National Defence University
Department of Military Technology
Series 2: Reasearch Reports No. 5

Critical Infrastructure Protection

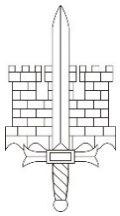
Jouko Vankka (ed.)

MAANPUOLUSTUSKORKEAKOULU
SOTATEKNIIKAN LAITOS
JULKAISUSARJA 2: TUTKIMUSSELOSTEITA NRO 5

NATIONAL DEFENCE UNIVERSITY
DEPARTMENT OF MILITARY TECHNOLOGY
SERIES 2: RESEARCH REPORTS NO. 5

Critical Infrastructure Protection

Jouko Vankka (ed.)



NATIONAL DEFENCE UNIVERSITY
DEPARTMENT OF MILITARY TECHNOLOGY
HELSINKI 2023

Jouko Vankka (ed.): *Critical Infrastructure Protection*
Maanpuolustuskorkeakoulu
Sotatekniikan laitos
Julkaisusarja 2: Tutkimusselosteita nro 5
National Defence University
Department of Military Technology
Series 2: Research Reports No. 5

DISCLAIMER

Working papers are preliminary works in progress that have been posted or published to stimulate discussion and comments. In addition, they provide information to the general public.

The views, opinions, findings, and conclusions expressed in these papers are strictly those of the author(s) and do not necessarily represent the views of the National Defence University. Therefore, they should not be reported as such. The National Defence University assumes no responsibility or liability for the statements, opinions or conclusions expressed in these papers.

Although checked by departmental publishing boards of the National Defence University, these working papers have not been through a process of blind peer review.

Uusimmat julkaisut pdf-muodossa: http://www.doria.fi/handle/10024/73990

ISBN 978-951-25-3380-0 (pbk.)
ISBN 978-951-25-3381-7 (pdf)
ISSN 2737-0615 (online)

Maanpuolustuskorkeakoulu – Sotatekniikan laitos
National Defence University – Department of Military Technology

PunaMusta Oy
Tampere 2023

PREFACE

Postgraduate seminar series with a title Critical Infrastructure Protection held at the Department of Military Technology of the National Defence University. This book is a collection of some of talks that were presented in the seminar. The papers address threat intelligence, a protection of critical supply chains, cyber security in the management of an electricity company, and privacy preserving data mining. This set of papers tries to give some insight to current issues of the critical infrastructure protection.

The seminar has always made a publication of the papers but this has been an internal publication of the Finnish Defence Forces and has not hindered publication of the papers in international conferences. Publication of these papers in peer reviewed conferences has indeed been always the goal of the seminar, since it teaches writing conference level papers. We still hope that an internal publication in the department series is useful to the Finnish Defence Forces by offering an easy access to these papers.

Editor

CONTENTS

STUDY ON THE RELEVANCE OF THREAT INTELLIGENCE ON THE PROTECTION OF CRITICAL INFRASTRUCTURE

Juhani Eronen

Oulu University Secure Programming Group

exec@ee.oulu.fi

Abstract

Threat intelligence is one of the latest buzzwords within the security industry. Various claims have been made by vendors and practitioners on the benefits of sharing and utilising different types of data about attacks and attackers. This is reflected in threat intelligence standards that enable sharing high-level data as well as low-level indicators. For example, it is commonly expressed that sharing efforts should be focused on attack methodologies instead of attacker infrastructure, as the latter set of data ages quickly and can be varied between victims. However, very few studies exist that attempt to quantifiably justify these claims.

In this study we evaluated the usage of threat intelligence by NCSC-FI, the Finnish National Cyber Security Centre, and within the Finnish national sensors network HAVARO. The study is focused on data that can be used in large-scale passive network monitoring.

First, we evaluated the sources of threat intelligence utilised by, or readily available to NCSC-FI. We measured overlaps among the sources, and categorised the data. Although the evaluated sources were seen to have some overlapping information, using more sources increases the detection capabilities of detection systems. Evaluating threat intelligence is a multifaceted field that is only partially covered by current research.

Despite the discussion related to sharing data on more high-level attacker data such as tactics, techniques, and procedures, the sharing of actionable high-level network threat data seems elusive in practice.

Purpose

The evaluation of the threat intelligence available for the HAVARO system of the National Cyber Security Centre of Finland (NCSC-FI).

Design/methodology/approach

Manual review and partial consolidation of threat taxonomies. Data set comparison with computer scripts.

Findings

The threat intelligence that is available in a standardized format is tactical in nature, and thus useful for detection systems. Using more sources increases the detection

capabilities of detection systems. Evaluating threat intelligence is a multifaceted field that is only partially covered by current research.

Originality/value

Introducing threat intelligence and its context. Review of the state of the art in the evaluation of threat intelligence. Using a large production data set of a detection system for evaluating threat data.

Keywords

Computer attack, advanced persistent threat, intrusion detection, threat intelligence

Paper type
Research paper

1 Introduction

The functioning of modern society relies on the resilience of its critical infrastructure. Most of the infrastructure providers are private companies that rely extensively on IT systems in the operation of their businesses. This combination makes critical commercial entities prime targets for a variety of attackers with motivation ranging from economic gain to geopolitics [1]. The following chapters summarize methods used to perform these attacks as well as to detect them. This leads to introducing the concept of threat intelligence. The section concludes by introducing the case study of this work, the evaluation of the threat intelligence available for the HAVARO system of the National Cyber Security Centre of Finland (NCSC-FI).

1.1 Common attacks

The internet in general is rife with attacks of differing scope and sophistication. There are, however, some aspects that are common to most if not all attacks. Most attackers spread quite simple attacks to a large number of recipients, with the goal of quick monetary gains for the attacker. The attack methods that are utilized are either hoaxes, software vulnerabilities, malicious software component (malware), or a combination of the methods.

Software vulnerabilities are weaknesses or errors introduced during the software development process that can have security implications. Common vulnerabilities include buffer overflows that enable attackers to run their code on a target system, and SQL injection that can lead to the compromise or modification of data from a service. Vulnerabilities are often used to run an exploit code that downloads and runs malware from systems controlled by the attacker.

Malware is a general term for a variety of undesired software components. Malware can be used for tasks such as sending unsolicited messages (spam), performing denial of service attacks, gathering credentials saved on a system, logging user activity such as keystrokes, sending documents found on the target system to the attacker, encrypting critical files and demanding ransom, and so on. Most current malware contact a command and control server (C2 or C&C), from which they obtain commands from the attacker. [2] Malware using vulnerabilities to spread directly between systems has

become rarer as the robustness of software has increased in many application areas. Current malware usually requires a human intermediary to spread itself to new systems.

Hoaxes are often encountered as spam via email, different messaging tools, forums or social media services. Some hoaxes are traditional fraud such as advance-fee scams performed via electronic means. A major portion of hoaxes are related to phishing, i.e. gathering service credentials or payment details. Some hoaxes, however, are used to spread malware. The simplest form of these hoaxes try to entice the used into installing malware, e.g. by presenting it as a legitimate application. Typical hoaxes try to persuade users to compromised websites serving exploit kits, i.e. software that tries to identify any unfixed vulnerabilities from the user host, and to exploit them to spread malware.

Most current attacks utilize exploit kits to spread malware, either by compromising popular websites and inserting exploit kits, or by using spam to send links to exploit kits. Developments in prevalent exploit kits are actively followed up and reported by various security researchers.

1.2 Sophisticated attacks

Some publicly reported attacks have increased in scope and sophistication, which has prompted the use of the moniker advanced persistent threats (APT) for these attackers. In short, APT attacks differ from common attacks in the sheer amount of work involved in customizing the attack as well as remaining undetected on the target system for as long as is needed to reach the intended goals. The sophistication used in APT attacks is reflected in the fact that most of them are detected only months if not years after the fact.

A commonly accepted model for APT attacks is the cyber kill chain presented in [3]. It divides attacks into seven phases: reconnaissance, weaponization, delivery, exploitation, installation, command and control, and actions on objectives. Not all phases are necessarily included in all attacks, but the model captures the complexity that might be involved in these attacks.

Reconnaissance might be as simple as looking up websites and search engines, making phone calls, and sending emails to find a suitable target person within the selected organization. As an example, technical support forums and professional social networking systems can be used to identify key personnel and the people they work with. The success of an attack can be easily enhanced by utilizing tools to identify technical details, such as networks and software used by the target organization.

The reconnaissance phase might indicate that the target uses a software with known vulnerabilities. Depending on the vulnerability and the protections of the target system, exploitation of these vulnerabilities might be either straightforward or very demanding. In many cases, exploit codes for known vulnerabilities are either publicized or available for purchase. In these cases, it is usually straightforward for the attacker to use this information to package an exploit for this vulnerability with selected malware, an attack phase called weaponization in [3]. If no known vulnerabilities are identified in the reconnaissance phase, new vulnerabilities and respective exploit codes may need to be identified in the target system. Finding these new, so-called 0-day vulnerabilities can involve extensive costs, which limits both their users and targets.

In Lockheed's kill chain model, weaponization is only relevant for attacks that exploit software vulnerabilities. In some cases, the attackers might only need access to a particular system such as email server or document repository. A sufficiently elaborate and targeted hoax (spear phishing) might be all that is needed in order to obtain the necessary credentials for these systems. Similarly, the victim can be lured into installing malware on their system, which eliminates the need for software vulnerabilities.

In any case, the attack needs to be delivered to the target. As is the case with common attacks, email attachments and compromised websites are commonly used. Input from the reconnaissance phase is critical for successful delivery. If the communication partners and subject of the target person are known, it is simple to write a convincing email purporting to originate from a known sender with the weaponized payload as an attached document. Similarly, if the target is known to visit a poorly protected website, it can be compromised to include an exploit kit that spreads the weaponized payload [4]. Attackers have been known to use elaborate methods for delivery, such as the use of malicious USB sticks in the well-publicized Stuxnet case [5], or compromising software distribution sites to replace update files with malware [6].

Once the malware has been installed, it usually reports to a command and control server. The command channel can either be a job queue, which the malware periodically polls for actions added by the attacker, or interactive, which allows the attacker to run commands in real time. Command channels used in APT attacks are usually encrypted, so that the commands and their responses cannot be observer by outsiders. In some publicized APT attacks, command and control channels have used multiple compromised systems relaying messages between the target and the attacker in an attempt to hide to hide the infrastructure used by the attacker. Some attackers have even been reported to use satellite internet connections to further obfuscate the target of the communications, as in this case the target could be anyone within the vast geographical area served by the satellite [7].

Once attackers can command their malware on the target system, they can start to take the actions necessary to obtain their original objectives. In many observed cases, the objective is to gain access to sensitive data. Compromised systems are scoured for documents, which are sent out to internet, often to other systems compromised by the attackers. This process, called exfiltration, needs to be performed stealthily, as avoiding being detected is a priority for the attacker.

If the compromised system does not contain the necessary data, the attacker needs to compromise further systems within the target organization, a process commonly called lateral movement. This is usually easier than compromising the first system, as most protections are geared toward outward threats rather than internal compromises. Current attackers often stop using malware after they have established some access to a target system. Instead, they try to gain administrative credentials so that they can continue their attacks using the kinds of tools normally used by administrators, in order to avoid raising the suspicion of the defenders.

The central idea of the cyber kill chain is that defenders should consider all the different phases of attacks in the design of their defenses, and try to detect attacks in the earliest possible phase to minimize the damage they cause. An intelligent and resourceful attacker can bypass common protection mechanisms, for example by

crafting an email message that does not trigger spam protections and using unique malware that antivirus software does not recognize. Thus, prudent defenders should strive to utilize multiple the detection and prevention mechanisms available, which both increases the protection coverage of kill chain phases and conforms to the practices of defense in depth [8].

1.3 Detecting attacks

A variety of technologies have been developed for detecting attacks. The US Institute of Standards and Technology (NIST) divides these technologies into the categories of host-based, network-based, wireless, and network behavior analysis systems. The usage of more than one category of systems is recommended, which in most cases means employing at least host and network-based systems. NIST further divides the detection methodologies used to signature-based, anomaly-based and stateful protocol analysis. Signature based detection looks for signs of known attacks, whereas anomaly-based detection tries to find deviations from the normal operations of the observed system. Stateful protocol analysis strives to find abnormal usages of the protocols used for network communication. [9]

While signature-based systems can be seen to be limited in their effectiveness, using them is a part of the security due diligence. The amount of data shared and publicized on observed attacks has vastly increased in recent years [10]. Much of that data can be directly used by signature-based solutions. According to the kill chain model, using data about known attacks enables the detection of new attacks in earlier phases [11]. The term threat intelligence has been adopted for actionable data related to scans attacks or attackers [3].

Attack detection is by no means a simple technical issue. Competent human operators are needed for the monitoring of technical detection systems. Serious incidents are often discovered by system administrators and users who report of suspicious events.

1.4 Threat intelligence

The most common form of threat intelligence results from the immediate observations from attacks, such as the malware found on compromised systems and the network infrastructure they communicated to. Sharing this intelligence forces the attackers to stop using exactly the same malware and network infrastructure in different attacks in order to avoid being discovered. It has been argued that these changes are usually trivial for the attackers, and that threat intelligence sharing efforts should focus on the kinds of aspects of the attacks that are harder to change, such as attacker tactics and tools [12]. Others suggest that instead of collecting and sharing technical indicators of compromise (IOC) related to attacks the focus should be on following the attackers and their attack campaigns [13]. As a result, a variety of threat intelligence feeds and services have emerged [10; 14].

Diverse tools and standards have been developed to capture the different data sets as well as the related use cases. Prominent standards include OpenIOC [15], Structured Threat Information eXpression (STIX) [11], and Malware Information Sharing Project (MISP) [16], which, while primarily being a tool, uses its own data format. STIX is both the most versatile and the most complex of these standards. It covers a

wide scope of threat intelligence, ranging from technical details to attacker motivations.

The UK Centre for the Protection of National Infrastructure (CPNI) divides threat intelligence into four distinct subtypes: strategic, tactical, operational, and technical [10]. As its name suggests, technical threat intelligence covers the technical details of attacks. Technical and operational intelligence is short-lived and useful for detection. Tactical threat intelligence can be thought of as a summary on the commonalities between different attacks. If, for example, multiple attacks involve the compromise of a single workstation followed by network logons to other workstations, it might be prudent to prevent all network logons that are not strictly necessary. This type of tactical threat intelligence is useful for the design and administration of systems.

Strategic threat intelligence is defined as high-level information related to attackers as well as their motivations and targets. This kind of information is useful for risk management and decision making related to the priorities and resources for defenses. News on breach victims can be thought of as rudimentary strategic intelligence: if a number of your peers have been reported to be breached, it is likely that you are on also on the target list. Operational threat intelligence answers to the more immediate question of who is being attacked right now. This kind of information may be hard to come by outside the law enforcement and intelligence communities. A simple example of operational intelligence are the recipient lists of spear phishing emails.

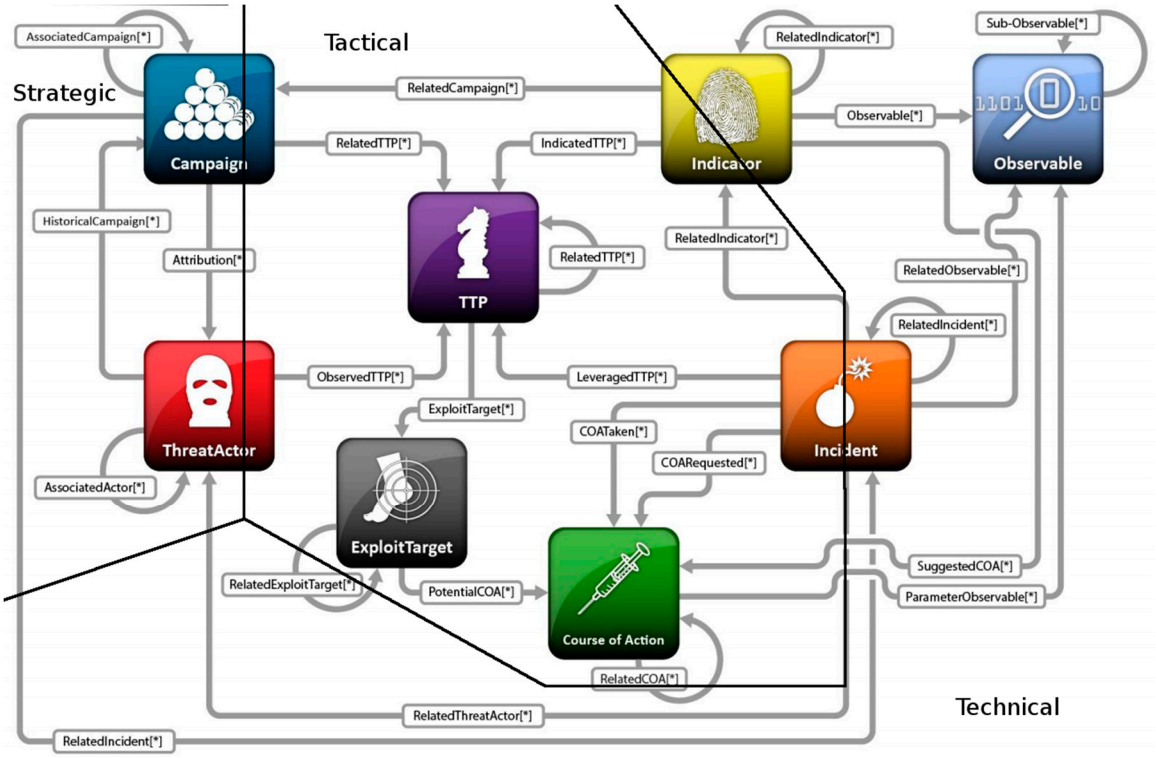


Figure 1: Threat intelligence categories of STIX constructs

Figure 1 categorizes the different constructs defined within the STIX model according to the CPNI subtypes. In the STIX terminology, Observables are the technical details such as IP addresses, which can be combined to form an Indicator of an attack. Both clearly represent technical threat intelligence. Tools, Tactics and Procedures (TTP) describe attacker actions, Exploit Target details their toolbox, and

Courses of Actions (COA) the steps that can be taken to prevent attacks or recover from them. These constructs are within the realm of tactical threat intelligence. Incidents as well as incident Campaigns can include some operational threat intelligence, whereas most of the information related to Campaigns and Threat Actors can be used as strategic threat intelligence.

