Information Management and Sharing for National Cyber Situational Awareness

Florian Skopik · Thomas Bleier · Roman Fiedler

AIT Austrian Institute of Technology GmbH
2444 Seibersdorf, Austria – http://www.ait.ac.at/it-security/
{florian.skopik | thomas.bleier | roman.fiedler}@ait.ac.at

Abstract

ICT has been integrated massively in business processes in recent years, thus producing an enormous dependency on these technologies. The potential impact of these dependencies (for example if the IT systems are lacking appropriate security levels) are remarkable – the malfunction or total loss of public energy grids, the banking system, supply chains or public administration can cause enormous economic damage and massively affect entire nations. This paper describes the concepts and development of a system to improve the national situational awareness in complex ICT infrastructures which is being carried out in the Austrian national research project CAIS (Cyber Attack Information System). The core of this system consists of two methods and derived prototypical software implementations: a modelling and simulation tool for analysing the structure of large ICT systems in terms of their security and resilience against cyber attacks, and an analysis and evaluation tool for the investigation of the current threat situation in networks. This paper particularly focuses on distributed anomaly detection and evaluation, and demonstrates how these tools can be applied in course of a sophisticated methodology in order to build a national information system that allows efficient information sharing and collaborative mitigation of threats in the cyberspace.

1 Introduction

The Internet has grown to a massive economic sphere of activity – not only for the new economy, where Internet-based businesses have grown from startups to multinational and billon-dollar companies faster than any businesses before, but also for the "dark side" of entrepreneurship where moral behaviour is not the first on the list of a company's objectives. In the beginning of IT attacking other computers was mostly done for self-expression or competition between hackers – nowadays this has become big business. The reason for this is that there is a market for personal data such as credit card information, bank account credentials, e-mail addresses or even software vulnerabilities. Simple goods like a credit card number (including all details needed to pay with it) or the credentials to access another one's bank account are sold for a few dollars up to a few hundred dollars [Sym10], but information about a previously unknown vulnerability in a software application are sold for thousands of dollars [Rad09] [Kin07]. There is no clear picture of the volume of these markets, but damage is huge – a recent Europol report for example talks about around € 750 billion of losses every year [Eur11].

The technology behind these cybercrime black markets is what we have to deal with every day in information security. Spam mails are used to advertise goods and spread phishing links or malware, viruses spread their infection and carry dangerous payloads, drive-by downloads are used

N. Pohlmann, H. Reimer, W. Schneider (Editors): Securing Electronic Business Processes, Vieweg (2012), 1-11

to infect victims when they are accessing an unsuspicious website, rootkits hide the existence of other malware on a system to be able to act undetected as long as possible, and botnets are used to control a large number of victim's systems for malicious purposes. Antivirus software is helpful, but for itself cannot provide an adequate level of security since a long time ago. Thus a complex ecosystem of detective, preventive and corrective controls (antivirus and filter software, patching and updating, security functionality in operating systems, incident response organizations, etc.) is necessary to provide an environment with a satisfactory security level.

Despite these problems IT has become a ubiquitous part of our life. This does not only refer to the "always connected" style of living nowadays – even if IT is not used obviously many activities and processes are underpinned with information technology. The basic infrastructures of our society like energy, transportation or the financial system have leveraged efficiency and cost benefits by supporting processes with ICT systems and therefore adopted those systems in a rapid manner. On the other hand this makes these systems susceptible to many of the security problems illustrated before. A deliberate or unintentional disruption as a result of technical or human failure or due to natural causes could lead to social destabilization. As a consequence, IT security measures are rapidly being adopted in almost all areas utilizing ICT. These efforts are hampered by the complexity and networking of modern ICT facilities and the rise of mobile data traffic and cloud computing pose further challenges to securing today's infrastructures. A number of recent high-profile incidents have shown the vulnerability of critical infrastructures, which depend on ICT, to sophisticated cyber attacks. For example, the Stuxnet virus [Fal10], explicitly designed to attack industrial process automation facilities, demonstrated the vulnerability of SCADA systems in critical infrastructures. Other examples of targeted attacks on nation-wide ICT infrastructures with enormous effects were the cyber attacks against Estonia in 2007 [Ott07] and Georgia in 2009 [Tik08].

