CYBERSANE

D5.1

Prevention and Response to Advanced Threats

CYBERSANE

Project number:	833683
Project acronym:	CyberSANE
Project title:	Cyber Security Incident Handling, Warning and Response System for the European Critical Infrastructures
Start date of the project:	1st September, 2019
Duration:	36 months
Programme:	H2020-SU-ICT-2018

Deliverable type:	Report
Deliverable reference number:	DS-01-833683 / D5.1 / N.1
Work package contributing to the deliverable:	WP 5
Due date:	08 2020 – M12
Actual submission date:	<14 September 2020>

Responsible organisation:	STS
Editor:	Kontakis Konstantinos, Spanoudakis Georgios
Dissemination level:	PU
Revision:	N.1

Abstract:	This report reviews the state-of-the-art on proactive detection and response approaches, which make use of artificial intelligence, deep learning or machine learning algorithms. The outcomes of this report shall provide an automated detection and response approach to efficiently protect the CIIs.
Keywords:	ADS, AI, DL, IPS, Incident Handling, Incident Response, ML, RA Methodologies, Simulation Environments
	The project CyberSANE has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 833683.

Editor

Kontakis Konstantinos, Spanoudakis Georgios (STS)

Contributors (ordered according to beneficiary numbers)

Luis Landeiro Ribeiro, Luis Miguel Campos (PDMFC)

Ruiz Jose Francisco (ATOS)

Martinelli Fabio (CNR)

Zamarripa Sergio (S2)

Mitton Nathalie, Staddon Edward (INRIA)

Papastergiou Spyros, Karantjias Thanos, Hatzikou Menia (MAG)

Karagiorgou Sophia (UBI)

Athanatos Manos, Papadogiannaki Eva (FORTH)

Kontakis Konstantinos, Spanoudakis Georgios (STS)

Karypidis Paris-Alexandros (SID)

Disclaimer

The information in this document is provided “as is”, and no guarantee or warranty is given that the information is fit for any particular purpose. The content of this document reflects only the author`s view – the European Commission is not responsible for any use that may be made of the information it contains. The users use the information at their sole risk and liability.

Executive Summary

The protection of critical infrastructures is a complex activity that requires work at different levels and correlation of events in order to understand what is happening and how to better protect a system. In this way, we can differentiate between different strategies: pre-emptive solutions, active solutions, reaction capabilities, analysis and adaptability, self-healing, etc. As we know it doesn't exist a perfect strategy for protecting systems, one of the initial activities we have performed in CyberSANE is to list and study different protection mechanisms that will be adapted in our approach in order to understand how to improve and advance them.

Due to the criticality of the CIIs we focus here in proactive detection and response of cyber-threats as from our point of view it is key for the successful development of the CyberSANE platform. Therefore, an extensive study on the latest works regarding the proactive detection and response methodologies has been carried out in order to identify the most prominent solutions in this area. The areas of research were selected after several discussions in the project and aiming to cover the tools and approaches we have in the technical WPs, with special focus in WP5 (hybrid-net). The findings of this study aim to serve as the basis for the rest of the tasks in order to pick out the most suitable approaches for the implementation of the Cyber Fusion Models. This could be done by either integrating one or more of the presented methodologies or by enhancing the capabilities of the currently owned consortium tools through the development of additional features in need.

Contents

List of Figures

List of Tables

Chapter 1 Introduction

This deliverable presents the findings and work performed in Task 5.1. We describe a detailed state-of-the-art on proactive detection and response approaches that will be used as basis for the enhancement of the other activities done in CyberSANE, more concretely in WP5. The different areas covered here were selected after discussing at technical level of the more important elements we need to cover in the project and information we need for improving some tasks and activities, such as the models or the algorithms for the anomaly detection tools.

This document provides a desk research on the latest methodologies used to counterattack and handle cybersecurity threats with respect to Critical Infrastructures (CIs) and Critical Information Infrastructures (CIIs). The rest of this document is structured as follows:

- Chapter 2 describes the state-of-the-art on incident handling frameworks, incident response strategies, and approaches of digital chains of evidence.
- Chapter 3 briefly presents CyberSANE's threat taxonomy as a mean to identify the necessary set of threats that should be addressed in this deliverable, followed by a number of tools and techniques for the efficient proactive detection and response.
- Chapter 4 includes both the most widely known and latest works on anomaly-based detection methodologies across various technological backgrounds like statistical, machine learning, data mining, deep learning, and genetic algorithm techniques.
- Chapter 5 documents the risk assessment methodologies, the attack and simulation environments found in literature, as well as the data visualization techniques of an attack.
- Chapter 6 features the concluding remarks of this deliverable.
- Chapter 7 includes a glossary of the most commonly used abbreviations.
- Chapter 8 concludes with all the bibliography of this deliverable.

Chapter 2 Incident Handling & Response Approaches

This chapter describes the latest works on incident handling and response approaches which could be of potential use in CyberSANE project, focusing on both widely adopted strategies, and several frameworks tailored to the needs of specific industries. A short review on the existing standards, guidelines and best practices for the prevention, detection, response and mitigation of threats on CIs has been already described in deliverable "D2.1 - Cyber Incident Handling Trend Analysis". However, since the analysation and prioritisation of such incidents is perhaps the most critical point of decision within the security incident process (Cichonski, et al., 2012, p. 32), each approach shall be thoroughly analysed in the upcoming Tasks of WP5 to end up with a properly coordinated and distributed incident handling solution. This solution aims to provide a set of predefined procedural actions that could effectively manage any security, covering thus most of today's requirements and challenges regarding the protection of CIs (Alcaraz & Zeadally, 2015).

2.1 Incident Handling Frameworks

The relatively recent discovery of Stuxnet worm and its potential impact on today's CIs (Farwell & Rohozinski, 2011) emerged the necessity to address cyber-security and privacy-by-design aspects in a holistic manner, in order to efficiently shield a CI's hardware and software components against both existing and future cyber-threats. At first, several initiatives were implemented based on well-defined frameworks which followed business-related guidelines and organizational risk management processes. However, this high-level perspective in detection and prevention phases, quickly denoted the need of a deeper and more robust analysis, giving birth to a next generation of frameworks. The initial framework that introduced cybersecurity improvements to CIs (NIST, 2014) was issued by the National Institute of Standards and Technology (NIST) back in 2014 and acted as the cornerstone for several other frameworks. A typical example found on literature lies to the definition of the Italian National Cybersecurity Framework (Baldoni & Montanari, 2016) which may conform with NIST's guidelines, but it is heavily oriented towards the Italian law and enterprises. However, a purely practical point of view at limiting the effects of cyber-attacks was addressed in (Bottazzi, et al., 2017), where a recording of the suggested procedures which have to be followed by the IT team of an organization like a CI took place. This study produced a minimal set of arrangements that must be adopted in order to efficiently handle a security incident across time, space, and data domains. All the well-defined and tested prior (ex-ante) and post (ex-post) actions to a cyber-attack are displayed in Table 1. The proposed planning and deployment could be indeed characterized as resourceful and time-consuming measures, but on the other hand, the expected time and cost gains from the sufficient handling of a potential security incident overcome these factors.

Domain	Actions	
	Ex-ante	Ex-post
Time	Minimize Internet border gateways	Close all or part of the Internet border gateways
	Logging of the critical information in a usable format	Scan logs starting from the ones related with the compromised sector

Space	Initiate a deep hardware and software probing of the infrastructure	Limit or restrict traffic coming from compromised sectors
	Proceed to network's segmentation	Denial of services to the compromised network links
	Identify and segregate each application's data	Prevent any communication between the compromised application and the rest of infrastructure
	Support of a CDN (Content Delivery Network) service	Switch to the CDN (Content Delivery Network) service
	Access application services through limited and well-known only terminals	Isolate access to application services
	Access infrastructure services through limited and well-known only terminals	Isolate access to infrastructure services
Data	Configure backup of services and data	Initiate a recovery process from an existing offline-backup
	Employ data encryption capabilities	Reduce data access to only a few entities
	Set-up and put offline workstations for emergency cases	Deploy to production the already prepared offline workstations

Table 1: Incident Handling Actions Across Time, Space, and Data domains

Another incident handling framework which lied the foundation for the overall security management at the Computer Emergency Response Team (CERT) of the Republic of Mauritius was introduced in (Usmani, et al., 2013). The presented framework denoted the high organizational value of the information exchanged over a network like the Internet and succeeded in easing the task of traffic monitoring conducted by a local Internet Service Provider. The key concept behind their incident handling was a three interdependent layered architecture of the framework, where stakeholders are actively involved in its overall implementation. The core layer encapsulates all the necessary security mechanisms, and the last layer displays the results of the monitoring process. Their solution was able to handle security incidents and breaches in a timely manner, allowing the design of a security policy and the set-up of an effective business continuity plan (BCP). The potential invasiveness to sensitive information was detected by an automated monitoring mechanism, preventing thus its improper or criminal usage by third-party tools or threat-actors (Bottazzia & Mea, 2017). However, this framework as well as most of current approaches make use of linear procedures optimized for the handling of a single incident, lacking thus on the sufficient handling of complex or simultaneous cyber-attacks. The alarming rise of such kind of attacks denotes more than ever the need of a coordinated security incident handling methodology across several and different types of CIs, in order to enhance their threat-knowledge and assign timely a set of responsibilities and actions to the appropriate persons and resources. (Daley, et al., 2011) developed a coordination model based on a cooperative operations' structure for the efficient handling of cyber-security incidents. Doing so, they were able to detect faster a cyber-threat and share its knowledge with the rest of peers compared with the traditional linear approaches. The presented incident handling operationalization took also into consideration the necessary autonomy, and customization capabilities of each CI, allowing them to finetune and scale their system according to specific needs.

Another domain of interest lies to the CIs that belong to industrial sector and are usually controlled by a combination of specialized hardware and software solutions known as Supervisory Control and Data Acquisition (SCADA) systems. Since the efficient handling of a security incident by their authorized personnel presupposes a necessary set of knowledge and skills, it is evident that security mechanisms for controlling human-made activities will be provisioned, without interfering though with the normal functionality of the CI. The research community shown quite limited interest towards this area since an agent-based framework (Bigham, et al., 2004) for the monitoring of system and the automated relocation of personnel privileges was empirically the most noteworthy approach for several years. However, a most recent work by (Alcaraz, et al., 2009) came up with a reputation-based mechanism which allows a SCADA system to identify the best possible operator to handle an incident, and a supervisor to monitor the operator and submit its feedback for assessment purposes. The presented reputation module made use of the “reward” and “fear” incentives to store human-related information, as well as an Adaptive Assignment Manager (AAM) mechanism to assign the security incident to a certain operator, while the latter component was also responsible to find the ideal supervisor for the monitoring of the overall incident handling process.

2.2 Incident Response Strategies

An Incident Response (IR) strategy is nowadays deemed necessary to efficiently deal with the risks associated with the confidentiality, integrity and availability policies set up by a CI. From a technical management perspective, an incident response consists of several phases that aim to efficiently isolate and handle an incident in a timely manner. Even though that various methodologies have been proposed to reduce the impact of an incident’s effects and trace their origin (DePaul University, 2002; Schreck, 2018), no universally adopted standard has been yet adopted. NIST has successfully identified the complexity and necessity of dealing with a security incident within an organisation, and they proposed a non-linear approach for the efficient “respond” to such incidents (Cichonski, et al., 2012). They have defined the four incident response life cycle phases shown in Figure 1, which could be addressed in any future incident response framework or contingency plan.

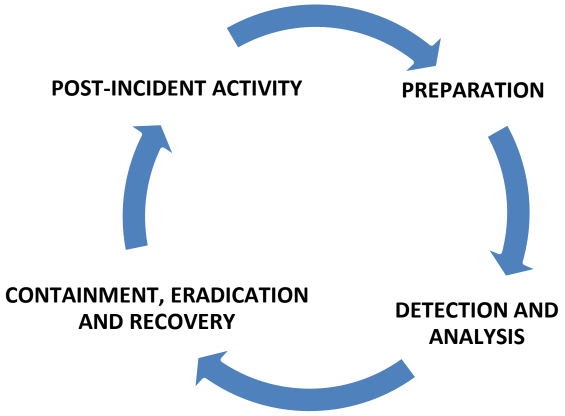
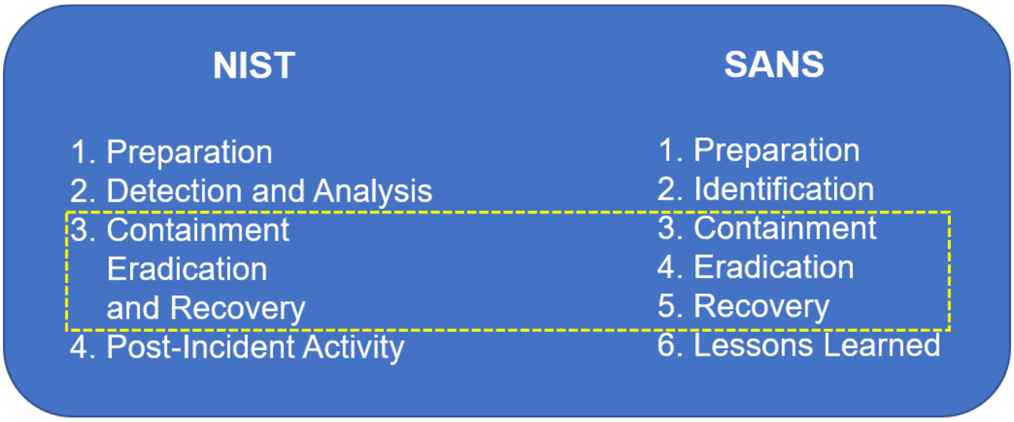


Figure 1: Incident Response Life Cycle

The first phase lies to the preparation and actions undertaken to both prevent and handle a possible security incident. The preparation of an incident response involves the facilitation of communication and reporting mechanisms, integration of incident analysis hardware solutions, and adoption of mitigation software capabilities. On the other hand, the prevention of incidents takes place by providing some of the best practices on cyber security domain, like the execution of periodic risk assessments, improving network security with Virtual Private Network (VPN) solutions, enabling malware detection and protection tools, etc. The second phase of Detection and Analysis takes advantage of a predefined set of attack vectors, which classify incidents based on the attack methodologies used along with the source of the most common precursors and indicators. The incident is then captured and documented by various techniques (e.g. network/system profiling, event correlation, packet sniffing, etc.), it is attributed a prioritisation and the appropriate individuals are notified according to the organisation's policy. Afterwards comes the third phase of this life cycle, the Containment, Eradication and Recovery. At first it is chosen a containment strategy to decrease the effect of the security incident, followed by the collection of evidence across all systems of interest in order to identify the attacking hosts, and ultimately set up an eradication and recovery procedure where it is deemed necessary. Last but not least, the fourth phase comes to the fore, where Post-Incident Activities take into account incident-related metrics and evidence retention methodologies to learn about new threats and further improve the cyber security of an organisation like a CI.