The threat intelligence standards heavily emphasize the sharing of refined indicators, much in the spirit of [12], with claims such as "STIX will enable the sharing of comprehensive, rich, "high-fidelity" cyber threat information" [11], and "MISP focuses on the exchange of the most valuable indicators selected and annotated by analysts" [17]. According to a common maxim it is always useful to add new feeds of threat intelligence data to a detection system, implying that there is very little overlap between different feeds [10; 11]. This work evaluates these claims with the threat intelligence used by, or available for the HAVARO system.

1.5 Case study: HAVARO

HAVARO is the early warning and detection system for Finnish critical infrastructure and governmental systems offered by the NCSC-FI. It consists of passive network sensors that are placed within the client infrastructure, but operated and monitored by NCSC-FI personnel. The sensors are placed outside the network boundary, usually on the internet-facing side of the firewall, to minimize privacy implications of monitoring. Events related to suspected incidents are reported to a central system, where they are categorized by the system operator according to their observed severity into red, yellow, green and white, red being the most severe of those, and white meaning "not handled". The events are compiled into a continuously updating weekly report that the client can access at any time. Clients are also separately notified of red events.

HAVARO is similar in many aspects to the National Cybersecurity Protection System operated by the United States Department of Homeland Security for US Federal agencies. Operationally the system is called Einstein. The US Government Accountability Office (GAO) criticized Einstein for only utilizing signature-based detection-methods [4]. The GAO report quotes NIST on the need for multiple detection methodologies, and recommends DHS to include anomaly-based detection methods in Einstein. However, NIST also states that the combination of detection methods within network-based sensors include considerable tuning and customization to take into account the characteristics of the monitored environment [9]. This would require extensive manpower in systems that involve dozens of constantly changing and diverse networks. Although the area of anomaly-based detection has been under extensive academic research, many of the resulting systems have not been deployed operationally [18]. Other considerations on anomaly-based systems include scalability issues [19] and difficulties in employing machine learning methods in the field of attack detection [18]. Further, the GAO report did not address the role of threat intelligence within detection systems.

As part of its operations, NCSC-FI has formed an extensive network of international trusted partners. This network is a good source of diverse threat intelligence. NCSC-FI shares this intelligence with domestic partners as much as feasible, but some of the received intelligence has been marked sensitive, which limits its sharing. The most common classification system is the Traffic Light Protocol [20], although national

classifications are not unheard of. The intent of the classification is to avoid giving the attacker any indications that details of their operations are known. Limiting the distribution of intelligence to directly trusted partners is a simple way to accomplish this goal, but it also limits the usefulness of the intelligence. Without detection systems, sensitive threat intelligence can only be used for providing context to reported incidents.

A rarely discussed property of anomaly-based detection systems is that they may produce alerts that are difficult to analyze by the system operators, limiting the actionability of the alerts [9; 18]. Signature-based detection systems may suffer from similar problems if the used threat intelligence does not contain enough context on the attacks related to the signature. HAVARO has received acclaim from clients for producing detection reports that are easy to understand and act upon. Tens of red events are observed and reported by HAVARO monthly.

1.6 Evaluation of tactical threat intelligence

A good overview on current research on evaluating threat intelligence is given in [14]. All publicized evaluation approaches seem to be related to tactical threat intelligence feeds. A major portion of the other forms of threat intelligence seems to be found in the form of discussions, white papers, reports, or other forms of prose [10].

The Necoma project classifies the properties used in different tests into quality and scope properties [14]. The TIQ-test project uses the properties of novelty, overlap and population [21]. Table 1 summarizes the properties.

Table 1: Properties in threat intelligence tests

Property	Definition	Comments
Relevance	Is the data relevant to my use case?	
Accuracy	How much of the data is useful? How many false positives are involved?	
Timeliness	How current is the data? Is it still valid?	The complexity of investigations, reporting formats and systems may hinder timely sharing of data.
Novelty	Are the feeds updating regularly? Is data being removed and not only added?	
Scope	What is the coverage of the data?	Usually defined in terms of volume.
Overlap	How much of the data is common with other feeds?	
Population	Which networks or countries are represented in the data? Are there any trends in this data?	Need to be baselined to the sizes of the networks or countries. Need to avoid jumping to conclusions about the attackers.

Most of the previous research on feed evaluation has been concentrating on the overlaps between different data sources. The results have been unanimous: the threat data is mostly unique to a single feed [14]. The TIQ-test project saw overlaps between some feeds, but these were deemed to be the result of aggregate feeds that include other evaluated feeds [21].

2 Research work

I identified six relevant sources of threat intelligence for evaluation within this study. Five of the sources are from trusted partners of NCSC-FI. Three of the sources use MISP as their data format, one uses STIX, and one of the sources uses a custom CSV (Comma Separated Values) format. The STIX data was originally in the Extensible Markup Language (XML) format, which I converted with tools published by the STIX project to the JavaScript Object Notation (JSON) format for easier evaluation [22]. As the STIX data was in multiple versions of the STIX format, working with the data proved somewhat challenging. The sixth feed is the open source indicator feed in the OpenIOC format, provided by the security company FireEye [23]. Table 2 summarizes some basic scope properties of the feeds.

Table 2: Basic properties of the source feeds

Source	Format	Timespan	Indicator volume	Data volume
MISP1	JSON	940 days	300000	239MB
MISP2	JSON	484 days	90000	29MB
MISP3	JSON	195 days	120000	94MB
Custom	CSV	1085 days	120000	37MB
OpenIOC	XML	638 days	1000	1MB
STIX	JSON	776 days	50000	59MB

For comparison, similar properties of the NCSC-FI IOC feed are given in Table 3. NCSC-FI IOC is an automatically generated feed of indicator data related to common attacks, which explains its massive volume. As one would expect, there is much less data on the less prevalent sophisticated attacks as there is on common attacks.

Table 3: Properties of a common attack feed, for comparison

Source	Format	Timespan	Indicator volume	Data volume
NCSC-FI IOC	JSON	1040 days	5200000	318MB

The TIQ-test project provided statistical tools for evaluating threat intelligence [24]. However, the tools only utilize IP address data within the feeds. I wanted to evaluate all of the intelligence, which is why I ended up writing my own evaluation tools. TIQ-test also divides intelligence into outbound and inbound threats. This data was not always readily available in my source data. Thus, using the TIQ-test tools would probably have required manual classification work.

First, I present an overview on the contents of the feeds. As detailed in the previous chapters, STIX provides constructs for most aspects of different threat intelligence. The volumes of different STIX constructs within our STIX source are given in Table 4. As seen in prior research, nearly all of the intelligence is related to technical threat intelligence. Even all of the intelligence on TTPs was related to malware details, which makes it technical rather than tactical.

Table 4: Observed STIX constructs

STIX construct	Volume
Observable	50000
Indicator	30000
Incident	300
TTP	800
Exploit Target	0
Course of Action	0
Campaign	0
Threat Actor	0

As I inspected the intelligence manually, I noticed that analyst notes were added to some of the events. In many cases, these notes contained information that should have been inserted in other fields of the formats. Also, the notes themselves were scattered among a number of different fields. This suggests that the formats are complex and their usage is completely clear to all analysts. I categorized the fields to host-based indicators, network-based indicators and analysis. Figures 2 and 3 present category breakdown of our sources respectively in absolute numbers and percentages. Although only network-based indicators can be used in HAVARO, there are quantities of useful indicators in all of the evaluated sources.

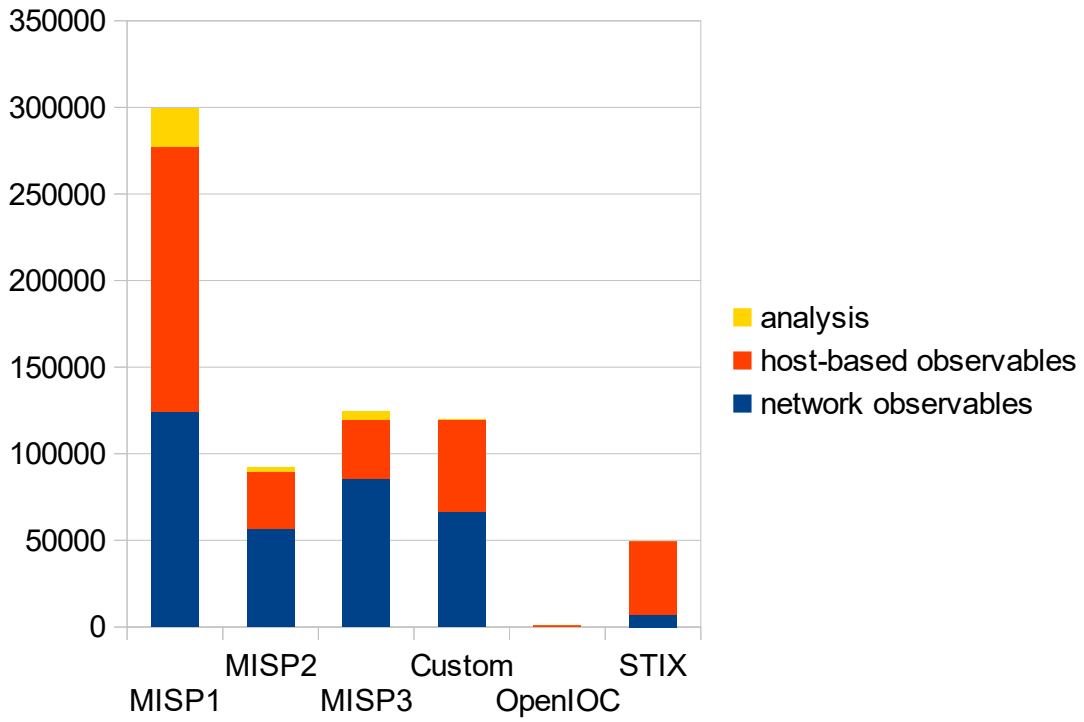


Figure 2: Category breakdown in numbers

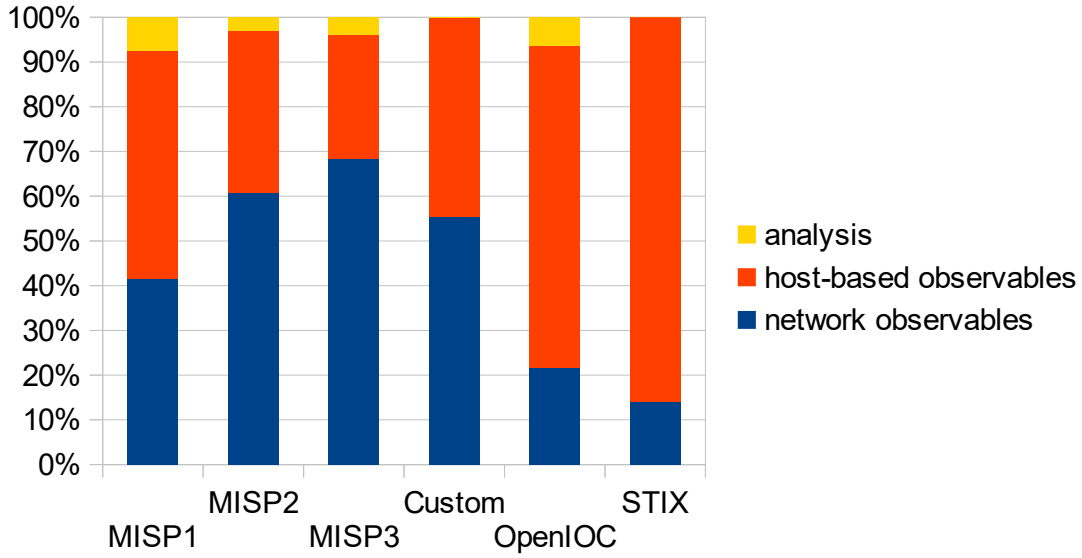


Figure 3: Category breakdown in percentages

I investigated the overlap between the sources with a simple indicator value matching script. The results, presented in Figure 4, shows considerable overlaps between some of the feeds. Other feeds seem to have most in common with misp1, the feed with the greatest data volume. According to my observations, the overlaps are to a great extent resulting from publicly available threat intelligence. Many of the sources try to ensure that they have covered all the relevant publicized attacks.

	misp1	misp2	misp3	custom	openioc	stix
misp1	100%	25%	25%	25%	0%	1%
misp2	81%	100%	70%	25%	0%	0%
misp3	61%	52%	100%	11%	0%	0%
custom	57%	23%	14%	100%	0%	1%
openioc	45%	19%	20%	35%	100%	1%
stix	6%	1%	1%	2%	0%	100%

Figure 4: Overlaps among the feeds

According to the cyber kill chain model, threat intelligence should ideally cover all the phases of attacks. Figure 5 shows the kill chain phase distribution within our sources. Some of the involved formats did not include the kill chain phase. In these cases, I categorized the indicators according to the kill chain model whenever possible. Most of the intelligence is related to the delivery, installation and command and control phases. My hypothesis is that this is both what the sources are mostly observing, and what they are most comfortable sharing to others.

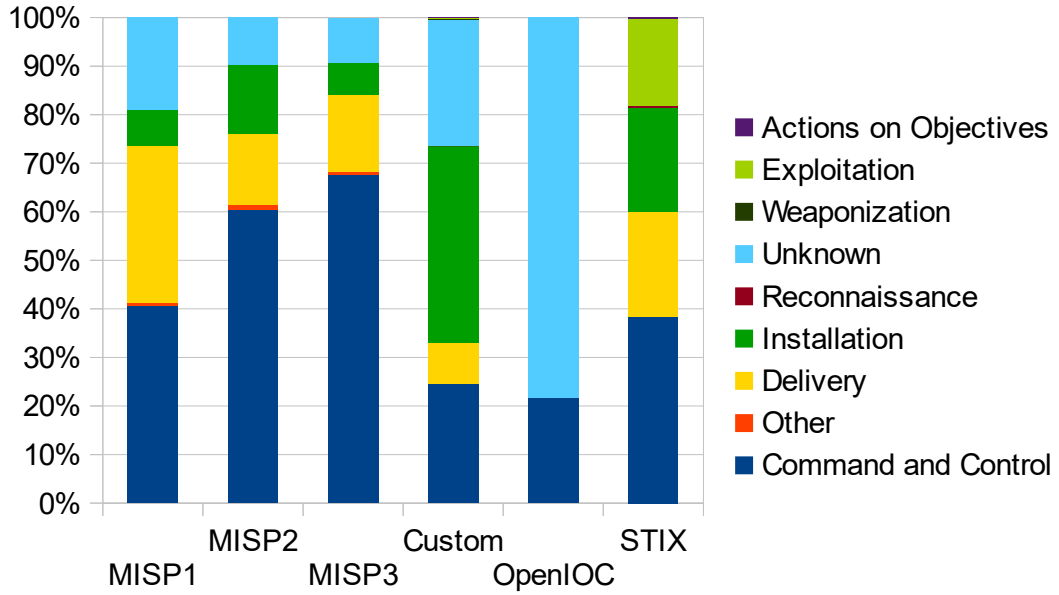


Figure 5: Kill chain phase coverage

Providing the necessary context to indicators is a critical quality issue for threat intelligence. Linking different indicators and attacks is a good way of providing context. Table 5 enumerates the links found in our sources. As a conclusion, the found proportion of links as well as analyst notes may provide a decent amount of context for the users of our sources.

The figures 6 and 7 provide compare the threat intelligence feeds evaluated in this work to the NCSC-FI IOC feed. While the volume of common attacks is much greater, the automatically generated feed is clearly lacking in fidelity related to kill chain phases.