All these events have shown the emerging need of cooperation in the defence of cyber incidents, as no involved party can handle those threats alone. The ENISA carried out a study on the feasibility of an EU-wide information sharing and alert system [Eni07], which will eventually result in a European system for information sharing called EISAS (European Information Sharing and Alert System for citizens and SMEs) [Eni11a]. In the context of the study the first step recommended is the establishment of thorough national information sharing and situational awareness capabilities that can then be aggregated on a European level.

This paper describes the concepts and rationale behind the project CAIS which research focus is on the technological foundations for a national "Cyber Attack Information System" in Austria. In particular, the project develops a novel anomaly detection approach that enables organizations to collaboratively discover and defend against attacks. Section 2 provides a rough overview about the on-going CAIS project, its objectives, and involved parties. Section 3 gives an overview about the CAIS architecture and motivates our work on fundamental building blocks. Section 4 outlines a collaborative and agile anomaly detection approach which is one of the highlights developed within the CAIS project. Section 5 discusses related work and finally Section 6 concludes the paper.

2 The CAIS Project

The ultimate goal of the CAIS project is to provide technologies for strengthening the resilience of today's interdependent networked services, and increasing their overall availability and trustworthiness. In particular, the research focus of the project is on the following challenges:

- Study of future cyber risks and emerging threats, particularly having the changing political and economic landscape in mind. Here, the well proven Delphi method – a systematic, interactive forecasting technique – is applied together with an extended group of subject matter experts.
- Evaluation of novel anomaly detection techniques by composing available network tools for log file management with new pattern mining approaches inspired by models from the domain of bio informatics. Here, high performance and scalability is of paramount importance.
- Creation of highly modular infrastructure models on multiple layers, spanning hardware-centric physical aspects, over data flow and service deployment perspectives, to abstract inter-organizational dependencies.
- Innovative tools for attack simulations, using aforementioned infrastructure models and applying game-theoretic approaches as well as agent-based simulations [Mac10] in order to forecast the effects of attacks on interconnected infrastructures, and the impact of countermeasures on various levels and from multiple viewpoints.
- Investigate the deployment and instantiation of a CAIS that connects single organizations, links and coordinates isolated anomaly detection efforts, and facilitates information sharing and mutual aid between organizations.

In order to reach these ambitious goals and finally ensure the wide applicability of developed tools, major stakeholders of Austria's security domain are involved. First, research institutions, such as the Austrian Institute of Technology and the University of Applied Sciences St. Poelten contribute their scientific expertise regarding anomaly detection techniques, and infrastructure modeling and simulation. Furthermore, the OIIP Austrian Institute for International Affairs studies cyber threats and risks to national critical infrastructures caused by cyber crime. The major telecommunication service providers T-Mobile Austria and T-Systems Austria, as well as the national Austrian Computer Emergency Response Team (CERT) ensure a sound implementation on a technical layer and practical applicability and validation. The Austrian Federal Chancellery (BKA), Federal Ministry of Interior (BMI) and the Federal Ministry of Defense (BMLVS) bring in requirements from a national security perspective.

3 A Cyber Attack Information System Architecture

This section describes the overall architecture of the system and motivates our overall work, including models for establishing situational awareness, effective attack prevention, and reporting and information sharing between the CAIS stakeholders.

3.1 Architectural Overview

Our concept of the CAIS [Sko12], as depicted by Fig. 1, involves two main types of stakeholders: the National Cyber Centre, responsible for the coordination of activities on a national level, and the individual organizations participating in the system. Single organizations typically run critical infrastructures and enable a vital information flow to the national coordination for both increasing the efficiency and effectiveness of national cyber defence activities but also for increasing their own abilities to defend against cyber attacks. Regarding the latter aspect, the national cyber centre provides, besides Computer Emergency Response Teams (CERTs) and other international initiatives, valuable feedback to single organizations, e.g., on effective countermeasures deployed by other organizations against large-scale attacks. Numerous challenges need to be addressed in order to set up such a CAIS. Some of the most urgent needs are discussed in the following parts of this work.