Following a similar approach but focusing into providing a well-defined management framework with a complete methodology capable of efficiently handling a security incident, (Mitropoulos, et al., 2006) combined all the necessary actions that should be taken once such an incident takes place. Their framework not only included the best practices and technological implementations on research and applied techniques in the form of passive incident response mechanisms, but they also examined and integrated several active incident response mechanisms to identify attacker's origin by deploying concrete software forensics traceback techniques (Mandia, 2001). At first, they stressed the requirement of marking and dealing with the critical security components of the system during the Preparation phase, which include the operating systems, the boot disks, the backup media and their procedures, the cryptographic checksums of critical application files, and the audit trail of system's components. In the upcoming Identification phase, two approaches are followed depending on the severity of the security incident, where either the attacker's point of entry is violently terminated to prevent further exposure of the system, or the system remains open as long as possible on purpose in order to collect as much information as possible. Such information is later taken into consideration in the form of evidence from the system's log examination and analysis processes. The second approach is also closely related with the Containment phase of their framework, where it aims to reduce the impact of an incident to a desired only context by disabling specific services, changing breached user accounts, temporarily disconnecting -or even- shutting down the compromised system. Afterwards, the Eradication phase comes into play to apply techniques that aim to prevent the reoccurrence of this specific incident attack by reinstalling/rebuilding one or more components of the infected system. Once all the necessary attacker elimination actions have been completed, the Recovery phase focus on the restoration of the appropriate services after carefully reviewing their configuration and the adopted protective and detective mechanisms. Finally, the overall incident response process concludes with the Follow-Up phase which documents all the afore-mentioned information and submits it for an evidence forensically sound analysis. Common mechanisms used to analyse and trace back an incident involve ICMP-based, IP marking and IP tunnelling techniques, but they also used host-based and application-based approaches like the Intruder Detection and Isolation Protocol (IDIP) integrated already in various intrusion detection systems (Schnackengerg, et al., 2001).

Except from the widely adopted NIST guidelines and its varieties on both research and applied areas of interest, there are several other recommendations which are also related with the incident response management. The SANS Technology Institute has published a brief guideline (Kral, 2012) composed of a 6-step process similar to NIST for the identification and handling of a security incident as shown in Figure 2, while the European Union Agency for Cybersecurity (ENISA) in (Maj, et al., 2010) skips the Preparation phase and enhances the actions that must be undertaken during and after the incident takes place. The common point between all those guidelines lies to the fact that they tend to comply with the structure defined in ISO/IEC 27035, a standard which stands out for its recommended information security incident management practices (ISO/IEC JTC 1/SC 27, 2016). The application of such practices has been reported before in literature across a variety of domains ranging from financial and research institutions, to IT service and power industries. An extensive survey on a finite set of incident management aspects for various organizations, along with their relative collection methodologies took place in (Tøndel, et al., 2014). This study provided suggestions for addressing the challenging practices in incident response by creating plans and classification of incidents, promoting collaboration and sharing lessons learnt between distinct teams and disciplines, as well as improving the existing incident management tools in the context of usability and false positive alerts. The same work also noted that those CIs which fall into power industry had not established a sufficient incident management plan for their smart grids, the users' information awareness was quite limited, and the network monitoring for abnormalities was done in a manual way (Line, 2013). In these cases, it was proposed the composition of a special team with extended decision-making power to fill the gap between Information and Communication Technology (ICT) and power automation staff, a Computer Security Incident Response Team (CSIRT) capable of detecting in time a cyber-security incident and countering its further impact upon the system. Specific actions that could be performed by such CSIRTs during and after a cyber-attack have been addressed in (Ruefle, et al., 2014; Steinke, et al., 2015), where a set of the necessary skills, as well as the collaboration of teams involved in security incidents' management, are thoroughly explained at an organization-level.



NIST	SANS
1. Preparation	1. Preparation
2. Detection and Analysis	2. Identification
3. Containment	3. Containment
Eradication	4. Eradication
and Recovery	5. Recovery
4. Post-Incident Activity	6. Lessons Learned

Figure 2. NIST & SANS Response Steps

All the aforementioned phase-based process of an incident along with its collaborative nature has been attributed as a pervasive feature of the security management that lies within a CI. Thus, handling a typical security incident is mainly based on an intensive diagnostic work which involves empirical data, analytical skills, communication skills, and the use of various strategies and security tools depending on the phase and the task as shown in Table 2. The results reflected in that table was the outcome of a qualitative analysis of the tasks performed by several IT security specialists across various high reliability organizations (Werlinger, et al., 2010). During the diagnosis of a security incident, several tasks were run by deploying various security tools in conjunction with a set of predefined strategies like the pattern recognition, hypothesis generation, communication and the dynamic integration of distinct security tool, a process also known as bricolage. However, since the cyber-threats field and zero-day attacks keep evolving over time, so must the incident response process. All relative tools and strategies followed, require a continuously monitoring and mixture of a solid technological and communicational background. Doing so, it is feasible not only to efficiently diagnose a security incident, but also deal sufficiently with its effects regardless of the task complexity hidden beneath and the triggering of a -sometimes- vast amount of false positive alarms in the context of an organization like a CI.

Phase	Task	Strategies	Security Tools
Preparation	Vulnerability Assessment	Tacit Knowledge	Scanners (e.g. Nessus, ISS)
			Community Lists
	Tool Configuration	Communication	Complex Organization Tools
Anomaly Detection	Monitoring	Pattern Recognition	Customizable Scripts
		Collaboration	
		Simulation	IT Tools (e.g. IDSs and Antivirus Software)
Anomaly Analysis	Receiving Notifications	Tacit Knowledge	Incident Ticketing System
		Communication	
	Verification	Hypothesis Generation	Analytical Skills
		Tacit Knowledge	Customizable Scripts
	Assessment		Administration Applications (e.g. Firewall Solutions)
		Pattern Recognition	
	Traceback Anomaly Source		IT Tools (e.g. IDSs and Antivirus Software)
		Hypothesis Generation	Analytical Skills
		Bricolage	Scripts and Tools Combination

Table 2: Security Incident Response Strategies and Tools

2.3 Digital Chains of Evidence

Nowadays, every embedded hardware or software component which plays a key role to the expected and uninterrupted functionality of a CI makes use of a data logging mechanism. Such mechanisms automatically generate and store data that provide crucial information about a security incident, like a network breach or a system compromise. According to (Kerr, 2005) this available information could be further correlated with one or more events, creating thus digital chains of evidence. A digital chain of evidence consists of digital evidence records which have to be secured against a predefined set of parameters necessary for the validation of the information. In the near past (Olsson & Boldt, 2009) presented the CyberForensic Timelab, a prototype scanner implemented with various file handlers written in Perl language. Their approach was able to scan and supplied with a collection of evidence, which in turn they were encoded in Document Type Description (DTD) format for further evidence correlation and event display purposes.

On the other hand, (Kumar, et al., 2011) addressed and tried to overcome the difficulties met in the interoperability of evidence collection with most Intrusion Detection Systems, and proposed a real-time forensic analysis methodology to obtain enough evidence to detect any damages occurred and trace back the origin of an attacker. They applied a time-lining technique to obtain a chronological sequence of events of interest and gain insight regarding a cyber-threat's cause, impact, and interconnections.

(Kuntze & Rudolph, 2011) presented a high-level architecture for collecting secure digital evidence by creating and storing a graph-based linking of each evidence record which ultimately represent one or more chains of evidence. Their approach made use of evidence generators based on hardware-based digital signatures (Richter, et al., 2010), evident collectors for the attribution of semantic information as well as the distribution of evidence records (Reith, et al., 2002), and a forensic database for the storage of those records in a graph-based structure (i.e. Resource Description Framework[1]). The required event correlation was possible by conducting either a graphical visualization of the relations between each evidence record, or by executing special purpose queries[2] to search for possible matching evidence records.

The need of collection and examination of digital evidence also gave birth to several proprietary tools and frameworks which treat evidence as binary objects, in order to proceed with metadata extraction and identify potential correlations. The most prominent solutions in this area are the EnCase Forensic[3], AccessData Forensic Toolkit[4], Sleuthkit , PyFlag (Cohen, 2008), and Wireshark (Ndatinya, et al., 2015). However, one crucial disadvantage of these tools lies to the fact that each one of them tends to specialize on certain only types of evidence, lacking thus the ability to conduct a holistic investigation and correlation of events. A typical comparison between their features, as well as the challenges that should be addressed regarding their diversity takes place in (Raghavan, 2013). It is worth noticing that the latter study also embraces the opinion expressed by (Garfinkel, 2010), where most digital forensic and chains of evidence solutions have to adopt standardized and modular approaches, in order to efficiently deal with the current threat landscape.

[1] https://www.w3.org/RDF/
[2] https://www.w3.org/TR/sparql11-overview/
[3] https://www.guidancesoftware.com/encase-forensic
[4] https://accessdata.com/products-services/forensic-toolkit-ftk

Chapter 3 Proactive Detection & Response

CyberSANE and any CI must adopt mechanisms able to perceive and treat a cyber-attack before it causes significant impact to system's functionality. For that reason, several proactive intelligence and proactive information security approaches are being continuously deployed and tested across numerous organizations. However, due to the variety of threat-landscape, distinct approaches have been developed over the years. Such approaches tend to either enable defence point and policy enforcement for each system's element, enhance the collaboration between those elements, or promote advanced adaptive technologies that automatically prevent a threat.

3.1 CyberSANE Threat Taxonomy

With the increase of crime in the cyberspace, it is necessary to possess a means of identification and countermeasure against these illegal business cases. The invention of threat classification mechanisms, called Threat Taxonomies, allow the association between the attacker's actions and a given threat category, granting security experts to identify patterns and react accordingly, increasing response time. Together coupled with Threat Models, it is possible to identify and detect numerous threats perceived against a specific cyber infrastructure. These products are presented in more detail in deliverable "D3.1 - Taxonomy of Threat Landscape for CIIs".

Before teaching a machine the capabilities of threat detection, it is necessary to list all potential threats against the system. Using threat classification methodologies, cyber-security specialists can list and classify all potential threats, also including where necessary the vulnerabilities which are generally exploited as well as critical effects but also remediation steps.

There are many different taxonomical approaches, each with their different strengths and weaknesses, such as NIST's CSRC taxonomy (NIST, 2020), ENISA's threat taxonomy (Marinos, 2016), the Taxonomy of Operational Cyber Security Risks (TOCSR) proposed by (Cebula, et al., 2014), and the Open Threat Taxonomy (OTT) developed from Enclave Security[5]. These different approaches are adapted to many different uses since they can take different forms, such as a mind mapping tree or a classical table. Moreover, taxonomies are designed to be applicable to the system which they are protecting, hence their elaboration by that systems cyber-security experts. This means that a taxonomy applied to specific system architecture with identifiable specifications and requirements, will not necessarily be exploitable on another system which could possess both different architectural decisions as well as specifications. An evaluation of such comprehensive taxonomies for information technology threats has been recently conducted by (Launius, 2018).

Since CyberSANE's purpose is to cover multiple CIIs and share detection information between units, multiple taxonomies are not feasible. A single taxonomy was, therefore, created to cover CIIs as a single system, decomposed into multiple sectors. The resulting product was based upon the format used by ENISA (Marinos, 2016), thus giving a solid ground to begin construction of CyberSANE's core detection taxonomy. However, ENISAs taxonomical structure possessed certain limitations regarding its categorical choices. Indeed, certain threats such as Denial-of-Service or Man-in-the-Middle attacks were categorised as single individual attacks, whereas in practice there are numerous methods to perform either attack. It was thus decided to expand upon the existing threats presented

[5] https://www.auditscripts.com/resources/open_threat_taxonomy_v1.1a.pdf

by ENISA, presented in Table 3, to expand into multiple subcategories, thus differentiating between the high level threat, and the threat type.

High-Level Threats
Physical
Accidental
Environmental Disaster
Failure
Outage
Eavesdropping / Interception / Hijacking
Nefarious Activity

Table 3: High-Level Threat Categories

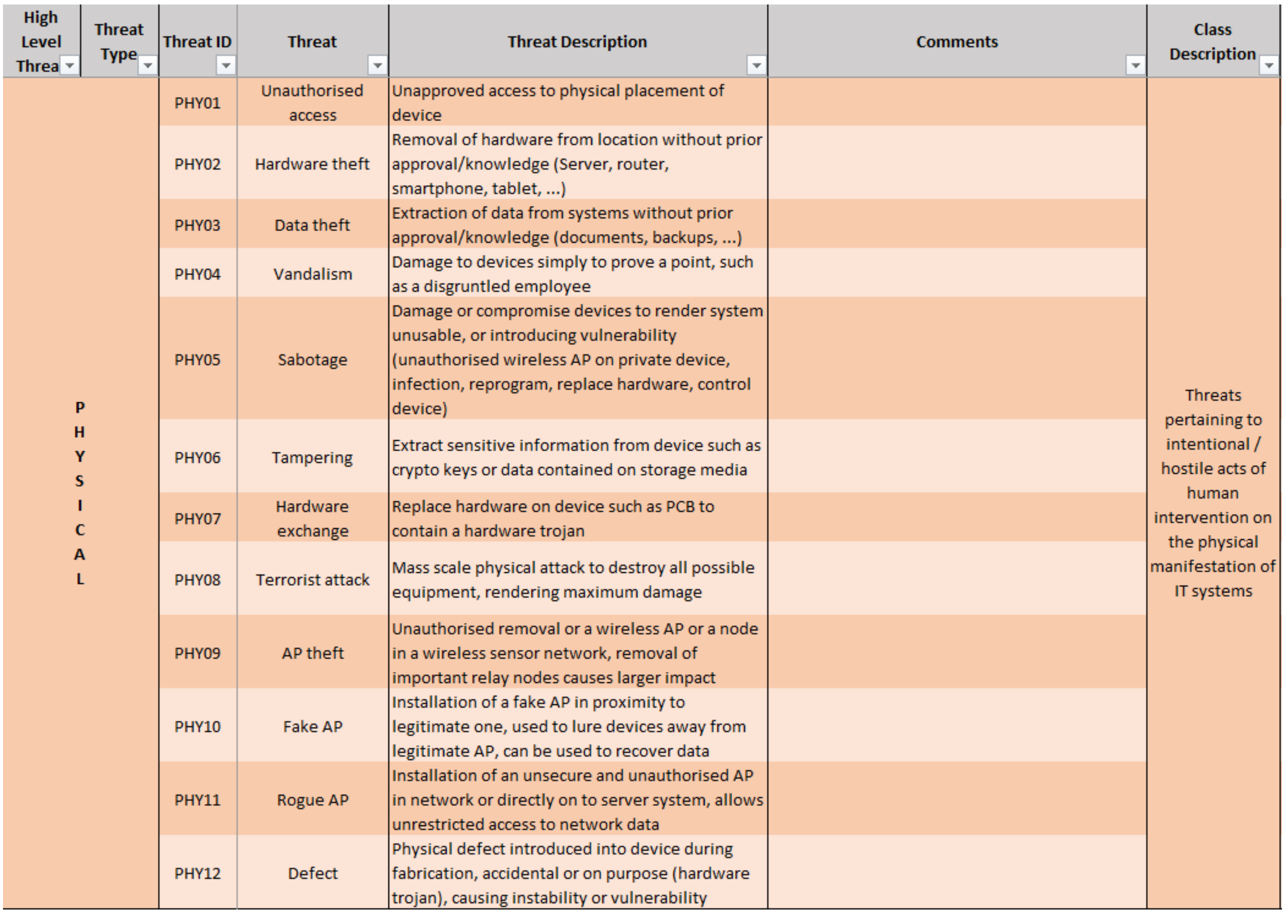
It is immediately possible to identify that certain categories, such as “Eavesdropping” and Nefarious Activity” are vast areas, regrouping practically all cyber-threats. It was here that our Taxonomy was able to expand upon these categories, associating more specific threat types, as presented in Table 4.