Table 5: Links found in sources

Source	References between attacks	References between indicators
MISP1	18000	120000
MISP2	5000	0
MISP3	4200	32000
Custom	0	0
OpenIOC	0	0
STIX	10	4000

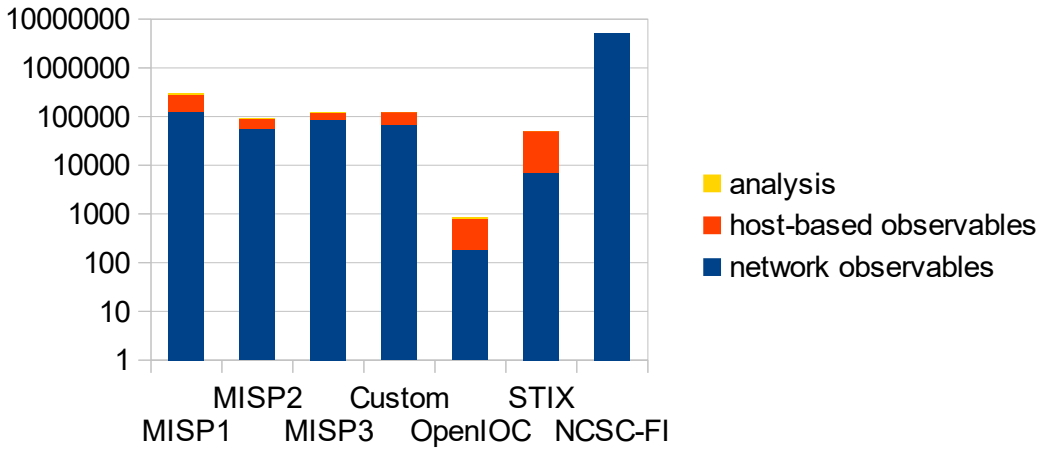


Figure 6: Volume of common vs sophisticated attacks (logaritmic scale)

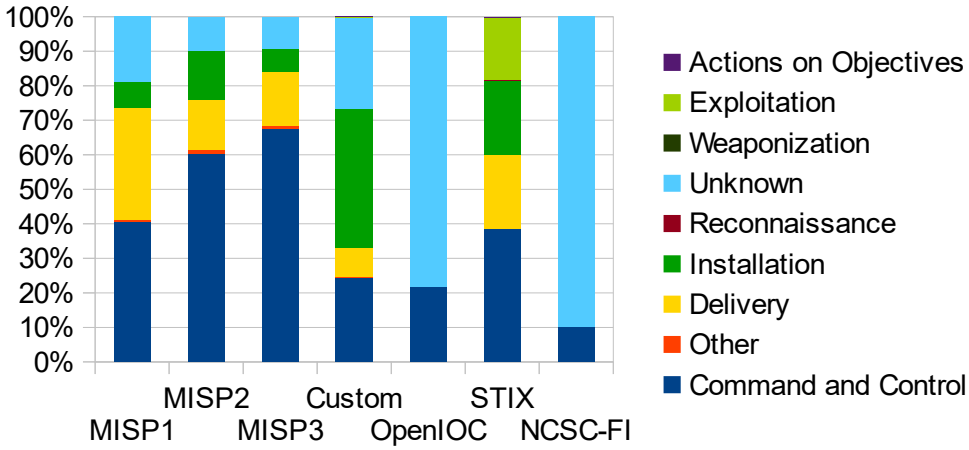


Figure 7: Attack phase coverage in a feed of common attacks

3 Conclusion

This work evaluated the threat intelligence in the context of HAVARO, the early warning and detection system for Finnish critical infrastructure and governmental systems. The threat intelligence that is available in a standardized format is tactical in nature, and thus useful for detection systems. The threat intelligence sources used in this study only contained intelligence on roughly half of the phases of the cyber kill chain. Although the sources were seen to have some overlapping information, using more sources increases the detection capabilities of detection systems. Evaluating threat intelligence is a multifaceted field that is only partially covered by current research.

4 Discussion

As the volumes of shared tactical threat intelligence are massive, automating the generation, sharing, deployment and withdrawal of threat intelligence is paramount.

However, this emphasizes the evaluation of the intelligence, as false positives are costly for the owners of the detection systems [18]. Future research should strive to establish a methodology for this evaluation work.

Some of the evaluation methods in earlier research were found to be better suited for data on systems that are vulnerable or infected with malware. This data is useful for reporting to the owners of these systems, but usually not relevant for detection systems. These evaluation methods should be employed on the information used by national reporting systems, such as the Autoreporter system operated by NCSC-FI.

There are multiple research issues related to the timeliness of threat intelligence. Most research, including this work, does not consider the publication times of indicators when evaluating source overlap. When using domain names in threat intelligence, the resolved IP addresses corresponding to the domains should be from the same time frame as the publicized threat intelligence. According to popular belief, timely threat intelligence results in more accurate attack detection. This should be verified by research on actual alerts from detection systems. Some aspects of attacker operations might be more lasting than expected.

References

[1] Aaltola, M. Cyber Attacks Go Beyond Espionage. FIIA Briefing Paper 200. [cited 18.9.2016] http://www.fiia.fi/assets/publications/bp200.pdf

[2] Jakobsson, M. Crimeware: Understanding New Attacks and Defenses" 1st edition. Boston: Addison-Wesley Professional, 2008. 608 pages. ISBN 978-032-150195-0

[3] Hutchins, M & Cloppert†, M & Amin, R. Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains. [cited 18.9.2016] http://www.lockheedmartin.com/content/dam/lockheed/data/corporate/documents/LM-White-Paper-Intel-Driven-Defense.pdf

[4] United States Government Accountability Office. DHS Needs to Enhance Capabilities, Improve Planning, and Support Greater Adoption of Its National Cybersecurity Protection System. [cited 18.9.2016] http://www.gao.gov/assets/680/674829.pdf

[5] IEEE. The Real Story of Stuxnet. [cited 18.9.2016] http://spectrum.ieee.org/telecom/security/the-real-story-of-stuxnet

[6] Symantec. Dragonfly: Western Energy Companies Under Sabotage Threat. [cited 18.9.2016] http://www.symantec.com/connect/blogs/dragonfly-western-energy-companies-under-sabotage-threat-energetic-bear

[7] Kaspersky. Russian-speaking cyber spies exploit satellites. [cited 18.9.2016] https://blog.kaspersky.com/turla-apt-exploiting-satellites/9771/

[8] SANS. Defense In Depth. [cited 18.9.2016] https://www.sans.org/reading-room/whitepapers/basics/defense-in-depth-525

[9] National Institute of Standards and Technology. Guide to Intrusion Detection and Prevention Systems (IDPS). [cited 18.9.2016] http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-94.pdf

[10] Chismon, D & Ruks, M. Threat Intelligence: Collecting, Analysing, Evaluating. https://www.cpni.gov.uk/Documents/Publications/2015/23-March-2015-MWR_Threat_Intelligence_whitepaper-2015.pdf

[11] Barnum, S. Standardizing Cyber Threat Intelligence Information with the Structured Threat Information eXpression (STIX). [cited 18.9.2016] http://stixproject.github.io/getting-started/whitepaper/

[12] Bianco, D. The Pyramid of Pain. [cited 18.9.2016] http://detect-respond.blogspot.co.uk/2013/03/the-pyramid-of-pain.html

[13] Crowdstrike. You Have an Adversary Problem. [cited 18.9.2016] http://www.slideshare.net/CrowdStrike/crowd-casts-monthly-you-have-an-adversary-problem

[14] Pawlinski, P & Kompanek, A. Evaluating Threat Intelligence Feeds. FIRST Technical Colloquium for Threat Intelligence, 2016. [cited 18.9.2016] http://www.necoma-project.eu/m/filer_public/b9/da/b9dafadd-adf8-4875-afd5-2d188dd96449/pawel-pawlinski-evaluating-ti-feeds.pdf

[15] Mandiant. OpenIOC - An Open Framework for Sharing Threat Intelligence. [cited 18.9.2016] http://www.openioc.org/

[16] The MISP Project. MISP Malware Information Sharing Platform and Threat Sharing. [cited 18.9.2016] http://www.misp-project.org/

[17] European Union Agency for Network and Information Security. Standards and tools for exchange and processing of actionable information. November 2014. [cited 18.9.2016] https://www.enisa.europa.eu/publications/standards-and-tools-for-exchange-and-processing-of-actionable-information/at_download/fullReport

[18] Sommer, R & Paxson, V. Outside the Closed World: On Using Machine Learning For Network Intrusion Detection. [cited 18.9.2016] http://www.utdallas.edu/~muratk/courses/dmsec_files/oakland10-ml.pdf

[19] Bellovin, S & Bradner, S & Diffie, W & Landau, S & Rexford, J. Can It Really Work? Problems with Extending EINSTEIN 3 to Critical Infrastructure. Harvard National Security Journal / Vol. 3, 2011. http://harvardnsj.org/wp-content/uploads/2012/01/Vol.-3_Bellovin_Bradner_Diffie_Landau_Rexford.pdf

[20] US-CERT. Traffic Light Protocol (TLP) Matrix and Frequently Asked Questions. [cited 18.9.2016] https://www.us-cert.gov/tlp

[21] Pinto, A. From Threat Intelligence to Defense Cleverness: A Data Science Approach. SANS Digital Forensics Summit 2015. [cited 18.9.2016] https://digital-forensics.sans.org/summit-archives/cti2015/From-Threat-Intelligence-to-Defense-Cleverness--A-Data-Science-Approach--Alex-Pinto-Niddel.pdf

[22] STIX Project. A Python library for parsing, manipulating, and generating STIX content. [cited 18.9.2016] https://github.com/STIXProject/python-stix

[23] FireEye. FireEye Publicly Shared Indicators of Compromise (IOCs). [cited 18.9.2016] https://github.com/fireeye/iocs

[24] MLSec Project. tiq-test - Threat Intelligence Quotient Test. [cited 18.9.2016] https://github.com/mlsecproject/tiq-test

CRITICAL INFRASTRUCTURE PROTECTION IN HOMELAND SECURITY – SUPPLY CHAINS

Markus Häyhtiö

Finnish National Defence University

Markus.hayhtio@kolumbus.fi

Abstract

Importance of international trade is tremendous for modern economies. A study conducted by the Bertelsmann Foundation's Global Economic Dynamics program [1] reveals the fact that one of the largest beneficiaries of the global trade was Finland with the annual gain in the income per capita of about € 1500,-.

International business demands seamless service and IT-infrastructure throughout the whole supply chain. This network of systems creates a very efficient method to run day-to-day operations, and material streams. However, dependences between different parts of this vulnerable ecosystem form a fragile web. This web is a lifeline of modern societies and requires protection in its all forms. While this lifeline pumps services and necessary information to all the necessary web members it is simultaneously vulnerable to malfunctions. These malfunctions have direct and indirect effects to all network participants. None of the economies is immune to these effects and even more, none of the economies can turn a blind eye to the existing threat.

Purpose

Purpose of this this article is to show to a reader, that a protection of critical supply chains demands a systems wide approach to the issue. It also shows that there is a need to a management approach, which:

1) explains the supply chain domain
2) gives a method to manage individual parts of a supply chain

Design/methodology/approach

This paper is a comprehensive literature review. It is based on the Ted G. Lewis' book "Critical Infrastructure protection in Homeland Security" chapter 16, Supply Chains.

Findings

This paper's results indicate, that there is a need to analyse all the systems and subsystems, which can affect supply chain efficiency. Existing life-cycle management systems enable a comprehensive management of critical components in the supply chain. At the same time a thorough comprehension of the capability management regarding these systems and the domain they create is needed.

Originality/value

This paper can be used as a guiding source for more detailed work on the areas described in the text.

Keywords

Capability management, critical infrastructure, supply chain management

Paper type

Research paper

1 Introduction

History of international trade is long. Role of globalization has steered development toward increasing global alignment of activities across countries, operations, and market offerings [2].

Despite the fact that international trade has long roots its significance has never been as big as now. Clear positive effects of globalization as a mechanism to spread wealth cross borders have made it possible to create a web of enterprises who work closely together across the globe. The main reason why organizations become international is the fact that they are simply able to operate cross border activities more efficiently than existing markets. [3]

As any other discipline, also international trade and international supply chain management has developed over the time. Few of the main authors and their approaches are listed below.

Monopolistic advantage theory [4]

Internalization theory [5;6]

Eclectic theory (OLI) [7]

Internationalization process theories [8;9]

Internationalization of new ventures [10;11]

These more of a macro level theories describe hows and whys of internationalization. On a more micro level there has a clear division between resource and competence-based views.

Resource based view of the internationalization emphasizes so-called firm specific advantages [12]. These resources should be valuable, rare, imperfectly imitable and strategically important. The more important role of knowledge and information has challenged this view. Competence based view [13] as an approach of the competitive advantage has increased its importance among scholars and industries.

Despite their differences, still variables common to both approaches are information and knowledge, which are transformed across the globe every second, every day. Both of these variables play a crucial role in the network-based organization, which basic assumption is, that firms are dependent on the resources controlled by other companies.

Positive effects are not without their downsides or risks. Macro level risks, which need to be taken into account can be divided in to following sub-groups:

Different cultures, political and economical systems and development

Interaction with national governments

Work within the limits of international trade and investment systems

2 Vulnerability and risk

Uncertainty is defined as "the difference between the amount of information required to perform the task and the amount of information already possessed by that organization." [14, p.5.] With risk we refer to the possible outcomes of an action, specifically to the loss that might be incurred if a given action is not taken [15]. A risk combines two attributes i.e. probability and impact. Probability is a measure of how often a detrimental event that results in a loss occurs. Impact refers to the significance of that loss to the organization. The level of risk is then perceived as the likelihood of occurrence of a detrimental event and the significance (impact) of that event [16, p.397].

But how to define a relevant risk? How do we decide which part of the supply chain creates a critical nod? How do we divide a system and its sub-systems into manageable components, without sacrificing the overall purpose of the system? How do define the capabilities, which need to be met? How do we prevent a situation in which a "tail wags the dog", meaning the risk preventing process defines the outcome and not the vice versa? There has to be a managerial approach, a methodology, which sets a framework for the capability management.

Protection of the critical components on a supply chain has to cover critical, recognized nods and most important production systems and their sub-systems. Systems' operations have to be analyzed on the four functions across the organizations. These four functions are according to Beer [17] implementation, coordination, control and intelligence.

In more detail these functions consist of:

1) Implementation consists of daily operations which enable production of physical products and services
2) Coordination function consists of the regulating system (task, authority, responsibilities), which is used to manage production operations.
3) Control function consists of supervision and management of the operations related to implementation and coordination of production of physical goods and services.
4) Intelligence consists of functions relating to the adaptation of environmental changes

Each one of these functions is built and run as a set of predefined processes. These processes are vulnerable to both uncertainties and risks. Protection of critical infrastructure requires thorough assessment of the vulnerability and risks at a process level.

Risk management strategies are not as straightforward as they may seem to be at the first sight. Firstly, because supply chains are, as stated earlier, dependent on several systems there is a need to analyze each system thoroughly in order to assess the correct

approach to the risks and vulnerability on each of the systems. Secondly, each of the above mentioned four processes must be analyzed and assessed within each of the organizations and their business domains. Thirdly, time should be considered as a variable on each of the analysis and its effect on the both vulnerability and risks should be analyzed thoroughly.

2 Network theories

This structure of network organizations and processes is referred as a service ecosystem. It describes the inter-functional and multidiscipline nature of service-oriented industries and operations. One has to bear in mind that this a chosen approach to the subject in question. Author follows the approach of Nordic School of Marketing [18] and Service-Dominant Logic (S-D logic)[19]. This was done due to emphasis on approaches' end-user preferences, which is a widely accepted way to develop services.

Firstly there are few assumptions we have to make in order to discuss about a matter. The first, a well-defined assumption we make is, that supply chain management is a part of the service industry. In their widely cited article Vargo and Lush [19] introduce a theory which main point is a transmission from goods-based exchange to more specialized skills and knowledge based economy, service dominant logic.

Other valid approaches would be Social Network Theory. This approach, which is widely used among the social sciences, is interested in the relationships between individuals and larger groups. Then again, much of the supply chain management is run and managed through automated systems without social interaction. On the other hand, these automated systems are created by humans, whose approach is connected to the social environment they are developed in.

Systems based approach to one's identity [20] has been a topic influencing both educational and social sciences. Obviously research on the area as complex as supply chain management, cannot ignore this topic either since we are not immune to the effects of either cultural or social environment surrounding us. Thus, it could be argued, that there is a need for comprehensive systems wide approach in the supply chain development work.

Social networks have significant importance to the success in supply chain operations. Uncertainty is defined, managed and accepted within the boundaries of a specific social network. Therefore every organization can reduce uncertainty by getting the possession of critical assets and forming ties with stakeholders who are more specialized on a specific operation within their social network. [21;22].

3 Evaluation of supply chain's capability to manage vulnerability and risk

As risk and uncertainty management are capabilities, there has to be a way to evaluate organization's or larger system's capability to manage both vulnerabilities and risks.

At the core of the S-D logic is the shift from an emphasis on the traditional goods based, tangible resources to dynamic resources, which act together with other resources. Vargo and Lusch refer to these resources as operand and operant resources, respectively.

As one of the foundational premises (FPs) of service dominant logic (S-D) is:

> "Value cocreation is coordinated through actor-generated institutions and institutional arrengements"
>
> [19, p.7]

Because supply chain management is highly dependent on the IT-infrastructure, there is an obvious need to manage capabilities running the whole system on supply chain value creation. As Vargo and Lusch state these arrangements need coordination and co-creation.