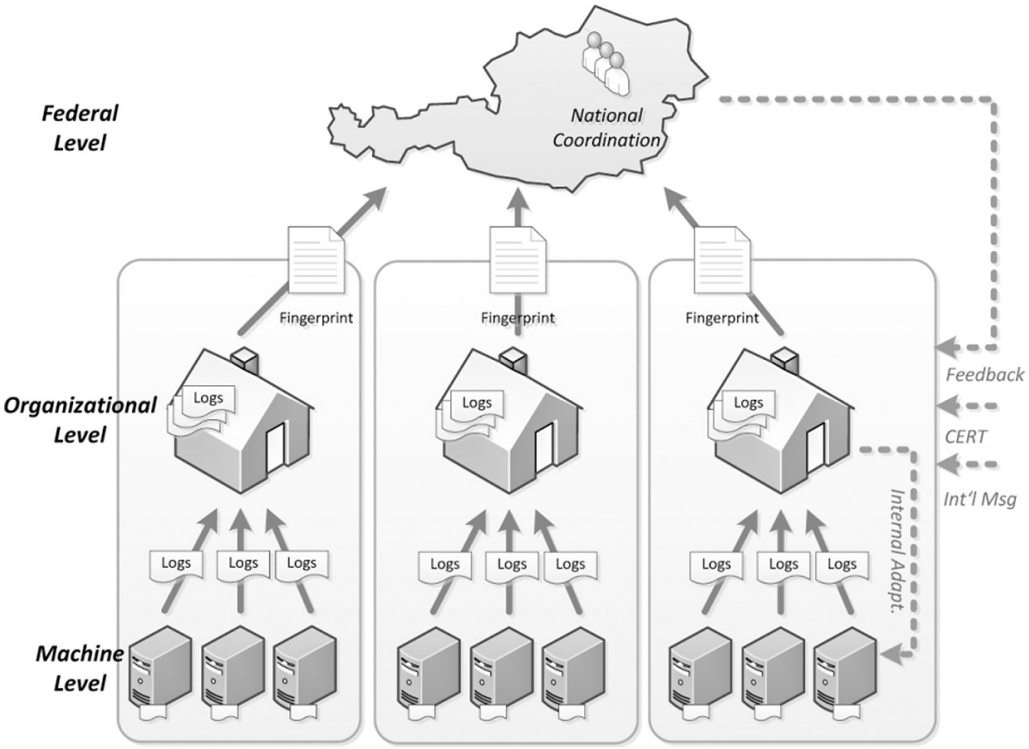

Fig. 1: Collaboration between stakeholders of the CAIS

3.2 Establishing Situational Awareness

The first precondition in taking any actions against malicious cyber activities is having situational awareness (SA) within the cyber space. Achieving this is far more complicated as it might seem, because of the multitude of actors and influences in this area. From a scientific point of view, a number of models of SA exist, e.g. [End95] or [Sar91]. In our work we roughly follow the model proposed by Endsley [End95] because of its concept of separation into different levels which correlates to the situation we find in cyberspace. Endsley describes three increasing levels of aware-

ness: perception, comprehension, and projection. As one advances through these levels, decision making capabilities are improved.

The EU FP7 project ResumeNet proposed a mapping of information sources and mechanisms to the first two levels of SA [Smi11] for identifying challenges, e.g., attacks, to computer networks. There are two key sources of information for situational perception: multilevel net- work measurement information and context information, which is external to the network under scrutiny, such as news items about an ongoing situation. These two forms of information – network and context – are used as inputs to various techniques that are used to build situational comprehension. There are three proposed main approaches to comprehension: (1) detection of the presence of a challenge, e.g., provided by anomaly and intrusion detection systems; (2) identification of the characteristics of the challenge, e.g., provided by classification [Ngu08] and data fusion [Tad06] techniques; and (3) the impact an attack is having on the network and associated services. In this paper we especially describe techniques for the first two approaches, which are further prerequisites for the third one, realized with advanced simulation techniques.

