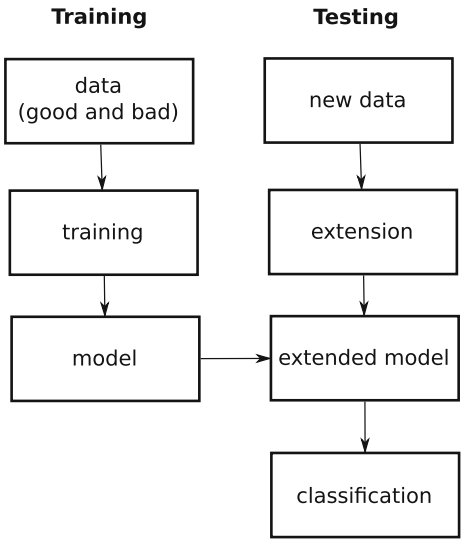

Fig. 3 Machine learning: training creates a model, while testing classifies new data using the model

low-dimensional presentation of the structure of the data, which can be clustered and classified to identify the anomalous traffic.

The goal of Sipola et al. (2011) is to find security attacks from network data. The proposed anomaly detection scheme includes *n*-gram feature extraction (Damashek 1995), dimensionality reduction and spectral clustering style linear clustering (Meila and Shi 2001; Shi and Malik 2000). It could be used for query log analysis in real situations. In practice the boundary between normal and anomalous might not be as clear as in this example. However, the relative strangeness of the sample could indicate how severe an alert is. The data in question is rather sparse and the discriminating features are quite evident from the feature matrix. This is the merit of the *n*-gram feature extraction which creates a feature space that separates the normal behavior in a good manner. The features describe the data clearly, and they are easy to process afterwards. The presented anomaly detection method performs well on real data. As an unsupervised algorithm this approach is well suited to finding previously unknown intrusions. This method could be applied to offline clustering as well as extended to a real-time intrusion detection system.

These results are elaborated in Sipola et al. (2012). The dimensionality reduction framework adapts to the log data. It assumes that only few variables are needed to express the interesting information, and finds a coordinate system that describes the global structure of the data. These coordinates could be used for further analysis of characteristics of anomalous activities. The practical results show that abnormal behavior can be found from HTTP logs. The main benefits of this framework include:

- The amount of log lines that needs to be inspected is reduced. This is useful for system administrators trying to identify intrusions. The number of interesting log lines is low compared to the total number of lines in the log file.
- The unsupervised nature and adaptiveness of the framework. The proposed methods adapt to the structure of the data without training or previous knowledge. This makes it suitable for exploration and analysis of data without prior examples or attack signatures. This means that the framework may also detect zero-day attacks.
- It works on the application layer in the network. The attacks themselves must in some way target the actual applications running on the computer. These logs might be more available than pure low-level network packet data.
- Visualization of text log data. It is much easier to analyze the structure of traffic using visualizations than it is to read raw textual logs.

The feature extraction from the web log is currently done with *n*-grams. However, this is only one method for it and other text-focused features might better describe the dataset. Furthermore, the dimensionality reduction scheme could be developed to adapt to this kind of data more efficiently, and the quality of the reduction could also be evaluated. Finally, automated root cause detection would make the system more usable in practice.

In Juvonen and Sipola (2012) a framework for preprocessing, clustering and visualizing web server log data is presented and used for anomaly detection, visualization and explorative data analysis. The results indicate that there are traffic structures that can be visualized from HTTP query information. Traffic clustering can give new information about the users. They could be categorized with more accuracy, and individual advertising or content could be offered. Using data mining methods, underlying structure and anomalies are found from HTTP logs and these results can be visualized and analyzed to find patterns and anomalies.

Article Juvonen and Sipola (2013) deals with extracting rules from the clustering results provided by a diffusion map training framework. Modern data mining technology in network security context does not always create understandable results for the end users. Therefore, this so-called black box system is not a desirable end goal. Simple conjunctive rules (Craven and Shavlik 1994; Ryman-Tubb and d'Avila Garcez 2010) are easier to understand, and rule extraction from the complex data mining techniques might facilitate user acceptance. The main benefit of this framework is that the final output is a set of rules. No black box implementation is needed as the end result is a simple and easy to understand rule matching system. The training data may contain intrusions and anomalies, provided that the clustering step can differentiate them. In addition, rule matching is a fast operation compared to more complex algorithms. The proposed framework is useful in situations where high-dimensional datasets need to be used as a basis for anomaly detection and quick classification. Such datasets are common nowadays in research environments as well as in industry, because collecting data is widespread. Our example case has been network security, which bears real benefits to anyone

using modern communication networks. The provided tools are useful for network administrators who are trying to understand anomalous behavior in their networks.

In Shmueli et al. (2013) another approach is taken to create a more online system. The training phase is computationally expensive in machine learning algorithms. Evolving datasets require updating the training. The proposed method updates the training profile using the recursive power iterations algorithm (Shmueli et al. 2012) and a sliding window algorithm for online processing. The algorithms assume that the data is modeled by a kernel method that includes spectral decomposition. A web server request log where an actual intrusion attack is known to happen is used to illustrate the online processing. Continuous update of the kernel prevents the problem of multiple costly trainings.

5 Conclusion

Knowledge discovery from network logs is a step forwards when trying to prevent attackers from using previously unknown attack types. This approach complements the traditional technologies that use fingerprints to detect the attackers. Anomaly detection finds the unusual traffic from networks and alerts the operators or uses automated countermeasures that prevent the attack from happening. This protection enhances the security of the web and creates a more efficient communications network for all participants.

References

Brachman RJ, Anand T (1996) In: Fayyad UM, Piatetsky-Shapiro G, Smyth P, Uthurusamy R (eds) Advances in knowledge discovery and data mining, chap. The process of knowledge discovery in databases. American Association Artificial Intelligence, pp 37–57. http://dl.acm.org/citation.cfm?id=257938.257944

Chandola V, Banerjee A, Kumar V (2009) Anomaly detection: a survey. ACM Comput Surv (CSUR) 41(3):15

Craven M, Shavlik JW (1994) Using sampling and queries to extract rules from trained neural networks. In: Proceedings of the eleventh international conference on machine learning, Morgan Kaufmann, pp 37–45

Damashek M (1995) Gauging similarity with n-grams: language-independent categorization of text. Science 267(5199):843–848

David G (2009) Anomaly detection and classification via diffusion processes in hyper-networks. PhD thesis, Tel-Aviv University

David G, Averbuch A (2012) Hierarchical data organization, clustering and denoising via localized diffusion folders. Appl Comput Harmon Anal 33(1):1–23

David G, Averbuch A, Coifman R (2010) Hierarchical clustering via localized diffusion folders. In: Manifold learning and its applications: papers from the AAAI fall symposium (FS-10-06). Association for the Advancement of Artificial Intelligence (AAAI), pp 28–31

Di Pietro R, Mancini LV (eds) (2008) Intrusion detection systems. Springer, Berlin

Fayyad U, Piatetsky-Shapiro G, Smyth P (1996a) From data mining to knowledge discovery in databases. AI Maga 17(3):37–54

Fayyad U, Piatetsky-Shapiro G, Smyth P (1996b) The KDD process for extracting useful knowledge from volumes of data. Commun ACM 39(11):27–34

Fayyad UM, Piatetsky-Shapiro G, Smyth P (1996c) Knowledge discovery and data mining: towards a unifying framework. In: KDD-96 proceedings of Association for the Advancement of Artificial Intelligence (AAAI), pp 82–88

Juvonen A, Sipola T (2012) Adaptive framework for network traffic classification using dimensionality reduction and clustering. In: proceedings of the 2012 4th IEEE international congress on ultra modern telecommunications and control systems and workshops (ICUMT), New York, pp 274–279

Juvonen A, Sipola T (2013) Combining conjunctive rule extraction with diffusion maps for network intrusion detection. In: Proceedings of the 2013 IEEE symposium on computers and communications (ISCC), New York, pp 411–416

Meila M, Shi J (2001) A random walks view of spectral segmentation. In: AI and STATISTICS (AISTATS) 2001

Mukkamala S, Sung AH (2003) A comparative study of techniques for intrusion detection. In: Proceedings of the 15th IEEE international conference on tools with artificial intelligence, New York, pp 570–577

Ryman-Tubb NF, d'Avila Garcez A (2010) Soar—sparse oracle-based adaptive rule extraction: knowledge extraction from large-scale datasets to detect credit card fraud. In: Proceedings of the 2010 IEEE international joint conference on neural networks (IJCNN), New York, pp 1–9

Shi J, Malik J (2000) Normalized cuts and image segmentation. IEEE Trans Pattern Anal Mach Intell 22(8):888–905

Shmueli Y, Wolf G, Averbuch A (2012) Updating kernel methods in spectral decomposition by affinity perturbations. Linear Algebra Appl 437(6):1356–1365

Shmueli Y, Sipola T, Shabat G, Averbuch A (2013) Using affinity perturbations to detect web traffic anomalies. In: Proceedings of the 10th international conference on sampling theory and applications (SampTA 2013), EURASIP, Bremen, pp 444–447

Sipola T (2013) Knowledge discovery using diffusion maps. Ph.D. thesis, University of Jyväskylä

Sipola T, Juvonen A, Lehtonen J (2011) Anomaly detection from network logs using diffusion maps. In: Iliadis L, Jayne C (eds) Engineering applications of neural networks, IFIP advances in information and communication technology, vol 363. Springer, Boston, pp 172–181

Sipola T, Juvonen A, Lehtonen J (2012) Dimensionality reduction framework for detecting anomalies from network logs. Eng Intell Syst 20(1–2):87–97

Trusted Computing and DRM

Nezer Zaidenberg, Pekka Neittaanmäki, Michael Kiperberg and Amit Resh

Abstract Trusted Computing is a special branch of computer security. One branch of computer security involves protection of systems against external attacks. In that branch we include all methods that are used by system owners against external attackers, for example Firewalls, IDS, IPS etc. In all those cases the system owner installs software that uses its own means to determine if a remote user is malicious and terminates the attack. (Such means can be very simple such as detecting signatures of attacks or very complex such as machine learning and detecting anomalies in the usage pattern of the remote user). Another branch of attacks requires protection by the system owner against internal users. Such attacks include prevention of users to read each other's data, use more than their allotted share of resources etc. To some extent anti-virus/anti-spam software is also included here. All password protection and used management software are included in this branch. The third branch, *Trusted Computing*, involves the verification of a remote host that the user machine will behave in a certain predictable way, i.e. protection against the current owner of the machine. The most common example for this kind of requirement is distribution of digital media. Digital media is distributed in some conditional access mode (rented, pay per view, sold for personal use, etc.). Obtaining digital media usually does not entitle the user to unlimited rights. The user usually may not redistribute or edit the digital media and may not even be allowed to consume it himself after a certain date. (Media rentals, pay per view)

N. Zaidenberg (✉) · P. Neittaanmäki · M. Kiperberg · A. Resh
Department of Mathematical Information Technology, University of Jyväskylä, Jyväskylä, Finland
e-mail: nezer@trulyprotect.com

P. Neittaanmäki
e-mail: pekka.neittaanmaki@jyu.fi

M. Kiperberg
e-mail: michael@trulyprotect.com

A. Resh
e-mail: amit@trulyprotect.com

M. Lehto and P. Neittaanmäki (eds.), *Cyber Security: Analytics, Technology and Automation*, Intelligent Systems, Control and Automation: Science and Engineering 78, DOI 10.1007/978-3-319-18302-2_13

However, as the user is consuming media on his private machine. How can the media provider assure himself that a malicious user does not tamper with the machine so that contents are not replicated? The problem of security against the owner of the machine is the problem region of Trusted Computing. In trusted computing as opposed to other branches of security the "attacker" is not limited to some attack surface that was exposed to him but can also use a soldering iron to tap into busses, replace chips and other system parts etc. Trusted computing also includes other protection tools against the current owner (or possessor of the machine if not the legal owner). For example protection of sensitive data or disk encryption solutions for laptops and mobile phones that can potentially be stolen. Trusted computing can also be used on the cloud to ensure that the host does not inspect a cloud server and the software running on the server is not stolen. Latest trusted computing technology involves means to ensure commands are sane and are not malicious, for example in computers on cars and avionics. In this chapter we will review DRM and Trusted computing solutions from multiple sources.

1 Ethics—Trusted or Treacherous Computing

Users don't like trusted computing. First and foremost, the concept of conditional access leads to numerous digital rights debates. For example, if I legally purchased contents, shouldn't I be allowed to make backups of said contents? Especially as no media vendors are currently proposing to offer free (or even cheap) replacements of corrupted media contents! However, if we allow media to be replicated then how can we disallow illegal copies? What is stopping the user from redistributing copies or "backups"? How can we distinguish legal use of copies (backups) and illegal copies of the same content?

Secondary, as many trusted computing devices requires the user to actively install something on his machine (a TPM chip, EFI firmware, etc.). And the said hardware component does not contribute to the end-user system features at all (if anything trusted computing only limits the user). Why would the user willingly spend money and install some piece of hardware in his computer that only serves to limit what she can and cannot do? All these reasons have lead Richard Stallman to call trusted computing treacherous computing and numerous hackers to try attacks on TPM chips and trusted platforms.

As of writing this chapter there is no clear cut winner in this technology battle. On the one hand there is still no massive install base for trusted computing solutions and on the other hand the trusted computing group is still alive and still releasing new trusted platform modules and specs.

2 The Trusted Processing Module by TCG

The trusted platform module, as demonstrated in Fig. 1 is a separate computer chip that is added to the computer motherboard and is frequently connected to the CPU using the LPC (low pin count) bus. The TPM is a cryptographic co-processor that is able to perform several cryptographic functions as well as generate and store keys. The TPM can also verify the hardware and software that the system runs on and attest for the system's sanity.

When a remote host is querying the system for sanity it can use the TPM to verify that the software that it runs on was not tampered. TPM supports two attestation methods: Remote Attestation and Direct Anonymous Attestation.

2.1 Remote Attestation

The attestation solution proposed by the TCG (TPM specification v1.1) requires a trusted third-party, namely a *privacy certificate authority* (privacy CA). Each TPM has an RSA key pair called an Endorsement Key (EK), embedded inside the TPM (that the user cannot access—at least not easily.) In order to attest itself, the TPM generates a new RSA key pair (Attestation Identity Key or AIK).

Remote Attestation is a typical trusted 3rd party process. Assuming Alice wants to attest Bob and Bob wants to be recognized by Alice but neither wants to reveal

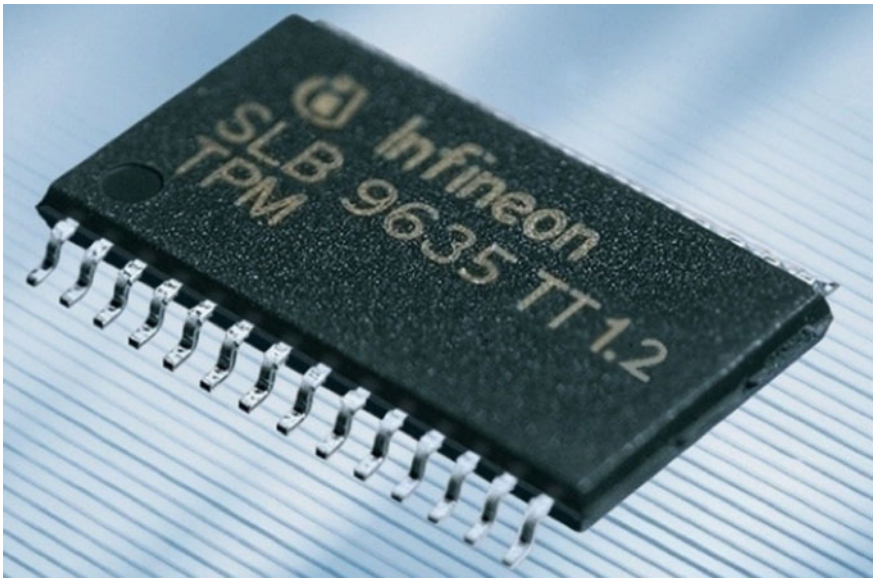

Fig. 1 Trusted platform module

his private identification code to each other, remote attestation is suggested. It requires a trusted party that both can trust.

Bob attests himself by signing the public AIK using the EK, to the trusted 3rd party. The trusted 3rd party then verifies Bob by using Bob's public EK. Of course the CA may and should blacklist TPMs if it receives too many requests using the same key simultaneously. Alice can later verify with the CA Bob has is indeed attested.

2.2 Direct Anonymous Attestation

The Direct Anonymous Attestation (hereby DAA) protocol was only added to the TPM standard in version 1.2. DAA is based on three entities and two steps. The entities are the TPM platform, the DAA issuer and the DAA verifier. The issuer is charged to verify the TPM platform during the Join step and to issue DAA credentials to the platform. The platform uses the DAA credentials with the verifier during the Sign step. The verifier can verify the credentials without attempting to violate the platform's privacy [zero knowledge proof (Quisquater et al. 1990; Blum et al. 1988)]. The protocol also supports a blacklisting capability, so that verifiers can identify attestations from TPMs that have been compromised.