High-Level Threat	**Threat Type**
Eavesdropping/ Interception / Hijacking	Reconnaissance
	Eavesdropping
	Man-in-the-Middle
Nefarious Activity	Denial-of-Service
	Disruption
	Side-Channel
	Transmission Control Protocol
	Routing
	Authentication
	Confidentiality
	Wireless Sensor Networks
	Data Integrity / Breach

High-Level Threat	Threat Type
	Software
	Malware
	Equipment
	Protocol
	Information Leak / False Information

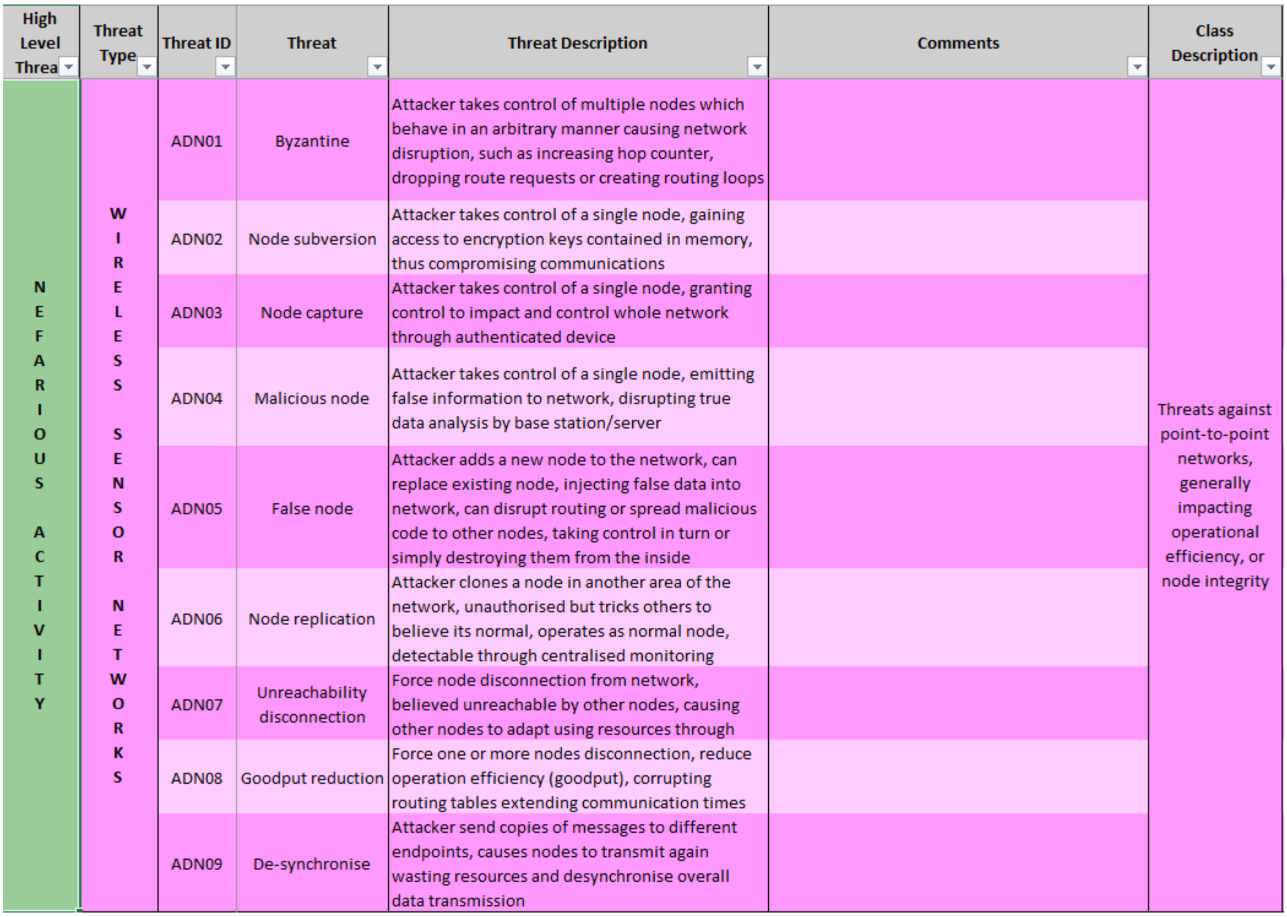
Table 4: Threat Types

This identification allowed the ability to categorise threats more precisely, increasing ease of use towards adding new threats or even new threat type categories. This expanded the taxonomy into more specific areas, such as IoT based Wireless Sensor Networks (WSN), and even listing threats towards specific exchange protocols in control systems. The resulting taxonomy contains currently 248 listed threats spread across a total of 22 threat categories. An extract of Physical level threats can be seen in Figure 3 and the threats towards WSNs can be seen in Figure 4.



High Level Threa	Threat Type	Threat ID	Threat	Threat Description	Comments	Class Description
PHYSICAL		PHY01	Unauthorised access	Unapproved access to physical placement of device		Threats pertaining to intentional / hostile acts of human intervention on the physical manifestation of IT systems
		PHY02	Hardware theft	Removal of hardware from location without prior approval/knowledge (Server, router, smartphone, tablet, ...)		
		PHY03	Data theft	Extraction of data from systems without prior approval/knowledge (documents, backups, ...)		
		PHY04	Vandalism	Damage to devices simply to prove a point, such as a disgruntled employee		
		PHY05	Sabotage	Damage or compromise devices to render system unusable, or introducing vulnerability (unauthorised wireless AP on private device, infection, reprogram, replace hardware, control device)		
		PHY06	Tampering	Extract sensitive information from device such as crypto keys or data contained on storage media		
		PHY07	Hardware exchange	Replace hardware on device such as PCB to contain a hardware trojan		
		PHY08	Terrorist attack	Mass scale physical attack to destroy all possible equipment, rendering maximum damage		
		PHY09	AP theft	Unauthorised removal or a wireless AP or a node in a wireless sensor network, removal of important relay nodes causes larger impact		
		PHY10	Fake AP	Installation of a fake AP in proximity to legitimate one, used to lure devices away from legitimate AP, can be used to recover data		
		PHY11	Rogue AP	Installation of an unsecure and unauthorised AP in network or directly on to server system, allows unrestricted access to network data		
		PHY12	Defect	Physical defect introduced into device during fabrication, accidental or on purpose (hardware trojan), causing instability or vulnerability		

Figure 3: CyberSANE Taxonomy - Physical Threats



High Level Threa	Threat Type	Threat ID	Threat	Threat Description	Comments	Class Description
NEFARIOUS ACTIVITY	WIRELESS SENSOR NETWORKS	ADN01	Byzantine	Attacker takes control of multiple nodes which behave in an arbitrary manner causing network disruption, such as increasing hop counter, dropping route requests or creating routing loops		Threats against point-to-point networks, generally impacting operational efficiency, or node integrity
		ADN02	Node subversion	Attacker takes control of a single node, gaining access to encryption keys contained in memory, thus compromising communications		
		ADN03	Node capture	Attacker takes control of a single node, granting control to impact and control whole network through authenticated device		
		ADN04	Malicious node	Attacker takes control of a single node, emitting false information to network, disrupting true data analysis by base station/server		
		ADN05	False node	Attacker adds a new node to the network, can replace existing node, injecting false data into network, can disrupt routing or spread malicious code to other nodes, taking control in turn or simply destroying them from the inside		
		ADN06	Node replication	Attacker clones a node in another area of the network, unauthorised but tricks others to believe its normal, operates as normal node, detectable through centralised monitoring		
		ADN07	Unreachability disconnection	Force node disconnection from network, believed unreachable by other nodes, causing other nodes to adapt using resources through		
		ADN08	Goodput reduction	Force one or more nodes disconnection, reduce operation efficiency (goodput), corrupting routing tables extending communication times		
		ADN09	De-synchronise	Attacker send copies of messages to different endpoints, causes nodes to transmit again wasting resources and desynchronise overall data transmission		

Figure 4: CyberSANE Taxonomy - WSN Threats

3.2 Proactive Detection & Response

Proactive detection of a network security incident is the process of the early identification of an imminent cyber threat before the affected parties become aware of the problem itself. Proactive response of a network security incident is the action that occurs in order to prevent or mitigate the threat before the actual infection. The approaches, which are widely used for proactive detection and response by organizations, can be divided into two categories (Gorzelak, et al., 2011):

i. the tools and techniques that can be deployed by an organization to internally monitor, detect and respond to cyber incidents
ii. the services that offer information interexchange across organizations, related to already detected network security incidents.

The most typical tools used for the proactive detection of network security incidents are firewalls, antivirus systems, system logs and intrusion detection/prevention systems, which are part of an organization's network infrastructure. More advanced tools that usually require extra resources (e.g., dedicated hardware and system configurations, available IP address space) are honeypots, sensor networks, network telescopes, sandboxes and Domain Name Service (DNS) monitoring and analysis techniques. These tools are mainly used as early warning systems that gather information based on current attack scenarios all over the network. Other tools serve for network flow monitoring, full packet capture, sinkholing, monitoring of Internet routing, log and event aggregation, correlation and visualization, industrial control systems monitoring, cloud services monitoring, endpoint monitoring, certificates monitoring, vulnerability scanning, automated malware analysis or information leakage monitoring. In the next paragraphs, we present some of the most popular and typical techniques and tools used for proactive detection of cyber security incidents.

Firewall: A typical firewall is a device or software that serves as a monitor and as a filter that permits or denies certain network connections. Firewalls are categorized either as network-based or host-based firewalls and can operate on different layers (e.g., application layer). Firewalls can be used in proactive incident detection by creating alerts when suspicious events occur. For instance, a suspicious event could be the communication of a known, blacklisted C&C server's IP address with an organization's internal IP address space. When such event is recognized, a firewall could raise an alert, prohibit the connection and then log the incident for further analysis. There are plenty of firewalls solutions; open-source or commercial, software-based or deployed on dedicated hardware devices. Popular firewalls are the following: iptables[6], IPFire[7], pfSense[8], OpenWRT[9].

Intrusion Detection and Prevention Systems: An Intrusion Detection System (IDS) is not only able to monitor and analyse the network traffic in depth, but also inspect an operating system's behaviour and files to identify malicious or abnormal activities. Depending on where the detection feature takes place, IDSs are classified as host-based or network-based systems. In both occasions however, an IDS can be also defined as an Intrusion Prevention System (IPS) thanks to its capability to respond to a security incident. For instance, an IPS is able to block a specific network connection that is characterized as malicious by employing a signature-based or anomaly-based methodology. More specifically, signature-based IDSs search for predefined attack patterns and heuristics (e.g., known malicious URLs inside the packet contents), while anomaly-based IDSs are able to learn how to distinguish normal activity from deviants. Popular intrusion detection/prevention systems are: Snort[10], Suricata[11], Zeek/Bro[12].

Honeypots: A honeypot aims to detect and monitor attempts of unauthorized use at machine or network level, by acting as a trap that attracts malicious agents. A honeypot (also known as server-honeypot) can be an entire virtualized system or just an isolated part of the operating system. Advanced honeypot techniques have been reported in the literature (Moore, 2016), where a number of specific thresholds are used as the triggering mechanism for a staged response to a cyber-attack. Another type of honeypot is the client-honeypot, which aims for communication with malicious servers in order to determine if an attack occurs and gather information based on this attack. Any client software that can interact with servers is able to become a client honeypot (e.g., web browsers, FTP clients, e-mail clients) (Grudziecki, et al., 2012). A honeynet serves as a combination of several honeypots and offers the simultaneous monitoring of multiple networks. Popular honeypots are the following: Kippo[13] (SSH honeypot), Cowrie[14] (SSH+Telnet honeypot), GHH[15] (web honeypot),

[6] http://freshmeat.sourceforge.net/projects/iptables/
[7] https://www.ipfire.org
[8] https://www.pfsense.org
[9] https://openwrt.org/docs/guide-user/firewall/start
[10] https://www.snort.org
[11] https://suricata-ids.org
[12] https://zeek.org
[13] https://www.honeynet.org/projects/old/kippo/
[14] https://github.com/cowrie/cowrie
[15] http://ghh.sourceforge.net

HoneyMail[16] (SMTP honeypot), Dionaea[17] (low-Interaction honeypot), IoTPOT[18](IoT honeypot), Multiple Honeypot Solution (MHN)[19], CONPOT[20](ICS/SCADA honeypot).

Sinkholes: A sinkhole is a server or network component to which malicious traffic is intentionally directed for further analysis. For example, after the activity termination of a C&C server address, a sinkhole server can be used as a replacement to track every connection for bot detection and tracking.

Sensor networks: Another approach to monitor a network and collect information that represents the communications of multiple and heterogeneous vantage points, is the deployment of a sensor network. Sensor networks' applications include but are not limited to healthcare monitoring environmental monitoring, industrial monitoring and threat detection. Since the attack surface is really wide in such networks (Alzaid, et al., 2008), specialized setups of sensors with attacks' simulations can be used to collect meaningful information that can be later used for proactive detection of security incidents.

Sandboxes: A sandbox is an isolated environment, where untrusted or unverified applications are able to run without affecting the host system or the internal network. Consequently, the behaviour of the untrusted application can be inspected in order to determine whether it is malicious or not (e.g., contacts a known malicious C&C server). If the application is malicious, then it can be further analysed in order to gain behavioural information about the malware for proactive security incident detection.

Network telescopes: A network telescope monitors network traffic that targets unused IP address spaces (Durumeric, et al., 2014). The network traffic that is destined to those IP addresses is considered suspicious, since it indicates automated scanning, DDoS backscatter or misconfigured/vulnerable devices, and can enrich mechanisms for proactive detection of security incidents.

Passive DNS analysis: Passive DNS analysis can serve for the investigation of possible network security incidents, such as threats that originate from botnets, by detecting domains that are involved in malicious activities (Bilge, et al., 2011). The analysis of DNS logs could be used to detect infected IP addresses that could be then used to build blacklists of known malicious domains.

Automated malware analysis and information leakage monitoring: Automated static malware analysis systems aim for the analysis of malicious files without using dynamic methods. These mechanisms can operate on binaries and memory dumps in order to extract static configurations of malware. Other relevant approaches focus on mobile malware identification. In addition, there are tools that monitor possible personally identifiable information from mobile applications.