3.3 Effective Attack Protection for the 21st Century

Today, mainly two forms of protective mechanisms are used. First, proactive means aim at avoiding breaches and infections as far as possible, i.e., by deploying firewalls and filters, implementing efficient software update policies of corporate ICT systems, analysis of early warning information (e.g., provided by CERTs), and performing end user trainings regarding the secure handling of computer systems. Second, reactive mechanisms, such as anti-virus programs and intrusion detection systems, try to detect and contain already ranging infections. Both forms have been proven useful, however, since attacks become more and more sophisticated, traditional approaches sometimes fail at protecting ICT systems. For instance, signature-based anti-virus systems cannot properly handle zero-day exploits and their heuristics can be by-passed with customized malware. Firewalls and IDS/IPS software are knocked out through obfuscation methods – in short, many threats are not preventable. Thus, these days we observe a major paradigm shift from prevention and remediation-focused approaches to response and containment strategies. This shift also requires organizations to move from traditional policy-based security approaches towards intelligent approaches, incorporating identification of anomalies, analysis and reasoning, and in-time response strategies. As recently pointed out [Emc12] basic properties of such mechanisms are:

- **Risk-based:** Determine the most important assets to be secured, since an organization cannot cost-efficiently secure all assets with a maximum strength.
- **Contextual:** Collect huge amounts of data and use analytics to identify relevant data sources for anomaly detection. A 'context space' is created by aggregating and correlating a wide variety of events, even if an attacker partially deleted his traces.
- **Agile:** Enable near real-time responses to minimize the exploitable attack window and to keep (financial) losses to a minimum.

Although attacks using customized tools are unique in each case, their manifestation and impact across the network is often very similar, for instance calling home functionalities might be visible in DNS logs, database accesses in SQL query logs, and probing causes events in Firewall- or IDS logs. Thus, in this paper, we introduce a reactive approach that discovers these manifestations and fuzzily determines deviations from a healthy state. Here we do not investigate widely used net

flow analysis on a network packet layer. Due to the increasing interconnections of network devices, the complexity of holistic monitoring on this layer is growing at an exponential rate. Thus, we rather deploy detection mechanisms on a much higher level of the network stack, in particular on the application layer, were we carefully select relevant service logs that promise to be most expressive in terms of anomaly detection. Using these log files, we are able to classify and correlate events, and thus, detect usage patterns *across different ICT systems* within an organization.

3.4 Event Correlation and Information Sharing

One important feature of the proposed CAIS is a facility to correlate information of detected anomalies and thus, learn about potential on-going attacks from different sources. Furthermore, it is essential to derive early warning information about the state and progression of those attacks. On an organizational level, individual IT systems collect information about their current operational status in log files. These log files are typically collected within the organization and evaluated with different levels of sophistication. In its simplest form, a simple matching of specific log entries leads to alerting operations personnel in case of any exceptional events. More sophisticated solutions include automatic detection and correlation of different events and log entries with complex algorithmic processing, as it is available in commercial log management and SIEM (Security information and event management) solutions. Nevertheless this information is typically handled within the organization and only on a case-by-case basis exchanged and analysed between organizations.

In the proposed system however, information from these internal organizational systems gets processed according to a set of rules to achieve anonymization and hide any critical information that is not allowed to be transferred outside the organization. This "fingerprint" of the actual situation (cf. Fig. 1) is then transferred to the national cyber centre on a federal level. This allows correlation and analysis of attacks on a national scale, and finally to comprehensively establish situational awareness. Ultimately, the national cyber centre is able to either reply with sophisticated recommendations to organizations under attack on efficient mitigation strategies, or to facilitate the set-up of direct collaboration links between organizations to provide mutual aid.

4 Agile and Collaborative Anomaly Detection

4.1 Requirements and Approach Outline

In contrast to many common net flow analysis approaches, our anomaly detection mechanism relies on log file analysis. Here, we do not only utilize one source, but a multitude of logging sources that are distributed across an organization. Additionally to attacks to single machines, this way, we aim at discovering distributed and coordinated attacks, especially if they manifest in different log sources, such as DNS lookups, firewall logs, and application server events.

Roughly, our defence approach consists of the following steps:

- identification of service logs of critical assets
- development of detailed monitoring techniques for these assets (e.g., services)
- reporting of monitoring results to a national cyber centre

- discovery of attacks by applying foundational anomaly detection mechanisms spanning the whole nation
- periodic adaptation of employed mechanisms to respond to new threats and to control risks

Focusing on the actual implementation, we identified numerous requirements within single organizations that need to be covered in order to allow an effective system operation:

Log File Management: (1) efficient aggregation of logging data with various source formats; (2) time-based and source-based event aggregation; (3) timestamp correction in order to maintain a consistent time base; (4) configurable compression in order to deal with large-scale data sets.