DAA allows differing levels of privacy. Using DAA Interactions is always anonymous, but the user/verifier may negotiate as to whether the verifier is able to link transactions (with the same user but not a specific user). Verifying transactions would allow persistent data to be saved over sessions and would allow profiling and tracing multiple logins.

3 Intel TXT and AMD/ARM Trustzone

Intel TXT technology defines unique extensions for the CPU instruction set to allow trusted execution. Using intel TXT, one can attest the hardware, OS and software currently running and ensure a stable (as opposed to tampered) system state. Intel TXT uses the TPM for measurements and cryptographic functions to attest to a 3rd party and ensure that system software or the OS that is currently running is indeed trustworthy and non-tampered with.

The PCR registers on the TPM contain measurements and SHA-1 hashes of various system stages and code and by checking and verifying these measurements the system can be trusted to boot a non-tampered software.

AMD/ARM Trustzone is the ARM/AMD implementation of trusted computing. It is roughly corresponding to Intel TXT. The Trustzone implementation is used by both AMD and ARM. Trustzone allows signed secure OSs to be loaded, for example, by using AMD/ARM SVM.

4 Other Architectures for "Trusted Computing"

These architectures provide means to prevent replication of data and thus introduce trust on various systems. We focused mainly on Video content delivery in this chapter. Different systems for preventing homebrew and pirated software on game consoles (which is another form of trusted computing) are covered in Chap. 3.

4.1 HDMI and HDCP and Its Predecessors

The Video industry has always been interested in mixed goals:

1. It searched for ways to deliver high quality video to the user's home. Generating a new revenue stream from videos that no longer appeared in cinemas (Video/DVD rentals).
2. It searched for ways to prevent the user from obtaining permanent access to the video equipment she rented by making illegal copies.

To some extent the battle was a lost cause to begin with because the user could always point a standard camera to the screen and just record using the camera (or create low quality copies using older, already broken technology). However, the industry was interested in preventing the user from making high quality copies (for example, digital quality copies in the case of HDMI). This approach led to several technologies whose purpose was to circumvent the user's ability to create illegal copies.

4.2 Macrovision, CSS and DeCSS

Old VHS video devices had a macrovision device that prevented direct creation of copies of VHS media by connecting two video devices to each other. The Macrovision devices modified the output stream in a way that was unnoticeable to users but prevented VHS devices to create VHS cassettes copies by daisy chaining devices.

CSS or Content Scrambling System is an encryption system that is used on all major DVDs. CSS was 40 bit encryption system. The use of CSS was supposed to make it impossible to copy video content directly from DVD to a video. This was done as the encryption keys were kept in unreadable (by data DVD players) location. CSS also allowed for DVD regions, Macrovision etc.

DVD CSS was broken at 1999, about 3 years after it was introduced with the introduction of DeCSS software. An inherent bug was used to reduce the keys from 40 bit to only 16 bit long and most players were able to break this encryption in less than 1 min by brute force.

Two of DeCSS authors remain unknown even today. The 3rd was a Norwegian teenager: Jon Lech Johansen. Mr. Johansen was brought to trial and acquitted by the Norwegian court. The prosecution appealed and Mr. Johansen was acquitted for the second time. When DVD was superseded by Blu-Ray and HD-DVD CSS was replaced with the AACS (Advanced access contents system, which was broken using leaked keys).

4.3 HDMI and HDCP

HDMI or High Definition Media Interface is a high quality media interface allowing high quality media transfer to monitors and screens. HDMI raised the problem of creating exact or near exact high quality replicas of video content.

To avoid copying the contents, HDCP will encrypt the content travelling between two end points of HDMI and will only provide contents to devices with trusted keys. These keys can later be revoked if they are stolen. By 2010 the master key for HDCP had been leaked, rendering all revocation list useless. It is possible that the revocation key was used too many times and provided sufficient data that made breaking the key easier.

5 Other Uses for Trusted Computing

Several attacks on the user can be done after the attacker has obtained full or even partial control on the end user device. For example the attacker may be interested in the contents of the user hard drive after obtaining control of the user laptop (for example, by stealing it) Such trusted computing content protection methods involve the usage of protected (encrypted) storage were the keys are saved on the TPM.

Microsoft Bitlocker is a full disk encryption solution that can be used on computers (especially laptops) to ensure that the disk contents are unreadable to an attacker, even if the computer/laptop was stolen. The complete disk is encrypted and the key to decipher the disk content is unreadable and saved on the TPM.

Mobile phones contain private data that can be exposed if the phone is lost or stolen. Numerous technologies have been generated by various sources from using TPM and encryption on the device to a more biometric approach. Examples include apple usage of fingerprint reading devices on the iPhone device that are required to unlock a stolen phone.

Other technologies include a kill code that is used to wipe the device and prevent it from connecting to the network ever again.

6 Attacks on Trusted Computing

The TPM is often connected to the LPC (low pin count) bus. A legacy slows bus that exist on virtually all PCs. Attacks on this exist for over 10 years. One of the first cases of attacks on this bus occured on the first XBox (Stiel 2005). By connecting and eavesdropping to the LPC bug several hackers have been able to intercept and reset the TPM.

In 2010 (Tarnovsky 2012) and later in 2012 (Tarnovsky 2010) Chris Tarnovsky demonstrated physical attacks against TPM chips by Infineon and ST microelectronic. Tarnovsky attacked the TPMs by eliminating parts of the TPM chips thermal casing and attacking (i.e. connecting external devices) to the chips itself.

Tarnovsky demonstrated that ST and Infineon chips are made of older processors chips from their past. He demonstrated that by physically attacking the chips itself he could expose and modify content on the TPM chip itself. Tarnovsky's methods require a special lab, chemicals and equipment which may not be in every hacker's reach. But it is definitely not beyond the reach of professional attackers and hackers.

One of the features of the Intel CPU is SMM or software maintenance mode. SMM is used to allow updating CPU code (microcode). SMM code is executed at higher permissions then user, kernel or hypervisor code on the Intel platform. Therefore, SMM is considered to run at permission ring-2 (if Ring 3 is userspace code, ring 0 is kernel code and ring-1 is hypervisor).

Rafal Wojtczuk and Joanna Rutkowska have demonstrated breaking TXT limitations using SMM (Wojtczuk and Rutkowska 2009). These attacks may have been Intel's main reason for devising the SGX extension. These attacks are possible because TXT protection blocks execution and permission in rings 3 (user space), 0 (kernel) and 1 (hypervisor) but TXT memory defense is still vulnerable to attacks on Ring-2 using SMM permission level which does not require any special permissions and can be used even after the OS has been attested by the TPM.

7 Beyond Trust—SGX

SGX or Software Guard Extension is an innovative technology from Intel that will be implemented in future chips. SGX provides a solution to the trusted computing problem on Intel platforms. SGX technology allows creating an execution container for each process in which the process memory is contained. This approach is similar to the approach taken by Qubes OS development to create separation using hypervisor code between applications so different applications are running on different virtual OSs (Blum et al. 1988) and by Trusted computing software such as TrulyProtect, which keeps secrets in the hypervisor layer (Zaidenberg 2013). At the time of writing this chapter SGX is not available with any Intel CPU on the market (thus there are no known attacks on SGX).

References

Blum M, Feldman P, Micali S (1988) Non-interactive zero-knowledge and its applications. In: Proceedings of the twentieth annual ACM symposium on theory of computing (STOC '88). ACM, New York, pp 103–112

Intel Trusted execution Technology—whitepaper hardware based technology for advanced server protection. http://www.intel.com/content/www/us/en/trusted-execution-technology/trusted-execution-technology-security-paper.html, http://www.intel.com/content/dam/www/public/us/en/documents/white-papers/trusted-execution-technology-security-paper.pdf

Quisquater J-J, Guillou LC, Berson TA (1990) How to explain zero-knowledge protocols to your children. In: Advances in cryptology—CRYPTO '89 proceedings, lecture notes in computer science, vol 435. Springer, Berlin, pp 628–631

Stiel M (2005) 17 mistakes Microsoft made with the Xbox security system. In: 22nd computer chaos club conference

Tarnovsky C (2010) Hacking the smartcard chip. Blackhat, DC

Tarnovsky C (2012) DEF CON 20: attacking TPM part 2

TPM Reset Attack Evan Sparks. http://www.cs.dartmouth.edu/~pkilab/sparks/

Wojtczuk Rl, Rutkowska J (2009) Attacking Intel® trusted execution technology. Invisible Things Lab, Blackhat, DC

Zaidenberg N (2013) TrulyProtect 2.0 and attacks on TrulyProtect 1.0. poster presentation. In: ECIW 2013—12th European conference on information warfare and security (Jyväskylä)

Part IV
Cyber Security and Automation

Cyber Security and Protection of ICS Systems: An Australian Example

Matthew J. Warren and Shona Leitch

Abstract Many aspects of our modern society now have either a direct or implicit dependence upon information technology. As such, a compromise of the availability or integrity in relation to these systems (which may encompass such diverse domains as banking, government, health care, and law enforcement) could have dramatic consequences from a societal perspective. These key systems are often referred to as critical infrastructure. Critical infrastructure can consist of corporate information systems or systems that control key industrial processes; these specific systems are referred to as ICS (Industry Control Systems) systems. ICS systems have devolved since the 1960s from standalone systems to networked architectures that communicate across large distances, utilise wireless network and can be controlled via the Internet. ICS systems form part of many countries' key critical infrastructure, including Australia. They are used to remotely monitor and control the delivery of essential services and products, such as electricity, gas, water, waste treatment and transport systems. The need for security measures within these systems was not anticipated in the early development stages as they were designed to be closed systems and not open systems to be accessible via the Internet. We are also seeing these ICS and their supporting systems being integrated into organisational corporate systems.

1 Introduction

This chapter will focus on cyber security and possible threats to Australia. Many critics dismiss the cyber security threat to Australia as being "hype or overstated" (Warren 2013) but, in fact, based on the available statistics, Australia has had

M.J. Warren (✉)
Faculty of Business and Law, Deakin University, Victoria, Australia
e-mail: matthew.warren@deakin.edu.au

S. Leitch
College of Business, RMIT University, Melbourne, Australia
e-mail: shona.leitch@rmit.edu.au

M. Lehto and P. Neittaanmäki (eds.), *Cyber Security: Analytics, Technology and Automation*, Intelligent Systems, Control and Automation: Science and Engineering 78, DOI 10.1007/978-3-319-18302-2_14

a significant number of cyber-attacks occur. Between the period 2011 and 2012, 438 cyber incidents occurred against Australia requiring a meaningful response by the Australian Federal Government and the Cyber Security Operations Centre (Australian Government 2013).

The chapter will describe the Cyber Security issues in relation to ICS systems (and their supporting components) the protection of these critical infrastructure systems in an Australian context, considering both the technical issues as well as the security policy considerations. It will reflect upon the Australian Maroochy water security incident as an example of the possible incidents that Australia faces. Also discussed will be the security strategy concept of *reliance* as a way of mitigating security threats, drawing upon examples relating to the Australian Federal Government. The chapter will conclude by looking at future security risks and issues in relation to ICS security.

2 ICS Security

ICS systems (and their components) have evolved since the 1960s from stand-alone systems to networked architectures that communicate across large distances. Their implementation has migrated from custom hardware and software to standard hardware and software platforms (Krutz 2006).

Industrial control system (ICS) is a general term that encompasses several types of control systems, including supervisory control and data acquisition (SCADA) systems; distributed control systems (DCS); and other control system configurations such as skid-mounted Programmable Logic Controllers (PLC) often found in the industrial sectors, and critical infrastructures (Stouffer et al. 2011). SCADA systems are highly distributed systems used to control geographically dispersed assets, often scattered over thousands of square kilometres, where centralized data acquisition and control are critical to system operation. Field devices control local operations such as opening and closing valves and breakers, collecting data from sensor systems, and monitoring the local environment for alarm conditions. DCS are used to control industrial processes such as electric power generation, oil refineries, water and wastewater treatment, and chemical, food, and automotive production. DCS are integrated as a control architecture containing a supervisory level of control overseeing multiple, integrated sub-systems that are responsible for controlling the details of a localized process (Stouffer et al. 2011).

Other key components of ICS systems, include (Stouffer et al. 2011):

- *Remote Terminal Unit* (*RTU*). The RTU, also called a remote telemetry unit, is a special purpose data acquisition and control unit designed to support SCADA remote stations. RTUs are field devices often equipped with wireless radio interfaces to support remote situations where wire-based communications are unavailable;

- *Programmable Logic Controller* (*PLC*). The PLC is a small industrial computer originally designed to perform the logic functions executed by electrical hardware (relays, switches, and mechanical timer/counters). PLCs have evolved into controllers with the capability of controlling complex processes, and they are used substantially in SCADA systems and DCS.

Within this paper we will refer to ICS, but this also includes SCADA and DCS systems, unless specially described.

The need for security measures within these systems was not anticipated in the early ICS development era as they were designed to be closed systems and not be accessible via the Internet. The increasingly networked and linked infrastructure of modern SCADA systems has changed those early security plans (Beggs and McGowan 2011).

The security risks and vulnerabilities of ICS systems have arisen because of system development based upon open based communications standards such as ethernet communications. ICS software companies have embraced the Transmission Control Protocol and Internet Protocol (TCP/IP) to improve integration across multiple systems. However, these developments have exposed the industrial sector to common Internet vulnerabilities within communication protocols which increase the risk of attacks (Pollet 2002).

Many of the security features used to protect ICS systems are also used to protect corporate networks. Table 1 gives an example of typical IT security controls found to protect ICS systems.

Other researchers argue that the key aspects of ICS security are based upon a holistic view of security including physical security, IT security and specific ICS security (Weiss 2010).

We are also seeing specific ICS security standards develop in relation to ICS security. The holistic view, ICS security, is also reflecting by the standards, Table 2 reflects the NIST view of ICS security.

Table 1 Typical IT ICS security controls (Krutz 2006)

Security areas	Description
Audit and monitoring logs	Forensic tools for incident analysis
Biometrics	Biometric authentication for system access
Firewalls	Controlling access between the SCADA and corporate networks
Intrusion detection systems	Monitoring network traffic and determining unauthorised access
Malicious code detection and elimination	Identification and removal of malicious code from SCADA systems
Passwords	Authentication mechanism for access to systems
Public key and symmetric key cryptography	Ensuring the integrity of data within the SCADA systems
Role based access control	Restricting access to part of the SCADA systems based upon job role, e.g. operator access

Table 2 NIST 800-82 view of ICS security (Stouffer et al. 2011)

Business case	Determine ICS vulnerabilities and risks
	Prepare business case
	Develop comprehensive security plan
Network architecture	Firewalls
	Locally separated control networks
	Network segregation
	Defence in depth architecture
ICS security controls	Management controls
	Operation controls
	Technical controls

The International Society of Automation (ISA) has 30,000 members worldwide and they develop standards in relation to Automation Systems (including ICS system). They have also specifically focused ICS and security, whereby they have developed a working party with the aim of developing a suite of ISA ICS specific security standards for their members. The type and status of the different ISA ICS security guidelines are shown in Table 3.

There is a major challenge with ICS security that many older legacy ICS systems are not compliant with new security technologies such as advanced encryption and intrusion detection devices (Beggs 2008). These can pose an issue for security in older ICS systems since it may slow down the ICS operations (Weiss 2010).

Table 3 ISA ICS security guidelines (ISA (International Society of Automation) 2013)

Type of guideline	Area
General	Terminology concepts and models (published as standard ISA-99.00.01-2007)
	Glossary of terms (in development)
	System security complain metrics (in development)
	Security lifecycle and use case (planned)
Policies and procedures	Requirements for a security management system (published as standard ISA-99.02.01-2009)
	Implementation guidance (planned)
	Patch management (in development)
	Requirement for solution supplies (in development)
System	Security technologies (published as Standard ISA-TR99.00.01-2007)
	Technical controls (in development)
	Security levels for zone and conduits (in development)
	System security requirements and security levels (in development)
Components	Product development requirements (in development)
	Technical security requirements for components (in development)

An interesting security difference between ICS systems and corporate systems, is that ICS systems are focused upon the continued availability of the connected industry process, e.g. power production and water processing. In contrast, corporate systems are focused on protecting the information within those systems from being attacked or corrupted (Weiss 2010). Historically, ICS systems were closed, stand-alone systems, now they are being linked in corporate networks and connected to the Internet to allow remote access and monitoring. These new technology innovations introduce new security risks (Shaw 2006).

3 Maroochy SCADA Security Case Study

The nature of SCADA systems in Australia is often very complex because of the vastness of the country and the remoteness of many of the utility plants and field stations (Slay and Miller 2008) within the SCADA systems. In 2000 a SCADA security incident occurred at Maroochy Water Services in Queensland, Australia. This was the first global example (that was declared publically) of an effective hacking incident against a SCADA critical infrastructure system.