During a value creation process, an IT-infrastructure (a cyber system) is dependent on five basic components:

1) Input information, which reflects the reality of the surrounding world
2) Stored data of the existing reality to help decision making processes
3) Information stimulating the "organ" (human, machine), which affects the and stimulates the system
4) Data referring to the desired future state of the system
5) Feedback information regarding the desired outcome of the system or parts of the subsystem [23, p.11.]

Each of the organisational functions, implementation, coordination, control and intelligence in the whole of the system and relating sub-systems is dependent on above mentioned five components. Each one of the five components is a subject to the vulnerability and risk. This requires a description of the capabilities, which are needed to manage both functions and basic components of system wide vulnerability and risk management.

As an example, how to reduce the existing risks and vulnerabilities on just a single part of the supply chain, we can use a project Finland has participated. On one of the EUs top IT projects during the past year, which started on a last quarter of 2015 a high-end telecommunications cable was built between Finland and Germany [24].

There are several up-sides on this project. First of all this cable opens a direct link between the mainland Europe and Finland and Finland's dependency on the third party service providers diminishes. Secondly, bilateral co-operation between Finnish and German operators enhances security. There are incentives for the both parties to monitor security environment and react in a manner, which reduces the likelihood of the service breakdowns.

There are still existing connections through Sweden and Baltic states and these connections play still a crucial role in the future. On the other hand, new telecommunications cable does not remove majority of the security issues related to the supply chains. Then again it reduces vulnerability significantly by increasing the telecommunications capacity. Another effect is the increase in number of foreign service providers in Finland. It is in their interest to secure and reduce all the risks affecting their long-term investments. Having said that, one must remember that supply chains are complex cross border systems and IT-systems create only a small, but crucial part of the operations.

4 Capability management

There are several perspectives to a capability management. One, which benefits our purpose is introduced by Anteroinen [25, p. 13].

"the ability or power to achieve a desired operational effect in a selected environment and to sustain this effect for a designated period".

There are obvious similarities to the earlier introduced S-D logic. Firstly, this definition does not define how objective should be achieved. Secondly, this definition takes in to account a domain operations are being run. This reflects to the risks in the international trade introduced earlier.

Also, cross-functional operations require a set of abilities, which enable efficient management of operations and minimalization of vulnerabilities and risks. As stated earlier, there is a need to asses all the parts of the supply chain in their business domain and reflect pre-determined outcomes. Capability management gives also a clear structure for definition process of risks. This helps to overview and concentrate on the relevant risks [26].

Table 1, annual value of imports and exports, 2000-2015

Year	Import € millions	Export € million	Balance of trade, € million
2000	36 837	49 484	12 647
2001	35 891	47 800	11 910
2002	35 611	47 245	11 634
2003	36 775	46 378	9 604
2004	40 730	48 917	8 187
2005	47 027	52 453	5 426
2006	55 253	61 489	6 237
2007	59 616	65 688	6 072
2008	62 402	65 580	3 178
2009	43 655	45 063	1 409
2010	51 899	52 439	539
2011	60 535	56 855	-3 680
2012	59 517	56 878	-2 639
2013	58 407	56 048	-2 359
2014	57 769	55 973	-1 796
2015*	54 256	53 829	-427
* prediction			

5 A fragmented discipline

As author has clearly illustrated the discipline under research, service and supply chain management as a part of it, is very fragmented. Risk and uncertainty management related services require an adaptive and open approach [27]. In military and defence industry a widely used approach is performance logistics (PBL) [28]. PBL is a combination of several logistics functions/services.

Basic idea of the PBL is, that a responsibility of the product / service system management is on a supplier of the system, unlike in the traditional end-user – supplier relationship [29] Berkowitz et al. define PBL [30, p 5]:

> *"…contractual mechanisms will include long-term relationships and appropriately structured incentives with service providers…, to support the end user's (warfighter's) objectives."*

In PBL a customer buys predetermined outcomes. These outcomes are dependent on and simultaneously vulnerable to outcomes of sub-systems. This interdependency is similar to the service systems and demands close co-operation among the whole supply chain [31].

Again, PBL operations need a set of easily quantifiable control point in critical locations [32]. They should support all the operations and take into account all the stakeholder groups participating to the service production [33]. Commercial service development follows the principles laid out by several scholars, but this vast amount of research does not suit Critical Infrastructure Protection (CIP). [34;18;35;36]

Then again, as Lewis [37] points, actors participating in supply chain management are commercial companies whose main purpose is to run commercially viable operations. Therefore CIP is not their first priority, but still an essential part of business due to its financial importance. Also, international trade expands its web so widely, that regional conflicts or crises are seldom a concern of another country from any other than commercial point-of-view. These actors have streamlined their operations to the point that no back-up systems exist [37]. If one would like to exaggerate slightly, any nod is crucial in this super streamlined model.

In stead of asking, how to protect well known critical nods, one should ask three questions:

1) How to increase carrying capacity for the short period of time for the critical imported goods?
2) How to monitor possible indicators, which affect supply chain operations?
3) How to protect supply chain in the abnormal situations?

6 Domain in supply chain CIP

Finland's geographical position dictates the transportation methods we are dependent on. There are only few opportunities other than sea-transportation to keep the Finnish economy running. Finnish economy is heavily dependent on international trade.

Lack of natural resources other than forestry products makes Finland highly dependent on resources enabling functions of modern society. Petrochemical product together with end-products of mining industry create almost 31% of the imports [38]. On the other side of the equitation is Finland's high dependcy on the forestry and chemical industry exports. These two categories create 40,3 % of all the exports.

There is an existing back-up storage for the most important items for the Finnish society. According to Österlund [39] these stored materials can keep Finland up and running for a short period of time, before the functionality of the society is put at the risk. Table 2 by Österlund [39] illustrates clearly critical time lines related to vital imports.

Table 2, Critical industry sectors

CRITICAL INDUSTRY SECTORS IN FINLAND AND THEIR MAIN IMPORTS

ACCORDING TO A INDUSTRY SURVEY 2011

Critical industry	*Main imported goods and materials*	*Rate of import dependency (%)*
Energy	Oil, gas, uranium, coal	Crude oil, uranium, coal, natural gas 100% - Share of imports in all energy production 65 % - Electricity: 15-20 %
Food sector	Pesticides, fertilizers and their raw materials, animal feed, agricultural machinery, chemicals, packaging materials Raw materials for the foodstuffs Packing materials	- Pesticides 100 % - Fertilizers (surplus approx. 50 %) - Ammonia % noble metal catalysts used in fertilizer production 100 % - Animal feed (soya protein) 70 % - Machinery 45 % - Raw materials for foodstuff 20 % - Food sold for customers 30 % - Packing materials
Health care	Pharmaceuticals, equipment, chemicals	- Raw materials for pharmaceutical production - Pharmaceuticals 85 % - Equipment 70 % - Packing materials
Forestry industry	Timber, fillers, coating pigments	- Timber 10-23 % - Fillers (kaolin) 70 % - Pigments
Chemical industry	Crude oil, basic chemicals, rubber	- Crude oil 100 % - Basic chemicals - Rubber 100 %
Technology industry	Components and parts, metals, minerals, fuels	- Components & Parts - Iron concentrate 100 % - Copper, nickel, & zinc concentrate - Components and other raw materials

This rather grim table shows the vulnerability of the Finnish society and its dependency on the imports and foreign supply chains. With the figures from the table 2 in mind, it is fairly obvious that the role of uncertainty screening is far greater in Finland's case than in countries less dependent on imports on all the sectors of the economy. Uncertainty monitoring enhances the opportunities to create alternative methods to fulfill the existing demand by the existing supply chains. It gives an opportunity to launch the contingency plans to the scale, which gives an opportunity to continue production of critical goods and services in the Finnish society.

7 Conclusions

This paper was started by the description of how dependent modern trade and societies are on information.

Looking at the protection of the critical infrastructure for the supply chains in Finland, we can divide the assessment work in to 5 domains.

1) Capability management on the recognized management areas:
    a. Creation of valid input information, which enables necessary vulnerability analysis
    b. Capability to store that valid information in a way that it meets the requirements for the protection of critical infrastructure
    c. Capability to manage the organ which uses the valid information
    d. Capability to predict and create scenarios, which require valid information
    e. Capability to manage feedback information and most of all, manage the pre-determined operations based on the scenarios
2) Management of the contractual environment by using the methods, which take in to account the needs of end-user and the needs of the whole of the supply chain as a system. One of the methods could be PBL.
3) Management of the carrying capacity and securing the alternative methods of transportation and suppliers of vital goods and services for the Finnish society
4) Management of the alternative and existing data transportation methods under all the conditions.
5) Enhancement of the capability to react rapidly changing political, social, environmental, legal and technological changes. This refers more to the first domain.

8 Discussion

Current global trading operates in an environment, which is highly vulnerable to abnormalities on the any part of the supply chain.

Looking at these domains it becomes evident, that assessment of an individual domain, process or actor is not adequate enough. There is the need to find those processes, which have the largest number of interfaces to each of the domains and just after that the assessment of these interfaces or nods needs to be done. Because supply chain consists of large number of multinational players, each one of these players is potentially a critical nod due to efficiency requirements defined by financial requirements. But unless individual nod does not affect pre-determined critical process in the critical domain a total collapse of a supply chain is not foreseeable. But there is a need for systematic method, which enables to assess whole systems and their subsystems.

Relationships between these parts and explaining theories are illustrated in the Figure 1.

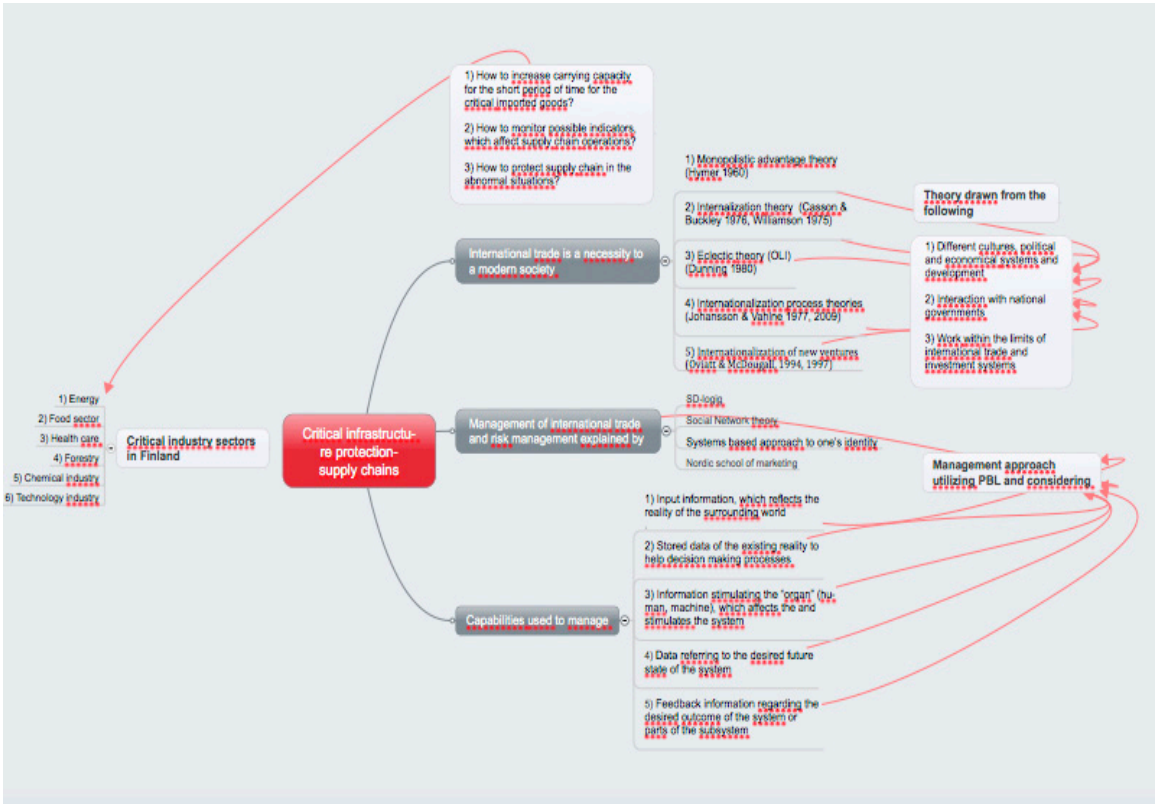


Figure 1, Supply chain and relationships between the main theories

This illustrated framework is a starting point for a systems wide analysis but it lacks a crucial aspect – a lifecycle of a system. There is an existing tool, ISO 15288, which gives an opportunity to manage the whole life cycle of a system. It is also compatible to quality management system ISO 9001, which is a risk preventative approach to supply chain management [40].

ISO 15288 division of process to enabling processes to lifecycle, infrastructure, project portfolio, HR and quality management bind the above illustrated areas to a manageable entirety. It also combines the human element, which plays an essential role in every service and capability creation and management. Adapting the main theories to the organizational project-enabling processes and utilizing the ISO 15288 can help to protect critical nods in the supply chains.

Combining the systems wide approach and explaining the theoretical background behind the various models creates a comprehensive model, which helps in the critical infrastructure protection. It shows that individual supply chains are collection of extremely complex systems and sub-systems backed-up with sometimes contradicting theories.

Dialogue between supply chain stakeholders does not jeopardize the risk management procedures of a supply chain, quite the opposite. It creates a solid base to understand system stakeholders and their needs throughout the different life cycles of a supply chain. Maglio, Srinivasan, Kreulen, and Spohrer [41] envision that service scientists could start understanding service systems by identifying stakeholders and their needs, and opportunities and problems in the environment. Theories behind the service science need to be opened during the development work. It should be done due to a fact that capability management requires open and multidiscipline dialogue between

different disciplines and functions. Especially, when the purpose of the system is being described following the basic principles of the PBL.

References

[1] Bertelsmann foundation. *Globalization Gains for Developed Countries Outpace Those for Emerging Nations [Online] Retrieved June 15, 2016, from http://www.bfna.org/article/globalization-gains-for-developed-countries-outpace-those-for-emerging-nations*

[2] Laanti, Riku, Mika Gabrielsson, and Peter Gabrielsson. "The globalization strategies of business-to-business born global firms in the wireless technology industry." *Industrial Marketing Management* 36.8 (2007): 1104-1117.

[3] Hymer, Stephen. "On multinational corporations and foreign direct investment." *The Theory of Transnational Corporations. London: Routledge for the United Nations* (1960).

[4] Buckley, Peter J., and Mark Casson. *Future of the multinational enterprise*. Springer, 1976.

[5] Williamson, Oliver E. "Markets and hierarchies." *New York* (1975): 26-30.

[6] Dunning, John H. "Toward an eclectic theory of international production: Some empirical tests." *Journal of international business studies* 11.1 (1980): 9-31.

[7] Johanson, Jan, and Jan-Erik Vahlne. "The internationalization process of the firm—a model of knowledge development and increasing foreign market commitments." *Journal of international business studies* 8.1 (1977): 23-32.

[8] Johanson, Jan, and Jan-Erik Vahlne. "The Uppsala internationalization process model revisited: From liability of foreignness to liability of outsidership." *Journal of international business studies* 40.9 (2009): 1411-1431.

[9] Oviatt, Benjamin M., and Patricia Phillips McDougall. "Toward a theory of international new ventures." *Journal of international business studies* 25.1 (1994): 45-64.

[10] Oviatt, Benjamin M., and Patricia Phillips McDougall. "Challenges for internationalization process theory: The case of international new ventures." *MIR: Management International Review* (1997): 85-99.

[11] Johnson, Gerry, Kevan Scholes, and Richard Whittington. *Exploring corporate strategy: Text and cases*. Pearson Education, 2008.

[12] Heene, Aimé, and Ron Sanchez, eds. *Competence-based strategic management*. London: Wiley, 1997.

[13] Fiol, C. Marlene, and Marjorie A. Lyles. "Organizational learning." *Academy of management review* 10.4 (1985): 803-813.

[14] Galbraith, Jay R. *Designing complex organizations*. Addison-Wesley Longman Publishing Co., Inc., 1973.

[15] Liesch, Peter W., Lawrence S. Welch, and Peter J. Buckley. "Risk and uncertainty in internationalisation and international entrepreneurship studies." *Management International Review* 51.6 (2011): 851-873.

[16] Zsidisin, George A., et al. "An analysis of supply risk assessment techniques." *International Journal of Physical Distribution & Logistics Management* 34.5 (2004): 397-413.

[17] Beer, Randall D. "A dynamical systems perspective on agent-environment interaction." *Artificial intelligence* 72.1 (1995): 173-215.

[18] Grönroos, Christian. "Marketing as promise management: regaining customer management for marketing." *Journal of Business & Industrial Marketing* 24.5/6 (2009): 351-359.

[19] Vargo, Stephen L., and Robert F. Lusch. "Service-dominant logic: continuing the evolution." *Journal of the Academy of marketing Science* 36.1 (2008): 1-10.

[20] Bronfenbrenner, Urie. "Ecology of the family as a context for human development: Research perspectives." *Developmental psychology* 22.6 (1986): 723.