Data Analysis: (5) adaptive event model with dynamic event classification (no predefined events, but application of learning models); (6) event clustering (correlation, co-occurrence) and grading; (7) rule-based anomaly detection with a mixture of static user-defined rules and dynamic rule sets gathered through machine learning techniques.

Reasoning and Reporting: (8) human-assisted reasoning and prioritization of detected anomalies; (9) privacy-preserving compression and abstraction techniques for secure and efficient reporting to the national cyber centre.

On the national level, the following requirements must be addressed:

Data Aggregation: (10) massive data collection from single organizations, (11) management of a fingerprint database, (12) data fusion

Simulation and Evaluation: (13) proper infrastructure models for simulations, (14) processing of fingerprints to feed the simulation, (15) powerful agent-based simulation matching reality.

Support and Advice: (16) decision making on a national level how to deal with attacks, (17) rapid but coordinated notification of organizations in case of attacks about efficient counter measures, (18) providing actual support in mitigating effects, e.g., through enabling resource sharing between organizations.

4.2 Process on the Organizational Level

The features of the CAIS anomaly detection approach on the organizational level can be grouped in three basic categories, according to the fundamental challenges described before:

1. **Log File Management and Refactoring** is about collecting and harmonizing log files from various sources following numerous protocols and formats.
2. **Data Analysis** is about discovering events that differ from every-day situations, i.e., anomalies, with minimal human intervention and configuration effort.
3. **Reporting and Configuration** is about an administration interface that allows tuning the whole system on the one side and delivering reports about significant events to the national cyber centre on the other side.

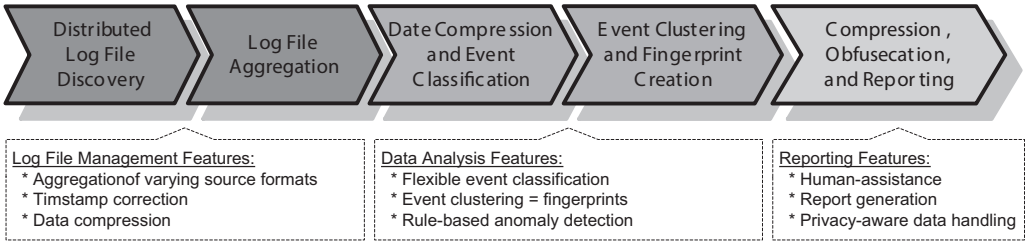

Fig. 2: Anomaly detection process within an organization in CAIS

The implemented system follows the process depicted in Fig. 2. Each single organizations runs through the given five steps and thus collects valuable information to assess the health status of its own ICT infrastructure, i.e., establishes situational awareness on an organizational level. In short, significant events manifested in log files, such as login attempts, service operation calls, resource allocations etc., are extracted and counted. Certain events are clustered, since similar events might manifest differently when coming from systems of different version and deployment models. Finally, a fingerprint is compiled (and optionally compared to fingerprints of previous time intervals) which is – roughly spoken – an overview about occurred events in a certain time span (e.g., an hour or a day). This fingerprint is then interpreted and compared with historic results to capture changes in the infrastructure and detect anomalies already on an organizational level. Additionally, this fingerprint is then reported to the national cyber centre for further analysis.

Notice, since the system neither transmits raw logging data nor data including personally identifiable information, our approach also accounts for proper protection of the privacy of the system users.

4.3 Process in the National Cyber Centre

On the national level, we implement the following features:

1. **Data Aggregation** deals with the collection and management of fingerprints from single organizations. Scalable systems need to be deployed that can manage the expected large amounts of data.
2. **Simulation and Evaluation** is about the assessment of the current situation from a national perspective, thus, from a higher level than any single organization could possibly match. Through correlation and comparison of fingerprints on a national level, distributed attacks can be discovered, for instance, when numerous organizations running similar infrastructures, are facing the same problems caused through the application of a zero day exploit.
3. **Support and Advice:** Without collaboration, every organization would need to deal with such an aforementioned coordinated attack on its own. However, with the national cyber centre in place, attacks can be discovered faster and mitigation actions can be effectively coordinated, e.g., experiences shared.