The following is an assessment of the Maroochy Water SCADA hacking incident (Slay and Miller 2008; Hughes 2003; Supreme Court of Queensland 2002). In 2000, Maroochy Shire Council experienced problems with its new wastewater system. Communications being sent by radio links to 150 wastewater pumping stations were being lost; pumps were not working properly; and alarms put in place to alert staff to faults were not going off. It was initially thought there were teething problems with the new system. Sometime later, an engineer who was monitoring every signal passing through the system, discovered that someone was hacking into the system and deliberately causing the problems. It was later determined that Vitek Boden used a laptop computer and a radio transmitter to take control of the 150 sewage pumping stations in the SCADA system. Over a 3 month period, he released one million litres of untreated sewage into a stormwater drain from where it flowed to local waterways and $50,000 had to be spent cleaning up the sewage discharge.

The attack was motivated by revenge on the part of Vitek Boden after he failed to secure a job with the Maroochy Shire Council. He was arrested by police, close to a pump station and in possession of a laptop PC which he was hoping to use to control another pump station in order to release further sewage. In October 2001, Vitek Boden was found guilty and imprisoned for 2 years. A subsequent appeal against his prison sentence was rejected.

The case of Vitek Boden and Maroochy Water highlights the potential threats that SCADA systems can face and also the physical consequence of the attacks. The case also highlights how simple information security policies within Maroochy Water could have prevented the attack, for example, revoking an employee's access

rights when they leave the organisation. The example also demonstrates that attacks can occur internally within an organisation; attacks do not always have to come from external sources.

4 Australian Strategic Cyber Protection

Australia's state and territory governments have defined critical infrastructure as being:

> those physical facilities, supply chains, information technologies and communication networks which, if destroyed, degraded or rendered unavailable for an extended period, would significantly impact on the social or economic wellbeing of the nation or affect Australia's ability to conduct national defence and ensure national security (Australian Government 2010).

ICS and SCADA systems form a key part of Australia's critical infrastructure. They are used to remotely monitor and control the delivery of essential services and products, such as electricity, gas, water, waste treatment and transport systems (Trusted Information Sharing Network (TISN) 2007).

The initial focus of the Australian Federal Government critical infrastructure protection policy (including ICS/SCADA systems) was that critical infrastructure protection was a commercial consideration and related only to information security (Busuttil and Warren 2004). The Australian Federal Government had been aware of the problems that Australian corporations may have in dealing with these security issues and responded by offering advice for corporations. The initial Australian Government advice, in (1998), suggested ways in which organisations could reduce critical infrastructure risks in the following ways (Busuttil and Warren 2004):

- Organisations should implement protective security controls such as passwords, etc., in accordance to a defined security standard such as AS/NZS 4444 (now has been replaced with ISO/IEC 27002);
- Organisations should formally accredit themselves against information security standards;
- Organisations should raise awareness of security issues such as password security, awareness of online risks among their staff;
- Organisations should train their staff in how to use computer security systems efficiently and effectively.

This advice was subsequently updated, and in 2004 the Australian Government responded with new security advice (Australian Government 2004a):

- All Australian critical infrastructure should be evaluated using the Australian and New Zealand standard *AS/NZS 4360 Risk Management.* This process should be used to assist with the review of risk management plans for prevention (including security), preparedness, response and recovery.

In 2004 the Australian Federal Government formally defined the following; "critical infrastructure is defined as those physical facilities, supply chains, information technologies and communication networks which, if destroyed, degraded or rendered unavailable for an extended period, would significantly impact on the social or economic well-being of the nation, or affect Australia's ability to conduct national defence and ensure national security" (Australian Government 2004b). Many of the characteristics of critical infrastructure relates to Australian ICS/SCADA systems.

Historically, much of Australia's infrastructure was originally owned and operated by the public sector at the federal, state and local government levels (Smith 2004), however the majority of Australia's critical infrastructure has now been privatised and is under private sector ownership. Consequently, protecting Australia's critical infrastructure now requires a higher level of cooperation between all levels of government and the private sector owners, hence, the Australian Federal Government developed a policy for critical infrastructure protection that focuses broadly on addressing the following issues (Australian Government 2004a, b):

- Distinguishing critical infrastructures and ascertaining the risk areas;
- Aligning the strategies for reducing potential risk to critical infrastructure;
- Encouraging and developing effective partnerships with state and territory governments and the private sector;
- Advancing both domestic and international best practice for critical infrastructure protection.

The Australian critical infrastructure policies between 1998 and 2007 reinforced the importance of technical security controls for ICS/SCADA systems and also caused organisations to start to consider security risks in relation to ICS/SCADA systems.

As discussed by Warren and Leitch (Warren and Leitch 2010), the Australian Federal Government recognised the importance of crucial systems and the development of new industry support mechanisms, in particular, Trusted Information Sharing Network (TISN). The TISN is a forum in which the owners and operators of critical infrastructure work together by sharing information on security issues which affect critical infrastructure (Trusted Information Sharing Network (TISN) 2007). TISN requires the active participation of critical infrastructure protection owners and operators, regulators, professional bodies and industry associations, in cooperation with all levels of government, and the public. To ensure cooperation and coordination, all of these participants should commit to the following set of common fundamental principles of critical infrastructure protection (Trusted Information Sharing Network (TISN) 2007; Warren and Leitch 2010):

1. Critical infrastructure protection is centred on the need to minimise risks to public health, safety and economic confidence, ensure economic security, maintain Australia's international competitiveness and ensure the continuity of government and its services;

2. The objectives of critical infrastructure protection are to identify critical infrastructure, analyse vulnerability and interdependence, and protect from, and prepare for, all hazards;
3. As not all critical infrastructure can be protected from all threats, appropriate risk management techniques should be used to determine relative severity and duration, the level of protective security, set priorities for the allocation of resources and the application of the best mitigation strategies for business continuity;
4. The responsibility for managing risk within physical facilities, supply chains, information technologies and communication networks primarily rests with the owners and operators;
5. Critical infrastructure protection needs to be undertaken from an "all hazards approach" with full consideration of interdependencies between businesses, sectors, jurisdictions and government agencies;
6. Critical infrastructure protection requires a consistent, cooperative partnership between the owners and operators of critical infrastructure and governments;
7. The sharing of information relating to threats and vulnerabilities will assist governments, and owners and operators of critical infrastructure to better manage risk;
8. Care should be taken when referring to national security threats to critical infrastructure, including terrorism, so as to avoid undue concern in the Australian domestic community, as well as potential tourists and investors overseas;
9. Stronger research and analysis capabilities can ensure that risk mitigation strategies are tailored to Australia's unique critical infrastructure circumstances.

These TISN principles relate to the owners and operators of ICS/SCADA systems and provide a framework in which they could co-operate and coordinate their systems.

In 2008 the Australian Federal Government identified a number of new security challenges, "it is increasingly evident that the sophistication of our modern community is a source of vulnerability in itself. For example, we are highly dependent on computer and information technology to drive critical industries such as aviation; electricity and water supply; banking and finance; and telecommunications networks. This dependency on information technology makes us potentially vulnerable to cyber-attacks that may disrupt the information that increasingly lubricates our economy and system of government" (Warren and Leitch 2010).

The Australian critical infrastructure policies between 2004 and 2008 have defined the importance of the relationship between government and the private sector and in particular the setting up the TISN program to allow for the distribution of key information. The TISN program has allowed for special interest groups to be set up around different critical infrastructure areas, e.g. water and power, and this allowed for more information to be shared regarding ICS/SCADA security amongst utility operators using ICS and SCADA systems.

In 2009 the Australian Federal Government had responded to the issues regarding cyber security and critical infrastructure by proposing a new coherent and government led approach to critical infrastructure protection. The primary objectives identified focus on all areas of Australian society where there are security risks: that individuals should be aware and take steps to "protect their identities, privacy and finances online"; that businesses and the government operate "secure and resilient information and communication technologies"; and trusted electronic operating environments that supports Australia's national security and maximises the benefits of the digital economy (Rudd 2008). The Australian Federal Government also has developed a wide range of new strategic directions to focus Australia's critical infrastructure protection programs (Rudd 2008):

- Improve the detection, analysis, mitigation and response to sophisticated cyber threats, with a focus on government, critical infrastructure and other systems of national interest;
- 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 infrastructures, networks, products and services;
- Model best practice in the protection of government ICT systems, including the systems of those transacting with government online;
- Promote a secure, resilient and trusted global electronic operating environment that supports Australia's national interest;
- 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.

As part of the 2009 Australian Federal Government strategy, a new of number entities were developed. These include (Rudd 2008):

- *CERT* (*Computer Emergency Response Team*) *Australia.* This new Government body has moved to a national CERT strategy to enable a "more integrated, holistic approach to cyber security across the Australian community";
- *Cyber Security Operations Centre* (*CSOC*). The core functions of the CSOC are focused mainly on government, infrastructure and critical private sector systems and aims to be a source for all issues related to awareness (especially the detection of sophisticated threats) and a facility to respond to cyber security risks and problems which are of national importance. Another key aspect of CSOC is that it provides Australian Defences with a cyber warfare capability and provides a resource designed to service all government agencies (Australian Government 2009).

The 2009 developments again increased the opportunities of the owners and operators of ICS/SCADA systems to co-operate and coordinate with each other as well as the Australian Federal Government.

Since 2010, the Australian Federal Government had started to refocus away from critical infrastructure protection to critical infrastructure resilience. The Australian Attorney General Robert McClelland announced that "The time has come for the protection mindset to be broadened, to embrace the broader concept of resilience" (Australian Government 2010). "The aim is to build a more resilient nation, one where all Australians are better able to adapt to change, where we have reduced exposure to risks, and where we are all better able to bounce back from disaster" (Cherry 2010).

The Australian Federal Government in 2010 launched the new critical infrastructure resilience strategy. The aim of this new strategy is the continued operation of critical infrastructure in the face of all hazards as this critical infrastructure supports Australia's national defence and national security and underpins our economic prosperity and social wellbeing (2011). More resilient critical infrastructure will also help to achieve the continued provision of essential services to the community (Australian Government 2010). The strategy introduces a number of new concepts (Warren and Leitch 2011):

- *Critical infrastructure resilience*—the ability to reduce the magnitude, impact or duration of a disruption to critical infrastructure whatever its cause so if essential services are damaged or destroyed, they can recover as quickly as possible. This is an important part of creating a nation where all Australians are better able to adapt to change, have reduced exposure to risks, and are better able to bounce back from disaster;
- *Mutual responsibility*—the responsibility of critical infrastructure resilience management and operation is shared between the owners and operators of critical infrastructure, and all levels of Australian government (federal, state/territory and local). The owners and operators of critical infrastructure are primarily responsible for ensuring the security of their assets.

The 2010 developments has a direct impact upon owners and operator of ICS/SCADA systems as they had to ensure that their systems were resilient enough to recover from cyber attacks.

In January 2013, Australia's first National Security Strategy, *Strong and Secure: A Strategy for Australia's National Security,* was released. This strategy reinforced the importance of the protection of Australians against many security threats including cyber threats. The strategy proposed that within the next 5 years (2013–2018) the Australian government will develop an integrated cyber policy and operations approach, resulting in a unified cyber security policy approach covering a range of areas such as cyber safety, critical infrastructure protection and resilience managed under a single strategic policy theme (Australian Government 2013).

The strategy defined the need to "strengthen the resilience of Australia's people, assets, infrastructure and institutions" against cyber-attacks. This means the issue is not just one of protecting against cyber-attacks, but also the ability to rebuild systems quickly after a cyber-attack, as well as minimising the impact of a cyber-attack (Australian Government 2013).

A key aspect of the strategy was the creation of the new Australian Cyber Security Centre, that will be in operation by the end of 2013 and aims to improve partnerships between government and industry. This will combine the existing intelligence, defence and law enforcement cyber entities into a single centre allowing for a faster sharing of information between government and industry.

The Australian critical infrastructure policies post 2008 have built upon the partnership, but have introduced the new security concept of resilience for ICS and SCADA systems, that have the ability to quickly rebound from any incident and minimise any impact. The Australian government had also developed the Australian Cyber Security Operations Centre that will allow the real time assessment of cyber security and the sharing of that information between government and the private sector (including utility operators using ICS and SCADA systems).

5 Discussion

In terms of protecting ICS and SCADA systems, there are three key areas that have to be considered; these are policy, business drivers, and technical issues.

5.1 Policy

Historically, the major issue facing Australia with regards to the previously distributed model of decision making regarding critical infrastructure protection was the ineffective management in securing Australia's critical infrastructure in real time. The introduction of the new Australian Cyber Security Operations Centre, however, will streamline decision making and information sharing. What is key in an area such as cyber security is that speed is often of the utmost importance in decision making; the interaction of a large number security and intelligence agencies would slow this process down (Warren and Leitch 2010). A pivotal aspect of the Australian Cyber Security Operations Centre would be the faster decision making time and the quicker distribution of information with corporate partners operating critical infrastructure, and ICS and SCADA systems.

5.2 Business Drivers

A key driver for ICS and SCADA development in the replacement of older legacy systems with newer technology is the requirement to reduce associated ICS and SCADA operating costs. Key business drivers for integration with enterprise management systems had meant that ICS and SCADA systems have become interconnected with corporate networks and directly, or indirectly, with the Internet.

This high level of integration can extend to remote access by operational staff, suppliers and external organisations, further increasing the exposure of these systems to network vulnerabilities associated with Internet threats (Trusted Information Sharing Network (TISN) 2010b; Trusted Information Sharing Network (TISN) 2012). This means that the systems are being developed and modified to reduce business operational cost imperatives.

5.3 Technical Issues

ICS and SCADA systems face new and sophisticated security risks. The most recent has been the development of sophisticated new malware such as Stuxnet. Stuxnet is unique malware; in the fact it will only do damage to particular industrial control system made by Siemens. These "Programmable Logic Controllers (PLCs) as they're known, are used to control automated processes in some key industrial settings, including chemical plants, oil refineries, pipelines, and nuclear power plants" (Trusted Information Sharing Network (TISN) 2008).

The aim of Stuxnet is to reprogram lower level aspects of ICS/SCADA systems by modifying code on PLCs to make them work in a manner the attacker intended and to hide those changes from the operator of the equipment. In order to achieve this goal the malware creators amassed a vast array of components to increase their chances of success. This includes zero-day exploits, a Windows rootkit, the first ever PLC rootkit, antivirus evasion techniques, complex process injection and hooking code, network infection routines, peer-to-peer updates and a command and control interface (Trusted Information Sharing Network (TISN) 2008; Murchu and Chien 2011).

There have been claims made that the target of the Stuxnet attack were the Siemens PLCs that were being used by Iran to control thousands of centrifuges to enrich uranium (Trusted Information Sharing Network (TISN) 2010a). It is claimed with the Iranian Stuxnet example that the malware increased the frequency of the electrical current supplying the centrifuges, causing the centrifuges to spin faster and faster, eventually making them spin at 1,410 cycles per second and thereby critically damaging the centrifuges, and therefore impacting the ability of Iran to enrich uranium (Broad 2010).

It does not matter who developed Stuxnet, or who the target was, the concern is that ICS and SCADA systems are now vulnerable to malware attacks. Historically, SCADA systems have not focused on malware detection and removal for fear that it could slow down the industrial operation in question (Falliere et al. 2011).

As mentioned before, many new ICS and SCADA systems are Internet enabled which allows remote access and operations of the ICS and SCADA infrastructure. This has raised a number of new potential threats and vulnerabilities such as malware infection.

6 Conclusion

Like all developed western countries over the last decade, Australia has taken major steps in the protection of its national critical infrastructure (including ICS and SCADA systems) against new and developing cyber risks.

The Australian cyber security protection model has been an effective model that has developed over a ten year period, becoming increasingly more sophisticated during that time. As part of the sophistication of strategies, the concept of cyber resilience has developed and been integrated into Australia's national security strategies.

An unknown issue is whether the current policies and strategies for protecting Australia against the growing future cyber risks are enough to protect Australia in an ever changing cyber world of security threats (in particular those related to SCADA systems) and whether additional controls and security mechanisms will be needed.