Except for the techniques that are used to detect and respond to cyber threats that exist in the wild, there is another aspect, which is fundamental for the proactive detection of network security incidents; the services that offer interexchange and knowledge sharing of cyber threat related information among organizations. These services and data feeds used for the proactive detection of network security incidents are either publicly available, commercial or require a user's subscription. These threat intelligence data feeds provide users with up-to-date information regarding potential attack sources and malicious entities. Such information sources are feeds of malware URLs, phishing sites, botnet command and control (C&C) servers, infected machines (bots), sources of

[16] https://github.com/sec51/honeymail
[17] http://dionaea.carnivore.it
[18] https://github.com/IoTPOT/IoTPOT
[19] https://github.com/pwnlandia/mhn
[20] http://conpot.org/

abuse (e.g., spam), defaced websites, vulnerable services (ENISA, 2020). In the following paragraphs, we present some of the most popular services that aim for information sharing across organizations. Other relevant feeds can be found in ENISA's GitHub repository that contains a comprehensive list with external information sources for proactive detection of incidents[21].

MeliCERTes: The Cyber Security Platform MeliCERTes is part of the European Strategy for Cyber Security and is a network that focuses on establishing confidence and trust among national CSIRTs of the member states by promoting operational cooperation and sharing focused on computer security incidents[22]. MeliCERTes is a platform that offers a security incident management solution, while it makes use of open-source projects; IntelMQ[23], Malware Information Sharing Platform (MISP) (Wagner, et al., 2016), Viper[24], OwnCloud[25], Jitsi[26]. IntelMQ collects and analyses security vulnerability events from multiple sources. MISP organizes the collected information, presenting each distinct vulnerability report as an event, which is then available for exchange among CSIRTs. Viper receives the events that are produced by MISP for critical malware analysis. OwnCloud is used to securely exchange files within Trust Circles, while Jitsi offers real-time communication channels meant for quick response and collaboration.

OASIS Cyber Threat Intelligence: The OASIS Cyber Threat Intelligence[27] supports automated information sharing for cybersecurity situational awareness and real-time network threat analysis. The platform is based on the Structured Threat Information Expression (STIX) language and serialization format for cyber threat intelligence exchange (Barnum, 2014). In addition, Trusted Automated Exchange of Intelligence Information (TAXII) is used, which serves as an application layer protocol for the secure communication of the cyber threat information over HTTPS (Connolly, et al., 2014).

The ShadowServer Foundation: The ShadowServer Foundation is a non-profit security organization that collects and shares threat related data (e.g., botnets, C&C and DoS reports). The data cover multiple types of network security incidents and the reports produced are shared across network providers, national governments and law enforcement[28].

Malware Domain List: The Malware Domain List offers data that concern malicious domain names, which are known to propagate malware[29].

Abuse.ch: Abuse.ch[30] shares malware-related data via different tools and services, such as URLhaus[31], SSL blacklist (SSLBL)[32] and MalwareBazaar[33]. URLhaus is a project that aims for sharing malicious URLs that are being used for malware distribution. SSLBL detects malicious SSL connections, by identifying and blacklisting SSL certificates used by botnet C&C servers.

[21] External information sources for proactive detection of incidents. Available in: https://github.com/enisaeu/IRtools/blob/master/information_sources.md
[22] https://github.com/melicertes/csp
[23] https://github.com/certtools/intelmq
[24] https://github.com/viper-framework/viper
[25] https://github.com/owncloud
[26] https://github.com/jitsi
[27] https://oasis-open.github.io/cti-documentation/
[28] https://www.shadowserver.org
[29] http://www.malwaredomainlist.com
[30] https://abuse.ch/
[31] https://urlhaus.abuse.ch
[32] https://sslbl.abuse.ch
[33] https://bazaar.abuse.ch

MalwareBazaar shares malware samples with the infosec community, AV vendors and threat intelligence providers.

The SpamHaus Project: The Spamhaus Project is an international non-profit organization that tracks cyber threats such as spam, phishing, malware and botnets. The SpamHaus project shares threat intelligence with networks, corporations and security vendors, and works with law enforcement agencies to identify and pursue spam and malware sources worldwide[34].

All of the aforementioned tools, techniques and services have been widely adopted within the cyber-security domain, aiming to timely detect a threat and efficiently prevent its spread. Over the last few years, a couple of works have presented some novel detection methodologies based on them, which require significant lower computational cost and perform faster comparative analysis in respect with older classical approaches. In (Naik, et al., 2019), it was addressed the necessity of dealing with the ransomware attacks across all kind of enterprises, mainly due to their polymorphic behaviour and their dominance over the rest of cybersecurity threats. The proposed methodology utilised a range of fuzzy hashing algorithms and clustering methods for the detection of ransomwares by attributing them with a similarity factor based on a sample dataset of such threats. On the other hand, (Cruz, et al., 2016) presented a domain-specific solution for SCADA systems, where a multi-layered Distributed IDS was deployed and evaluated against an electrical distribution grid with satisfactory results in the context of CIs.

[34] https://www.spamhaus.org/

Chapter 4 Anomaly-Based Detection Methods

Anomaly-based detection methods aim to identify unusual patterns that do not conform to the expected pattern of a target group. Anything that deviates from this expected behaviour is considered as an anomaly and could denote a serious cybersecurity threat to the system. An Anomaly Detection System (ADS) can utilize several techniques for the recognition of anomaly, which range from statistical and rule-based approaches, to machine learning and data mining approaches.

4.1 Statistical Approaches

One of the simplest approaches to identify anomalies in a dataset lies in the usage of a statistical data analysis procedure. Such approaches are capable of flagging those data points that deviate from a statistical distribution parameter, denoting in that way a potential threat to the system. Typical examples of such parameters vary from the arithmetic mean and standard deviation to regression models and hypothesis testing. Statistics-based techniques in an ADS, monitor every network flow of the system and are employed to search for activities that are not considered normal and compare them with normal ones. The normal activities, also known as profiles, can be user, network connections, network traffic, etc. The characteristics of this approach are simple to implement, it could be deployed in real-time, but it requires very good knowledge on statistics. The most widely used statistical approaches can be classified into Principal Component Analysis, Clustering, and Entropy Analysis.

4.1.1 Principal Component Analysis

A Principal Component Analysis (PCA) employs dimension reduction approaches that aim to make simpler the detection of potential cybersecurity attacks. It is a widely used procedure to reduce dimensions and summarize the data in large datasets for further analysis and visualization. It processes the data and tries to interpret their structure by means of a number of components which are the linear combinations of the original variables. PCA is often not a standalone approach rather the first step of the data analysis that could be later used by other techniques that are more multivariate. In the classical PCA approach, the first principal component corresponds to the direction in which the projected data points have the largest variance. The second component is then taken orthogonal to the first and must again maximize the variance of the data points projected on it. The procedure happens until all the principal components are produced. Unfortunately, it is very sensitive to anomalous observations, thus the outlying points can be captured early on and have a negative impact on the analysis (Rousseeuw & Hubert, 2018). In addition, it may not capture the variation of regular observations. The steps of the PCA anomaly detection algorithm are:

i. Provide a data set of normal activities and operations of the network that could be used as a training set for the PCA model in the next steps
ii. Pre-processes the provided data so they have zero mean and unit variance
iii. Train and deploy the PCA model using the training set consisted of the normal data flows
iv. Test the new data after scaling them down with the mean and standard deviation acquired from the training set
v. Use the trained model to check for anomalies of the new data

PCA has a reputation for being successful in monitoring systems with highly correlated variables. In (Harrou, et al., 2015) it was developed a PCA-based Multivariate CUmulative SUM (MCUSUM) strategy for the detection of small anomalies on the emergency department of a hospital centre. The results of the proposed algorithm showed that this MCUSUM solution was able to improve the anomaly detection capabilities of the tested system compared to the classical PCA-related approaches.

4.1.2 Clustering

Clustering methodologies aim to describe an available dataset by grouping similar items into several, but distinct categories known as clusters. These clusters afterward determine the key behaviour elements that could be used as a similarity measurement to detect various types of attacks. Cluster analysis is very useful when handling large datasets since it finds groups with similar characteristics in the data that can be analysed separately. Cluster analysis can be divided into two major categories (Bhuyan, et al., 2014).

The first category is known as partitioning clustering, which means that the method is breaking the dataset into separate groups. The methods that belong in this category search for the best available clustering in k groups. The most popular method is k-means which is used to find groups that have not been explicitly labelled in the data. It uses the squared Euclidean distances and minimizes the sum of those distances. An advantage of k-means is that it exceeds the hierarchical clustering in computational speed if the variables are huge and the k is relatively small. On the other hand, it is difficult to predict the k value thus k-means could not consider robust because it uses averages rather than specific values. A partitioning clustering that is more robust is the Partitioning Around Medoids (PAM) or k-medoids method. It is similar to the k-means method (both try to minimize the distance between points labelled to be in a cluster and a point designated as the centre of that cluster) but PAM uses data points as clusters' centres (medoids) while in k-means the centre of the cluster does not necessarily to be within the input data points. For a partitioning approach, if k (number of clusters) can be provided accurately then the task is considered easier. Incremental clustering (in supervised mode) techniques are effective for fast response generation while it is also advantageous in cases of large datasets that grouped into a similar number of classes for detecting network anomalies because it reduces the computational complexity during intrusion detection. It provides stable performance in comparison to classifiers or statistical methods.

The other category of clustering is hierarchical clustering that builds clusters in the form of levels. Unlike the previous category, the hierarchical clustering does not require a number of clusters during the initialization phase. Hierarchical clustering algorithms make use of either divisive methodologies, or agglomerative methodologies. Practical tests between those two types in both random and real-life data sets showed that no algorithm can be clearly declared as superior, but their performance and accuracy are rather based on the dataset used (Roux, 2018). In the divisive (or top-down) clustering method, the first step is to assign all the observations to a single cluster and then partition the cluster to two least similar clusters. This happens recursively for each cluster until there is one cluster for each observation. The inverse of the divisive method is the agglomerative method (or bottom-up) clustering method where each observation is assigned to its own cluster. The procedure is again recursive; thus, the similarity of the clusters is computed, and the two most similar clusters are joined. This happens until only a single cluster is left. In generic, a drawback of clustering-based methods is that in the majority the proposed techniques have been used to handle continuous attributes only. Also, clustering-based intrusion detection techniques make the assumption that the larger clusters are considered normal and smaller clusters as an anomaly (e.g. an intrusion). This assumption eases the evaluation procedures of the technique otherwise it would be harder to do so.

Clustering approaches like the k-means and the Expectation Maximization techniques have been also used for observation validation purposes in other studies. (Bou-Harb, et al., 2013) motivated by the lack of accurate enough scanning detection systems, proposed a novel fingerprinting statistical approach which combined several statistical techniques and probabilistic distribution methodologies like the Bhattacharyya distance (Kailath, 1967), Mann-Kendall (Kendall, 1948), and Wald-Wolfowitz (Friedman & Rafsky, 1979). Doing so, they were able to efficiently analyse probing activities and identify the scanning technique, the software tool or the worm/botnet used, as well as certain predefined patterns followed. Their empirical evaluation was performed on the analysis of a massive 55GB darknet traffic dataset, where the extracted inferences were not only promising in terms of accuracy, but they could be potentially used for threats' mitigation as well. Other works (Ye, et al., 2001) deployed clustering techniques for the efficient host-based intrusion detection, where a Statistical Process Control (SPC) methodology was responsible for the process stability and the reduction of variability. Several years later, a similar SPC approach was also incorporated as the core detection technique and threshold verification mean of a framework able to detect fast attacks from the victim perspective (Abdollah, et al., 2009). Their study proposed the usage of a dynamic threshold value to differentiate the normal and abnormal behavior in a network, depending on both real network traffic data and data coming from an experimental setup.

4.1.3 Entropy Analysis

Today, several traffic anomaly detection applications make use of an entropy-based approach to measure the randomness or diversity of their data-generated functions. Entropy is defined as the level of irregularities that occur in a system. There are many forms of entropy, but only a few have been applied to network anomaly detection. The most used form of entropy is the Shannon entropy (Shannon, 2001), that measures the entropy of the information content. In a few words, entropy-based anomaly detection consists of detecting sudden changes in the time series of the empirical entropy of specific traffic characteristics that are related to the specific anomaly. The entropy of a characteristic shows the scattering of the corresponding probability distribution in a single number, that helps the analysis. However, such a compression necessarily loses relevant information about the distribution of the analyzed feature. Besides Shannon, Titchener (Titchener, 1998) and parameterized Rényi (Yan, et al., 2008) and Tsallis entropies (Basicevic, et al., 2015), more generalized or specialized forms of entropy can be also found in network anomaly detection research. In the meanwhile, (Nychis, et al., 2008) had already proceeded to an empirical evaluation of several entropy-based metrics used for network anomaly detection purposes. The results of their study showed that port and address distributions do not suffice alone in a fine-grained anomaly detection system, while at the same time, traffic features distributions should be computed using bi-directional flow abstractions in order to avoid false positives. Over the last few year, (Bereziński, et al., 2015) presented a parameterized entropy combined with a supervised learning solution, which was able to outperform the classical Shannon-based, as well as the volume-based limitations referenced in the previous work.

4.2 Machine Learning Techniques

The fast-growing field of cybersecurity and the need of efficiently dealing with advanced threats in optimal times, led to the development and deployment of diverse Machine Learning (ML) methodologies. All studies in this area adopt the key properties of ML algorithms, which are no other than the scalability and adaptability. These features are of course deemed necessary in the development of CyberSANE's Cyber Fusion Models, in order to guarantee system's rapid response to new and uncharted security threats. ML algorithms can be categorized into four distinct categories depending on the approach followed during the process of their input and output data (Ayodele, 2010). Supervised learning approaches tend to make use of labelled instances as training data with the corresponding desired output. In those occasions, ML takes advantage of various supervised learning techniques to identify previously unknown events that differ significantly compared with the target investigated dataset. On the other hand, unsupervised learning approaches look for unidentified patterns in a dataset with non-existing labels in order to properly classify new instances with minimum human interaction. Semi-supervised learning lies between supervised and unsupervised learning by combining both labelled and unlabelled data during training. Semi-supervised learning approaches tend to find use in the detection of network anomalies with satisfactory results regarding threats' detection rate and the overall false positive rate. Last but not least, reinforcement learning approaches take into consideration the software agents of a system and the actions that have to be taken by them in conjunction with their environment. Their ultimate goal is to generate experience which could be used to maximize long-term rewards.

4.2.1 Intrusion Detection

One of the most important elements in a network, from the point of view of network management and cybersecurity, is to keep it safe from malicious intrusions. If a malicious attacker is able to access your network, it could have a massive impact in your system, ranging from economic losses, to downtime of services, data breaches, and loss of customer trust.

An Intrusion Detection System (IDS) is a cybersecurity tool that works with your network to keep it secure and inform when somebody is trying to break into the system. There are exist different types of IDS on the market and figuring out which one fits better with the needs a user may have is complex. An intrusion detection solution can be defined as the anomaly detection problem where a null hypothesis of no intrusion is assumed, the network is continuously trained with new data to deduct a “normal” behaviour, and once an intrusion hypothesis is applied then a certain set of anomalies should be evident in the model (Rubin-Delanchy, et al., 2016). Today, CIIs are composed of several network security systems where each one of them comes with its own IDS. An IDS aims to discover and identify external and internal intrusions, preventing in this way the leakage or alteration of sensitive information to unauthorized users. For that purpose, various ML methods (Buczak & Guven, 2016) have been proposed in the near past to deal with the cyber security intrusion domain. Their complexity and outcomes were heavily based on the datasets used for training and learning purposes (Tavallaee, et al., 2009; Yavanoglu & Aydos, 2017), where either a packet-level or flow-level dataset was selected and had its headers employed for the needs of intrusion detection. A couple of key challenges addressed in this study were the necessity of partitioning the input data streams, employing less difficult and time-consuming learning datasets, and collect as much results as possible working in a parallel mode.