[21] Hoffman, Martin L. *Empathy and moral development: Implications for caring and justice*. Cambridge University Press, 2001.

[22] Henriques, Irene, and Sanjay Sharma. "Pathways of stakeholder influence in the Canadian forestry industry." *Business Strategy and the Environment* 14.6 (2005): 384-398.

[23] Kuusisto, Tuija. "Kybertaistelu 2020." *Julkaisusarja 2: Asiatietoa, No. 1/2014* (2014).

[24] Virtanen, S., Suomen ja Saksan välisen datakaapelin laskeminen Itämeren pohjaan alkaa Santahaminassa – 100 miljoonan euron hanke [Online] Retrieved June 15, 2016, from http://www.tekniikkatalous.fi/tekniikka/ict/suomen-ja-saksan-valisen-datakaapelin-laskeminen-itameren-pohjaan-alkaa-santahaminassa-100-miljoonan-euron-hanke-6057452

[25] Anteroinen, Jukka. *Enhancing the development of military capabilities by a systems approach*. Maanpuolustuskorkeakoulu, 2013.

[26] Teller, Juliane, Alexander Kock, and Hans Georg Gemünden. "Risk management in project portfolios is more than managing project risks: A contingency perspective on risk management." *Project Management Journal* 45.4 (2014): 67-80.

[27] Ng, Irene CL, Roger Maull, and Nick Yip. "Outcome-based contracts as a driver for systems thinking and service-dominant logic in service science: Evidence from the defence industry." *European Management Journal* 27.6 (2009): 377-387.

[28] Gansler, Jacques, and William Lucyshyn. "Evaluation of performance based logistics." (2006).

[29] Randall, Wesley S., Terrance L. Pohlen, and Joe B. Hanna. "Evolving a theory of performance-based logistics using insights from service dominant logic." *Journal of Business Logistics* 31.2 (2010): 35-61.

[30] Berkowitz, David, et al. "Performance Based Logistics." *Center for the Management of Science and Technology, Huntsville, AL* (2003).

[31] Kleemann, Florian C., and Michael Essig. "A providers' perspective on supplier relationships in performance-based contracting." *Journal of Purchasing and Supply Management* 19.3 (2013): 185-198.

[32] Gansler, Jacques, and William Lucyshyn. "Evaluation of performance based logistics." (2006).

[33] Brown, Trevor L., Matthew Potoski, and David M. Van Slyke. "Managing public service contracts: Aligning values, institutions, and markets." *Public Administration Review* 66.3 (2006): 323-331.

[34] Pine, B. Joseph, and James H. Gilmore. *The experience economy*. Harvard Business Press, 2011.

[35] Gummesson, Evert. "Extending the service-dominant logic: from customer centricity to balanced centricity." *Journal of the Academy of Marketing Science* 36.1 (2008): 15-17.

[36] Vargo, Stephen L., and Robert F. Lusch. "Evolving to a new dominant logic for marketing." *Journal of marketing* 68.1 (2004): 1-17.

[37] Lewis, Ted G. *Critical infrastructure protection in homeland security: defending a networked nation*. John Wiley & Sons, 2014.

[38] Tulli, *"Tullin valvonnan vuosijulkaisu", [Online]* Retrived July 5, 2016, from: http://www.tulli.fi/fi/suomen_tulli/julkaisut_ja_esitteet/vuosikertomukset/index.jsp

[39] Österlund, Bo. "Does Finland have enough sea transport capacity?." (2015).

[40] ISO, ISO/IEC/IEEE 15288:2015, [Online] Retrieved July 5, 2016, from: http://www.iso.org/iso/rss.xml?csnumber=63711&rss=detail

Systems and software engineering -- System life cycle processes

[41] Maglio, Paul P., et al. "Service systems, service scientists, SSME, and innovation." *Communications of the ACM* 49.7 (2006): 81-85.

CYBER SECURITY IN THE MANAGEMENT OF AN ELECTRICITY COMPANY

Jouni Pöyhönen

University of Jyväskylä

jouni.a.poyhonen@jyu.fi

Abstract

The functioning of a modern society is based on the cooperation of several critical infrastructures, whose joint efficiency depends increasingly on a reliable national electric power system. Reliability is based on functional data transmission networks in the organizations that belong to the power system. Furthermore, reliability is linked to the usability, reliability and integrity of system data in the operating environment, whose cyber security risks are continuously augmented by threatening scenarios of the digital world.

In Finland, the production of electricity is in various ways decentralized, which contributes to the reliability of the power system. Finland has about 120 enterprises that produce electricity and about 400 power plants, in which electricity is produced using various production methods. Power system process control is highly automated and networked. This report focuses on the procedures applied to cyber security management in the processes of electricity companies, whereby different standards will also be utilized.

The major contributions of the article are that it integrates cyber security management and risk analysis into the process structures of individual electricity companies and that it utilizes the PDCA (Plan, Do, Check, Act) method in developing a company's cyber security management practices.

In order to put the measures into practice, the leadership of an electricity company must regard trust-enhancing measures related to cyber security as a strategic goal, maintain efficient processes and communicate their implementation with a policy that supports the strategy.

Keywords

Critical infrastructure, Electricity company, Cyber security management, Trust

Paper type

Research paper

1 Introduction

Finland's electric power system – comprising power plants, a nationwide transmission grid, regional networks, distribution networks and electricity consumers – is part of an inter-Nordic power system together with the systems of Sweden, Norway and

Eastern Denmark. In addition, there are direct current transmission links to Finland from Russia and Estonia in order to connect the Nordic system to the power systems of Russia and the Baltic countries. The inter-Nordic system is furthermore connected to the system in continental Europe via direct current transmission links. Fingrid Plc. is responsible for balance management, in other words, for maintaining the momentary power balance between power production and consumption in Finland. Monitoring as part of balance management is handled round the clock at the Fingrid Main Grid Control Centre in Helsinki. Within the inter-Nordic system, the main purpose of balance management is to maintain the frequency of the power system, which represents the balance between electricity production and consumption. The better the balance is maintained, the less does the frequency vary and the better is the quality of electricity. [3]

Electricity is produced at Finnish power plants in various ways, using several energy sources and production methods. The major sources of energy include nuclear power, water power, coal, natural gas, wood fuels and peat. In addition to the sources of energy, production can be classified according to the production method. In Finland there are about 120 enterprises that produce electricity as well as around 400 power plants, over half of them hydroelectric power plants. Nearly a third of electricity is produced in connection with heat production. Compared with many other European countries, Finland's electricity production is decentralized. A diverse and decentralized electricity production structure increases the security of the national energy supply. [4]

The functioning of a modern society is based on the cooperation of several critical infrastructures, whose joint efficiency depends increasingly on a reliable national electric power system. Crucial in the cyber environment are functional data transmission networks and the usability, reliability and integrity of system data in the operating environment, whose cyber security risks are continuously augmented by threatening scenarios of the digital world. A modern society depends entirely on a cyber environment that provides dynamic services.

In Finland's Cyber Security Strategy, the cyber environment (domain) is defined as an electronic information (data) processing environment that consists of one or more information technology infrastructures. According to the Strategy, cyber security refers to a desired end state in which the cyber environment is reliable and in which its functioning is ensured. Critical infrastructure, furthermore, refers to the structures and functions that are indispensable for the vital functions of society. They include physical facilities and structures as well as electronic functions and services. [16]

The global threats within the cyber environment have remained at a high level over the past few years, as stated in the annual international business world surveys by the World Economic Forum. They are seen to be among the major global threats based on the probability and impact of their realization. [18]

This finding by the World Economic Forum is supported by continuous news in different media that shake our trust in cyber security. For example, an extensive cyberattack to the power grid caused power failures in Ukraine in December 2015 [8].

The electric power system with all its components belongs to critical national infrastructure: it is vital for the operations of the country and its outage or destruction would weaken national security, the economy, public health and safety as well as make the operations of state administration less effective.

The criticality of the power system is expressed clearly in the seminar presentation 'The power system as a basis for a functioning society' (Sähköjärjestelmä yhteiskunnan toimivuuden perustana) given by the former chief executive officer of the National Emergency Supply Agency. The table below is an extract from the presentation. It describes the effects of power failure on the operations of society as a function of the duration of the failure. Endangered cyber security has been regarded as one of the most significant threats to the functioning of energy supply and energy networks. [10]

Table 1: The consequences of power failure [10]

Interruption time	**Consequences**
1 second	Sensitive industrial processes may stop. Data in information systems may be lost.
1 minute	Some industry and hospital processes will stop.
15 minutes	Shops will be closed. The failure may harm people's daily activities and cause traffic delays.
2–3 hours	Industrial processes may undergo significant damage. Mobile phone networks will face problems. Domestic animal production will be disturbed.
12–24 hours	Water supply to homes and offices will stop. Buildings will start to become cold in the winter. Frozen goods will begin to melt.
Several days	The operations of society will be seriously harmed. Industry and services will not function. Workplaces and schools will be closed. Buildings will suffer from frost damage.

Because the electric power system provides a basis for almost all services in society, its operation must be as uninterrupted as possible. Even short power failures are broadly visible as disturbances in other critical services. Therefore, achieving and maintaining a high availability level in the operation of various processes within the power system is the primary goal of the organizations responsible for them. For this purpose, the continuity of operation must be ensured and recovery from disturbances

must be quick. Creating and maintaining situational awareness related to cyber security in individual organizations play a key role in these activities. Controlling process operations and taking coordinated situation-specific decisions call for real-time, comprehensive situation awareness regarding the organizations' cyber readiness and the factors that affect it in a dynamic operating environment.

A diversified and decentralized power production structure increases the national security of power supply. Considering cyber security in the different parts of the infrastructure further enhances trust in the services of our society.

The national significance of an electric power system is very similar irrespective of the country. For example, in the USA the power system is considered to be a critical infrastructure and a key resource for the functioning of the entire society. The basic structures of the power system are similar to those in Finland. However, in the USA the structural and technical implementation of grid load balance management differs significantly from the corresponding procedure in Finland. Grid management has traditionally been affected by administrative regulation, and an individual electricity company has owned a broad regional production and distribution chain through vertical integration. The electricity company has thus been in charge of power production, transmission and distribution to consumers. The deregulation that has occurred in the past two decades has increased the number of companies in the field, particularly in transmission and distribution, and thus led to the abolishment of vertical integration. As a consequence, competition between enterprises has increased. These events have brought challenges to grid balance management. In the USA the aim is to perform balance management by directing production to consumption in different parts of the grid. This is done by balancing the sums of input and output currents at nodes, in accordance with Kirchoff's circuit law. In the USA it can be seen that the grid represents a technologically highly advanced system entity and that its solutions call for the use of the most demanding technologies. Grid technology and its control procedures constitute the principal areas in examining cyber security. [13]

This article focuses on factors related to cyber security management in an individual electricity company that is part of Finland's power system. We will also examine how these factors are taken into account in the company's process structures while creating trust in its operation within a dynamic cyber environment.

2 An electricity company's cyber environment and its main cyber security threats

2.1 The structure of an electricity company's cyber environment

The transmission network of Finland's power grid is owned by Fingrid Plc. The distribution network consists of dozens of enterprises, and electricity is produced by about 120 enterprises and 400 power plants in different parts of the country. The system structure is thus highly decentralized, and there is no comprehensive vertical integration of ownership in the Finnish power system. Every company is responsible for managing its own working processes. Balance management in the Finnish system, however, is centralized, which also compensates for the benefits of vertical integration

in system management and control. From the perspective of the entire power system, the major threats to physical safety and cyber security concern the transmission and distribution networks, switching and transforming substations, and power plants. A decentralized structure limits the potential consequences of these threats in the power system. On the other hand, decentralized electricity production requires good overall management of the system, effective distribution systems in the electricity companies as well as the capability to manage and control power plants. The companies that own power plants must also have a well-functioning logistics control system. The aim is to optimize the size of raw material stocks in each power plant according to their consumption as cost-effectively as possible. Therefore, the correct timing of raw material deliveries plays a significant role for the continuity of production.

The general networks and working processes involved in the operation of an electricity company can be illustrated with a logistics framework that comprises a supplier network, a production process, a client network, and information and material flows that connect them. Information technology (IT) systems are part of a company's infrastructure and thus constitute a significant part of the operations that support a company's core processes. Corporate-level IT systems are related to administration and to the management of information and material flows in the network. The production level includes industrial automation systems (industrial control systems, ICS). Figure 1 presents the structure of a company's logistics framework and common IT and industrial automation systems.

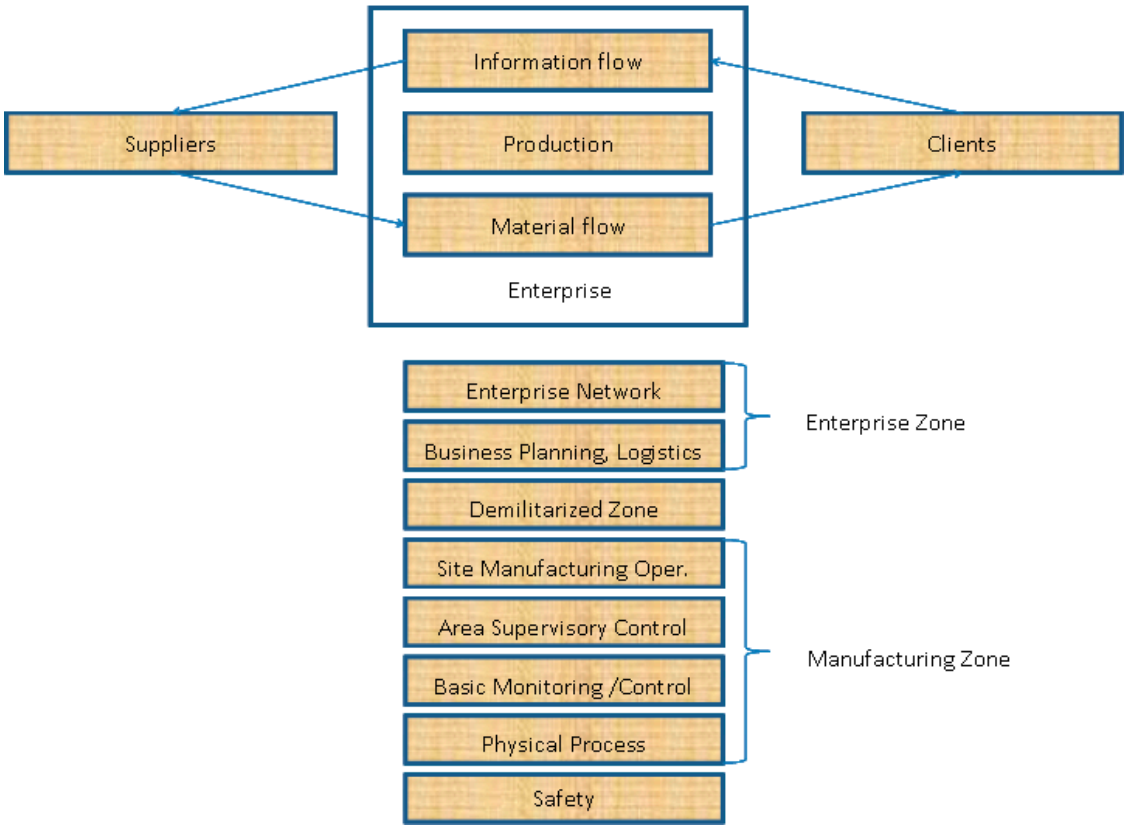


Figure 1: The logistics framework of an electricity company (adapted) and common IT and industrial automation systems [1 adapted; 11]

The highest levels of IT system hierarchy include the general information systems of administration and the enterprise resource planning (ERP) system. The top level of a typical ERP system includes overall process management by, for example, guiding the production volume. It also covers the restocking of raw materials, storing, distribution, payment traffic and human resources. If needed, between ERP software and control rooms there may be a manufacturing execution system (MES), which makes it possible to transfer the information obtained from the control room to the ERP system.

The industrial automation systems of production within an electricity company comprise their own hierarchy levels. Topmost of them is the control room, from which the operation of the entire process is presented to the supervisors in graphic form. Based on the information, process alarms are handled and the operation of the process is monitored and controlled. The next level consists of process stations, which house devices for process control, measuring and regulation. The same level also includes the actions taken to monitor faults and interferences in devices. The lowest level comprises the field equipment used to control and monitor process actuators and to gather measurement data.