References

Australian Government (1998) Protecting Australia's national information infrastructure. Report of the interdepartmental committee on protection of the national information infrastructure, Attorney-General's Department, Barton, ACT

Australian Government (2004a) Critical infrastructure protection national strategy. Attorney-General's Department, Barton, ACT

Australian Government (2004b) Protecting Australia against terrorism. Department of the Prime Minister and Cabinet, Barton, ACT

Australian Government (2009) Cyber security strategy. Attorney-General's Department, Barton, ACT

Australian Government (2010) Critical infrastructure resilience strategy. Attorney-General's Department, Barton, ACT

Australian Government (2011) CSOC—cyber security operations centre. Defence signals directorate (DSD), http://www.dsd.gov.au/infosec/csoc.htm. Accessed 10 Jan 2011

Australian Government (2013) Strong and secure: a strategy for Australia's national security. Department of Prime Minster and Cabinet, Barton, ACT

Beggs C (2008) A holistic SCADA security standard for the Australian context. In: Proceedings of 2008 Australian information warfare and security conference (Perth), paper 27

Beggs C, McGowan R (2011) Fostering SCADA and IT relationships: an industry perspective. Int J Cyber Warfare Terrorism 1(3):1–11

Broad WJ, Sanger DE (2010) Worm was perfect for sabotaging centrifuges. New York Times. http://www.nytimes.com/2010/11/19/world/middleeast/19stuxnet.html?_r=0. Accessed 15 Oct 2013

Busuttil T, Warren MJ (2004) A risk analysis approach to critical information infrastructure protection. In: Proceedings of the 5th Australian information warfare and security conference, Perth

Cherry S (2010) How Stuxnet is rewriting the cyber terrorism playbook. IEEE spectrum online. http://spectrum.ieee.org/podcast/telecom/security/how-stuxnet-is-rewriting-the-cyberterrorism-playbook. Accessed 15 Oct 2013

Falliere N, Murchu LO, Chien E (2011) W32.Stuxnet dossier. Symantec. http://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/w32_stuxnet_dossier.pdf. Accessed 15 Oct 2013

Hughes G (2003) The cyberspace invaders, the sunday age, 22 June 2003

ISA (International Society of Automation) (2013) ISA99 committee on industrial automation and control systems security, http://isa99.isa.org/ISA99%20Wiki/Home.aspx. Accessed 10 Feb 2014

Krutz R (2006) Securing SCADA systems. Wiley, Indianapolis

Pollet J (2002) Developing a solid SCADA security strategy. In: Proceedings of the 2nd ISA/IEEE sensors for industry conference, Houston, pp 148–156

Rudd K (2008) The first national security statement to the parliament address by the prime minister of Australia, the Hon Kevin Rudd MP. http://www.pm.gov.au/media/speech/2008/speech_0659.cfm. Accessed 10 Dec 2008

Shaw W (2006) Cyber security for SCADA systems. PennWell Press, Tulsa, OK

Slay J, Miller M (2008) Lessons learned from the Maroochy water breach. In: Goetz E, Shenoi S (eds) IFIP international federation for information processing, vol 253, Critical Infrastructure Protection. Springer, Boston, pp 73–82

Smith S (2004) Infrastructure, http://www.parliament.nsw.gov.au/prod/parlment/publications.nsf/0/C6389C30B-0383F9ACA256ECF0006F610. Accessed 10 Nov 2010

Stouffer K, Falco J, Scarfone K (2011) Guide to industrial control systems (ICS) security. Special Publication 800–82, NIST (National Institute of Standards and Technology)

Supreme Court of Queensland, Boden RV (2002) Appeal against conviction and sentence, QCA 164, Brisbane

Trusted Information Sharing Network (TISN) (2007) About critical infrastructure. http://www.tisn.gov.au. Accessed 15 July 2009

Trusted Information Sharing Network (TISN) (2008) What is SCADA? http://www.tisn.gov.au/www/tisn/tisn.nsf/Page/e-Security#_What_is_SCADA. Accessed 3 July 2010

Trusted Information Sharing Network (TISN) (2010a) The shift to resilience. CIR News, vol 7, no 1, Barton, ACT

Trusted Information Sharing Network (TISN) (2010b) Fact sheet: critical infrastructure and resilience: whose responsibility is it? Barton, ACT

Trusted Information Sharing Network (TISN) (2012) Risk management for industrial control systems (ICS) and supervisory control systems (SCADA) information for senior executives, Barton, ACT

Warren MJ (2013) A major step forward on cybersecurity. ABC. http://www.abc.net.au/unleashed/4484508.html. Accessed 10 Oct 2013

Warren MJ, Leitch S (2010) Commercial critical systems and critical infrastructure protection: a future research Agenda. In: Proceedings of the 2010 European information warfare conference, Thessaloniki, Greece

Warren MJ, Leitch S (2010) Development of a supply chain management security risk management method: a conceptual model. In: Proceedings of the 9th European conference on information warfare and security (Thessaloniki). Academic Publishing, Reading, pp 327–333

Warren MJ, Leitch S (2011) Protection of Australia in the cyber age. Int J Cyber Warfare Terrorism 1(1):35–40

Weiss J (2010) Protecting industrial control systems from electronic threats. Momentum Press, New York

Towards Dependable Automation

Jari Seppälä and Mikko Salmenperä

Abstract Automation runs the modern society and it's critical systems. It is a networked software product depending on the co-operation of old and new technologies. Information security for automation systems should be regarded in light of the most important quality required from automation—dependability. This chapter focuses on process of developing dependable solutions for the entire lifecycle of automation systems. The approach includes a guideline for securing automation and a dependability model that is a data flow model extended with security and automation requirements. Results of this analysis should be used in final requirement specification for implementation. Dependability model is the key tool in secure development lifecycle. It can be used in new product development, improving an old automation system and also during the active lifecycle of automation to manage inevitable changes occurring during the entire lifespan of automation system.

1 Introduction

Automation runs the modern society. It is integrated into all processes and devices such as electricity production, distributed heat production, every day goods manufacturing, cars, washing machines and mobile networks. These automation system examples are built on top of hardware, communication networks and software.

Automation is a distributed software product. It consists of measurement, control decisions, acting upon those decisions and integration. It can be distributed over a machine, like a washing machine, or over a country like an electric grid. Hardware in automation environment is a combination of regular IT hardware well known

J. Seppälä (✉) · M. Salmenperä
Department of Automation Science and Engineering, Tampere University of Technology, Tampere, Finland
e-mail: jari.seppala@tut.fi

M. Salmenperä
e-mail: mikko.salmenpera@tut.fi

M. Lehto and P. Neittaanmäki (eds.), *Cyber Security: Analytics, Technology and Automation*, Intelligent Systems, Control and Automation: Science and Engineering 78, DOI 10.1007/978-3-319-18302-2_15

from office environment and domain specific devices and systems. Most high-level automation functionalities, covering areas like Manufacturing Execution Systems (MES), Enterprise Resource Planning (ERP) and system design rely on regular Commercial Off-The-Shelf (COTS) hardware for both software and networks. Remaining hardcore automation has requirements that can only be satisfied by special domain specific solutions. This special hardware includes for example embedded controllers in field devices, Supervisory Control and Data Acquisition (SCADA) systems, Distributed Control Systems (DCS), Programmable Logic Controllers (PLC) and automation substations. Same division occurs also in the network equipment. Part of the network is created with standard Ethernet components. However, the core automation network, providing for example Quality of Service (QoS) properties, uses automation specific field buses, protocols or Industrial Ethernet solutions. Although the latter are based on IEC 61784-2 standard, the reality is that different profiles in the standard have poor interoperability. The automation specific solutions result many times in vendor lock and affect the available security measures.

Typical automation is a rather heterogeneous environment in comparison to regular office. From the hardware point of view, the increasing need for integration forces the automation systems to open up and provide support for Internet protocols, namely TCP/IP protocol family. Integration is so essential in modern automation that TCP/IP is pervasive throughout the automation system. One could say that Internet has reached the entire automation from the lowest field device all the way up to MES and ERP systems. From information security point of view, modern automation can be characterized as a highly networked and widely distributed software product exposed to all the benefits and threats present in modern Internet environment.

The opening of automation systems, wide variation in automation applications and the integration of Internet communication technologies into the heart of automation networks make automation the most complicated target for information security.

Automation is always integrated into business processes (Fig. 1) On the bottom of the pyramid are automation devices, control devices and other field devices. Layer by layer information systems are integrated with communication networks to both higher and lower level systems. In essence the whole automation is one big integrated, distributed system with special requirements rising from the automation aspect.

The most fundamental requirement of any automation application, the core essence of automation, is dependability. Dependability is all the separate means such as robustness, security, safety, usability and all other aspects that affect the overall stability and reliability of the automation system and it's environment. All of these aspects are crucial. This chapter focuses on information security as a tool for improving dependability while taking into account these aspects.

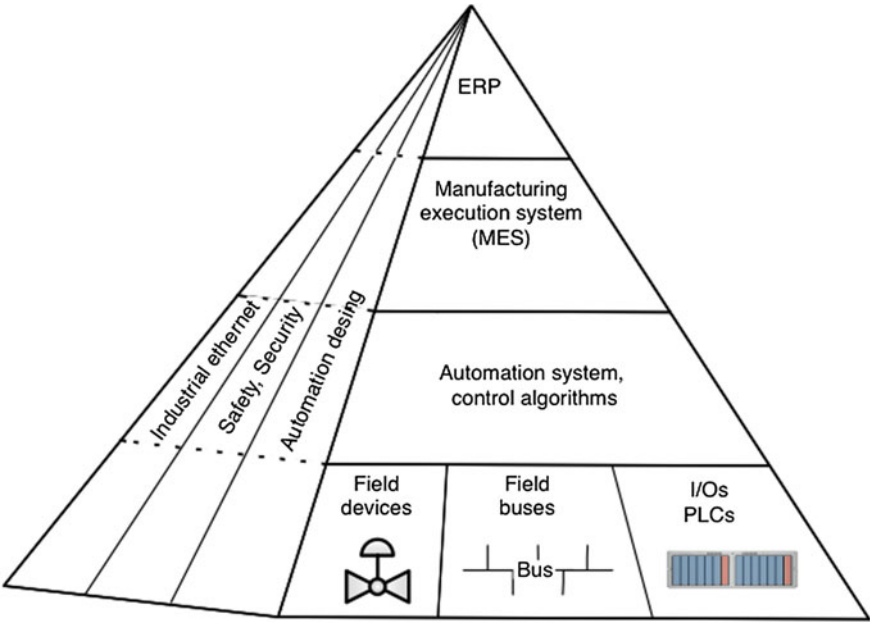

Fig. 1 The general view of a typical business process integrated automation

2 Towards Dependable Automation

Information security for automation systems should be regarded in light of main quality of automation—dependability. Dependable automation systems are built on various technological solutions. For example, safety systems try to keep environment and people out of harms reach and control designs are used to stabilize the complex cyber-physical processes (Lee 2008). Modern Human-Machine-Interfaces (HMI) and different visualization approaches are used for enabling the user to make intelligent decisions in situations where information flow can overwhelm the human operator. The information security, however, is currently not considered a viable technological solution in automation systems. It is still regarded as more of an additional cost factor than enabler for better quality of production and uptime.

This view of security in automation has begun to change. Recent security related events that were presented to the general public are acting as a catalyst in increasing the significance of security for automation. One such event was Stuxnet, (Lagner 2011). It showed how vulnerable the infrastructure in the heart of modern society is against attacks to automation systems running for example energy production and how the threat environment is changing (Bennet 2011).

The knowledge of designed hard coded passwords and lack of adequate security measures was not news to automation engineers. This had been common practice intended to protect against accidental changes. The automation systems were

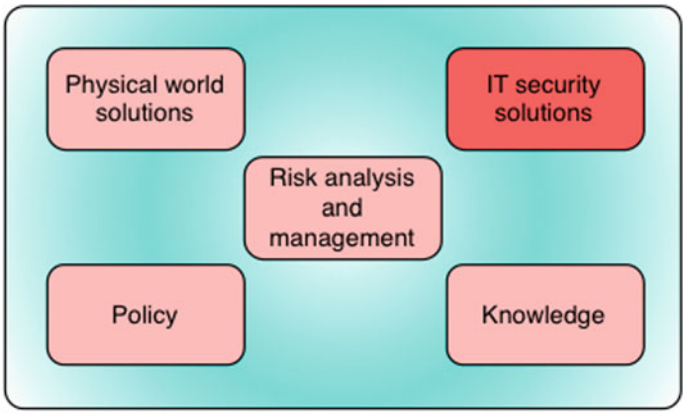

Fig. 2 The enablers of dependable automation

designed with functionality in mind, not with security in mind. However, the main result of Stuxnet with its' predecessors and successors was the rise of perceived risk of breached automation systems for companies and even nations.

IT security solutions can no longer be an additional feature of automation, a nice to have, but it is absolutely necessary. IT security is and will continue to be one core enabler of dependability in modern networked automation systems (Fig. 2).

3 Time Dependence in Automation

Time dependence in automation is a design decision. Control systems are time dependent and require measurements and control signals to be provided with predetermined time frames. Alarms, however, are events with high priority and should take precedence in the network communication. These messaging requirements affect not only the automation software but also the networks and protocols used in automation environments. Until event based control (Åström 2008) proves to be viable solution in automation the operations in time critical loops must be periodic and deterministic. Hence, for networks and protocols there must be a possibility to prioritize messages and enable cyclic messaging. These are called real-time requirements.

Real-time is commonly misunderstood concept. For most of people it means "something very fast". The correct definition uses utility function. The utility function defines the usability of the information related to time. Therefore, a slow control loop with 10 min cycle time might require real-time communication. Control loops in automation systems usually require isochronous real-time, whereas alarm have typically hard real-time requirement (Fig. 3b, d).

The security measures in automation must take these requirements into account. For example, scanning network messages for viruses must not break the real time requirement. Neither the heuristic virus scanning can be allowed to slow the

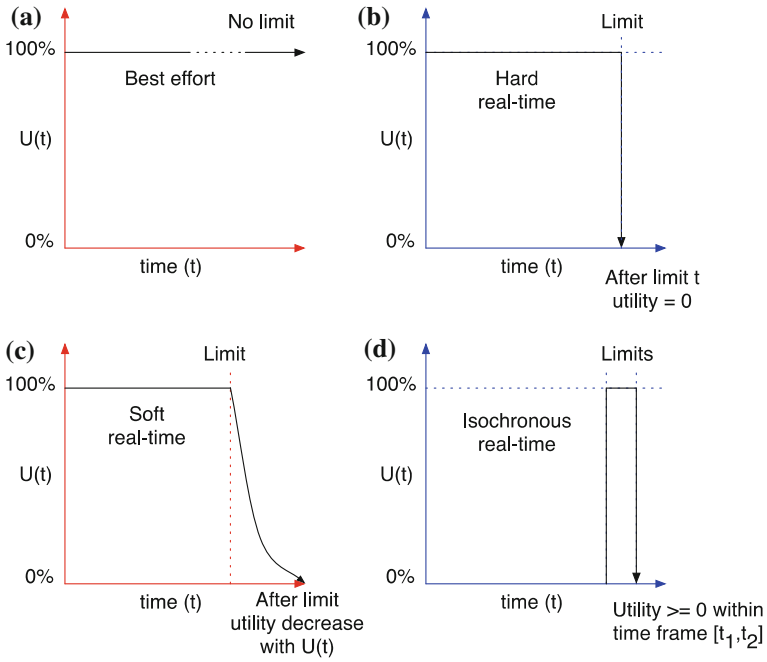

Fig. 3 The normal IT functions usually at *left hand side* (**a**, **c**), whereas automation always contains functions from the *right hand side* (**b**, **d**)

supervisory station(s) in such way that alarm signals are not presented to the system operator. The common IT security measures are designed for IT requirements (Fig. 3a, c). Therefore, applying them to the automation environment requires careful considerations, planning and deep knowledge of the automation system.

Automation systems have long lifespan. It is usually measured in decades than years. This assumption is true however only for the core automation system. The core system consists of valves, actuators, PLCs, pipes that are the hardware in the system. The core hardware contains also the operating stations and integrated information systems. The upgrade cycle of the automation hardware can be for example 5 years. This means that every 5 years some part of the process is updated. The upgrade cycle time frame depends on the system design, the system itself, the maintenance contracts, the state of business and so forth.

The core automation system can be considered stable. However, the environment around automation system changes more rapidly (Fig. 4). The typical lifecycle of modern computer systems accessing the automation is around 3 years. From automation system point of view, this means for example that computers used for maintenance tasks can lose the essential hardware support necessary for connecting to the automation. Good example of this was the removal of serial port from laptops. This change in the environment can be characterized as COTS effect (Fig. 5).