As aforementioned, network intrusion detection can be performed in a high number of ways. This is one of the issues of why it is important to have an updated and strong intrusion detection solution that can cover the large number of attacks that could possibly occur in a system. Additionally, this is one of the most important issues that should be dealt in CIs. Therefore, it is quite crucial to understand the type of attacks that can be used in order to prepare an effective prevention. Usually, network intrusion is done using flooding techniques or overloading the network in order to gather data from it, so that it can be later attacked from a weak point or inserting malicious data in the system for gaining access. Among others, some of the most typical attacks that could target a system and be detected by an intrusion detection solution are the following ones.

Malware: There exist several different types of malware that could affect a system. Among others, the most commonly used are viruses, trojans, bots, etc. Each type has a different way of working and how it compromises a system, ranging from simply blocking a system, to running in the background and compiling information of users or the system for years without no one noticing it.

Scanning attacks: This attack involves sending data packages (information) to the network in order to obtain knowledge about the network topology, ports (open or closed), type of traffic allowed in the network, etc. This type of attacks usually scans ports in order to introduce malicious applications (virus) or code.

Asymmetric routing: Attackers use asymmetric routing (when the data in a network uses a different path for sending and receiving) in order to send malicious data to the network so it can bypass cybersecurity setups (e.g. firewalls).

Buffer overflow attacks*:* This type of attack is very common and is used for penetrating sections of memory of devices on a network in order to replace the regular data with malicious ones that could subsequently used for executing an attack.

Traffic flooding: This type of attacks is more commonly known as Denial of Service (DoS) and Distributed Denial of Service (DDoS). Its main objective is to saturate the network so it can be disabled, or make it easier to be penetrated by other type of attacks.

Protocol-specific attacks: Specific attacks targeting protocols used for network communication such as TCP, ICMP, etc.

CIs like power infrastructure control systems have been proved to be a common target of various cyber-security threats, mainly due to absence of physical security in the context of substations (Govindarasu, et al., 2012), as well as the lack of mature anomaly detection technologies for those substations. Simultaneous and continuous cyber-attacks could even lead to catastrophic cascading events like a power outage of the infrastructure. (Hong, et al., 2014) proposed a novel integrated anomaly detection algorithm, capable of monitoring critical devices and communications (intelligent electronic devices, transformers, circuit breakers, etc.) for intrusion detection. Their approach was based on both host-based and network-based detection methodologies, where the first is optimized for attack similarity and threats' determination according to discrepancies found in logs from different time periods (Ten, et al., 2011), and the second is responsible for the multicasting of messages regarding a malicious behaviour with the usage of Generic Object Oriented Substation Event (GOOSE) and Sampled Measured Value (SMV) techniques.

On the other hand, (Yan & Zhang, 2013) presented a novel intrusion detection system based on a language-based approach for the early detection of various types of cyber-threats, including sophisticated versions of Advanced Persistent Threats. At first, their system converted the low-level network traffic into a trace sequence using the DBSCAN algorithm (Ester, et al., 1996). Afterwards, the Helix model (Peng, et al., 2011) was deployed as the medium of grammar induction, upon where grammar rules were authored to parse new network trace sequences into structural representations. The experimental results of this structured modelling of network traffic behaviour showed considerable higher precision and recall with regard to the KDD99 dataset (Lee & Stolfo, 2000). Such datasets are continuously used in literature and enhanced on a regular basis to provide a sufficient set of both normal and simulated attack activities. Doing so, the intrusion detection of existing or novel approaches is evaluated and benchmarked against both known and unknown cyber-threats. (Meira, et al., 2018) took advantage of the publicly available datasets of NSL-KDD (Noto, et al., 2012; Dua & Graff, 2019) and ISCX (Shiravi, et al., 2012) to measure the performance and novelty detection capabilities of one-class unsupervised algorithms like the autoencoding neural network, k-means, nearest neighbour and Isolation Forest. The results showed that all these techniques generate a lot of false positive alerts, but once their data are processed before the execution of outlier detection, then they are able to detect most of the anomaly instances. For that purpose, it was proposed the application of a holdout methodology to sufficiently train each algorithm, followed by the discretisation of continuous features using an equal frequency technique, and finally the provision of a data normalisation methodology to uniformly scale all features.

Another set of complex components integrated today on several CIs which have to be also shielded against cyber-attacks are the Industrial Control Systems (ICS). The protocols designed for ICS suffer from the lack of encryption on application layer, are prone to network packets' interception tactics, and are commonly compromised to perform DoS attacks (Brenner, 2013; Long, et al., 2005). A novel approach for the detection and prevention of cyber-threats on ICS was presented in (Brugman, et al., 2019), where Software Defined Networking (SDN) was deployed to route their network traffic on a cloud infrastructure for further security checking. Their Cloud Based Intrusion Detection and Prevention System (CB-IDPS) took advantage of the cloud computing services provided by Amazon[35] for the inspection of data in a virtual private cloud using Network Function Virtualization (NFV) and service function chaining techniques. Their architecture also employed several other tools like the OpenDaylight[36] for network's SDN controller, as well as the Zeek[12] and Snort[10] for the monitoring of traffic. The results of their work were quite promising in the context of scalability and resilience, improving at the same time any delay constraints and routing issues reported in previous works related with SDN and NFV techniques (Yu, et al., 2017; Kumar, et al., 2017).

4.2.2 Advanced Persistent Threats & Response

Advanced Persistent Threats (APTs) is a class of threats that keeps growing exponentially (Singh, et al., 2019) across public sectors, including of course those of CIIs. APT differentiate with the traditional cyber security attacks since they operate in a low and slow profile of attendance, making their detection and mitigation of effects very difficult to identify. Such threats encompass a set of sophisticated attack techniques and are usually well-resourced by organisations and NGOs or even state sponsored. An APT can be classified into two stages, the intrusive and disruptive stage. During the former, an attacker tries to identify the security tools used and collects as much data as possible, while during the latter the attacker aims to disrupt the operational state of one or more components of the CI (Tankard, 2011). Even though dealing with APT is a challenging task, the latest advances in AI and ML has made possible the faster detection and mitigation of their effects (Oluwasegun & Aminat, 2019; de Abreu, et al., 2020).

The likelihood of the early detection and analysis of an APT attack was thoroughly investigated in (Bhatt, et al., 2014), who presented a research framework capable of handling complex multi-stage APT attacks. The core of the framework was composed of a layered security architecture where the access to a layer was possible only by processes and applications of its immediately outermost layer. The treatment of a possible attack across any layer was utilized by a hypothesis attack model that adopted the seven phases of an Intrusion Kill Chain (IKS) proposed in (Hutchins, et al., 2011) and depicted in Figure 5.

The effective detection and analysis of a threat was based on an IDS backed by an Apache Hadoop[37] infrastructure. The IDS was responsible to collect all alerts and logs coming from the system's sensors, which were configured and triggered in accordance with a predefined set of malicious behaviours. Hadoop was chosen as the storage and data correlation solution thanks to its high availability and fault tolerance (Cowsalya & Mugunthan, 2015), and it was divided into five modules to sufficiently deal with the logging management, malware analysis and system administration tasks. The APT intrusion kill chain and the need to detect and prevent specifically the insider threats in a timely manner also concerned (Liu, et al., 2018). Their study was based on a data analytic perspective of the host, network, or contextual source of extracted information, and they managed to correlate the available detection or prevention algorithms with one or more data sources in distinct taxonomies. However, this research showed that most of defenders have limited APT attack tactics understanding, leading to poor network traffic analysis, classification, and detection capabilities. A

[35] https://aws.amazon.com/
[36] https://www.opendaylight.org/
[37] https://hadoop.apache.org/

potential solution to this issue was proposed in (Chizoba & Kyari, 2020), where APT attack tactics are first modelled, and afterwards both their dataset and a real-world “clean” dataset are generated through simulation. Upon them an ensemble of classifiers composed of Support Vector Machine, Random Forest, and Decision Tree algorithms battled for majority voting to provide the best possible attack classification and detection accuracy.

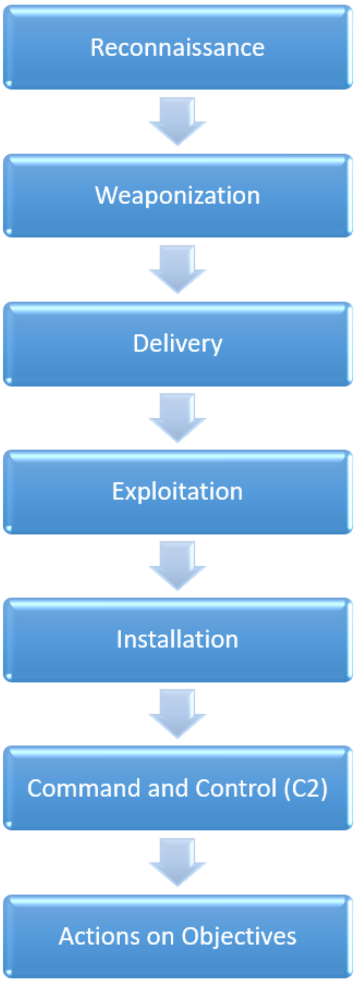


Figure 5. Intrusion Kill Chain (IKS)

Once more, (Sinha, et al., 2020) clearly noted the vital importance of energy CIs to our society, as well as the disastrous consequences in case of their disruption due to an APT cyber-attack among others. The necessity of guarding a power utility infrastructure was also addressed in the near past by (Hasan, et al., 2019), where they deployed ML techniques for the mining of network data, in order to detect the different stages of a potential APT attack. Last but not least, taken into account that power grids are one of the most crucial CIs which are continuously targeted by APTs, (Tian, et al., 2020) introduced an APT-honeypot game to study the offensive and defensive interactions made between their attackers and defenders. Their results were based on the prospect theory instead of the classic utility theory and showed that honeypots can be used for both offense and defense purposes on such CIs, while the bounded rationality affects the Bayesian-Nash equilibrium strategy followed, reducing thus an attacker’s payoffs.

4.3 Data Mining Techniques

The proactive detection of cyber-security threats has been also transformed into a big data problem since it involves the analysis of an increasing volume of data. The latest works in this area employ data mining, pattern matching algorithms or other reasoning approaches (Tianfield, 2017) to detect unexpected behaviour and prevent a possible compromise of the system. There are two types of intrusion attacks which can be detected using data mining methods. First are the host-based attacks when the intruder focuses on a particular network entity or a group of them, and the second are the network-based attacks, when the intruder attacks network components and tries to change their performance and status. The data mining techniques can be divided into three categories (Agrawal & Agrawal, 2015) which are the clustering, classification and the combination of these two, the hybrid technique.

The clustering technique which divides the data into similar groups is used because there is no need for prior knowledge. Some approaches are:

i. k–means: the user defines k-clusters and the method groups the data into k groups. This method can be used for fast anomaly detection of new incoming data;
ii. k–medoids: similar to k-means but more robust. It is not so tolerant of outliers and noise. This method performs better than the k-means and can detect network anomalies that have unknown intrusion;
iii. EM Clustering: something of an extension of the k-means since it considers the mean of the cluster when it decides the cluster assignment. It is a weighted approach of the previous two and it outranks them in performance accuracy;
iv. Outlier Detection Algorithms: this method tries to find patterns in data that are not expected (outliers). Some examples are the Distance-based Approach that is based on the Nearest Neighbour. This approach examines the distances of data points to calculate how far they are from their neighbours. It is used and proved to be effective in detecting DoS attacks. Another method is the density-based local outlier approach that provides each data point with a degree of being outlier based on its local neighbours.

The classification technique for anomaly detection identifies the categories of new instances after the training of some known observations. The observations in this technique are divided into normal and abnormal. Some known classification techniques are:

i. Classification Tree: follows the structure of a flow-chart, also known as a decision tree. The most common algorithms are ID3 and C4.5 which follow the top-down construction scheme;
ii. Fuzzy Logic: calculates the degree of membership for each input of the data and based on pre-defined rules, the output is produced (normal or malicious);
iii. Naïve Bayes network: it provides efficiency in large datasets and takes advantage of the correlation between the variables. It calculates conditional probabilities and is typically accompanied by a directed acyclic graph where there are weighted links between the variables;
iv. Genetic Algorithms: are used for optimization problems and are inspired by natural evolution (usage of mutation, selection, etc.). They provide classification rules from the data and they fit those rules for optimal solutions. Also, they are robust against noise, they have a high detection rate and low false-positive rate. For the purposes of this report though, Genetic Algorithms shall be described in section 4.5 below;
v. Neural Networks (NNs): are inspired by the structure of the human brain. The neural models that are used for anomaly detection can form any classification decision and can solve any problem as long they have the right structure (enough hidden layers);

vi. Support Vector Machines (SVMs): are supervised learning methods for classification and are widely used. For anomaly base detection, the one-class SVM is based on normal behaviour that detects rare events (e.g. attacks). SVM is a max-margin method, which tries to find a function that is positive for regions with a high density of points, and negative for small densities.

The final technique is the hybrid approach that is kind of a combination of the previous two. That allows combinations of supervised and unsupervised techniques (like k-means and ID3, Naive Bayes and decision tree, k-medoids and Naïve Bayes, etc.). Hybrid approaches have better results in classification, because they can leverage the best performing features of each of the previous techniques and combines them resulting in higher accuracy of anomaly detection. A set of advanced anomaly detections systems that make use of data mining techniques to detect novel network intrusions based on both known and unknown attack patterns have been introduced before in (Barbara, et al., 2001; Ektefa, et al., 2010). Taken into account their anomaly detection capabilities and the number of false positives, ADAM, IDDM, and MINDS tend to be some of the most prominent solutions on cyber-security domain.

4.4 Deep Learning Techniques

Deep Learning (DL) is a subset of ML that uses Artificial Neural Networks (ANNs) as algorithms to solve the problems they are presented with. ANNs are logic structures inspired by the way the human brain works. ANNs consist of an input layer, an output layer and one or more hidden layers between them. The ANNs with only one hidden layer are known as Shallow Neural Networks, as opposed to Deep Neural Networks, which have several hidden layers. The advantage of Deep Learning (DL) over traditional ML techniques, it is that it does not require labelled data for training. The disadvantage is that it requires a much larger amount of high-quality training data to achieve proficiency in resolving a certain problem.

Cyberdefence mechanisms can be met today across numerous hardware and software levels of a CII, ranging from the application and network level, to the host and data level. Such mechanisms aim to efficiently deal with APT and prevent attacks by detecting intrusion attempts as well as other security breaches. Over the last few years several DL techniques on cybersecurity have been emerged as advanced and alternative solutions to the traditionally ML techniques. Most of those techniques are enhancements to existing ANN solutions with Deep Autoencoders, Restricted Boltzmann Machines, Recurrent Neural Networks, or Generative Adversarial Networks (Berman, et al., 2019; Imamverdiyev & Abdullayeva, 2020).