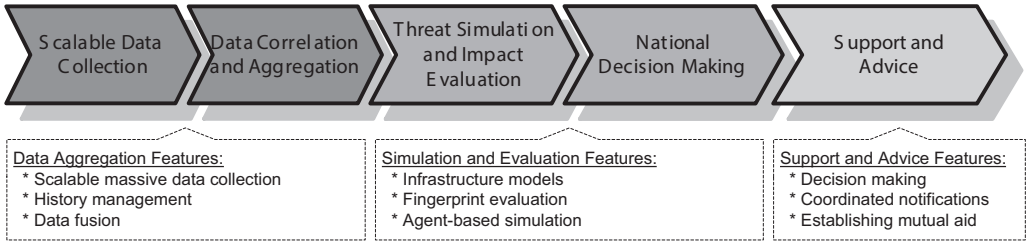

Fig. 3: Anomaly detection process on a national level in CAIS

In detail, as depicted in Fig. 3, the cyber centre collects and compares fingerprints of different organizations, possibly operating in the same domain. This way, the centre is able to compare the situations across organizational boundaries and thus to identify distributed attacks that simultaneously target systems of different companies, e.g., major telecommunication providers or large banks of a certain country. In this step, the second tool that is developed in context of the CAIS project comes into play. Because having a characterization of the organizational situations, we are able to predict the consequence of attacks and forecast – at least to a certain degree – future problems and weaknesses of the national ICT assets. After national decision making, e.g., through governmental authorities, optional support and advice is initiated to help threatened organizations if required.

5 Related Work

As ICT systems are being applied in a greater number of critical areas, cyber attacks are becoming more frequent and have an increasing impact. Situational awareness plays an important role in the defence and survival of ICT infrastructures against a cyber attack. Attack detection relies on cyber sensors, such as intrusion detection systems (IDS), log file sensors, anti-virus systems, malware detectors, and firewalls [Jaj09]. Many of the sensor techniques used today are based on sophisticated anomaly detection techniques, i.e., finding non-conforming patterns in data [Cha09]. The results from various research fields, such as data mining, statistical analysis, machine learning, as well as information theory are applied to anomaly detection. Since many of the attack detection tasks are performed at a local level, within a single organization, such as an Internet Service Providers (ISP), cross-domain security information sharing is a crucial step to correctly understanding the situation for national cyber defence. However, in practice, security information sharing is usually accomplished via ad-hoc and informal relationships [Dhs09]. Often, national Computer Emergency Response Teams (CERTs) assume the role of national contact points for coordinating and aggregating security incidence reports via communication channels such as email, instant messaging, file exchange/storage, VoIP, IRC and the Web [Eni11b]. Other means exist for information sharing. Internet forums such as the Internet Storm Center from SANS (http://isc.sans.org/) collect and provide data about malicious activities on the Internet. Commercial service providers, such as Arbor Networks (http://www.arbornetworks.com/), offer network-wide threat information updates and analysis services.

6 Conclusion and Future Work

Because of the increasingly sophisticated and distributed nature of cyber attacks, e.g., that use botnets as a platform, and our dependence on ICT-coupled critical infrastructures, a coordinated multi-domain approach to cyber incident response is required. This paper has motivated the realization of a national incident response cycle and implementation of a cyber attack information system (CAIS). The main goal of such a CAIS is to strengthen the resilience and trustworthiness of today's national ICT infrastructure through collaborative cyber defence. Our design aims at linking existing initiatives, maintaining organizational responsibility, and activating inter-organizational collaboration on a national level.

The cyber attack information system introduced in this paper is the first step towards establishing national cyber situational awareness and enabling mutual aid of independent organizations. For that purpose, information sharing about intra-organizational situations, i.e., status of technical infrastructures, is key to effectively cope with today's attacks on a large scale. Thus, one future aim of the CAIS project is to study incentive models which eventually motivate organizations to share information even regarding their critical assets. We consider such models, where organizations provide internal data voluntarily, more promising than legal enforcement models. However, in order to reach this ambitious goal, we need to make sure that CAIS stakeholders establish trust in the CAIS platform in terms of data privacy and security. Thus, participating in the CAIS must provide a clear benefit for companies whilst keeping the risk for data exploitation low. Advanced security technologies and rigorous processes need to be established to ensure the implementation of an efficient cyber attack information systems in Austria.