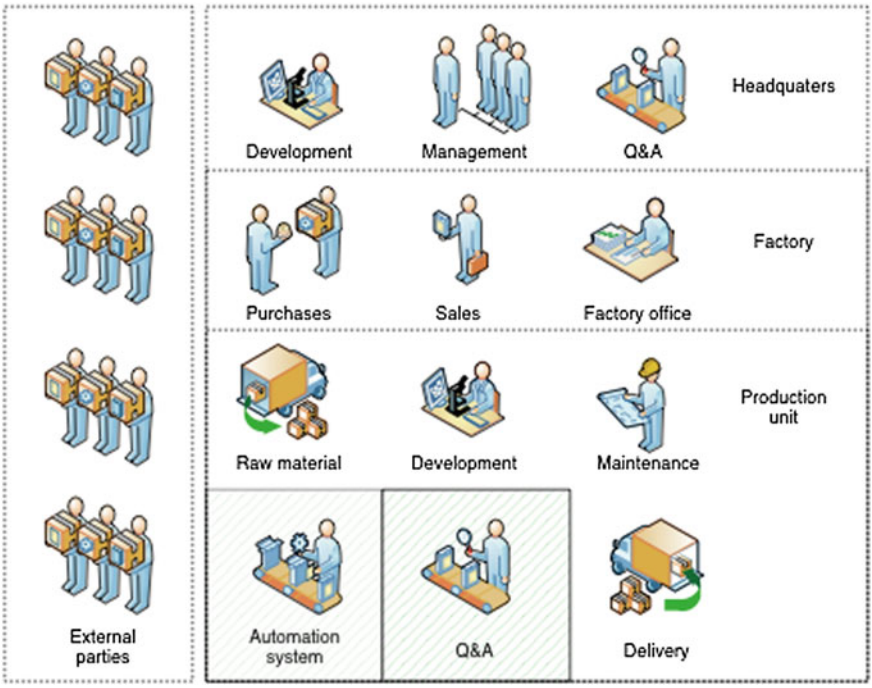

Fig. 4 Core automation (automation system, Q&A), internal (development, management, purchases etc.) and external (subcontractors, raw material providers etc.) parties

Replacing vendor specific hardware with COTS has been one of the cost savings methods for that last decade. Due to harsh automation operation environment, this is not straightforward to apply. For example, the Ethernet connectors in automation may look totally different from their COTS siblings.

COTS also changes the security risks for automation. Tools available to any interested party are rapidly becoming easier to use. Main reason for this is ever increasing presence of TCP/IP protocols in automation systems. (Un)fortunately this also means that security testing toolkits such as KALI Linux distribution can and should be also used in automation environment—with great caution (Linux 2013).

Another rapidly changing area related to automation systems is development tools. These tools follow the normal fast lifecycle of office tools causing problems for automations systems. For example, a PLC might still require old DOS based tool used via serial port to configure the IP address for the device. For automation system owner this means saving an old computer for maintenance purposes.

Many security professionals criticize the existence of such old technologies. As many times, they must be told that it is a business decision. If the system is functioning as expected, the cost of renewing the devices is not meaningful. This planning is usually done already in early purchase phase, when the lifecycle

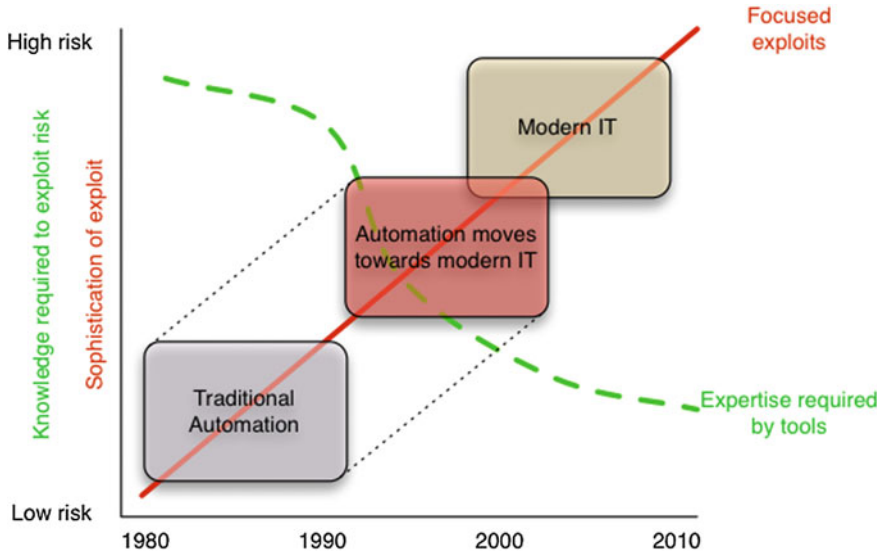

Fig. 5 COTS effect that is the transformation from traditional to modern IT solutions and evolution of security tools and knowledge

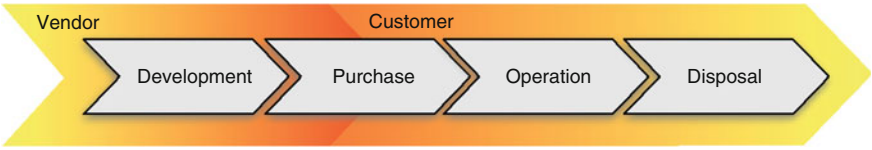

Fig. 6 The lifecycle of an automation system

management is decided. It should always be remembered that automation system is purchased as part of production system or machine. The most important goal for the system or the machine is to produce the product or do the task it was designed for.

The automation lifecycle starts from development and ends with disposal of the system (Fig. 6). The security lifecycle follows the same cycle. The notable difference is that security lifecycle consists of main responsibility change and/or division between vendor and the customer. Also other parties may hold some part of the security responsibilities.

In light of these challenges, it is vital that security is included into whole lifecycle of the automation product starting from development, including the purchase and ending at the disposal of the system. None of these phases is more important than the other. This chapter focuses on development and purchases. Operations and disposal are covered in short.

4 Security Challenges

Modern day information security threats have evolved from simple amateur hacker made viruses into a whole ecosystem of online crime. Latest addition to this ever-changing threat scenario is the governmental operators running large scale targeted operations of industrial espionage and sabotage. Modern information security must consider all of these threats in regard to motivation and goals of operators. Three archetypes can be identified, namely hacktivists, organized crime and governmental operators. Each of these types has their own agenda.

Hacktivists are more or less traditional amateurs trying out their skills and showing off to their peers. Their resources are typically limited and motivations obscure. Damage caused is more often accidental than intended. When considering the nature of automation they pose a serious threat nonetheless as automation is a fragile real-time system. Any breach in automation security might have serious and unpredictable results even if they are unintended.

Organized crime has formed a whole business ecosystem around online crime. Criminals have a clear purpose in their operations; making money. They are prepared to finance operations with significant capital. They are hiring expensive experts and the resources at hand far exceed those of hacktivists. Even if the threat they pose is today mainly focused on online shopping and banking it is only a matter of time until someone invents how to make money by compromising automation security. There are also fringe threats caused by their operations. Such threats include ever growing botnets and similar constructs that might cause serious unintentional problems to automation systems.

The Stuxnet showed the arrival of governmental operators to the online era. Their operations are well funded, coordinated and extremely difficult to defend against. The Stuxnet was developed for years; it was well tested and then deployed against a single facility. It contained several zero day exploits, counterfeit certificates used to sign code so that it appeared to be originating from acceptable sources and advanced remote control systems used to steer the attack towards intended target. Other attacks parallel to Stuxnet were also able to gather relevant information about all systems they infected and transmitted this information back to controller.

5 Securing Automation Lifecycle

The goal of modeling in automation security is to extend formal requirement process analysis and requirements specifications to span over the entire automation lifecycle; starting from business needs and processes through software design and implementation to active use of automation systems. Long lifecycle and resistance to upgrades in automation makes this kind of approach necessity. Contrary to office application automation remains in use decades after being deployed. Updates are

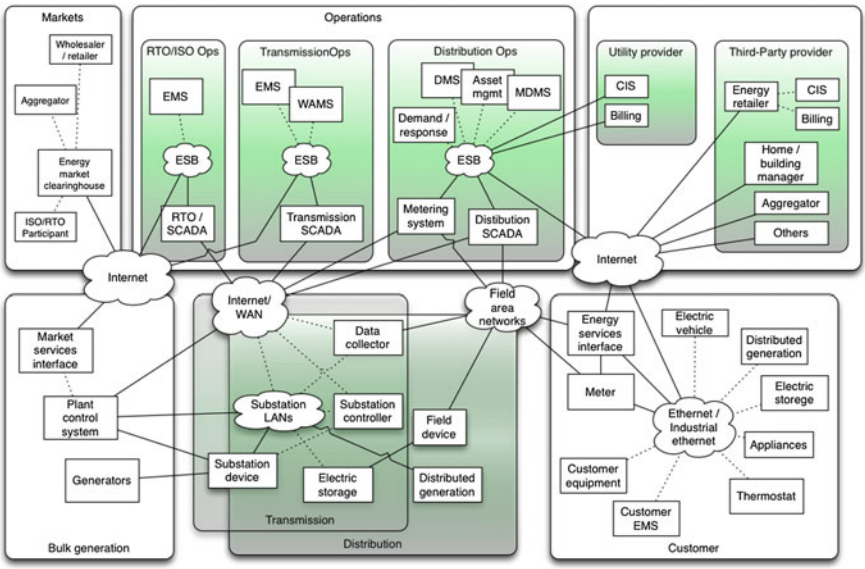

Fig. 7 Conceptual diagram for Smart Grid information networks

possible, but major changes to software architecture and design patterns used in both design and implementation are impossible or expensive. This requires long-term approaches and practices for dependability analysis and design for the automation systems.

The advancements in IT during the recent decade are one of the main reasons why a smarter automation system is now a possibility. The key concept is to acquire more information for better decision-making. The key technologies for securing automation already exist; challenge is only about applying them successfully—and in a scale never seen before. Typical key qualities are dependability, reliability and resiliency, which are in the core of the system. The modern automation is also a key part of society's critical infrastructure elevating requirements for information security. Third distinguishing feature is the sheer complexity of large-scale integrated automation application, for example Smart Grid: they are a 'system of systems', a mega systems formed of interconnected subsystems (National Institute of Standards and Technology 2012). It is difficult to comprehend all the connections and cause-effect relationships. Even relatively simple faults can cascade or unexpectedly propagate from one system to another, possibly causing massive failures or blackouts. Smart Grid is an example of the enormous scale of a large scale system (Fig. 7).

Integration of information systems in such a large-scale system is challenging. The most common legacy integration architecture is point-to-point integration where each connection between systems is maintained separately. This leads into scalability issues as number of connections between systems grows rapidly by the number of nodes. The system quickly becomes impossible to administer. Common

integration architectures are Enterprise Integration Architectures (EAI) and Enterprise Service Buses (ESB). Fundamental idea of ESB is to separate application and integration logic in a distributed manner, whereas EAI is centralized hub solution (Chappel 2004).

ESB is well-suited ideology for large-scale environments. It defines a set of core functionality forming a messaging fabric and a common set of components used to build the actual integration solutions. It has no single point of failure, compared to centralized hub, making the architecture more dependable.

6 Guideline for Dependable Automation

Integration needs of automation environment are rather heterogeneous. However, the core of information systems and the roles of these systems are rather well defined and thus it is possible to provide guide for the implementation of the dependable automation. These general steps relating to automation integration cases can be identified. This model for developing automation applications is derived from Eerola (2013) and Salmenperä et al. (2013).

6.1 Create an Essential Model Through Business Analysis

An essential model (business/abstract model) defines the purpose of the use case without technical details. It answers the "what" is done, but does not state "how" it is done. The essential model clarifies what is the business need that the use case is solving, explaining it in terms that are understandable to experts from various fields. The result could be expressed as a simple data flow diagram or use case diagram, combined with text (Fig. 8). This step should also define high-level security and automation dependability requirements and business risks (Brown 2008).

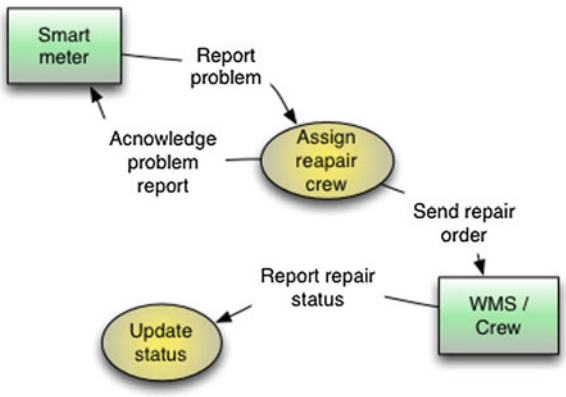

Fig. 8 The illustration shows a DFD (Data Flow Diagram) representation of the essential model of a fault repair use case in Smart Grids (Flick and Morehouse 2011)

The use case starts with smart meter reporting about a problem. The business need is to assign a workforce crew to repair the problem. The Work Management System (WMS) will receive a repair order from the system and dispatch a crew to fix the problem. WMS then reports back to the system about the status of the repair work. The smart meter and WMS are technical details, but in this scenario and from the integration solution point of view, they are external systems. This model is intuitive and easy to grasp for non-electricity professionals as well, and does not indicate how the system will function internally.

6.2 Define the Use Case Explicitly

After answering "what" is done, the next step is to describe "how" things are done, including some technical details about the system. The use of middleware or the integration architecture in general should not be defined yet. This can be achieved, for example, using data flow diagrams and ignoring (for now) how the flows will actually be implemented. Thus, this step should contain some level of technical detail and participating systems.

Smart meter fault use case data flows start from alarm event at a smart meter leading to dispatch of repair crew (Fig. 9). It includes already some technical details

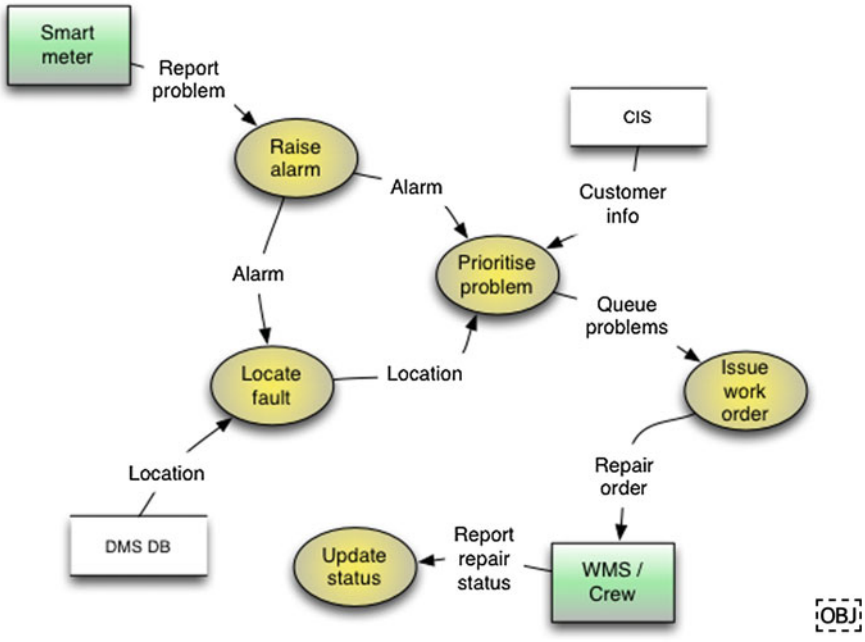

Fig. 9 Data flow diagram defining the Smart meter fault use case in detail

of the sub-processes and systems. This step can benefit from utilizing general use case modeling principles and rules. Questions that should be answered are for example: who or what initiates the use case? It could be triggered by an event, a user action, or it could be time-based or a continuous "housekeeping routine", for example. What type of coordination or orchestration is required? This depends a lot on how complex the case is. At this point, it is not necessary to determine which system will take care of the coordination or orchestration. What is the degree of automation? That is, does the process require human intervention at some point?

6.3 Determine the Participating Information Systems

This is very logical: it is important to list the systems that need to be integrated. After defining the use case, it is clear which systems will participate. However, it is important to actually list them. This helps to clarify that all required systems are identified, and whether there are systems mentioned that are not necessary. The use case definition shows what information and functionality is required. Based on that it is easy to see which systems contain the necessary information and can perform the required functions. In our example case, the Distribution Management System (DMS) and Customer Information System (CIM) are participating in the case (and also the Smart meter and WMS, which are external systems).

6.4 Define the Orchestration of the Process

The general use case definition defined what type of orchestration is needed to manage the entire transaction. This step will concentrate on how the orchestration is implemented. Possible error conditions and situations should be considered in this step. Simple data transfers require very little orchestration contrary to more complex long running transactions. It is important to define explicitly how much complexity of the transaction is handled. This is also a design decision: should for example the middleware handle the transaction, or is the intelligence within the endpoint systems? This has major consequences on how the integration will be implemented. If the orchestration logic is mostly within the endpoint systems, the integration solution's role is mainly to function as a message broker.

In the example case, the orchestration could be handled by the DMS. The DMS would receive the alarm, and it is then responsible for locating the fault, consulting the CIS for customer information, prioritizing the problem, and sending a repair order to the WMS. It will also update the status based on the reports from the WMS.

6.5 Define and Implement Processes

Analyze the use case and the participating systems and their functions regarding selected paradigm. Most important software products within the electricity industry will likely remain as large, monolithic structures for the foreseeable future. However, certain functionality or information within these systems can be exposed, so that other systems have easy access to them.

6.6 Define Data Flows

The earlier diagrams define many data flows between processes and information storages. Some flows are internal to the endpoint systems, and some are inter-system flows. The latter are the ones that the selected implementation framework will need to handle. This step can identify critical failure points in the system and shows how the earlier design may have a major dependability impact on automation.

6.7 Define the Information Content of Data Flows

Based on the data flows defined on the previous step, it is straightforward to define what information each flow contains. The result of this step is a comprehensive definition of the information content of each flow.