An early work that adopted an RNN approach for anomaly detection purposes took place in (Debar, et al., 1992). The RNN was continuously trained with Unix command-line arguments to monitor user activity and predict a potential network-level intrusion. However, network's training was not able to keep up with the increasingly change of users' habits over time, limiting thus its intrusion detection capabilities greatly. Several years later, (Veeramachaneni, et al., 2016) developed an ANN deep autoencoder capable of aggregating numeric features over a time window from web and firewall logs. These features were used as input to train an unsupervised anomaly detection system composed of various techniques like the principal component reconstruction of signal, and a multivariate probabilistic model over the feature space. Their solution was also able to periodically incorporate analysts' feedback in order to keep up with the latest cyber-security advances and improve its detection accuracy. A few years ago, taken into account the aforementioned works and the key-difficulties met in cyber-security domain regarding the application of ML techniques (Sommer & Paxson, 2010), an online unsupervised DL approach for the detection of anomalous network activity in real-time based on system logs was presented (Tuor, et al., 2017). This approach made use of model variants of trained Deep Neural Networks (DNNs) and RNNs, in order to recognize and attribute network activity to specific users and consecutively assess if such activity is normal or malicious. The results were more than promising and showed that the novel decomposition of anomaly scores into the contributions of individual user behaviours, achieved improved anomaly detection in regards with insider threats, outperforming competitive PCA, SVM and Isolation Forest baselines.

4.4.1 Deep Learning Approaches for Anomaly Detection

DL models have been broadly used to detect outliners. This section describes some of the most relevant DL models applicable for anomaly detection.

4.4.1.1 Deep Neural Networks

A Deep Neural Network is an ANN with more than one hidden layer. They usually are feedforward networks in which data flows from the input layer to the output layer without looping back. This basic approach can be used for anomaly detection but is not as broadly used as other more complex ANN models, such as the Recurrent Neural Networks or the Convolutional Neural Networks, which are usually capable of achieving better performance.

4.4.1.2 Recurrent Neural Networks

Recurrent Neural Networks (RNNs) are ANNs where connections between units form a directed cycle. This enables the RNN to possess a dynamic behavior in time.

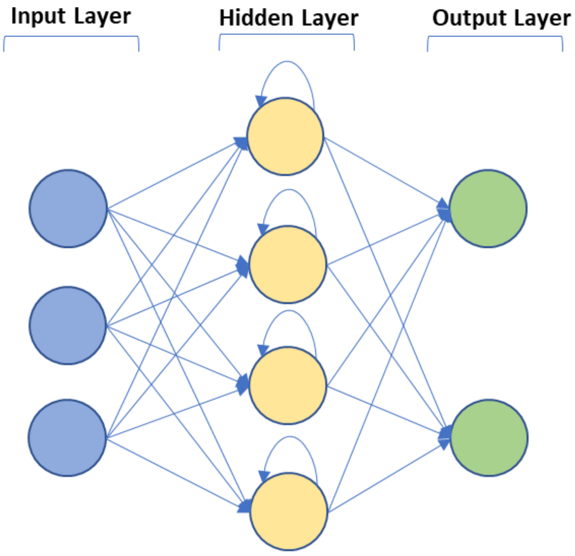


Figure 6: Example of a Recurrent Neural Network

RNNs have been proposed as a way detecting network traffic anomalies (Radford, et al., 2018) as an effective and unsupervised tool, and they propose combining it with other supervised and human-assisted methods in order to achieve performance improvements. (Goh, et al., 2017) used an RNN to detect network anomalies with the Secure Water Treatment (SWaT) testbed dataset[38], which represents an industrial water treatment plant. Compared to other models that obtain data in a single second, the presented RNN approach takes into consideration and correlates a sequence of time-series data, achieving thus a lower false positive rate. Their approach was also capable of identifying which sensor the anomaly is occurring at.

(Kim, et al., 2016) implemented the IDS classifier based on LSTM-RNN and evaluated the IDS model. They used some instances of the KDD Cup 1999 dataset[39] for training. They reported a 98,8% detection rate among the total attack instances, but with a slightly high false positive rate, namely 10%.

4.4.1.3 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are a subtype of deep neural networks, which are mostly used for image recognition. In an extremely simplified way to describe them, they are a subcase of DNNs formed by the following layers:

i. Convolutional layers apply a mathematical operation named convolution[40] multiple times to the input to generate feature maps
ii. Sub-sampling layers reduce the size of the feature maps to decrease the resource consumption and to avoid overfitting
iii. A fully connected layer makes the classification of the input

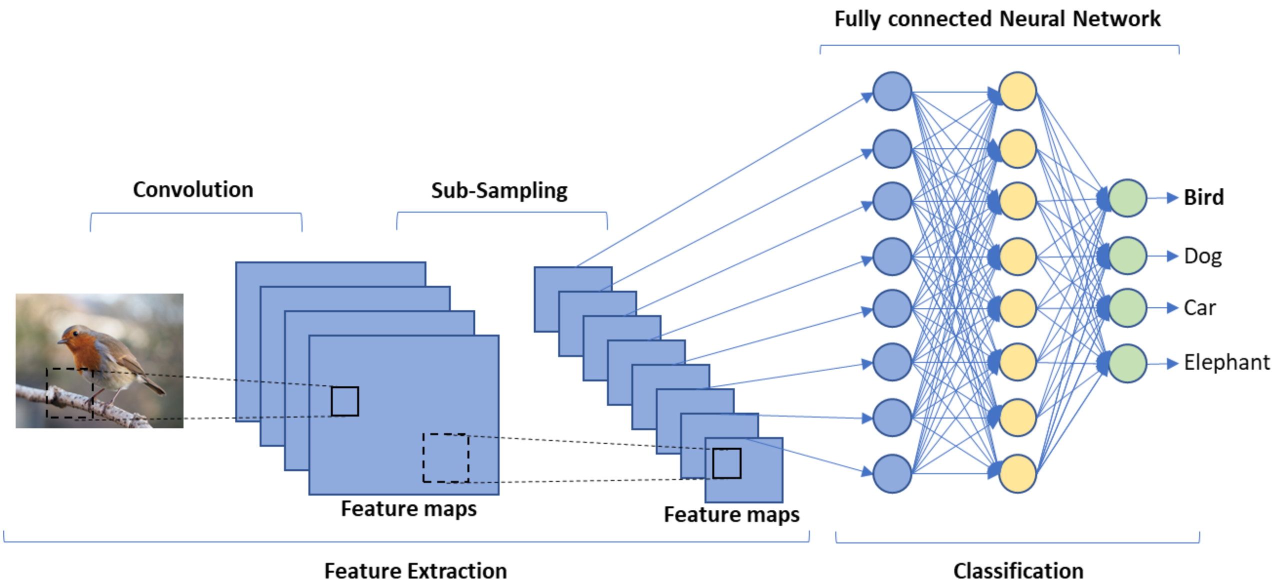


Figure 7: Example of a Convolutional Neural Network

[38] https://itrust.sutd.edu.sg/testbeds/secure-water-treatment-swat/
[39] http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html
[40] https://en.wikipedia.org/wiki/Convolution

CNNs have been successfully used for anomaly detection in several studies. In (Kravchik & Shabtai, 2018), CNNs are applied to detect cyberattacks in industrial control systems. The dataset used was the SWaT dataset -already mentioned in the RNN section- which represents an industrial water treatment plant. This dataset includes 36 different cyberattacks. The experimental results showed that the CNN was capable of detecting 32 of them. The remaining 4 were not being detected because as stated in the dataset description, they failed to have the expected impact on the system.

4.4.1.4 Restricted Boltzmann Machines

Boltzmann Machines (BMs)[41] are defined as *"a network of symmetrically connected, neuronlike units that make stochastic decisions about whether to be on or off"* (Hinton, et al., 1984). A Restricted Boltzmann Machine (RBM)[42] is a type of Boltzmann Machine in which a node is connected to all the nodes in the opposite layer by symmetric (non-directional) connections, but at the same time it is not connected to any other node in its own layer. This restriction significantly reduces the training costs, compared to the BMs.

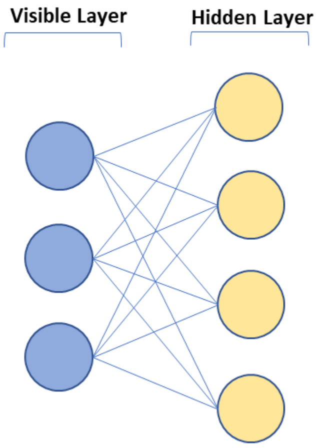


Figure 8: Example of a Restricted Boltzmann Machine

In the near past, RBDs have been successfully used for intrusion detection purposes. A study conducted by (Fiore, et al., 2013) proposed *"using the Discriminative Restricted Boltzmann Machine to combine the expressive power of generative models with good classification accuracy capabilities to infer part of its knowledge from incomplete training data"*. It is worth also mentioning that RBDs have been also presented as an alternative network anomaly detection approach in cloud-based infrastructures (Monni, et al., 2019).

4.4.1.5 Deep Belief Networks

A Deep Belief Network (DBN) is a type of DNN which consists of several layers of stacked RBDs.

[41] https://en.wikipedia.org/wiki/Boltzmann_machine
[42] https://en.wikipedia.org/wiki/Restricted_Boltzmann_machine

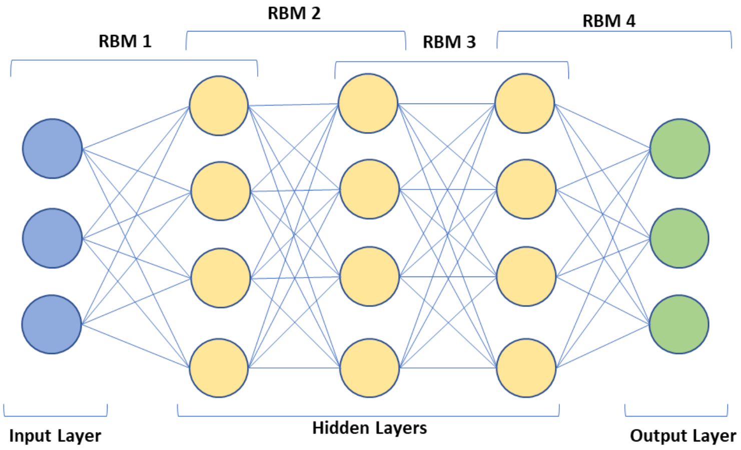


Figure 9: Example of a Deep Belief Network

(Sharma, et al., 2016) explored DBN for anomaly detection comparing their performance to that of a “classic” neural network. The conclusions of their study showed that the presented deep belief architecture performed better compared with the classic approach.

4.4.1.6 Deep Autoencoders

Deep autoencoders are a type of DNN intended to reproduce in their output a compressed copy of the data received in the input layer with the minimum possible loss of data. They are composed of two opposed deep-belief networks, one for encoding the input data (encoder) and a second to decode them (decoder). The encoder learns to translate the input to a low dimensional copy, while the decoder tries to restore it with as minimum as possible deviation. However, an amount of data loss during the encoding procedure is desirable, as it aims to preserve only the “key features” of the input, allowing thus the autoencoder to map other similar inputs to the same output and enabling pattern recognition.

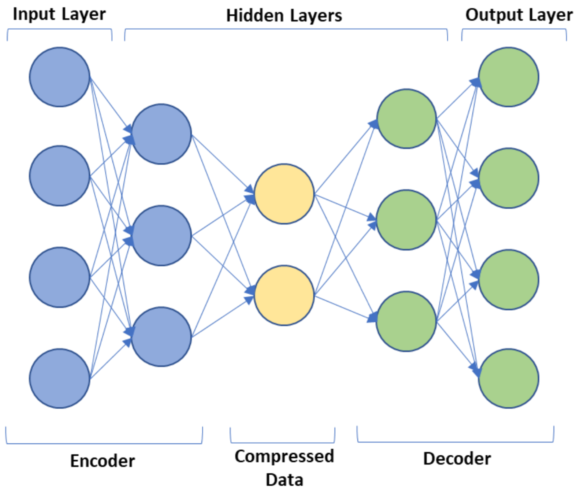


Figure 10: Example of a Deep Autoencoder

Deep Autoencoders have been broadly used for anomaly detection usually in combination with other models or with modifications, such as the MemAE (Gong, et al., 2019), a Deep Autoencoder augmented with a memory module. The presented modification was intended on purpose, since Autoencoder tended to reconstruct the anomalies so well that they were not recognizable anymore. Last but not least, (Amarbayasgalan, et al., 2018) proposed the combination of a Deep Autoencoder with the Density Based Spatial Clustering algorithm DBSCAN[43], as they observed that this combined solution obtained better results.

4.4.1.7 Generative Adversarial Networks

Generative Adversarial Networks (GANs) were firstly introduced by (Goodfellow, et al., 2014). In their work they proposed pitting a generative model (such as an RBM or a DBN) against a discriminative model (such as a RNN or CNN). They used the analogy of the generative model being like a team of counterfeiters, and the discriminative model being the police. This competition game makes both adversaries to continuously improve their methods. A few years later, (Di Mattia, et al., 2019) conducted an extensive survey on GANs and applied several state-of-the-art approaches for anomaly detection purposes, aiming to enhance the empirical validations across different datasets. Their application to cyber-security was also analyzed by Harris in a two-part article (Harris, 2018; Harris, 2018).

4.4.2 A Comparison Between DL Models

Many of the approaches listed in this section are described and compared in (Ferrag, et al., 2020). The authors performed a comparative experiment on DNNs, RNNs, CNNs and Deep Autoencoders using two real traffic datasets, the CSE-CIC-IDS2018[44] and the Bot-IoT dataset[45].

The DNN performance in classification was usually the lowest, but it required a lower training time compared with the rest of the approaches. RNN showed both good classification performance and satisfactory anomaly detection accuracy, but it also illustrated a slightly higher percentage of false positives than the CNN, DBN and Deep Autoencoder approaches. CNN presented in average the best performance for deep discriminative models, and its results were very close to those of the Deep Autoencoder, the best performance among all the generative models. RBMs and DBN shown an average performance higher than the average performance of the deep discriminative models, but lower than the Deep Autoencoder. The same study also compared the aforementioned DL approaches with three different ML techniques, namely the Random Forests, Naïve Bayes and SVM. According to their experimental results, the presented DL techniques were able to clearly outperform all three of them.

DL approaches have been proven suitable for cyber-security anomaly detection in several studies, with high detection rates and low rates regarding false positives results. CNNs and Deep Autoencoders are two of the most common approaches that show a higher performance. There is not a specific technique that clearly outperforms the rest in terms of performance and accuracy, since sensitivity, specificity, true positive, and true negative rates vary depending on the type of

[43] https://en.wikipedia.org/wiki/DBSCAN
[44] https://www.unb.ca/cic/datasets/ids-2018.html
[45] https://www.unsw.adfa.edu.au/unsw-canberra-cyber/cybersecurity/ADFA-NB15-Datasets/bot_iot.php

attack evaluated and the dataset used. In order to try to achieve a higher accuracy at the cost of simplicity, these models could be combined using a Committee Machine[46] approach, on which the output of several neural networks are combined to achieve a solution superior to those provided by the individuals.