2.2 The main threats to cyber security in an electricity company

When evaluating the role of electricity production systems in the cyber world as well as the factors that affect their cyber security, it is of primary importance to be aware of the most central features of the systems. For instance, the distributed industrial automation systems used in controlling production processes can be characterized by saying that their operation is highly established and that their life cycles are long compared with other IT systems in a company. The life cycles of industrial automation systems can even be several decades, as far as the basic systems are concerned. Moreover, the structure of the basic systems is changed infrequently. The changes are mainly carried out as system life-cycle updates in connection with larger maintenance or alteration works. The resources of industrial automation systems are also restricted, which is why it has not been possible to use typical technological information security solutions or cryptographies in them. Their user organizations are properly trained for their tasks and thus familiar with the devices as well as with the operating principles and operating environments of these devices. The data warehouses of industrial automation systems chiefly include process data, whereas administrative IT systems commonly include confidential business information. Unlike in administrative IT systems, no direct connection to the internet is usually needed in industrial automation systems. In the latter systems, IT devices are not used for purposes other than their decentralized tasks within the production process, its measurement and control tasks, and security functions. The monitoring of operations and staff in industrial automation systems is strictly controlled because of, for example, the availability and safety requirements of process operation. [5]

The aforementioned IT and industrial automation systems are part of the common cyber world, in which the primary risks are related to the loss of money, sensitive information and reputation as well as to business hindrance. Security solutions are hereby the key elements in risk management. The vulnerabilities behind the risks can

be analysed as insufficient technology in relation to attack technology, insufficient staff competence or inappropriate working methods, deficiencies in the management of organizations, and lacks in operating processes or their technologies. The most common motives of attackers are related to the aim of causing destructive effects on processes, making inquiries about process vulnerabilities, and anarchism or egoism. These attacks can even be carried out by state-level actors, but perhaps most commonly by organized activists, hackers or individuals acting independently. [12]

Harmful measures to the systems of an electricity company can be implemented by foisting mal- and spyware into the systems utilizing the staff; or they can include intruding or network attacks via wireless connections or the internet. The intruders' goals may be related to the prevention of network services, the complete paralyzation of operations, data theft or distortion, and the use of spyware. Components pre-infected with so-called backdoors or the programming of components intentionally for the purposes of attackers is also increasingly common in today's cyber world. [12]

In the USA the security threats to the electric power system concern power plant logistics. They involve interfering and harming raw material supply routes, doing physical damage to transmission and distribution networks as well as to the transformer and switching substations between them, or performing cyberattacks to the control and regulation systems of the power grid. [13]

Protecting the power system against threats implies measures taken based on risk assessment, and they ensure the availability of primarily digital information in the operating processes being examined. The measures are highly significant for the overall availability of the systems that support the processes. Availability plays a key role in achieving business results and promoting the reliability of activities. Further central goals include the reliability and content integrity of information within the processes and used by the processes. Overall trust should be built from these starting points, based on the target organization's realistic idea of its own capabilities to reliably manage the challenges involved in operations within the cyber world. The following section addresses the significance of trust in the cyber environment for the operations of an electricity company. Moreover, trust-enhancing measures applicable to a company will be mapped.

3 Trust in an electricity company's cyber security

3.1 The significance of trust for cyber security

Trust in the operation of organizations and its continuous maintenance with effective measures are central factors affecting cyber security. Security is based on trust. Without trust there is no security, and vice versa. It is also good to be conscious of the fact that perfect safety is in general hardly achievable, and this also applies to the cyber world, which is a dynamic environment difficult to anticipate. Therefore, it is particularly important to understand the great significance of trust in the cyber world and its security. The role of measures enhancing trust is emphasized. When we build operations in the cyber world on a foundation that is as sustainable as possible, we can utilize the diverse opportunities it offers. [15]

Finland's national cyber security strategy highlights the need to increase general cyber trust throughout the entire society. The strategy underlines the role of authorities, but at the same time it notes that in practice most production and service provision with bearing on the national product comes from the private sector. It further states that cooperation between the public and private sectors is an indispensable prerequisite for achieving the goals of the strategy. Citizens' activities also play a key role in enhancing security. An increase in citizens' cyber competences is immediately visible as competence at work and as other daily IT skills. Measures that promote a balanced cyber trust in all sectors of society improve the possibilities for safe operation in an information society, produce shared added value as well as ensure and increase the operational preconditions of both the public sector and the business sector. The strategy emphasises that all actors, from individuals to enterprises and public administration, are responsible for their own preparedness for cyber threats. Education and research also occupy an important role in maintaining and developing cyber security and in disseminating information throughout society. [16]

The ISO 9000 Standard states that an organization achieves success by acquiring and maintaining the trust of clients and other relevant interest groups. Understanding their present and future needs contributes to the organization's continuous success. The standard includes the central concepts of quality management and the principles for building trust. It can be applied by organizations that pursue ongoing success in their operation by utilizing a quality management system of their own. The quality of an organization's products and services is determined by how its clients experience that their needs and expectations are met. Clients also look for guarantees on the organization's ability to systematically produce products and services that correspond to their requirements. The ISO 9000 Standard comprises seven quality management principles, which constitute a commonly accepted basis for applying the standard series. The standard also specifies the benefits to an organization that has adopted the principles in its operation. The seven basic quality management principles are related to customer focus, leadership, the engagement of staff, a process approach, continuous improvement, evidence-based decision-making, and relationship management. [7]

3.2 Cyber trust and process management

Establishing measures that increase cyber world security and trust in a company is primarily the responsibility of corporate leadership. Integrating the necessary measures with the idea of ensured business activities increases their significance and benefits through better processes for the entire organization, interest groups and society. If security is not considered, risk analysis reveals potential damages as well as their costs and social consequences. The leadership's views and requirements brought out in the analysis play a central role in developing security planning for the operating process. The costs and other resources allocated to the activities are simultaneously specified. [17]

An organization has a management system generally suitable for its business environment when it is managed systematically and at a high level, taking into account customers, the significance of staff, the efficiency and guidance of processes, continuous

development of activities, and interest group communication. The management system can also be utilized in managing the processes of the cyber environment.

Process management theory has developed along with industrial production. The development of industrial mass production led to the use of variation theory while developing production process control: to perform control measures, uniform product quality was monitored with statistical methods. Statistical process analysis led to the observation that variation occurs everywhere in nature and in the processes and systems created by humans. After analysing distributions that involved variation, variation was classified into two types according to its causes: variation due to common causes (or the system itself) and variation due to special causes (i.e. named and assignable causes). Systemic variation has random causes and it is therefore often normally distributed, according to Gaussian distribution. Variation resulting from special causes does not follow any regularities. The common causes of variation are thus constantly present in the process. An individual cause produces only little deviation, but several causes together generate considerable variation. The causes of special variation, on the other hand, are not constantly present in the process. They come from outside of the process and usually generate more variation in the process than the common causes. In uncontrolled processes, deviation as a result of both types occurs simultaneously. [14]

In principle, Lillrank's theory on the causes of process variation can also be generalized to the processes of an electricity company. The measures taken by corporate leadership can be targeted at reducing variations resulting from both aforementioned types of causes. Proper planning and control of process performance reduce variation generated by random causes. At a general level, it is always recommended to aim at reducing this variation. If corporate leadership, in particular, concentrates too much on process changes resulting from random causes, it can lead to overreactions in process control due to the measures chosen. At its worst, this can lead to loss of control in managing the overall process. The actions of corporate leadership should indeed be targeted primarily at proactively preventing variation generated by special causes. Almost without exception, serious cyber security disturbances occurring in the operating process cause blackouts. They do not represent normal process variation but are deviations resulting from special causes. They are not in the normal range of variation. Taking these special causes into account in planning and proactively implementing managerial activities reduces related risks and improves the overall reliability of the company's operations.

3.3 Measures increasing cyber trust

The following measures related to cyber security management in an electricity company encompass the aforementioned seven principles of quality management.

In order to comprehensively build corporate cyber security, corporate leadership must define and guide actions at the strategic, operational and technological-tactical levels. The strategic level provides answers to 'why' and 'what' questions. The operational and tactical levels answer the 'how' question. The approach guided by questions ensures that the right things are done and that they are done in line with the set goal.

The technological-tactical level must implement the goal-oriented activities defined at the strategic level, not create it. The company's organizational capability in implementing the cyber security measures required by the technological-tactical level ultimately determines how the company manages potential disturbance situations. [15]

Building corporate cyber security management begins from the level of vision and strategy work. The visions created by corporate leadership to enhance cyber trust are translated into strategic goals, operational-level actions, guidelines and a policy. The practical measures derived from the strategy are realized at the technological-tactical level. Organizational capability factors enable the success of the measures.

In this article, creating a vision of cyber security in an electricity company is presented in the context of continuous development and maintenance of cyber trust as part of national critical infrastructure. The strategic choices supporting the creation of visions are primarily related to corporate social responsibility, company reputation, and ensuring business and its economic efficiency. The leadership is expected to make concrete strategic choices as well as support and guide the execution of the chosen measures throughout the organization. It is also important that the leadership ensures sufficient resource allocation to the measures. The chosen measures should be comprehensively communicated to the company's interest groups. [17; 7]

The measures at the operational level promote the strategic goals. Comprehensive measures that increase security and trust call for holistic cyber security management. It must be based on risk assessment and analyses of the measures based on the assessment. It is also important that the company declares and communicates the policy with which the leadership commits to the measures required to develop cyber security management. The declaration of a policy that ensures cyber security and the development of related procedures must be integrated with the organization's general policies. The highest organizational level is responsible for creating a policy that defines acceptable risk levels and the measures used in the reduction of risks [17]. The concrete measures at the operational level must be targeted at ensuring data security solutions and at creating business continuity and recovery plans [6]. The maintenance of situational awareness regarding the cyber environment of the electricity company's processes, furthermore, makes it possible to monitor the effects of the operational measures and, when needed, to react efficiently to events that constitute a threat within the company's operating environment. The aim must be to continuously monitor the availability of processes and to support decision-making in disturbance situations that require analyses and decisions [2].

The tactical corporate level encompasses the systems and processes that comply with the logistics framework. Consistent and predictable results are achieved more efficiently when operations are handled and managed as interrelated processes that function as a coherent system [7]. Cyber security threats set special requirements for these processes in addition to other operational requirements. At a general level, the performance of processes is determined according to their client-based demands. In an electricity company, uninterrupted production of electricity can be regarded as the most important requirement, and it is achieved through a high availability level of the processes. In the cyber environment, the target can be achieved by defining the processes to be protected, choosing process control mechanisms successfully, and by

using expedient technological solutions and services to protect the processes [17]. Successful operation also calls for the adoption of security-oriented values to guide the activities of staff [14]. The aforementioned solutions suitable for the cyber environment constitute an entity that can be called a technological-tactical level.

The continuous improvement of activities related to cyber security as well as the development of staff competence enhance the organization's capability to proactively prevent disturbances and tolerate potential changes in process operation caused by them. Taking the staff into account at all organizational levels, as well as focusing on competence and the possibilities it opens to fully influence in the organization, develops the overall operations of the company [7]. The continuous development of activities and staff competence support the measures taken at the strategic, operational and technological-tactical level.

Sufficient knowledge of the observed process is the starting point for continuous improvement. The measures taken rely on the idea that we observe process variation and reduce it by tackling variation that results from special causes. Addressing this type of variation often requires the use of different basic quality management tools in order to find the causes for deviations. In the context of cyber security development, the continuous improvement of organizational processes can also be seen as a proactive measure and thus as a measure that increases trust in the operating environment. Continuous process improvement is based on the continuous assessment of activities. In an electricity company, the main assessment criterion for an efficient process is its availability. When it comes to cyber security, the main constituent of overall availability is the availability of information in technological systems. High availability is achieved by continuously monitoring process meters and by adopting process performance improvement as an ongoing approach. It is typical that a learning and development-oriented organization is continuously looking for areas in which to improve.

In addition to performance measurement and different quality management tools, the organization can utilize feedback systems and benchmarking for the continuous improvement of processes. Traditional organizational feedback systems – such as internal self-evaluations, audits and reviews as well as external audits and their outcomes – can produce data for the development of operations and continuous improvement also regarding cyber security. For this purpose, cyber security and related trust-enhancing measures must be integrated into the feedback as one of its dimensions. Benchmarking, on the other hand, can be efficiently promoted by, for example, establishing branch-specific user groups among companies and maintaining regular exchange of information between them, particularly on measures that have been effective in solving disturbance situations and recovering from them.

Cyber security management calls for the constant maintenance of staff competence and consideration of their training needs. Staff competence is a crucial factor that determines the level of the entire company's activities. The capacity of human resources can be increased by developing employees' knowledge and skills related to cyber security and thus developing the company's capabilities. Challenges related to capabilities grow when an enterprise's cyber environment becomes more complex along with globalization and technological development. Investment in staff

competence can transform the enterprise's capability into core competence, which can be used to pursue competitive advantage through trust. This will provide unique added value to both the company and its customers. In the case of an electricity company, successful cyber activity development and maintenance through staff competence can ideally lead to long-lasting added value, in spite of rapid changes in the operating environment. Valuable capabilities are helpful in a company's threat and risk management and can consequently facilitate, in particular, the utilization of profitable opportunities.

Figure 2 summarizes the aforementioned measures taken to increase an electricity company's cyber trust.

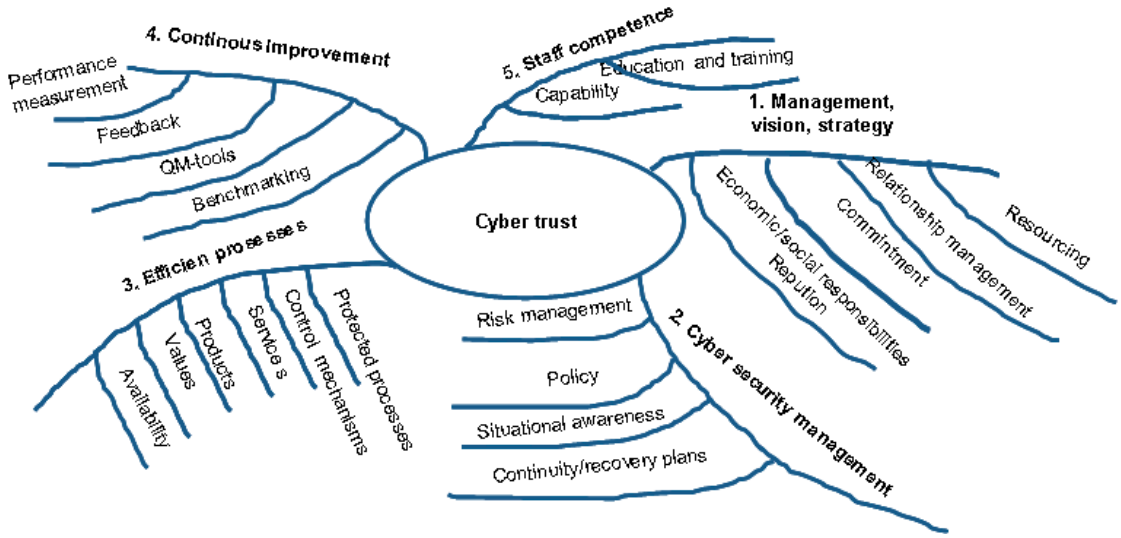


Figure 2: Measures increasing an electricity company's cyber trust

4 Implementing the measures that enhance cyber trust

4.1 An integrated management system and its components

When management in an organization is performed systematically, we talk about the organization's management system. A management system can comprise various control systems that comply with different standards, such as a quality management system, an information security management system and an environmental management system. In order to put into practice principles that comply with different standards, an organization may describe the required measures in its integrated management system (IMS). The IMS is a description of the procedures everyone should apply in the organization. With the help of guidelines and operations models jointly defined by the leadership and staff, the aim is to purposefully maintain a high level of activities and to develop the activities with an eye on set goals as well as the needs of clients and interest groups. The integrated management system compiles process descriptions, guidelines, recordings, indicators, tasks and feedback into a functional whole, which guides and supports the organization's mission and vision as well as the actions taken to realize them.

Management and the necessary measures related to cyber security in electricity production, including their objectives, must be documented in, for example, an organization's quality manual for the entire staff to see.

4.2 Trust-enhancing measures based on risk analysis

The vision for achieving a company's goals is the point of departure for trust-enhancing measures. The definition of strategy derived from the vision guides the actions taken in order to achieve the goals. At the first stage, it is most practical to facilitate the definition of strategy by performing risk analysis on cyber threats. When examining an electricity company, the targets of risk analysis are determined by the company's logistics framework and its IT processes. An electricity company's systems include a fuel logistic and feed system, a production system and its support processes, and the electricity distribution system. Because all the aforementioned components are needed in the operation of an electricity company, their mutual dependence as well as operations management and monitoring are crucial for the success of overall production. In managing cyber security, the different functions of the logistics framework must be treated as subjects of equal value.

If an organization is familiar with the factors affecting the operation of processes, their most vulnerable points in the cyber world and the cyberattack methods most probably threatening the processes, it possesses the most relevant information for creating protective plans for potential treats. Vulnerability analysis against attack methods is a systematic tool for identifying and assessing risks related to process operation as well as for choosing the most suitable measures to enhance cyber security trust. The analysis provides a comprehensive overall picture of the needs to develop the processes.