6.8 Create Dependability Models

The information content, each of the data flows has certain security and automation requirements; privacy, integrity, confidentiality, alarm, real-time and auditability. This step should define those requirements. Collaboration between an automation expert and an information security professional is strongly recommended for this step. Information security experts know the right questions that will bring up security requirements; automation experts understand the system and know how to answer these questions.

6.9 Choose Information Security Implementation Methods

Once the security requirements are defined, appropriate security implementation methods should be used. This includes both general decisions (e.g. “this data flow

needs to be encrypted in order to provide confidentiality"), and implementation specific details ("this design choice offers these technologies for encryption").

6.10 Implement the Solution/Orchestration

Last step is to implement the actual solution. Secure development process should be used to ensure that the requirements are met.

7 Securing Development

Creating secure automation starts from the design board. Taking into account the security challenges, threats, trust zones etc. must be integrated into normal development process. Information flow diagrams used together with threat modeling provide solid and clear base for securing development. After all, it is information and its' flow we are trying to secure.

Information flow, or data flow, diagrams were introduced in mid 1970s for structured analysis and design (Stevens et al. 1974; Yourdon and Constantine 1979). Since then information flow diagrams have been accepted as standard tool for various modeling tasks especially in the software engineering. Nowadays the Unified Modeling Language (UML) has replaced basic information flow diagrams tools as an engineering tool and it has various presentation formats that are powerful for many uses (Douglass 2004). However, UML tools are intended mainly for automatic code generation and other software domain specific purposes. UML tools are too deep in modeling the functionality in software products, making these tools unsuitable for creating common ground for discussions.

Automation engineers are used to handling information flows since the flow concept of input-output models with flows is built-into their education (Sterman 2000). They are also used in various areas of systems thinking and modeling. Software people are also familiar with data flow diagrams, a form of information flow representation, in their software or telecommunication education. The use of information flow concept in both areas leads to obvious conclusion that information flows could be used as common representation form for information security related discussions between customer and vendor or automation engineers and software developers. However, there must be a starting point for the discussions.

Communication between humans around shared problem should start with agreeing upon the definition of the problem. In information security this can be done by utilizing use case concept (Douglass 2004). For example, the IT personnel might describe file transfer task as information security problem "User attached USB stick to automation system PC. The stick might contain malware." The automation engineer might describe same task from the usability viewpoint "I have to transform firmware or system settings updates with USB stick since the computer

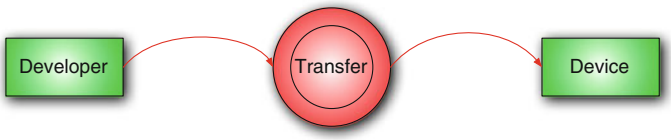

Fig. 10 Context diagram of the firmware transfer use case. Unreliable links should be clearly marked

operating the system is not attached to network." From these descriptions a use case, which require co-operation between automation and IT departments can be created. Let us call it "transferring new device firmware to automation system without network connection." This use case is an example, which can be solved without use of information flows. It is simple enough to explain the idea behind using information flows as security analysis tool (Fig. 10).

8 Dependability Model

Information flows can be used for analyzing security of a system. The traditional approaches are not suitable for automation environments, since they do not take into account the automation requirements.

Automation networks require capability for real-time traffic, possibility to prioritize and classify information, and capability to build fault tolerant networks. The dependability model extends traditional information flow to include the first two automation requirements as well as audit requirements based on the transferred information. The dependability model will consist of the requirements special for automation system and includes also human factor. It should be integrated into normal product development cycle (Fig. 11).

At first the dependability model is created describing all information flows related to use case and product as well as the automation requirements and the security requirements for the information items in the flow. This model can then be used as input for company's Secure Development Lifecycle (SDL) containing all the development steps company has taken to produce secure products. SDL processes contains for example used threat analysis methods, lists of accepted mitigations to threats, secure application programming interfaces and training for producing secure code. The SDL can be implemented as procedures and tools in development environments or plain lists for programmers. In practice, the implementation must be light enough to be used but rigid enough for auditing if needed. Good example of SDLs is Microsoft Security Development Lifecycle (Howard and LeBlanc 2003; Howard and Lipner 2006).

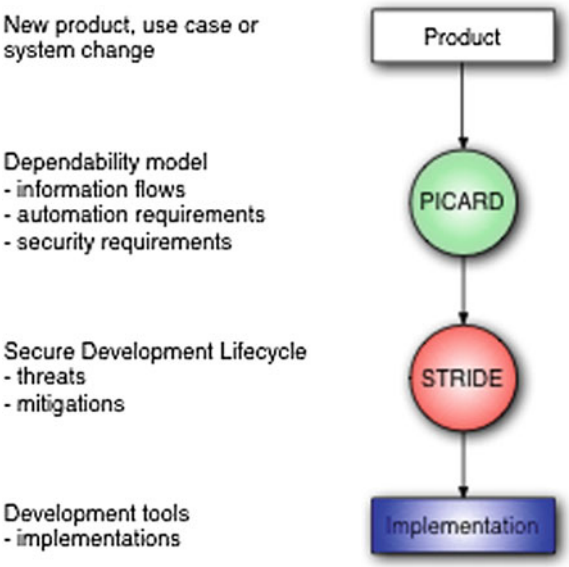

Fig. 11 Securing development starts from product concept via dependability model and ends with secure development supporting tools

9 PICARD Extension

The threat model categorization, like Microsoft's STRIDE approach (Howard and LeBlanc 2003), is good tool after the dependability model has been created. It is suitable for software development but not for high level information flows used as basis for discussions between parties related to the project.

The dependability model uses extensions required for automation. It contains two new features for the standard information flow diagrams. Compared to the traditional information flow, in our approach the flow is an object containing the information data of the flow. The first extension is security and automation requirements of the information inside the flow. This is marked as text PICARD where the interpretation of individual letters is described in Table 1.

Table 1 Description of PICARD extension

Letter	Short for	Description
P	Privacy	Personally identifiable information (PII)
I	Integrity	Integrity of the information is required
C	Confidentiality	Information is confidential e.g. process optimization configurations, or material requiring IPR protection
A	Alarm	Event based information requiring prioritization
R	Real-time	Cyclic information which must be deterministic i.e. delays are known and without unspecified variation
D	Auditing	Information has auditing requirements e.g. Sarbanes-Oxley (2002)

The requirement for information flow is presented as bold caps letters in flow diagram

The second extension is color-coding of the link, process or storage according to the trust that can be placed on it. For example, red can be untrusted link over Internet. Green can be trusted link inside the automation network. Black or blue can be link, which is not yet classified. The selected colors do not matter, as long as the same are used throughout the automation project.

The main idea of extended model is simple and can be divided into 10 phases:

1. Create context model of the use case with external entities. This can be extracted from the previously defined explicit use cases.
2. Identify the trust boundaries.
3. Color-code the components based on trusts.
4. Create next level information flow diagram.
5. Identify the trust boundaries.
6. Color-code the components based on trusts.
7. Add security requirements to flows (apply bold to letters P, I, C).
8. Add automation requirements to flows (apply bold to letters A, R, D).
9. Add red color code to untrusted/unsafe flows from automation perspective (e.g. over Internet, over office network). Leave unclassified to black.
10. Iterate from 4 until desired detail level is achieved.

Note! The auditing can be considered to be both security and automation requirement.

The STRIDE approach is used to analyze and decide mitigation strategies whereas the PICARD is more suitable for modeling information flows inside an automation system. Another difference between STRIDE and PICARD is that PICARD's target is not to prioritize the threats rather classify the trust and requirements of the paths, processes, storages and external entities related to the information flow.

Dependability models can be quite large since they present the whole information flow chain of a use case. For example, dependability model of general firmware transfer use case in Smart Grid environment can be divided to two transactions. First transaction is from trusted developer, using company's web server, to trusted internal network and second from configuration laptop to the end device (Fig. 12).

Security decisions are always based on assumptions. In the firmware update to device case the assumptions are that we can trust the developer and the storage of the distribution web server. This of course requires that security measures for trusted parties are functioning and for example web server is properly hardened.

USB sticks are unreliable transfer devices, due to the human factor. Therefore, it should be verified that nobody has been able to alter the contents of the firmware package. Good practice is to use different virus scanner provider than in downloading phase. In the last update phase where firmware data is uploaded to the device, there can be other communication links used than Ethernet. In this case the most vulnerable part of this flow is the laptop used for upload.

The dependability model is used as input into company's Secure Development Lifecycle. The models are as good as the information used to create them. They must be updated when new technology is presented into the environment or into the

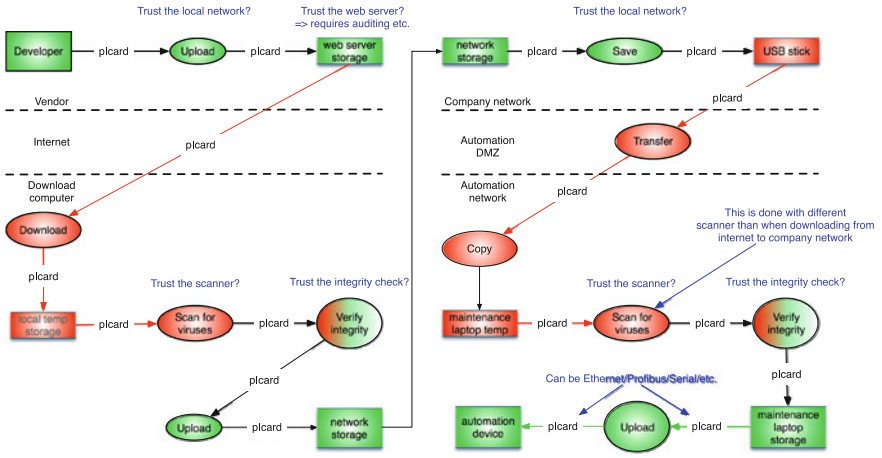

Fig. 12 Dependability model of general firmware transfer use case in Smart Grid environment. Note the change of flow content in the last phases

automation system. This is similar to updating the automation schematics when a process change is introduced. The dependability model should be viewed as another essential process diagram.

10 Securing the Purchases

Poorly orchestrated purchases have been the source for many problems in products containing software. This is especially the case in large software systems and automation is a large networked software system. This blame is valid. However, blaming the customer or the vendor is meaningless since the core problem lies with the security knowledge and interaction of these parties.

This communication challenge has been recognized for example in Finland where nation wide project "Common Requirements for Vendors" was started in 2011. The project was managed by Technical Research Centre of Finland (VTT) and funded by National Emergency Supply Agency. Participants to the project included over 40 different vendors, users, integrators, Tampere University of Technology and Finnish Communications Regulatory Authority (CERT-FI). The project resulted in requirement table where sheets presented specific requirements in different phases of purchase process.

The resulted table is exceptional since it not only provides requirement but also the goal and an implementation example for each requirement line, unlike common security requirements and standards. The approach is intended for (1) shrinking the discussion gap between the customer and vendor, (2) distribute the security

knowledge for each party and (3) include the lifecycle from purchase to maintenance. This kind of commonly-agreed-upon-approach is required to secure the purchases.

11 Securing the Operations

The key to securing operations is following the decisions made in development and purchase phases. The maintenance plans must contain update procedures for dependability and threat models. The best practice is to create a change procedure, which requires update of the models.

Automation systems cannot be scanned for security while in operation. The systems usually contain real-time control loops and heartbeat links, which monitor the operation of the system. Poorly implemented security scan may disturb the network or devices causing the safety system to initialize system shutdown. It must be noted that optimizing the performance of an automation system results in more delicate balance of the system. Therefore, the security testing must be integrated into normal maintenance scheduling.

12 Securing the Disposal

The difference between secure disposal of an automation system and IT systems is to know where critical information might be stored. In automation system the sensitive information can be found in odd places. It can be found for example in valves, PLCs and other embedded devices. Therefore, the "know your system" is essential in disposing a production line, automation device or even a whole automated factory.

Automation systems tend to change ownership during their long lifecycle. Secure disposal procedures must be considered not only in final disposal of the system but also in every change of ownership. From company A's point of view, the transfer of an automation system to company B is disposing an asset. Process optimization knowledge may or may not be part of the transfer and it is vital to know where this information can be found.

13 Conclusions

Automation is a networked cyber-physical software system with additional requirements rising from inherent relationship with real world. It can be found in all areas of modern society for example cars, energy and everyday goods production,

washing machines and mobile networks. The ultimate goal for automation is dependable production.

All properties of automation system and changes to them are business decision where many aspects are considered. One of these properties is information security. Proper information security can only be attained if the process for it starts from design board and lasts the whole lifecycle of the automation system.

The guideline for dependable automation starts from modeling the business cases for the product continues as more detailed presentations of the system and ends with implementations and development decisions. It is a systematic approach based on information flow models, and takes into account both the security requirements and automation requirements. The guideline aims to model dependability requirements in such way that it is easily updated and applied not only during initial development of the system but also during the rest of its lifespan.

Threat models are used in software engineering to secure the product. However, the traditional threat approach is not sufficient in automation domain. Dependability model is an extended information flow diagram where automation requirements are integrated into flows. It provides a common ground for discussions between vendor and customer, and provides knowledge of the essential requirements for system development. The business aspect of using information security as tool for improving dependability is providing more reliable and resilient system, therefore, enabling better uptime.

Dependable automation can be achieved by integrating security into the product. Information security complements the control, monitoring and safety features by taking into account the intentional, and also unintentional, human factors like hacktivists, criminals and governments. Dependable automation provides competitive edge for the customer by decreasing production downtime, and for the vendor by increasing trust for the product.

References

Åström KJ (2008) Event based control. In: Astolfi A, Marconi L (eds) Analysis and design of nonlinear control systems: in honor of Alberto Isidori. Springer, Berlin, pp 127–147

Bennett S (2011) Insecurity in the supply of electrical energy: an emerging threat. Electr J 24 (10):51–69

Brown D (2008) The how to of essential modelling. IRM Training—White Paper, Melbourne. http://www.irm.com.au/papers/How_To_of_Essential_Modelling.PDF. Accessed 28 Aug 2008

Lee EA (2008) Cyber physical systems: design challenges. In: 11th IEEE international symposium on object oriented real-time distributed computing (ISORC). IEEE Press, New York, pp 363–369

Chappell DA (2004) Enterprise service bus. O'Reilly, Sebastopol

Douglass BP (2004) Real time UML, 3rd edn. Addison-Wesley, Boston

Eerola R (2013) Analysing integration and information security: enterprise service bus for smart grid. Master's thesis, Tampere University of Technology, Tampere

Flick T, Morehouse J (2011) Securing the smart grid: next generation power grid security. Syngress, Burlington

Howard M, LeBlanc D (2003) Writing secure code, 2nd edn. Microsoft Press, Bellevue
Howard M, Lipner S (2006) The security development lifecycle. Microsoft Press, Bellevue
Linux K (2013) The most advanced penetration testing distribution. http://www.kali.org
Langner R (2011) Robust control system networks. How to achieve reliable control after Stuxnet. Momentum Press, New York
National Institute of Standards and Technology (2012) NIST framework and roadmap for smart grid interoperability standards, Release 2.0. NIST Special Publication 1108R2
Salmenperä M, Eerola R, Seppälä J, Koivisto H (2013) Design and analysis of secure integration solution for smart grids. In: Proceedings of Automaatio XX-seminaari, automation and systems without borders—beyond future (Helsinki, 2013). Finnish Automation Society
Sarbanes-Oxley Act of 2002. Corporate responsibility. In: 107th Congress Public Law 204. Accessible from http://www.gpo.gov/fdsys/pkg/PLAW-107publ204/html/PLAW-107publ204.htm
Sterman JD (2000) Business dynamics. Systems thinking and modeling for a complex world. McGraw-Hill, New York
Stevens WP, Myers GJ, Constantine LL (1974) Structured design. IBM Syst J 13(2):115–139
Yourdon E, Constantine LL (1979) Structured design: fundamentals of a discipline of computer program and systems design. Prentice Hall, New Jersey

Specialized Honeypots for SCADA Systems

Paulo Simões, Tiago Cruz, Jorge Proença and Edmundo Monteiro

Abstract In this chapter we examine the role of specialized honeypots for detecting and profiling cyber attacks on SCADA-based Industrial Control Systems, debate how to implement such honeypots and provide a complete example of such an appliance. The honeypot concept has been used in general-purpose intrusion detection systems for a long time, with well-recognized contributions in revealing and analysing cyber attacks. However, a number of specialized requirements associated with SCADA systems within Industrial Control Systems in general are not addressed by typical honeypots. In this paper we discuss how the different approaches to security of typical information systems and industrial control systems lead to the need of specialized SCADA honeypots for process control networks. Based on that discussion, we propose a reference architecture for a SCADA network honeypot, discuss possible implementation strategies—based on the lessons learned from the development of a proof-of-concept Modbus honeypot—and propose two alternative deployment strategies, one based on low cost hardware appliances physically and logically located in the automation or field networks and the other based on virtualized field network honeypots physically located in the datacentre and logically located in the field or automation network.