4.5 Genetic Algorithms

Genetic Algorithms (GAs) are metaheuristic procedures inspired by the process of natural selection and constitute a subset of the evolutionary computation domain. Their main functionality borrows concepts from biological operations such as the mutation and crossover, in order to generate optimized solutions or resolve search problems during the detection of cyber-security attacks. They tend to find application on problems that present several local optima and their solution requires more intelligent approaches compared with the standard formula-based algorithms. Multi-objective GA approaches have been proposed before as the median to deal with the identification and minimization problem in realistic case studies. Such a study takes place in (Zio & Golea, 2012), where GA-based algorithms take advantage of the Italian's high-voltage electrical transmission network's topology (HVIET) to analyse and identify the most critical group of edges among all the nodes participating in this specific CI. The results were able to imprint the critical points of the network's functional relationship with a minimum only information of the topology beneath, but they also marked the necessity to include the physical characteristics as well in order to provide more robust and realistic insights.

A few years later, (Hamamoto, et al., 2018) proposed a high-accuracy and low computational cost methodology which combined both GA and fuzzy logic approaches for network anomaly detection purposes. Their aim was to provide a Network Anomaly Detection System (NADS) capable of detecting network anomalies autonomously and alerting the appropriate individuals whenever a potential threat was discovered. The prediction and detection of anomalies was taken place in two phases and was based on six distinct attributes extracted from IP flows data. During the first phase, a GA-related implementation was responsible for the prediction of network's behaviour, based on the characterization of network's traffic and thresholds' calculation using data collected from network's assets. On the second phase, a fuzzy logic approach was used as the evaluation mechanism to determine whether an anomaly was present, using the afore-mentioned parameters and an exponentially weighted moving average statistic (Cisar & Cisar, 2007). At its core, the presented profile-based anomaly detection system was heavily relied on traffic behaviour characterisation using GA concepts like the fitness evolution at a specific time interval. This network profile was attributed with burst cycles which included characteristics related with the user activity and the daily workload of a system (Proença, et al., 2006), in order to ultimately create a Digital Signature of Network Segment using Flow analysis (DSNSF). This signature stored all the necessary information about the expected traffic behaviour of the system, which was checked for possible behavioural deviances that could indicate a DoS, DDoS, or other type of cyber-attacks on the premises of an organisation like a CII.

[46] https://en.wikipedia.org/wiki/Committee_machine

Chapter 5 Risk Assessment, Cascading Effects & Simulation Environments

This chapter aims to address both the most influential and the latest works on risk assessment methodologies, cascading effects of cyber-attacks, and the simulation environments that could provide a holistic approach to the modelling of such attacks. Several cyber-security approaches, techniques and algorithms are presented and thoroughly described in the context of SCADA, ICT systems, and CIIs.

5.1 Risk Assessment Methodologies

Risk assessment methodologies involve some of the best preventive activities to protect the CyberSANE system and its components. The periodic execution of risk assessments is able to unveil potential risks to the system, enabling their future monitoring and prompting the definition of an incident handling approach for them. Such practices can definitely reduce the number of cyber-security incidents, as well as the applicability of their cascading effects in the critical sectors of the platform.

5.1.1 Risk Assessment Methodologies for SCADA & ICT Systems

In recent years, SCADA and ICT systems are critical components of industrial automation systems designed to collect and store data, to delineate and control industrial processes (Mattioli & Moulinos, 2015). They play a vital role in the efficient operations of CIIs of most Industry sectors, such as Energy (e.g. smart grid, power plant, energy management system), Maritime Transport (e.g. Cargo handling System, Vessel Tracking system, PLC marine pump and valve control) and Healthcare (e.g. web-based patient monitoring system, implantable cardiac defibrillators, Wireless Insulin Pump). In fact, most physical processes of Industry services are executed with autonomous or semi-autonomous mechanical, physical systems and machineries under the supervision of such sophisticated systems (Kalogeraki, et al., 2018). The importance of SCADA and ICT systems in critical infrastructure operations and the high impact of a security breach in such systems, attracts the attention of adversaries to conduct malicious activities including data leaks, damage, corruption, cyber espionage, robbery and physical/cyber-attacks capable of interrupting Industry operations, that can cause economic loss even political disruption, environmental harm and human casualties (Mattioli & Moulinos, 2015). Due to this high and multi-level impact that can be activated by the implementation of such security challenges, there is an urgent, pressing requirement for IT specialists and security officers and Industry operators to protect their interconnected SCADA and ICT systems (e.g., telemetry systems, data controllers, RTUs, audio-visual systems, satellite networks, etc.) (Kalogeraki, et al., 2018).

ENISA (Cadzow, et al., 2015) defines the concept of security in CIIs as the need for security coverage to a bunch of processes, techniques, and technologies related to CIIs. Information security enfolds a set of measures to be undertaken by supply chain operators, in order to protect and defend their ICT systems and the information processed within a system (which can be both cyber and physical) from malicious activity (e.g. unauthorized access, information leakage, modification, etc.) or destruction (ENISA, 2016). The implementation of cybersecurity measures on the interconnected

CIIs is the primary action to maintain security on their performances, which should be tailored by the well-known CIA triad information security model, that was first quoted in the late ‘70s by NIST: Confidentiality, Integrity and Availability. In recent years, this model embeds additional concepts, such as authenticity, accountability, non-repudiation, and reliability which should all be taken into account in an efficient balance (ISO/IEC JTC 1/SC 27, 2018; ENISA, 2016).

Security measures could be adopted on CIIs amid security risk management processes and risk assessment implementations (ENISA, 2016). The ISO27000:2018 standard (ISO/IEC JTC 1/SC 27, 2018) defines risk assessment as the overall process of risk identification, risk analysis and risk evaluation. The risk assessment process can be analysed in terms of conducting assessment of the vulnerability level, threat level, risk level and impact level. The outcome of a risk analysis is in most cases a list of risks or threats to a system, together with the corresponding probabilities following national or transnational or International standards in the field of risk management. A systemic review has been carried out to gather relevant existing literature regarding risk management methodologies to adumbrate cutting-edge issues and elicit important challenges.

The underlying principles of the Risk Assessment (RA) process are captured in the National Academy of Science (Red Book) (National Academy Press, 1983) where assessment and decision-making are distinguished (Stouffer, et al., 2015). Risk reflects three basic concepts: event, likelihood, and severity. Nevertheless, the main focus is on undesirable events which pose a loss in a specific context (a set of negative circumstances). A risk event can be certain or uncertain and can be influenced by a single occurrence or a series of occurrences. Likelihood indicates the frequency of an event is probable to occur. An event is modelled via likelihood of uncertainty by several mathematical theories such as probability theory (Ross, 2014), expected utility theory (Hogarth, 1987), Dempster-Shaffer theory of evidence (Shafer, 1976) and fuzzy set (Zadeh, 1965). These theories are developed for different purposes and refer to different classes of uncertainties. There are three categories of uncertainty are defined: aleatoric (randomness), epistemic (incompleteness) (O’Hagan, et al., 2006) and imprecision (vagueness) (Smithson, 1989). Epistemic and imprecision, are often met in software projects and pose for a potential risk.

There is a variety of past, well-known, outstanding risk management methods and risk assessment tools, which can be found in ENISA’s inventory of risk management and RA methods (ENISA, 2020), such as the ISO 27001-, 27005- and 31000- compliant “EBIOS” method used by ANSSI (National Cybersecurity Agency of France), the “OCTAVE” method (Alberts & Dorofee, 2002), a-priori distribution of subjectively estimated probabilities utilizing the Bayesian approach using UML modelling language, the Magerit open methodology for risk analysis and risk management and the Mehari method for harmonized risk analysis. Most of these methods and tools apply the commonly known rule of thumb "risk = probability x potential damage" (Zambon, et al., 2011). Additionally, traditional risk assessment approached are the “BowTie” (qualitative risk analysis method) and “CORAS” method recognizing the probability of an attack (Djordjevic, et al., 2002). According to (Theocharidou & Giannopoulos, 2015), existing risk management policies are using their own disparate methodologies with the absence of a common methodology and terminology, especially concerning CI-related risk assessment, which hurdles the comparison of risk assessment results among the EU Member States and greatens the appearance of cross-border multi-risks across multiple sectors.

Estimation of security risks on SCADA and ICT systems, assumes deep analysis and comprehension of parameters, such as the causes of vulnerabilities. In fact, with respect to SCADA systems, risk is assumed “a function of the likelihood of a given threat-source exploiting a potential vulnerability and the resulting impact of a successful exploitation of the vulnerability” (Theocharidou & Giannopoulos, 2015). The NIST SP 800-82 Rev.2 publication is devoted on SCADA system topologies to identify typical threats and vulnerabilities on such systems, presenting recommended security countermeasures to mitigate the associated risks (Stouffer, et al., 2015). Most SCADA and ICTs began as proprietary, stand-alone systems that were separated from the rest of the world and isolated from most external threats. Nevertheless, more recent SCADA systems have moved to more interoperability and open standards for cost efficiency and integration into management IT systems. For example, communication is now common over Ethernet TCP-IP including more standardized control protocols and applications. Open standards for SCADA systems are sources for adversaries to gain knowledge regarding the SCADA network topology (Igure, et al., 2006). (Permann & Rohde, 2005), propose the following steps for assessing a SCADA system including reconnaissance procedures to gather information on the target system, perform vulnerability Scanning within the SCADA network, meet the targets of evaluation (TOEs) identified in the assessment plan. In addition, they are presenting a list of open source and commercial tools for assessing SCADA systems (i.e. NMAP, NESSUS, STAT SCANNER, ETHEREAL, ETTERCAP, DEBUGGERS, FUZZERS, etc). A quantifying vulnerability method for critical infrastructures is introduced using the Infrastructure Vulnerability Assessment Model (I-VAM) by (Ezell, 2007). (Cheminod, et al., 2013) have presented the Quantitative modelling SCADA vulnerabilities CRA. Hence, SCADA systems are subject to external attacks and IT-based vulnerabilities. In general, the underlying causes of vulnerabilities in SCADA system architectures are the following (Kalogeraki, et al., 2018):

i. the misconfiguration of wireless devices;
ii. the high level of interdependency among transportation infrastructure systems;
iii. deficiencies in security controls as lack of cryptography policies used in SCADA networks (Igure, et al., 2006) or unskilled, naive employees revealing passwords to colleagues ignoring the potential risk (Daryabar, et al., 2012);
iv. the accessibility of systems via networks, devices and software components either directly (wired) or remotely (wireless) for scheduled or corrective maintenance purposes.

A scenario-based approach to risk analysis in support of cybersecurity has been introduced (Gertman, et al., 2006). In 2009, a cyber-terrorism SCADA risk framework has been presented (Beggs & Warren, 2009). The Institute for Information Infrastructure Protection (I3P), founded by the Department of Homeland Security (DHS) is a research SCADA project for “Unifying Stakeholders and Security Programs to Address SCADA Vulnerability and Infrastructure Interdependencies” (Ericsson, 2009), which aims to raise the security awareness of process control systems. (Cherdantseva, et al., 2016) have highlighted considerable risk assessment approaches on SCADA systems ranging from 2004 to 2014, stemming from the following countries: USA, Korea, France, Canada, China, Australia, Serbia, Ireland, and Italy. (Cardenas, et al., 2011) cover the scope broader than RA and also describe modules for attack detection and automated response to an attack. (Tantawy, et al., 2019) introduce the LOPA methodology that illustrates key mathematical assumptions that are violated in view of security attacks. The methodology involves the probability of a security attack on a cyber-physical system and it evaluates it through a test-bed case study.

A Cyber Risk Assessment Model for CIIs is presented in (Kumar, et al., 2020) which refers to the concept of regression analysis. (Ten,, et al., 2010) is a considerable research work introducing the four components of the security framework for SCADA systems: Real-time monitoring, anomaly detection, impact analysis and mitigation strategies. (Haimes & Horowitz, 2004) describe the eight-phase process risk filtering, ranking, and management method (RFRM) which builds on an adaptive two-player Hierarchical Holographic Modelling (HHM) method to identify risks. The approach updates on the advances in probabilistic RA that can be applied to estimate the risk (exposure or expected loss) from SCADA and DCS installations. To delineate risk assessment processes, there are various attempts to structure ontologies for general risk assessments, such as the AURUM system (Ekelhart, et al., 2009). (Markovic-Petrovic, et al., 2019) propose a new method for security risk assessment in SCADA networks dividing it into three phases: the objective phase, the subjective phase and the final assessment phase using fuzzy logic in all phases and analytic hierarchy process (AHP) in the subjective phase. A unique framework which contributes to the growth of the CI operator readiness in critical situation is analysed in (Foglietta, et al., 2019). Furthermore, it produces a platform of a cyber-attack detection subsystem and a risk assessment framework. It embraces a range of capabilities, (cyber-attack detection, mitigation strategies, interdependency and risk evaluation). (Knapp & Langill, 2019) give a clear understanding on SCADA and Control System protocols and their operations, presenting implementation guidelines for security measures of critical infrastructures for system-specific compliance.

Various research has been performed implementing fuzzy logic. For instance, in (Wu, 2013) a comparison between SCADA systems with traditional IT systems is realized summarizing the main risks of a SCADA system in information and security, following a typical power information system topology, based on assets, threats, vulnerabilities and security measures. The current model is established through a fuzzy analytic hierarchy process (FAHP) to quantitatively evaluate risks of a power information system.

Numerous sectorial risk assessment methods have been developed for SCADA and ICT systems for a variety of Industries. For example, (Yang, et al., 2019) describe a SCADA security assessment based on causality analysis for oil and GAS SCADA systems utilizing the fuzzy Mamdani reasoning to estimate factor neurons in the proposed model. In (Lanzrath, et al., 2020) results of the implementation of a methodology for high-power EM based risk assessment of large structures are presented in accordance with an example of smart grid substations. Another indicative example regarding power systems is shown in (Meng, 2015) presenting a research work regarding dynamic and static risk assessment for power information systems.

According to (Kalogeraki, et al., 2018), a successful RA approach for SCADA and ICTs may have the following characteristics: a structured body of cybersecurity knowledge (Zio, 2018; Kalogeraki, et al., 2018), business modelling and simulation techniques adoption to carry out different real-life cyber-attack scenarios and experiment with the results (Theocharidou & Giannopoulos, 2015; Kalogeraki, et al., 2018) implementing rational decision-making techniques for probabilistic RAs of complex cyber-attack scenarios, identify common or cross-border scenarios throughout national and regional limits (Theocharidou & Giannopoulos, 2015), engage all CII operators, including entities of both public and private sector participating, to have a clear and detailed view of SCADA cyber-risks at the asset-individual level, identify the overall cyber dependencies[47] across SCADA Networks to detect the impact at the system level (Kalogeraki, et al., 2018), be compliant with regulations and directives or international standards applying to the underlined sector providing collaborative practices to facilitate the sharing and transfer of risk-related information over cross-sectorial CIIs operators.