7 Acknowledgments

This work was partly funded by the Austrian security-research programme KIRAS (operated by the FFG) and by the Austrian Ministry for Transport, Innovation and Technology (BMVIT).

References

[Cha09] Chandola, V., Banerjee, A., Kumar, V.: Anomaly detection: A survey. ACM Comput. Surv. 41(3), 2009

[Dhs09] U.S. Homeland Security Cyber Security R&D Center: A roadmap for cybersecurity research, November 2009

[End95] Endsley, M.: Toward a theory of situation awareness in dynamic systems. In Human Factors 37(1), 32–64, 1995

[Emc12] EMC Press Release. Rsa chief rallies industry to improve trust in the digital world, after year filled with cyber attacks. http://www.emc.com/about/news/press/2012/20120228-03.htm, 2012.

[Eni07] ENISA: EISAS – European Information Sharing and Alert System - A Feasibiliy Study, 2077

[Eni11a] ENISA: EISAS – European Information Sharing and Alert System for citizens and SMEs – A Roadmap for further development and deployment, February 2011

[Eni11b] ENISA: Practical guide/roadmap for a suitable channel for secure communication: secure communicatio with the CERTs & other statkeholders, December 2011

[Eur11] Europol, Threat Assessment – Internet Facilitated Organised Crime iOCTA, 2011

[Fal10] Falliere, N., Murchu, L.O., Chien, E.: W32.Stuxnet Dossier. Tech. rep., Symantec Security Response, Oct 2010

[Jaj09] Jajodia, S., Liu, P., Swarup, V., Wang, C.: Cyber Situational Awareness: Issues and Research. Springer Publishing Company, Incorporated, 1st edn., 2009

[Kin07] David McKinney: Vulnerability Bazaar. In: IEEE Security & Privacy, 2007

[Mac10] Macal, C.M., North, M.J.: Tutorial on agent-based modelling and simulation. Journal of Simulation 4, 151–162, 2010

[Ngu08] Nguyen, T.T.T., Armitage, G.J.: A survey of techniques for internet traffic classification using machine learning. In: IEEE Communications Surveys and Tutorials, 10(1–4), 56–76, 2008

[Ott07] Ottis, R.: Analysis of the 2007 cyber attacks against estonia from the information warfare perspective. In: Proceedings of the 7th European Conference on Information Warfare. p. 163. Academic Conferences Limited, April 2008

[Rad09] J. Radianti, E. Rich, J. Gonzalez: Vulnerability Black Markets: Empirical Evidence and Scenario Simulation. In: Proceedings of the 42nd Hawaii International Conference on System Sciences, 2009

[Sar91] Sarter, N., Woods, D.: Situation awareness: A critical but ill-defined phenomenon. In International Journal of Aviation Psychology 1, 45–57, 1991

[Sko12] Skopik, F., Ma, Z., Smith, P., Bleier, T.: Designing a Cyber Attack Information System for National Situational Awareness, In Proceedings of the 7th Future Security Conference 2012.

[Smi11] Smith, P., Hutchison, D., Sterbenz, J.P.G., Schöller, M., Fessi, A., Doerr, C., Lac, C.: D1.5c: Final strategy document for resilient networking. In: ResumeNet Project Deliverable, http://www.resumenet.eu, August 2011

[Sym10] Symantec Global Internet Security Threat Report XV, Page 15, April 2010, Symantec Corporation, http://eval.symantec.com/mktginfo/enterprise/white_papers/b-whitepaper_internet_security_threat_report_xv_04-2010.en-us.pdf

[Tad06] Tadda, G., Salerno, J.J., Boulware, D., Hinman, M., Gorton, S.: Realizing situation awareness within a cyber environment. In: Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications. Orlando, FL, USA, April 2006

[Tik08] Tikk, E., Kaska, K., Rnnimeri, K., Kert, M., Talih¨arm, A.M., Vihul, L.: Cyber attacks against Georgia: Legal lessons identified, Novermber 2008, http://www. carlisle.army.mil/dime/getDoc.cfm?fileID=167