P. Simões (✉) · T. Cruz · J. Proença · E. Monteiro
CISUC-DEI, University of Coimbra, Coimbra, Portugal
e-mail: psimoes@dei.uc.pt

T. Cruz
e-mail: tjcruz@dei.uc.pt

J. Proença
e-mail: jdgomes@student.dei.uc.pt

E. Monteiro
e-mail: edmundo@dei.uc.pt

M. Lehto and P. Neittaanmäki (eds.), *Cyber Security: Analytics, Technology and Automation*, Intelligent Systems, Control and Automation: Science and Engineering 78, DOI 10.1007/978-3-319-18302-2_16

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 power plants (nuclear, thermoelectric, wind farms), management of distribution grids (electricity, gas, oil, water), control of sewage treatment facilities and many other applications.

In the past SCADA systems were restricted to isolated environments, relatively safe from external intrusion. These original systems were very simple by nature, mainly because there were no formal data processing or memory mechanisms involved, being little more than reactive systems interconnecting sensors and visual indicators. However, that simplicity was also their main drawback, being unfeasible for usage anything other than small-scale and physically limited scenarios. Also, since data logging was impossible, error or failure-debugging capabilities were very limited.

With time, SCADA systems evolved to their present situation, with the use of distributed topologies, RTUs (Remote Terminal Unit), PLCs (Programmable Logic Controllers) and more evolved data processing and networking technologies, the latter replacing legacy telemetry system interconnects. However, some of their components (for instance, PLCs) have a lifecycle that frequently spans several decades (it is common to find devices based or architectures with 20 or more years in use). This longevity has to do with maturity—a feature favored in critical systems, traditionally associated with reliability.

One of areas where SCADA systems have evolved considerably is in terms of their communication and interoperability capabilities. While original systems were isolated and self-contained by nature, they progressively started to open to the exterior world, making use of data communication networks for its own internal purposes and to share information with the outside world or even other systems. These connections might exist for various reasons: with the general purpose corporate Local Area Network (LAN), to exchange information with performance auditing or stock management applications; a Wide Area Network (WAN) connection to connect to other facilities (for instance, two power stations) or to an operations control center, separated miles away. Such WAN connections might be ensured using leased lines, dial-up or, more recently, the Internet itself (Ten et al. 2008; Davis et al. 2006). Additionally, it is frequent for original device manufactures to provide remote assistance using such mechanisms.

Also, proprietary equipment and protocols were also the norm on older ICS/SCADA systems, limiting interoperability between devices from different manufacturers (and sometimes, albeit less frequently, between different models of the same manufacturer). This created a situation of vendor lock-in that forced the customer to remain attached to a specific device family form a particular manufacturer due to the cost of migration. Presently, equipment and protocols have been

standardized, with the adoption of COTS (*Commercial Off-The-Shelf)* equipment whenever possible—for instance for LAN communication (Igure et al. 2006).

As a consequence of the introduction of data processing capabilities to SCADA systems, together with the evolution of embedded systems, operating systems also became part of the SCADA ICS ecosystem that evolved with time. From proprietary systems, the situation evolved up to the point where Windows or Unix-derivatives (Creery and Byres 2005) are being used, together with real-time operating systems such as VXWorks or Real-Time Linux (Davis et al. 2006).

This evolution brought significant benefits to SCADA systems, in terms of functionality, rationality and cost. However, it is also closely related with some of the most important security issues that currently affect those systems, which over time 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 et al. 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. Insecure protocols such as Modbus (2006) for instance, are still widely used in production systems. Moreover, as pointed by Clarke and Reynders (2004), 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.

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, as discussed by 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:

- Several works, such as Fovino et al. (2010) and Zhu et al. (2011), identify a significant number of ICS components that need to work on a continuous basis, without interruptions, up to the point of working years without being reinitialized. Stopping or reinitializing these components, even for short periods of time and planned maintenance windows, may provoke considerable cost raises in the industrial processes they manage.
- 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 ICT 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). This European project focuses 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 this project, SCADA honeypots for process control networks play an important role, providing the means for preventing or detecting attacks that directly target SCADA equipment, like PLCs or 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.

This chapter addresses the main aspects of the proposed SCADA network honeypot architecture, including the discussion of architectural options and implementation aspects. The rest of the paper is organized as follows: Sect. 2 addresses the problem of security in ICS. Section 3 provides an analysis of existing honeypot implementations and other related work. The proposed SCADA network honeypot solution is presented in Sect. 4, and the issues related with its implementation and deployment are discussed in Sect. 5. The final conclusions are presented in Sect. 6, along with some discussion of future work.

2 Security of Industrial Control Systems

Despite apparent similarities, there are several and significant differences between ICT and ICS domains when it comes to security. These differences are deeply rooted in their own specific characteristics.

The fundamental reason for these differences relates with the different mindset these systems are built with. This mindset is clearly reflected on the different priorities of ICS and ICT systems. Confidentiality and security have maximum priority for ICT networks, 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, as noted by ISA-99 (2007): availability comes first, even if at the cost of integrity and confidentiality. This difference of priorities, illustrated in Fig. 1, 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.

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, due to these differences. The impossibility or high cost of stopping production can be pointed as an example. It is also common to have the system's manufacturers blocking these updates. A simple example is the fact that a critical facility operator cannot install an update on a host operating system such as Windows unless the manufacturer of the SCADA software certifies it for the update. This is logical, since there may be unknown and dangerous interferences between the new version of Windows and the SCADA software. However, the certification process can be painfully slow, 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 widely known dangerous vulnerabilities. The critical facility operator is left with the dilemma of maintaining operating systems he knows for sure are not safe, effectively putting the ICS system at risk, or loosing the SCADA software warranty and support (and risking possible interferences) by patching the operating system.

SCADA communication protocols, which are responsible for the interaction between field and automation network 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 protocols is 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 (Triangle MicroWorks Inc 2002). Despite these known issues, SCADA protocols such as Modbus have a long lifespan and are still being massively deployed and used.

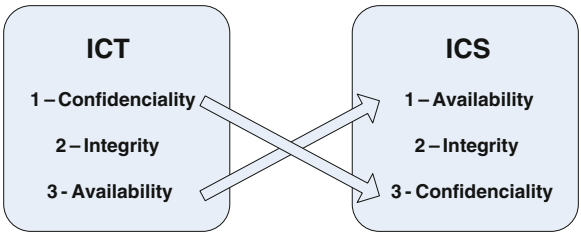

Fig. 1 ICT versus ICS priorities (adapted from ISA-99.00.01 2007)

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.

3 The Role of Honeypots Within SCADA Systems

This section introduces the general concept of security honeypots, identifies the various categories of honeypots (assuming a broad perspective derived mainly from previous work focused on the ICT field) and discusses the few known experiences of specifically designing honeypots for SCADA systems or ICS in general.

By definition, a Honeypot is a decoy or dummy target set up to attract and detect or 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.

In general Honeypots can be classified in two groups: research honeypots and production honeypots. Research honeypots are used to obtain intelligence information about attack methods, usually in generalist scenarios, while production honeypots are used to explicitly protect an ICT infrastructure by providing advance warning of attacks against the production infrastructure.

Honeypots can also be distinguished by the ability of the attacker to interact with the application or services, as discussed by 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, they require a lot of close monitoring and detailed analysis.
- Low-interaction honeypots emulate vulnerabilities rather than exposing real weaknesses, therefore restricting the attacker's ability to interact with it, as discussed by Provos (2004). Mainly used as decoys, they are also less flexible, albeit being more secure since there is little that the attacker can do. Honeyd (2008) and Nephentes Baecher et al. (2006) are good examples of low interaction honeypots.

Finally, honeypots may be classified as server honeypots or client honeypots, as pointed by Riden and Seifert (2010):

- Server honeypots are designed to passively wait for attacks.
- Client honeypots are able to actively search for malicious servers and behave like victims, a useful feature for detecting client-side browser exploits. Examples of client honeypots are Shelia, described in (Bos 2009), Honeymonkey (Wang et al. 2006) and CaptureHPC (Hes et al. 2009).

All these categories of honeypot have been identified based on a perspective that, whilst generalist, is mostly based on the ICT field. In the specific context of ICS, it is also necessary to consider the operation scope intended for the honeypot, since each of the three types of network that are present in typical ICS scenarios requires different approaches and technologies.

For our purposes, we consider a simplified ICS system infrastructure description that comprises tree levels: the first two are the *operations network* (where the supervisory components, such as master stations or Human-Machine Interface (HMI) devices, and historic process databases are—corresponding to Level 2 of ISA-95 (2000)) and the *field networks* (where basic control and process devices, such PLCs and RTUs are deployed, controlling and monitoring a critical process—corresponding to Level 1 of the ISA-95 recommendations). These two scopes roughly correspond to the process control network levels on ISA-95—it should be mentioned that there is not a distinction between the automation and device networks in this diagram, for the sake of simplicity (due to the evolving capabilities of PLC devices, the distinction between PLCs and RTUs is somehow vanishing, something that also contributes to somehow fuse the automation and device networks in a single network scope).

Apart from the first two levels, which are specific to process control, there is also the organization's *ICT network*, comprising other corporate devices (workstations, servers) and services (such as e-mail, accounting and stock or asset management).

In the *operations network*, for instance, a low-interaction honeypot might simulate the operation of a network server (e.g., SCADA control station). In the *field network* a honeypot may be implemented using a system capable of simulating the operation of an RTU (e.g., a SCADA protocol emulator). Finally, in the *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.

Figure 2 portrays the three distinct network levels and includes, as an example, a honeypot in the *field network*.

To the best of our knowledge, only two previous initiatives stand out as examples of ICS-specific honeypot implementations: the SCADA Honeynet Project (Cisco Critical Infrastructure Assurance Group—CIAG 2004) and the Digital Bond's SCADA Honeynets (Digital Bond 2006).

The SCADA HoneyNet Project aimed at creating a framework to simulate industrial networks. It started circa 2004, lead by the Cisco Critical Infrastructure Assurance Group, and it is able to simulate several levels of the system:

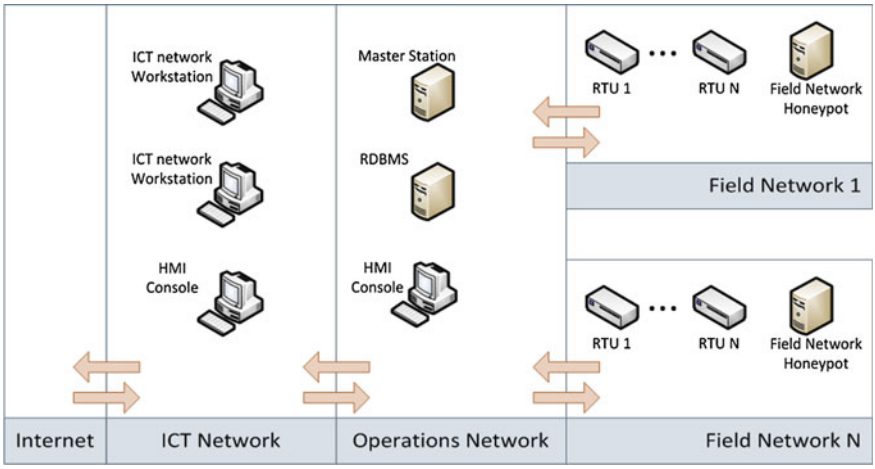

Fig. 2 Distinct network levels of the ICS environment

- Stack level—for simulation of the TCP/IP stack of a device.
- Protocol level—for simulation of specific industrial protocols, such as Modbus, and Ethernet/IP.
- Application level—for simulation several SCADA applications, such as web services and management consoles.
- Hardware—for simulation of serial ports and modems present in some SCADA devices.

The project is no longer maintained, but the software is still available for download at (Cisco Critical Infrastructure Assurance Group—CIAG 2004).

Digital Bond is another approach to SCADA honeynets (or honeypot networks). This solution uses at least two machines: one to monitor the network activity (in this specific case using a Generation III Honeywall) and the other to simulate a PLC with several services made available to the attacker (Modbus/TCP, FTP, Telnet, HTTP and SNMP).

The approach we now propose, described in the next section, distinguishes itself from these two initiatives by presenting a low cost, modular and highly configurable and manageable solution for an automation/field network honeypot. By bringing down the costs of honeypots and providing a flexible and modular approach to its deployment, we aim at large-scale deployment of honeypots in SCADA process control networks, turning them into commodity devices. Furthermore, by improving their manageability, we aim at effectively integrating them into the security management platforms already in place. The proposed SCADA network honeypot 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 (CockpitCI 2013)—but could be easily integrated into most existing security management platforms.

4 Proposed Architecture of the SCADA Network Honeypot

The honeypot presented in this paper is designed to operate in the process control network of 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 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 proposed honeypot is 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 honeypot is illegal and unauthorized (with the possible exception of management operations).

The proposed architecture is generic and compatible with the majority of SCADA protocols. However, Modbus was selected as the preferential protocol to be supported in the proof-of-concept prototype we developed, due to three factors: standardization, popularity and because it is based in an open specification, whose documentation is easily obtainable. For this reason, from now on we will specifically mention Modbus components, even though those components could be easily switched to add support for other SCADA protocols.

The architecture for the proposed honeypot for monitoring automation/field networks is presented in Fig. 3. This is a hybrid Modbus honeypot architecture, in the sense that it runs both simulated and complete implementations of services

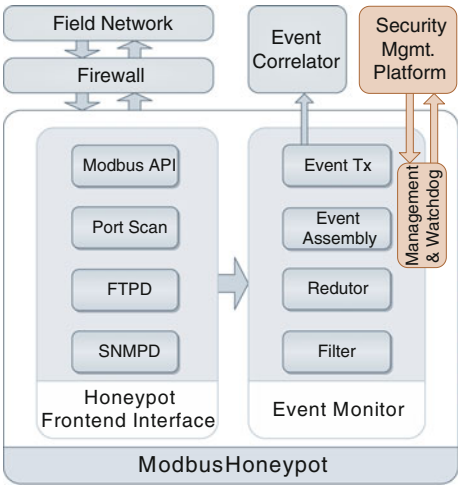

Fig. 3 Modbus honeypot software architecture

commonly available on PLC devices. Next, we describe in detail its main components and building blocks.

4.1 Honeypot Front-End Interface

The connection with the automation/field network is made through the *Honeypot Front-End Interface*. Within this block there are four main components:

- The Modbus API simulator (*Modbus API*, in Fig. 4), which accepts Modbus commands and behaves like a regular PLC, providing all the necessary protocol functionality, such as access to registers and Modbus operations.
- A File Transfer Protocol module (*FTPD*), providing a File Transfer Protocol (FTP) service like the ones commonly found on commercial PLCs.
- A Simple Network Management Protocol (SNMP) module that provides the SNMP (Case et al. 1990) device management interfaces and functionalities found on PLCs (*SNMPD*). This module and the usage of SNMP in SCADA environments, in general, will be further discussed in Sect. 4.5.
- 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 SCADA systems. 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 specific behaviour of the *Modbus API* is configurable (and, to a certain extent, programmable), in order to better mimic a wider ranger of real RTU/PLC devices. This is an important factor to prevent more sophisticated attackers from easily distinguishing honeypots from real devices: if all honeypots had the same predefined behaviour it would be possible to define a unique "behaviour profile" for the honeypots, thus compromising their stealthiness.

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. A more extensive discussion of the role of the SNMP module is provided in Sect. 4.5.

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 SCADA 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 *Management and 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), a standard data format designed for Intrusion Detection Systems (Debar et al. 2007). 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—thus

smoothing integration between the proposed SCADA honeypots and already existing intrusion detection applications.

IDMEF messages are to be sent using a secure channel between the honeypot (the sensor) and a higher level *Event Correlator* that is responsible for event processing and correlation. The exact nature of this processing node may vary from case to case, assuming for instance the form of a classical IDS application (the most likely scenario for generalized use of the SCADA honeypots) or the form of a specialized and distributed correlation engine (as is the case, for instance, for the CockpitCI project). One way or another, this secure message transmission is ensured by the *Event Tx* module.

4.3 Honeypot Management and Watchdog

The honeypot contains a *Management and Watchdog* module for remote management. This module allows security staff to modify the honeypot configurations (e.g. configuration of modules such as *Filter* and *Event reduction and aggregation*; configuration of the "Modbus" behaviour) 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. For cases were even out-of-band management is not an option, due to security restrictions, this communication may also use one-way communication devices such as the Waterfall unidirectional security gateways or data diodes (Waterfall Security Solutions Ltd. 2013), at the price of reduced functionality.

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. Connections from the honeypot to the attacker are the only outgoing connections that are allowed. It should be pointed out that this is the opposite of typical firewall configurations—where the interest is usually to protect the network node from outside attacks, whilst trusting internal applications—and this model of operation should be fully understood by the system operators.

4.5 Usage of SNMP in SCADA Environments

The importance of including an SNMP service on the honeypot is related with the fact that SNMP is extensively used, in SCADA environments, despite the fact that some authors recommend not allowing SNMP traffic in SCADA networks, for safety reasons (Krutz 2006).