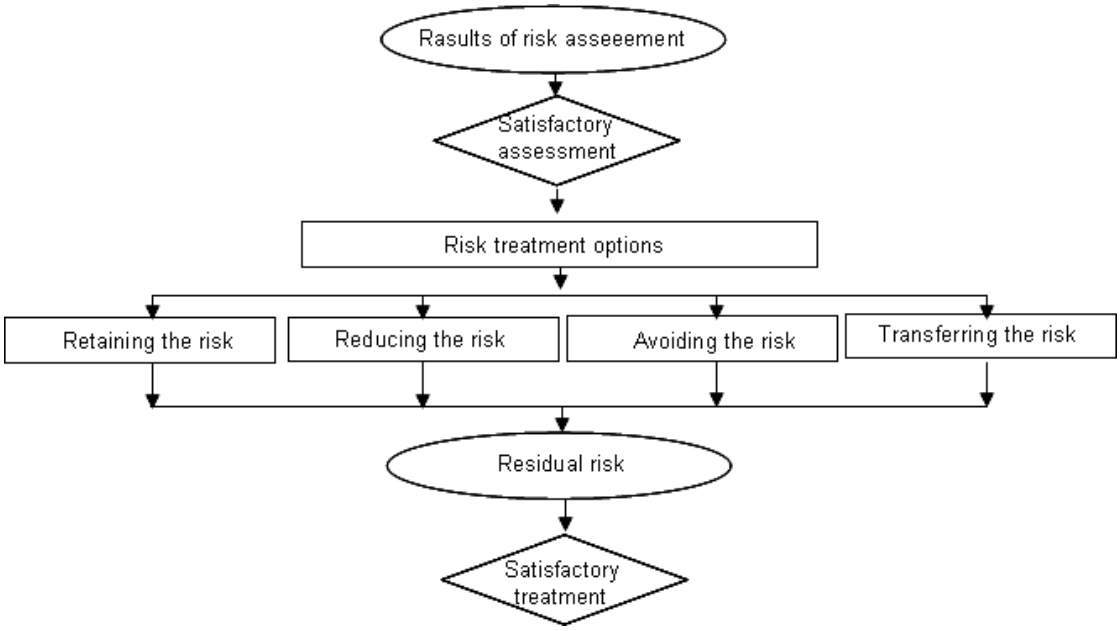


Figure 3: ISO27005: Risk treatment [6]

The risk management standard ISO27005 of the ISO27000 standard family includes the risk management process presented in Figure 3, which can be utilized in analysing the risks involved in the electricity production process.

Risks can be classified in a treatment process according to Figure 3. The aim should be to reduce or completely eliminate the most remarkable risks using different measures. Corporate leadership prioritizes the highest risks to the processes based on risk identification and chooses the measures that best suit risk management and development of proactive measures in the cyber environment. Less significant risks can be retained, aiming to manage them. Risk transfer in the cyber environment of an electricity company can be possible through its logistics network. This means that responsibility questions must be resolved using a clear internal operations model within the network.

4.3 The PDCA method as a tool for developing activities

An organization's policy demonstrates that its leadership is committed to implementing strategic measures. In the business world, general strategic measures are mainly targeted at promoting the business activities, which means that taking cyber security into account as part of the overall strategy supports the business development targets. Cyber security as part of company policy is a way of communicating to staff and interest groups on the necessity and significance of development projects. Operational goals are formed as processes derived from the policy, whereby risk analysis has been considered. In order to create the measures, the organization must have a systematic approach to developing its operations.

The ISO9000 Standard recommends the PDCA (Plan, Do, Check, Act) method for a systematic development of an organization's activities. The method is based on a cycle of four development phases. The first phase (Plan) comprises planning, during which the subject is analysed and alternative measures are created based on the analysis. In the realization (Do) phase, the chosen measures are put into practice. Thereafter the functionality, efficiency and appropriateness of the chosen measures are checked in practice (Check). At the last (Act) phase of the cycle, the chosen measures are improved, if necessary, and established as standard practice. After the cycle has been implemented once, one will return to the first phase and start a new cycle with improvement actions based on a new situation analysis. Development can thus proceed as an endless process, in which a new level of activities is achieved after each cycle. The method is based on the idea of continuous learning and continuous improvement of activities.

The measures during one round of the cycle usually require a lot of planning, so sufficient time should be reserved for them. It is important to select the measures in relation to the resources needed for their implementation. The maturity level of the organization's development activities affects the evaluation of the implemented measures. When developing cyber security, at an initial stage the aim can be to recognize the need for cyber security management and to define cyber security risks for business. Hereby, the PDCA cycle may comprise the administrative actions most necessary according to risk assessment, such as a coherent information security policy in production, practical guidelines for maintaining information security in production,

and potential preliminary system-specific cyber security checks. The targets for development must later be chosen according to risk prioritization.

The following is one possible process model for developing cyber security management with the PDCA method:

PLAN, planning phase

1. Choose the target for development based on risk assessment
   - schedule and goal
2. Create a picture of the current situation
   - earlier measures
   - disturbances in the branch resulting from special causes
3. Analyse the problems and define corrective actions
   - identify potential harms caused by the disturbances
   - choose the measures available to anticipate and manage the situation

DO, implementation phase

4. Implement the chosen measures
   - choose the actors responsible for implementation
   - organize information and training for staff

CHECK, checking phase

5. Check the impact of the measures
   - compare the results with the goals
   - return to phase 3 if the goals have not been achieved

ACT; regularize the measures

6. Regularize the chosen development measures
   - update necessary guidelines, technological solutions and services
   - continue staff training
7. Draw conclusions and make plans for the future
   - continue development according to new goals
   - update threat and risk analyses

In this section of the article, we have described one way of launching primary basic solutions related to cyber security management in an electricity company. These first steps provide a basis for later development activities and continuous improvement in a dynamic cyber environment.

5. Conclusion

The national power grid and its electricity production are part of a country's critical infrastructure – the operation of a modern society is based on a reliable electric power system. Ensuring the availability and reliability of processes in electricity companies in all environments is vital for the efficient functioning of critical infrastructure.

Therefore, the measures taken in electricity companies in order to manage and control the cyber security of processes are an essential component of the reliability of production.

The major cyber environment risks within the processes of an electricity company require that trust is enhanced and maintained at all levels of business activity. Comprehensive measures to increase cyber trust, together with the development of capabilities related to cyber activity, also improve a company's competitive edge.

The initial measures taken to develop cyber security management and trust in an electricity company can be summarized and prioritized as follows:

1. It is ensured that the company sees cyber security measures as strategic goals and that sufficient resources are allocated to the chosen measures.
2. Risk assessment is performed and the company's policy is updated to meet the requirements of cyber security.
3. The primary trust-enhancing development measures needed based on risk assessment are taken at the first development phase, using the PDCA method.
4. A continuous process is formed of the development actions by choosing the subjects of the next cycle, and the PDCA development cycle is repeated. This procedure will provide the organization with a culture of continuous learning and improvement. The organization's capabilities and competitive advantage are enhanced.
5. The impact of the measures is monitored as part of the company's audit and management procedures (e.g. as part of the ISO 9001 Standard procedures).

Investigations have revealed that the extensive power failure in Ukraine on 23 December 2015 was caused by a coordinated cyberattack by an external party to the control systems and data warehouses of three enterprises in charge of power distribution. One potential target of the attack is suspected to be the industrial automation system, which the hackers may have managed to enter via a remote access service. When preparing for cyberattacks to industrial automation systems and trying to improve their resistance, organizations are recommended, in the first place, to introduce the best practices of cyber security management. [9]

The power failure in Ukraine and other international experiences of disturbances in power grids highlight the crucial role of developing trust in electricity company operations as well as the importance of their organizational management in the cyber environment.

References

[1] Bowersox D., Closs D., Jessop D., Jones D., Logistical Management, New York, John Wiley & Sons, Ltd., 1986.

[2] Faber S. Flow Analytics for Cyber Situational Awareness. SEI Blog, 2015. https://insights.sei.cmu.edu/sei_blog/2015/12/flow-analytics-for-cyber-situational-awareness.html

[3] Fingrid Oyj, Voimajärjestelmän yleinen kuvaus. Retrieved on 6 August 2016 from http://www.fingrid.fi/fi/voimajarjestelma
[4] Finnish Energy. Retrieved on 25 October 2015 from http://energia.fi/energia-ja-ymparisto/sahkontuotanto
[5] Finnish Society of Automation. Teollisuusautomaation tietoturva. Verkottumisen riskit ja niiden hallinta. [online document], 2010. www.viestintavirasto.fi/attachments/tietoturva/TeollisuusautomaationTietoturva.pdf
[6] Finnish Standards Association SFS. SFS-käsikirja 327. Informaatioteknologia. Turvallisuus. Tietoturvallisuuden hallintajärjestelmät. SFS ry, Helsinki, 2012.
[7] Finnish Standards Association SFS. Johdanto laadunhallinnan ISO 9000 -standardeihin. [online document], 2016. www.sfsedu.fi/files/126/ISO_9000_kalvosarja_oppilaitoksille_2016.ppt
[8] Helsingin Sanomat (6 January 2016). Poikkeuksellinen kyberhyökkäys onnistui sammuttamaan ukrainalaisten sähköt. http://www.hs.fi/ulkomaat/a1452053903722
[9] ICS-CERT. Cyber-Attack Against Ukrainian Critical Infrastructure. Retrieved on 23 June 2016 from https://ics-cert.us-cert.gov/alerts/IR-ALERT-H-16-056-01
[10] Kananen I, National Emergency Supply Agency. Sähköjärjestelmä yhteiskunnan toimivuuden perustana. Seminar presentation on 2 December 2013. [online document] http://www.fingrid.fi/fi/asiakkaat/asiakasliitteet/Seminaarit/K%C3%A4ytt%C3%B6varmuusp%C3%A4iv%C3%A4/2013/K%C3%A4ytt%C3%B6varmuusp%C3%A4iv%C3%A4%20021213%20Kananen.pdf
[11] Knowles W., Prince D., Hutchison D., Ferdinand J., Disso P., Jones K. International journal of critical infrastructure protection 9. A survey of cyber security management in industrial control systems, 2015.
[12] Lehto M. Cyber Security: Analytics, Technology and Automation. Springer, 2015.
[13] Lewis T. Critical Infrastructure Protection in Homeland Security: Defending a Networked Nation. Second Edition, 2015.
[14] Lillrank P. Laatuajattelu. Laadun filosofia, tekniikka ja johtaminen tietoyhteiskunnassa. Otavan Kirjapaino Oy, Keuruu, 1998.
[15] Limnell J., Majewski K., Salminen M. Kyberturvallisuus, Docendo Oy, Jyväskylä, 2014.
[16] Secretariat of the Security Committee. Finland's Cyber Security Strategy. [online document], 2013. http://www.defmin.fi/files/2378/Finland_s_Cyber_Security_Strategy.pdf
[17] Stouffer K., Falco J., Scarfone K. NIST Special Publication 800-82. Guide to Industrial Control Systems (ICS) Security. Recommendations of the National Institute of Standards and Technology. U.S. Department of Commerce. [online document], 2011. http://csrc.nist.gov/publications/nistpubs/800-82/SP800-82-final.pdf
[18] World Economic Forum. Retrieved on 8 August 2016 from http://reports.weforum.org/global-risks-2015/part-1-global-risks-2015/technological-risks-back-to-the-future/#frame/20ad6

PRIVACY PRESERVING DATA MINING IN HIGH SECURITY PUBLIC AUTHORITY ENVIRONMENT

Klaus Zaerens

Finnish National Defence University

klaus.zaerens@iki.fi

Abstract

The awareness for capabilities of malicious actors such as cybercriminals, hactivists and foreign governments has increased during the last three years. The preparation phase of an intrusion or attack may have been ongoing for several years unobserved. Long term attack preparations are very difficult to detect and they need significant amount of computing power. Because high security environment used by the public authority is very expensive, interest arises on can lower cost environments utilized for intrusion detection when focusing on data privacy and separation of the operative environment. There has been discussion can intrusion detection analysis be done in less costly environment and separate from the operative environment by ensuring the privacy of data with current technology.

In this paper we examine the challenges in privacy preserving data mining techniques in high security public authority context. We describe current approaches for overcoming the challenges and find that none of the approaches solves the privacy and security requirements sufficiently in order to distribute secure analysis to public environment.

Keywords

Privacy, Data Mining, Cloud computing, Security, Public authority

1 Introduction

Computational systems have increasingly significant role in the infrastructure of our society. Computational systems are used in storing data, as a communications tools, as an aid for decision making and management systems. If we consider the public authorities such as security officials, today's computational systems include operational data, commands and directives, controls for infrastructure of the community as well as monitors of the vital elements of society and critical infrastructure. Systems are also used for communication and collaboration of different authorities. Systems with sensitive data are very attractive targets to any hostile acts or just alluring challenges to individual hackers. The more important specific system is for operations or the more sensitive data system contains, the more interesting target is formed. This draws attention from different actors with even malicious intentions such as hactivists, criminals and foreign governments.

The properties, features and requirements of the high security public authority environment called "VAHTI information security instructions for Finnish public authority environments" is defined by Finnish Government Information Security Management Board [1]. Finnish Ministry of Defence has published information security auditing criteria and instructions called KATAKRI for reviewing VAHTI aligned secure environments [2]. Information security requirements for high security environment are increased in relation the more sensitive or classified information system contains. In Figure 1 the classification for security levels is presented [2].

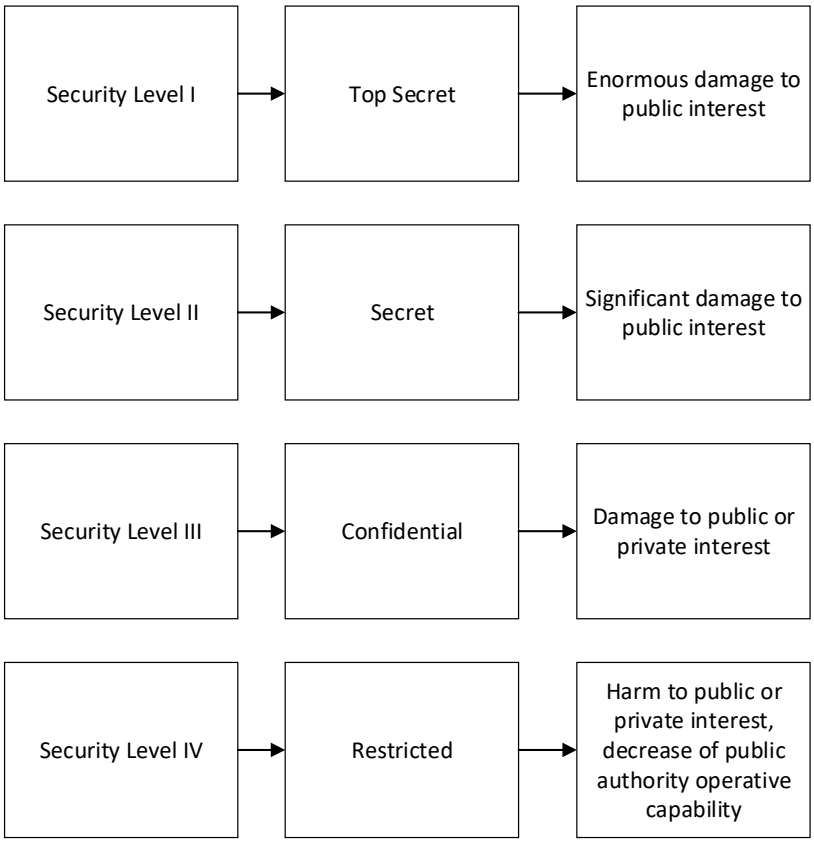


Figure 1: The classification of security levels.

The public authority high security environments are usually at least at Confidential level. Confidential level has quite demanding requirements for the environment such as all connections, data storages and technical equipment needs to be at least doubled, datacenters must operate without national electricity supply and datacenters must be protected against biochemical and radioactive attacks. These requirements ensure the availability and survivability of services during all kinds of catastrophic events. The technical excellence comes with a downside where the investment and operating costs for such an environment are vast. Therefore the environment should be designed and sized to the core operations execution and the idle time of the environment should be as small as possible.

Analysis of technically advanced malicious activity in the environment requires data from long period of time and from multiple data sources. The data analysis is performed by several known and previously unknown search patterns. This kind of analysis is computationally challenging and consumes lots of resources. Along with the

expenses of computation, heavy analysis should be separated from the operative system to ensure the availability and small latencies of critical real time systems.

It is argued that ensuring the privacy of the classified and sensitive data by the privacy preserving data mining technology we can share the computational challenge of the analysis with less secure environment. This would enable us saving the expenses of secure infrastructure and scale up the computational resources for sufficient level without disturbing the execution of operative systems. In this paper we have an overview on challenges considering privacy preservation data mining and approaches to overcome them.

The paper proceeds as follows. In Chapter 2, we examine the essence of privacy preserving data mining by defining its relevant terminology, characteristics and principles, as well as the scope of a public authority. In Chapter 3 we present the challenges identified in privacy preserving data mining and in Chapter 4, we focus our discussion to key directions overcoming the challenges. We will conclude in Chapter 5 with key findings and a description of the future of privacy in the public authority context.

2 Privacy preserving data mining

Improving privacy has become more essential issue to be considered since data in high security environments has become more openly and accessible. The possibilities and utilizations of new technology such as cloud computing have increased the amount and type of threat vectors in public authority environment. Such an environment contains sensitive and classified information on objects. The solution deployments based on concepts of pervasive and ubiquitous computing make sure that the environment contains also significant amount of information on users. Moreover modern mobile technology reveals real-time information of users such as location information. These issues make privacy preservation important, to secure the secret and classified information of the operations, events and infrastructure. Also there must be ensured the own safety and security of authority personnel.