In fact, SNMP is widely used not only to manage the RTU and PLC devices (the original purpose of SNMP) but also, in the case of some manufacturers, to poll process information from those control devices—thus complementing the SCADA operations. The manner of implementing such hybrid SCADA/SNMP varies from one manufacturer to another. Wingpath, for instance, uses a Modbus to SNMP converter (ModSnmp 2013) that acts as a gateway between Modbus slaves and an SNMP manager (see Fig. 4). To poll a value from a given PLC the SNMP manager sends a SNMP GET command to the ModSnmp server. This server then sends a Modbus request to the PLC, receives the corresponding response and sends a GET-RESPONSE message to the SNMP manager. Other approaches use SNMP traps to poll data from PLCs (Tsubakimoto 2011)

Moxa also uses SNMP to gather information from control devices, but instead of the gateway approach of Wingpath and (Tsubakimoto 2011), its control devices are SNMP capable (Moxa 2009). This solution aims at small tasks, such as monitoring

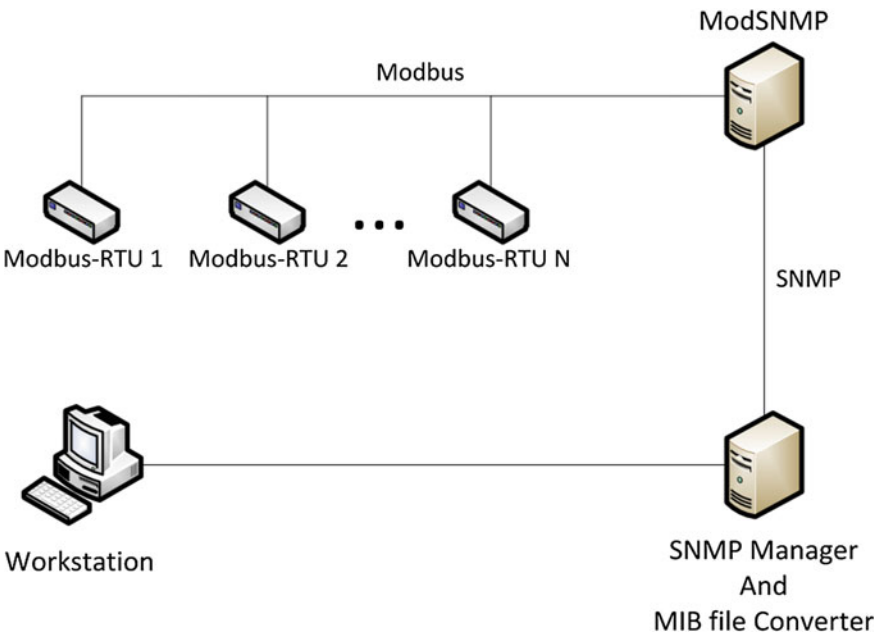

Fig. 4 Wingpath ModSnmp diagram (adapted from ModSnmp 2013)

environmental sensors to measure temperature and humidity, power regulators, or video cameras in server rooms.

For all these reasons, it was deemed as fundamental to include a full-fledged SNMP agent in the honeypot, in order to attract and mislead potential attackers (that often use SNMP as an attack vector for PLC and RTU devices). This SNMP agent is able to mimic the expected device management functionalities typical of SNMP and, if necessary, can also be configured to provide advanced access to industrial process information.

5 Implementation and Deployment Notes

The proposed honeypot may be deployed using two different approaches. The first one is the usage of low cost hardware appliances, physically and logically placed in strategic locations of the automation/field network. The second approach corresponds to the deployment of virtualized honeypots, logically placed in the automation/field network but physically placed in the datacentre. In this section we also discuss a third alternative, based on a significantly different approach to the implementation and deployment of the SCADA network honeypot.

While the basic concept for the SCADA honeypot is potentially effective in a simple deployment, with the device being placed on the automation or field network, actively monitoring scouting and attack preparation procedures for an intrusion attempt, there are ways to improve its capabilities. Those could be further enhanced and explored by creating a fake HMI system console (deployed in the process control network), configured to interact with a set of field/automation network honeypots—this configuration broadens the attack footprint detection capabilities, while exposing the honeypot as an interesting target for an attacker compromising the HMI as a preliminary intrusion vector.

5.1 SCADA Honeypot as a Low Cost Hardware Appliance

The proposed 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 Foundation (2012) 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. At current prices the bill of materials for our complete proof-of-concept is less than 50 Euro.

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 (2011) library. For parsing the Modbus messages a module based on the Pymodbus (2011) library 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 (2000) version 5.4.3, while the *FTPD* module uses VSFTPd (2011) version 2.3.5. 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 (2010) library and developed in-house for reduced footprint and computational requirements.

There were several potential pitfalls related with the usage of low cost SBC solutions as the basis for the honeypot, including reliability, performance and compatibility with industrial regulations.

Regarding performance, our experimental tests allow us to conclude that SBC solutions such as the Raspberry Pi easily provide more than enough resources. This is not a surprise, since the PLC and RTU devices the honeypots mimic typically have even less computational resources. Despite the increased processing overhead honeypots have to deal with, their response time is more than adequate (both from the perspective of the attacker and from the perspective of the IDS system it is connected with).

Regarding reliability we do not have definitive results, but our empirical tests show that by using carefully selected hardware components it is possible to have remarkably stable honeypots, functioning without interruption for at least several months, even when exposed to less than ideal environmental conditions. These tests obviously need to be confirmed by more extensive validation work, but they are nonetheless quite promising.

Regarding compatibility with industrial regulations, we have not conducted an extensive survey for every type of industry that uses SCADA systems. Nonetheless, based on preliminary analysis, most of the restrictions appear to be related with the physical components of the honeypot, and can be addressed by using compliant power sources and/or compliant casing. ATEX directives for explosive environments (Directive 94/9/EC 1994), for instance, can be addressed by using COTS ATEX-compliant cases. These cases may cost several times the price of the base honeypot hardware, but the final bill of materials is still quite attractive.

5.2 SCADA Honeypot as Virtualized Appliances

While in some cases the physical location of the honeypot in strategic locations of the automation/network is relevant from a security point of view, in other cases it is enough to have the honeypot logically located in the automation/field network—since most external attackers have no way to determine its physical location. For those situations an interesting alternative to physical honeypot appliances is the use of virtualized honeypots appliances.

Our proof-of-concept implementation is fully portable to a virtualized environment, being compatible with most commercial virtualization platforms. This means the critical infrastructure operator is able to deploy many virtualized honeypots

using a single server located in its datacentre, with reduced costs. Furthermore, the usage of the virtualization environment allows sophisticated network solutions (such as external monitoring of the traffic to/from the honeypot, advanced routing of non-legitimate traffic to the honeypot and advanced protection of the management traffic between the honeypot and the IDS platform).

5.3 Alternative Architecture for a High-Interaction Honeypot

The two previous approaches for deploying the SCADA network honeypot are directly in line with the proposed architecture and the proof-of-concept implementation we developed. We now introduce an alternative and substantially different approach for the design of the honeypot.

This approach involves the use of a real control device (PLC) to act as a decoy to attacker. In this solution the attacker interacts with a real PLC—that is, nonetheless, not associated with any industrial process—and his interactions are forwarded to an Intrusion Detection System (IDS). Since Modbus protocol communications are unencrypted, the interactions can be forwarded (for example with an Ethernet tap or port mirror switching) to a security adaptor, whose purpose is to detect and generate security events to be sent to a correlator (or, more generically, an IDS) for further analysis. This architecture is illustrated in Fig. 5.

The security adaptor can be built using a scaled-down version of the honeypot appliance hereby described, with only basic network traffic capture and event processing functionality.

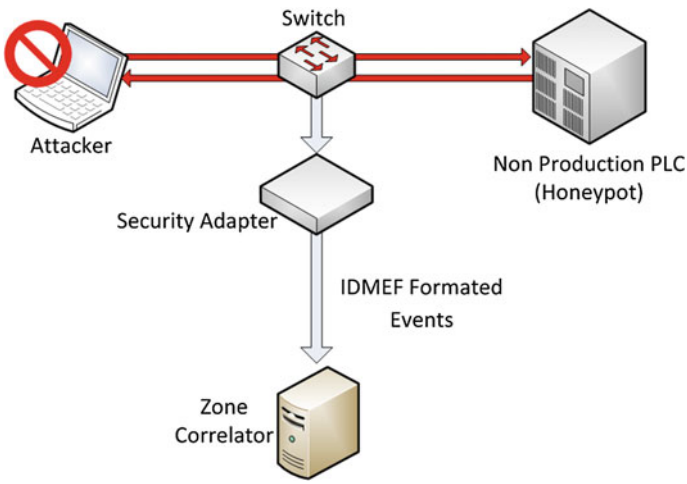

Fig. 5 SCADA Honeypot using a physical RTU/PLC device

This solution has advantages as well as disadvantages, when compared to the solution previously described. The main advantage of this solution has to do with the usage of a real, non-simulated, PLC. Since it uses real hardware, it is more difficult for the attacker to realize that he/she is dealing with a honeypot. It is also simpler to "develop" a honeypot supporting other SCADA protocols or mimicking other commercial PLC devices.

However, there are disadvantages when using real equipment to deploy a Modbus honeypot. First, the total cost of the solution will be substantially larger, because it requires a real PLC to be acquired. Moreover, the cost will increase with the number of honeypot deployments. In the previous solution, it is possible to simulate more than one PLC in one piece of hardware. With this solution, it is necessary to acquire one PLC for each honeypot deployed. Also, with a virtual PLC, it is possible to adjust configurations (e.g., available services and protocols, memory registers) to reproduce the hardware behavior from different vendors, which is more difficult, if not impossible, with real equipment. The complexity of the solution increases as well (especially in a multi honeypots scenario), but the cost is substantially lower in comparison with using real PLCs to build automation or field network honeypots.

6 Conclusion

This chapter presented the architecture for a hybrid honeypot for SCADA process control networks. It is able to detect attackers by simulating a complete Modbus TCP device, not only incorporating the protocol and control device logic but also other services such as SNMP and FTP services, commonly found on commercially available PLCs and RTU equipment. Moreover, it goes further by incorporating a *Port Scan* module capable of monitoring the remaining network ports in order to detect additional attacks or probes.

Other strong points of this honeypot device are the low cost hardware requirements and its configurability, allowing security managers to fine-tune the monitoring specifications through the different modules of the honeypot. Alternatively, the proposed honeypot can be virtualized and installed in the datacentre, with increased flexibility and reduction of operation costs.

Despite being a value proposition as a standalone device, the capabilities of the proposed SCADA honeypot can be explored to its full when deployed within an integrated security architecture such as the one proposed by the CockpitCI project, providing a distributed security probing mechanism with broad scope detection capabilities.

Acknowledgments 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 ongoing support and collaboration.

References

ANSI/ISA-95.00.01 (2000) Enterprise-control system integration part 1: models and terminology, ISA 2000

Baecher P, Koetter M, Holz T, Dornseif M, Freiling F (2006) The nepenthes platform: an efficient approach to collect malware. In: Recent advances in intrusion detection. Lecture Notes in Computer Science, vol 4219, pp 165–184. Springer, Berlin

Bos H (2009) Shelia: a client-side honeypot for attack detection. http://www.cs.vu.nl/~herbertb/misc/shelia/. Accessed 10 Sept 2013

Case J, Fedor M, Schoffstall M, Davin J (1990) A simple network management protocol (SNMP), IETF RFC 1157, May 1990

Cisco Critical Infrastructure Assurance Group—CIAG (2004) SCADA HoneyNet Project, http://scadahoneynet.sourceforge.net/. Accessed 10 Sept 2013

Clarke G, Reynders D (2004) Practical modern SCADA protocols: DNP3, 60870.5 and related systems. IDC Technologies, Oxford

CockpitCI (2013) Cybersecurity on SCADA: risk prediction, analysis and reaction tools for critical infrastructures, FP7 SEC-285647 project website. http://www.CockpitCI.eu/. Accessed 10 Sept 2013

Creery A, Byres E (2005) Industrial cybersecurity for power system and SCADA networks. In: Industry applications society 52nd annual petroleum and chemical industry conference. IEEE, pp 303–309

Davis CM, Tate JE, Okhravi H, Grier C, Overbye, TJ, Nicol D (2006) SCADA cyber security testbed development. In: NAPS 2006. 38th North American power symposium. IEEE, pp 483–488, doi:10.1109/NAPS.2006.359615

Debar H, Curry D, Feinstein B (2007) Rfc 4765: The intrusion detection message exchange format (IDMEF), March 2007. http://www.ietf.org/rfc/rfc4765.txt. Accessed 10 Sept 2013

Digital Bond (2006) SCADA honeynet. http://www.digitalbond.com/tools/scada-honeynet/. Accessed 10 Sept 2013

Directive 94/9/EC (1994) Equipment and protective systems intended for use in potentially explosive atmospheres (ATEX)

Fovino IN, Coletta A, Masera M (2010) Taxonomy of security solutions for SCADA sector. Project ESCORTS Deliverable 2.2, Version 1.1

Hes R, Komisarczuk P, Steenson R, Seifert C (2009) The Capture-HPC client architecture. Technical report, Victoria University of Wellington

Honeyd (2008) Developments of the honeyd virtual honeypot. http://www.honeyd.org/. Accessed 10 Sept 2013

Igure V, Laughter S, Williams R (2006) Security issues in SCADA networks. Comput Secur 25 (7):498–506. doi:10.1016/j.cose.2006.03.001

ISA-99.00.01 (2007) Security for industrial automation and control systems—part 1: terminology, concepts, and models. American National Standard, Research Triangle Park

Kang DJ, Lee JJ, Kim BH, Hur D (2011) Proposal strategies of key management for data encryption in SCADA network of electric power systems. Int J Electr Power Energy Syst 33 (9):1521–1526

Krutz RL (2006) Securing scada systems. Wiley, Indianapolis

Libpcap (2010) TCPDUMP/LIBPCAP public repository. http://www.tcpdump.org/. Accessed 10 Sept 2013

Modbus (2006) Modbus application protocol specification V1.1b. http://www.modbus.org/specs.php. Accessed 10 Sept 2013

Modbus-tk (2011) Implementation of the Modbus protocol in the Python programming language. http://code.google.com/p/modbus-tk/. Accessed 10 Sept 2013

ModSnmp (2013) ModSnmp: Modbus-SNMP converter. Wingpath software development. http://wingpath.co.uk/modbus/modsnmp.php. Accessed 15 Sept 2013

Moxa (2009) Why can't we be friends: monitoring the server room by introducing Modbus to SNMP. Moxa Americas Inc, Moxa White Paper
Net-SNMP (2000) Net-SNMP. http://www.net-snmp.org/. Accessed 10 Sept 2013
Provos N (2004) A virtual honeypot framework. In: Proceedings of the 13th conference on USENIX security symposium—(SSYM'04), vol 13. USENIX Association, Berkeley, p 1
Pymodbus (2011) A Modbus protocol stack in python. http://code.google.com/p/pymodbus/. Accessed 10 Sept 2013
Raspberry Pi Foundation (2012) What is Raspberry Pi? http://www.raspberrypi.org/. Accessed 10 Sept 2013
Riden J, Seifert C (2010) A guide to different kinds of honeypots. Symantec connect. http://www.symantec.com/connect/articles/guide-different-kinds-honeypots. Accessed 10 Sept 2013
Spitzner L (2002) Honeypots: tracking hackers. Addison-Wesley, Boston
Ten C, Liu C, Manimaran G (2008) Vulnerability assessment of cybersecurity for SCADA systems. IEEE Trans Power Syst 23(4):1836–1846. doi:10.1109/TPWRS.2008.2002298
Triangle MicroWorks Inc (2002) DNP3 overview, Raleigh, North Carolina. http://www.trianglemicroworks.com/documents/DNP3_Overview.pdf. Accessed 10 Sept 2013
Tsubakimoto (2011) SNMP agent module for tsubaki monitor maker mitaro 32. http://tsubakimoto.com/press/20110221.html. Accessed 15 Sept 2013
VSFTPd (2011) Vsftpd—secure, fast FTP server for UNIX-like systems. https://security.appspot.com/vsftpd.html. Accessed 10 Sept 2013
Wang Y, Beck D, Jiang X, Roussev R, Verbowski C, Chen S, King ST (2006) Automated web patrol with strider honeymonkeys: finding web sites that exploit browser vulnerabilities. In: NDSS
Waterfall Security Solutions Ltd. (2013) Waterfall one-way. http://www.waterfall-security.com/category/products/scada-protocols/. Accessed 10 Sept 2013
Zhu B, Joseph A, Sastry S (2011) A taxonomy of cyber attacks on SCADA systems. In: Proceedings of the 2011 international conference on internet of things and 4th international conference on cyber, physical and social computing (ITHINGSCPSCOM '11). IEEE Computer Society, Washington, pp 380–388