In this paper we adopt the philosophical collection of definitions on privacy presented by Schoeman [3] and Walters [4]. This collection states the privacy as the right of the person to determine which personal information about himself/herself may be communicated to others, as the control over access to information about oneself and as limited access to a person and to all the features related to the person.

We emphasize that even privacy by definition is strictly bound to human, same techniques of privacy preserving can be used for securing the classified information on any physical object. Approach presented by Gavison [5] is more flexible in this sense. According to her, privacy consists of secrecy anonymity and solitude. Secrecy implies what information is known, anonymity refers what attention is paid to an individual and solitude connotes the physical access to an individual. If the term an individual is changed to an object or target, we can utilize the approach to any classified or sensitive information. We consider the privacy violation as the event compromising any of the previous definitions.

As more data can be collected and stored than ever previously in history, system enables knowledge based on source data to be more truthful, accurate and comprehensive. The increasing amount of data stored in systems and development of data mining

technology has enabled novel approaches to business intelligence [6]. The U.S. Government definition on data mining in "Data-Mining Reporting Act of 2003" [7] enlightens the utilization of data mining capabilities in public authority context. This definition is presented in Figure 2.

DATA-MINING

The term 'data-mining' means a query or search or other analysis of 1 or more electronic databases, where

a) at least 1 of the databases was obtained from or remains under the control of a non-Federal entity, or the information was acquired initially by another department or agency of the Federal Government for purposes other than intelligence or law enforcement;
b) the search does not use a specific individual's personal identifiers to acquire information concerning that individual; and
c) a department or agency of the Federal Government is conducting the query or search or other analysis to find a pattern indicating terrorist or other criminal activity.

Figure 2. Data-Mining Reporting Act of 2003 by the U.S. Government [7].

Discovering new dependencies and relations from source data which contain sensitive personal information can have problematic consequences [6]. Instead having just details from an individual, we can model the behavior of that individual.

Even if the individual cannot be identified straight from the result, but a convenient combination in small result set can isolate the information and so forth reveal the individual. There has been a lot of research to overcome this. For example Doyle have presented in their paper approaches to overcome this disclosing individual information [8].

In more general, privacy preserving data mining technology tries to solve the problem. It combines the research areas of data mining and privacy. It can be considered as a knowledge extraction from large amount of data while ensuring the confidentiality and sensitivity of the source data [9]. Solutions try to present a novel approach to summary data where results (e.g., a set of association rules or a classification model) are shown not to inherently disclose individual information [6]. The right to privacy and the need of knowledge discovery are in some extent competing goals. The research tries also to determine this trade-off [9]. Illustration on the position and focus of potential malicious attackers in privacy preserving data mining process is presented in Figure 3.

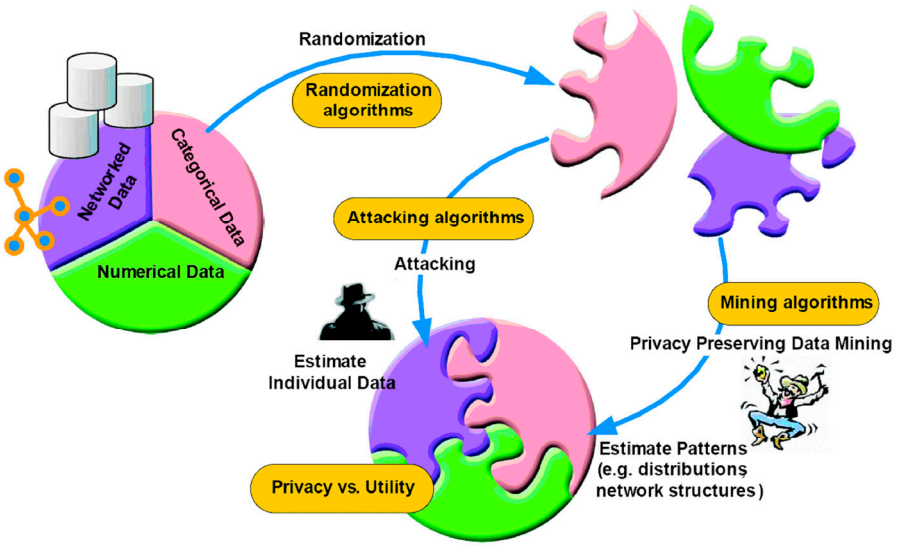


Figure 3: Example of privacy preserving framework presented by Wu et al. [10].

The privacy preserving data mining techniques can be classified according to the following features [11]:

- data distribution (centralized or distributed),
- the modification applied to the data (encryption, perturbation, generalization, and so on) in order to sanitize them,
- the data mining algorithm which the privacy preservation technique is designed for,
- the data type (single data items or complex data correlations) that needs to be protected from disclosure, and
- the approach adopted for preserving privacy (heuristic or cryptography-based approaches).

The evolution and research of these techniques are expected to continue strong since the both national and international legal requirements for protecting data are tightened, contractual obligations (liabilities) for enterprises require it, companies try to solve the tradeoff between efficiency with stakeholders through risks losing trade secrets to stakeholders and similar tradeoff between research potential in collaborating with competitors through revealing internal cost or operative information [6].

3 Privacy Preserving Data Mining Challenges in Public Authority Context

This chapter examines some challenges and obstacles presented in the area of privacy preserving data mining in more detail. In this chapter we narrow our observation to networking environment like the virtualized public authority environment [12]. The main privacy breaches in a network are identity disclosure, link disclosure and content disclosure [13]. In identity disclosure the identity of an individual associated with specific node is revealed. In link disclosure the relationship between two nodes are revealed. In content disclosure the data associated with the node is compromised.

To quantify and evaluate the breaches mentioned above properly is not an easy task. It is difficult to model the capability of a malicious attacker [13]. We cannot be fully aware or model the reconnaissance skills of an attacker. As a consequence we do not know how much information an attacker has before analyzing our extracted result on knowledge. Moreover it is challenging to quantify the value of lost information [13]. The preemptive measures should be in relation to the actual threat.

In networking environment it is noteworthy that the results of knowledge extraction can be correlated to each other [13]. This is different situation from the basic tabular result where each tuple is separate and individual. Therefore too much revealed information on one node can reveal also information on another related node. Moreover some nodes are more trustworthy than others [13] meaning that some source data can be more accurate and truthful than other and vice versa. The networked environment is dynamic and the behavior of the nodes can change according to situation.

Processes where multiple participants collaborate and share information to jointly achieve a goal, should be considered more content disclosure problem than identity or link disclosure problem [13]. In that case the behavior or the trustworthiness of the process participants or the task itself determines the severity of the breach.

4 Approaches Preserving Privacy in Data Mining

In this chapter we present the most relevant approaches to deploy the privacy preserving data mining. The main goals of a privacy preserving data mining algorithm are [13]:

It should have to prevent the discovery of sensible information.

It should be resistant to the various data mining techniques.

It should not compromise the access and the use of non-sensitive data.

It should not have an exponential computational complexity.

Criteria to evaluate privacy preserving data mining algorithm contain privacy level, hiding failure, data quality and complexity [9]. Privacy level states how well can hidden information be estimated. Hiding failure indicates the portion of sensitive information that is not hidden by the algorithm. Data quality refers to the data quality after application of algorithm and also to the quality of mining results. Complexity reflects the overall performance of the algorithm [9].

The key directions of privacy preserving algorithms or approaches are data publishing, changing the results, query auditing, cryptographic methods, solving high dimensionality, considering utility and process design.

Privacy-preserving data publishing contain different kinds of transformation methods like randomization, k-anonymity, l-diversity and how perturbed data can be used in conjunction with classical data mining methods such as association rule mining [14]. In some cases there is need for changing the results of data mining applications to preserve privacy. For example association rules can be suppressed and hidden [14]. Similar method to changing results is the query auditing. In query auditing the results of queries are either modified of restricted [14]. Cryptographic methods are useful

when data is distributed across multiple sites and the owners of the data need to work with a common function, process or task. Cryptographic protocols ensure the secure computation without compromising the sensitive content of the data. High dimensionality of the real data sets makes the privacy-preservation extremely difficult challenge from computational and effectiveness point of view. Meyerson and Williams showed that optimal k-anonymization is NP-hard problem [15]. This area tries to find theoretical approach to deal with the high dimensionality. Enhancing the data privacy there is tradeoff for utility of the data (See illustration in Figure 4). Because of the computational challenges, the amount of anonymized attributes should be small as possible. However the amount of attributes used for knowledge extraction should be as large as possible [16].

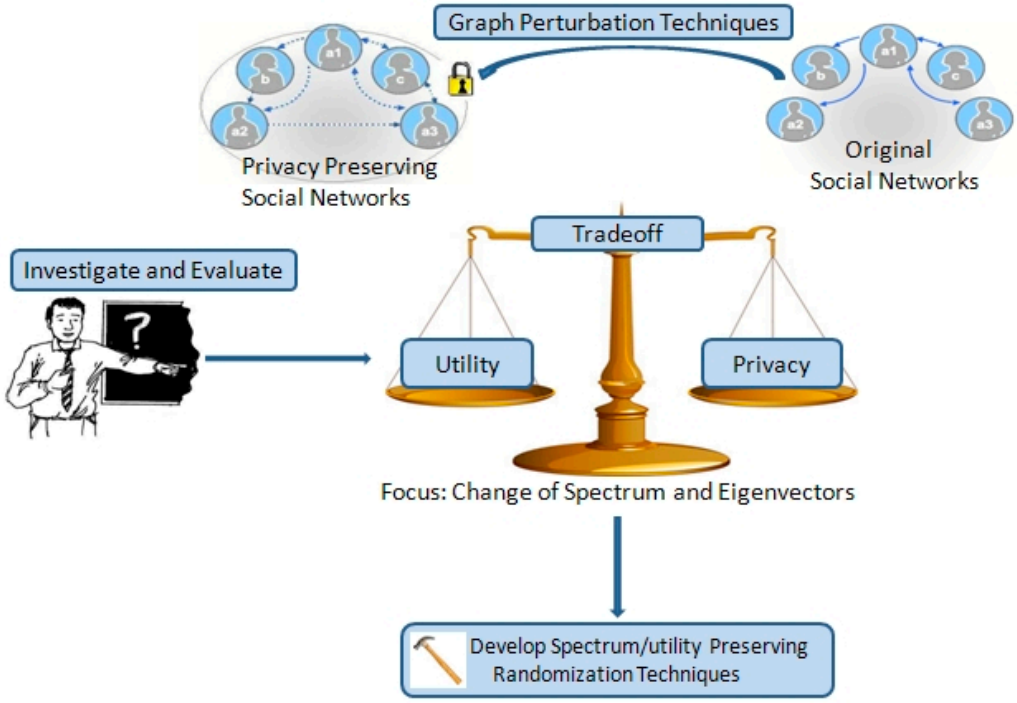


Figure 4: Illustration on tradeoff on utility an privacy presented by Wu et al. [17].

Privacy process design principle is well defined by the Article 29 Working Party (WP29) and the Working Party on Police and Justice (WPPJ) [18]. Working groups pointed out aspects and objectives that should be considered while designing the privacy preserving system. The considerations are presented in Table 1.

Table 1: Privacy by design [18].

Aspect / Objective	Description of principle in data processing system
Data minimization	The sensitive data should be avoided or the amount should be minimized in collecting, processing or using tasks.
Controllability	Data subjects should have effective means of control their personal data.
Transparency	Developers and operators should sufficiently inform about the means of operations.
User-friendliness	Privacy-related functions and facilities should be user friendly.
Data confidentiality	Only authorized entities have access to personal data.
Data quality	Data quality should be ensured by technical means.
Limitation of data use	Data and processes used in different tasks or purposes can be separated and isolated from each other in a secure way.

In public authority context we need to ensure privacy of classified information related to users, tasks, capabilities and data contents which can refer to humans or physical elements. Moreover the dynamic nature of environment and interaction with participants increase the challenge with privacy preservation. Obviously none of the key directions is sufficient by themselves to solve every aspect of the challenges presented.

For ensuring the privacy in environment presented cryptographic methods are the most effective and reliable. However, ciphering and deciphering cause latency to computation. That is not tolerable when dealing with operative system which has real-time processing requirements. In addition the whole environment can be considered secure, so the enhancement impact is only mostly to security. The high dimensionality of the data makes sufficient data publishing or changing only the mined results to a never-ending story. Query auditing techniques can be useful if the context can be limited. We suggest that the best results for privacy preserving can be achieved by considering utility aspects together with process design. We admit that the full privacy considering every aspect cannot be achieved with this approach, but the best possible without compromising the system operation. On design phase we emphasize the evaluation of trustworthiness of the participating partners, since that can be identified as a threat to the source data and privacy preservation of mined result.

In this paper there has been a lot of discussion on privacy breaches and threats to data mining. However we should remember that the data mining itself is not the risk, but the infrastructure that supports producing it. The more complete and accurate data is, the better results we gain from data mining. The consequence is that the better data results cause problems in privacy regardless of the intended use. For example it is easier to hack into the data storage built for the data mining purposes instead of hacking into each original source database [6]. As a consequence the data should not

be revealed from the secure environment to public environment even because of cost effective computation.

5 Conclusion

In this paper, we examined the privacy preserving data mining paradigm within the high security public authority context. Public authority environment is very challenging environment. There are several different types of sensitive and classified information such as user information, user capability, and physical as well as human object information. We analyzed the possibility to distribute intrusion detection analysis outside the high security environment.

We identified the primary obstacles to adopting privacy preserving data mining in public authority environment and examined possible solutions to overcoming them. We stated that the current technology on privacy preserving data mining is not sufficient in order to prevent revealing classified information if intrusion analysis is done in open and public environment.

After all we find privacy preservation data mining to be an emerging research area especially in networked public authority operations. Implementations for challenges and original idea presented in this paper need more research and work to be possible in practice.

References

[1] VAHTI 2/2013 Instructions, Finnish Ministry of Finance, https://www.vahtio-hje.fi/web/guest/home , Accessed 17-Apr-2016.

[2] National Information Security Auditing Criteria KATAKRI, Finnish Ministry of Defence. http://www.defmin.fi/files/1870/KATAKRI_versio_II.pdf , Accessed 17-Apr-2016.

[3] Schoeman, F.D.: Philosophical Dimensions of Privacy: An Anthology. Cambridge University Press. (1984)

[4] Walters, G.J.: Human Rights in an Information Age: A Philosophical Analysis, chap. 5. University of Toronto Press. (2001)

[5] GAVISON, R.: Privacy and the limits of the law. The Yale Law Journal 89, 3 (January 1980), pp 421–471.

[6] Vaidya, J., Clifton, C. W., Zhu, Y. M. Privacy preserving data mining. (Vol. 19). Springer Science & Business Media. (2006)

[7] Feingold, M., Jeffords, M., Leahy, M. Data-mining reporting act of 2003. U.S. Senate Bill, July 31 2003.

[8] Doyle, P., Lane, J., Theeuwes, J., Zayatz, L. Confidentiality, Disclosure and Data Access: Theory and Practical Applications for Statistical Agencies. Elsevier, Amsterdam, Holland, 2001.

[9] Bertino, E., Lin, D., & Jiang, W.. A survey of quantification of privacy preserving data mining algorithms. In Privacy-preserving data mining . Springer US. pp. 183-205. (2008)

[10] Wu et al. Towards Privacy and Confidentiality Preserving Databases http://webpages.uncc.edu/xwu/career/. Accessed 22-Apr-2016

[11] Verykios, V.S., Bertino, E., Nai Fovino, I., Parasiliti, L., Saygin, Y.,Theodoridis, Y.: State-of-the-art in privacy preserving data mining. SIGMOD Record 33(1), pp. 50–57 (2004)

[12] Zaerens, K., Mannonen, J. Concept for the Construction of High Security Environment in Public Authority Cloud, Lecture Notes in Electrical Engineering, Springer-Verlag, September 6, 2012.

[13] Liu, K., Das, K., Grandison, T., Kargupta, H. Privacy-Preserving Data Analysis on Graphs and Social Networks. (2008)

[14] Aggarwal, C., Yu, P. An Introduction to Privacy-Preserving Data Mining. In Privacy-preserving data mining. pp. 1-9. Springer US. (2008)

[15] Meyerson A., Williams R. On the complexity of optimal k-anonymity. ACM PODS Conference, 2004.

[16] Venkatasubramanian, S. Measures of Anonymity. Privacy-Preserving Data Mining Models and Algorithms. In Privacy-preserving data mining. pp. 81-104. Springer US. (2008)

[17] Wu et al. PSNet: Privacy and Spectral Analysis of Social Networks http://webpages.uncc.edu/xwu/PSNet/. Accessed 22-Apr-2016.

[18] Article 29 Data Protection Working Party, Working Party on Police and Justice.. The Future of Privacy. (2009) http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2009/wp168_en.pdf . Accessed 18-Apr-2016.

National Defence University
PO box 7
FI-00861 Helsinki
Finland

www.mpkk.fi

ISBN 978-951-25-3380-0 (pbk.)
ISBN 978-951-25-3381-7 (pdf)
ISSN 2737-0615 (online)

SOTATAIDON YTIMESSÄ

Puolustusvoimat
The Finnish Defence Forces