Summarizing, the literature shows that effective cost-benefit analysis and evaluation of SCADA and ICT cyber-risks are based on a straightforward approach combining a set of parameters and features, such as the likelihood of security events, the consequences of the event itself and the exploitation level of vulnerability (Zambon, et al., 2011). On this account, novel risk and resilience assessment approaches that may assess and demonstrate the ability to develop and implement effective risk assessment strategies and ensure SCADA systems resilience against aftermath cyber-incidents.

Credible approaches that can be useful to the current project's objective could be the Cyber/Physical Security Management System (CYSM) collaborative approach (Papastergiou, et al., 2015), the MEDUSA's research method (Papastergiou & Polemi , 2017) which sets a number of concepts, algorithms, and tools evolved from research, specially designed to protect the IT infrastructure and

[47] Cyber dependency can be defined the connection between two or more assets, where the current state of one asset is directly affected from the state of another asset

associated systems and the MITIGATE collaborative, dynamic, evidence-driven Maritime Supply Chain Risk Assessment approach (Kalogeraki, et al., 2018; Papastergiou & Polemi, 2018).

5.2 Attack & Simulation Environments

The evolving cyber-threats' landscape and the enormous effort needed to efficiently secure data in the context of a CII, denoted the necessity of adopting more advanced environments capable of providing a holistic picture of the system. For that reason, over the last years several types of attack modelling and simulation techniques have been developed to deal with a network's vulnerabilities, the behavioural analysis of a cyber-threat, and the potential objectives of an attacker. A proper utilisation of such techniques not only provides an improved planning of a rapid response to a security incident, but it also assists with the automation procedure of the threat modelling through simulation-driven approaches. Last but not least, their results could serve as the basis of an evaluation regarding the proposed security enhancements found in a CII, as well as benefit the improved understanding and management of stakeholders' risks (wherever such a scenario applies).

5.2.1 Attack Graph Methods & Algorithms

In recent years, attack modelling is considered a useful tool in risk assessment of cyber-physical systems (i.e. SCADA systems) (Kriaa, et al., 2012). On such systems, attack vectors are strongly dependent on considerations regarding the technical and operational environment where an attack takes place. On this account, attack graphs are data structures that are able to model all possible avenues of a network attack. Attacks on SCADA systems may cause disruption or damage of CIIs. The attacker's profile is a parameter of in-depth security risk analysis to identify security risks in SCADA systems. The attacker's profile appears to have one or a combination of the following characteristics.

As it has been already mentioned in the deliverable of CyberSANE D3.1 "Taxonomy of threat landscape for CIIs", remarkable cyber-attack vectors against SCADA systems are (Kalogeraki, et al., 2018) database pear on the transport layer, at application layer (lack of security control to many of the attacks; backdoors and holes in the network perimeter; Cinderella attack on time provision and synchronization, communications hijacking and man-in the middle attacks). These attacks fall in four categories:

i. on the Communication stack;
ii. on the UDP port (attacks onSCADA protocols);
iii. on the hardware;
iv. on the software.

In order to evaluate the vulnerability of SCADA and ICT networks the effects of interconnected relations must be considered. A typical process of vulnerability analysis can be conducted via scanning tools identifying individual vulnerabilities. Local vulnerability information together with network information (i.e. connectivity between hosts) are able to build attack graphs (Jha, et al., 49-63). Attack graph paths are considered a series of exploits, the so-called atomic attacks, which can drive the process to an undesirable state (e.g. an adversary gains administrative access to a critical host) (Jha, et al., 49-63). Attack graphs can be utilized for detection, defence and forensic analysis purposes (Jha, et al., 49-63).

Cyber-attack prevention technologies typically use attack graph generation and analysis methods to identify all possible paths that attackers can exploit to gain unauthorized access to a system (Ou & Singhal, 2011). There is considerable work for attack graph generation and analysis. The Model checking algorithm is a technique for checking whether a formal model M of a system satisfies a given property p (Jha, et al., 49-63). A typical example is the model checker NuSMV (Cimatti, et al., 1999), in which model M is a finite labelled transition system and p is a property expressed in Computation Tree Logic (CTL) (Jha, et al., 49-63). Attack graphs can be assumed direct graphs in the form of representing a network (nodes are states and edges are the application of an exploit that can transfer a network state into another, more compromised network state) (Chochliouros, et al., 2009). The ending states of the attack graph represent the network states in which the adversary has met his goals. In addition, an attack graph can be considered in the form of a dependency graph exploit (Chochliouros, et al., 2009).

Dependability can be characterized by the following attributes: availability of service (readiness for correctness), reliability (continuity) of service, safety (absence of negative/catastrophic consequences) of users and the environment, confidentiality (unauthorized disclosure), integrity and maintainability (repair) (Chochliouros, et al., 2009). The literature reviews numerous techniques to evaluate dependability in terms of security. Cyber-attacks can be considered either intentional or unintentional (accidental) and dependability can be evaluated through stochastic analysis, (sophisticated method to measure the probability/acceptability of faults (Chochliouros, et al., 2009). In case a large network analysis is either an explicit or strict requirement a quantitative, much more complex analysis is preferred, which can be achieved through probabilistic models (Chochliouros, et al., 2009). Sophisticated attacks aggregate multiple vulnerabilities of a system. Such advanced attacks can be modelled with probabilistic attack graphs that represent the different states of a system as nodes and the relations between different states as directed edges. In this way, they allow computation of all potential attack paths to a target of interest. One single path describes the various steps of an attack where exploiting one vulnerability grants access to other vulnerabilities, e.g. by gaining some privileges. Even for small networks, as presented in (Singhal & Ou, 2017), attack paths may become quite complex and several tools may be required to develop the attack graphs.

In terms of dependency evaluation, traditional techniques to ensure that a service is operating correctly are covering the “absolutely necessary” are Block Diagrams (BDs) and Fault Trees (FTs) whilst more sophisticated approaches rely on Markov Models (Chochliouros, et al., 2009). Remarkable examples of semi-quantitative risk assessment approaches for SCADA and ICT systems are found in the literature, such as the Fault Tree Events Analysis which estimates the frequency of event occurrence in an undesired (top/root) logical scale (Ralston, et al., 2007). The OBEST object-based event scenario tree illustrates combined features of event tree analysis and Monte-Carlo discrete event simulation along with concepts of object-oriented analysis for risk assessment (Wyss & Durán, 2001). (Schneier, 1999) introduced the attack trees as a method to formalize the security of systems and subsystems regarding varying attacks. A probabilistic-based RA Tool provides a foundation for the estimation of risk reduction when applied to SCADA security (McQueen, et al., 2006). Augmented vulnerability trees and two new indices for quantifying risks were introduced by (Ralston, et al., 2007). (Byres, et al., 2004), illustrate the use of attack trees for assessing vulnerabilities in SCADA systems and control hardware. Significant research is carried out on assessing the Byres attack trees, to estimate vulnerabilities in SCADA systems based on MODBUS and MODBUS/TCP communication protocols, and reckon the features of the topmost attack event investigating possible ways to achieve the final goal of the attack (Cherdantseva, et al., 2016). In (Poolsappasit, et al., 2012), they make use of Bayesian attack graph generation for dynamic security risk management.

Concerning Markov Model approaches, (Xiaolin, et al., 2008) present a Markov game theory-based risk assessment model for network information systems. Moreover, they utilize a Markov chain to analyse the spreading process of potential threats and to assess the system risk. (Kriaa, et al., 2012) Additionally, to analyse the Stuxnet attack that targets SCADA systems and model its fundamental mechanisms in unique and rigorous graphs following the BDMP (Boolean logic Driven Markov Processes) formalism to deliver quantification results for each possible attack sequence and to illustrate the advantages of such modelling.

Topological Analysis of Network Attack Vulnerability (TVA) builds a so-called exploit dependency graph that contains information about conditions of an exploit and then searches this graph to combine various vulnerabilities (Jajodia, et al., 2005; Ou & Singhal, 2011). In (Ou, et al., 2005) the authors developed a MulVAL, a logic-based network security analyser. This is a vulnerability analysis tool that models the interaction of software bugs along with network configurations. NetSPA is a network security planning architecture that very efficiently generates the worst-case attack graphs (Artz, 2002).

A component metric is attached to each attack node derived from the Common Vulnerability Scoring System (CVSS) metric vector (Mell, et al., 2007). Based on these component metrics, models for cumulative and propagated risks have been developed. In (Homer, et al., 2009) an algorithm is provided to compute the probability of success of a multi-step attack using probabilistic reasoning that takes into account the conditional dependencies between attack paths. The method has been applied in an empirical study in (Zhang, et al., 2011) and extended in (Homer, et al., 2013).

A simulation-driven approach is a composite process aiming to discover and execute possible attack plans, based on a suitable formalism (Johnson, et al., 2018), where attack graphs model the attack steps, executed by a set of threat agents to produce risk-related results and allow the subsequent simulation. A variety of approaches explore attack simulation and computation of attack graphs over IT infrastructures; agent model (Rybnicek, et al., 2014), the automatic attack-graph generation model (Al Ghazo, et al., 2019), the intelligent goal-oriented agents' mode of (Shen, et al., 2004), the formal model to calculate very large attack graphs, allowing attacks' simulation in the domain of interest (Johnson, et al., 2018). The minimum critical-attacks graph set is considered an NP-completeness problem (equivalently with the min Label-Cut problem, MLC) (Al Ghazo & Kumar, 2019). The NP-completeness is established by reducing the Hitting-Set problem to the MLC, and Jha et al. (Jha, et al., 2002) also presented a greedy algorithm to the Hitting-Set problem that picks the elements with the highest hits first.

5.2.2 Cascading Effects

Cascading effects are observed when one or more cybersecurity incidents are propagated throughout the components of a CI, or in case those incidents' derivatives lead to some type of security breach. In most cases this chain of effects is inevitable since the majority of a system's components are interdependent and interconnected with at least another one. Therefore, most of the studies in this area focus into disaster risk reduction (McGee, et al., 2016) by understanding the interconnected relationships of a CI, as well as providing a modelling or simulation insight regarding the expected response of a CI to a certain incident. This model-based insight is used to depict the cascading effects on CIIs, present the consequences of a cyber-attack, and improve the decision making based on the recorded attack path.

In the modern era, cascading effects occur in various domains with growing frequency among interdependent assets of CIIs. The ISO31000:2018 international standard (ISO/TC 262, 2018) considers cascading effects along with the identification of interdependencies, whitespace risks and events as critical factors for conducting a risk assessment process. General observations of cascading failures fall in two phases: the slow cascading phase, where things still seem to work properly, and the fast cascading phase where the system gets out of control.

There are multiple causes of cascading failures. In some domains, the cascading behaviour can be identified to model cascading failure adequately. A variety of the existing studies are focused on the modelling of interdependent networks and theoretical analysis of the cascading effects, which capitalizes mainly on the theory of random graphs. The theory of random graphs has been typically used to investigate single networks. Such graphs are capable of analysing behaviours in real-world networks (Shin, et al., 2014). (Shin, et al., 2014) develops a framework to identify important topological properties of large-scale networks (e.g. including average diameter, node clustering coefficient, network modularity, degree correlation etc.). Such topological properties can vary under different scenarios.

Another generic characteristic is the difficulty of measuring the impact of cascading failures because of the high complexity and many indirect costs. The literature reviews limited data available on incidents that caused cascading effects. Whenever possible, historical data is used in combination with expert knowledge to keep the number of assumptions limited. Existing studies as presented in (Ouyang, 2014) help to define failure patterns or to detect interdependencies that are not obvious at first glance. The uncertainties challenging the interdependencies among infrastructures are discussed in (Hasan & Foliente, 2015). Concerning cascading failures in Industry networks, most current studies involve only single network models because it is difficult to represent real-world network systems of an Industry supply chain engaging multiple attributes and functions and interdependent network models (Tang, et al., 2016). A proposed solution of time-varied functional equations to quantify the dynamic process of failed loads propagation in an interdependent network is provided in (Tang, et al., 2016) including a twofold simulation case. (Huang, et al., 2013) aim to estimate cascading failures in cyber-physical interdependent systems by calculating the fraction of nodes that is able to continue the performance after the cascading failure stops and obtain accurate results via simulation methods. They show that there is a critical threshold. In case the proportion of failing nodes exceeds this value, the entire system collapses. (Panda & Bower, 2020) investigate and analyse correlations between traditional risks of critical infrastructures (according to the Sendai framework) and cyber-security risks with the respective cascading effects to identify characteristic of today’s complex and interrelated shocks and stresses.

Regarding cascading effect measurements, structural, “Systems-of-Systems” (SoS) research approaches are rich of contributions, ranging from the categorization of interdependencies (Rinaldi, et al., 2001; Pederson, et al., 2006) to their utilization in vulnerability analysis (Zio & Sansavini, 2011; Bloomfield, et al., 2010). The study of the interdependencies is an attractive field for research concerning SCADA and ICTs along with the importance of potential failure propagation among CIIs that may lead to cascading effects within the supply networks. In this context, new powerful methods are required to model and describe such SoS in an holistic manner to provide security and reliability assessment considering various types of threats and failures developing “what-if” scenarios to analyse interdependencies (Eusgeld, et al., 2011). A SoS approach is proposed in (Eusgeld, et al., 2011) for SCADA systems taking into account their interdependencies with the underlined CIIs focusing on the coupling of these systems and introducing the HLA simulation standard for interdependencies. In (Hashemi & Zarif, 2020), a control structure SoS approach for power distribution networks is analysed.

Behavioural analysis relieves mechanisms of failure propagation, cascading effects that occur as consequence of complex interactions among systems that reflect the theory of resilience (Giannopoulos, et al., 2012). The most promising research contributions reflect the domain of control rather than risk (Giannopoulos, et al., 2012). In this vein, there is enough space for innovation (Filippini & Silva, 2011) adjusting the control concept to interdependent systems and conducting an evaluation of the resilience (Giannopoulos, et al., 2012).

“Consequence” is defined by the ISO27000:2018 standard as the “outcome of an event affecting objectives”. Assessment of consequences embeds worst case scenarios according to cross-cutting criteria (Theocharidou & Giannopoulos, 2015). Impact is considered the effect of the security state of a system due to an information system’s change. The key-concepts and impact measurement in SCADA systems, including system (asset), vulnerability, threat impact (consequence) and security control-countermeasure have been identified (Cherdantseva, et al., 2016; Ericsson, 2009; Francia III, et al., 2012; Markovic-Petrovic & Stojanovic, 2014; Verendel, 2009).

Despite the strong diversity of models of cascading effects, a rough classification helps to get an overview on the different lines of reasoning. A set of modelling techniques and simulation approaches for critical infrastructures are given in (Oliva, et al., 2012) and a compact comparison between the different models for cascading failures in power systems is argued in (Guo, et al., 2017). Existing approaches can be divided into different classes according to their main focus, such as topological models, stochastic models, dynamic simulation models and agent-based models. This list is not exhaustive since other models exist that do not belong to any of the five groups, but the ones presented in the following are the most significant approaches with respect to their applicability capabilities in both existing and future domains.

Many models of cascading failures are based on topological properties of the network. Node degrees are used as weights in (Wang & Chen, 2008; Wei, et al., 2012) and local flow distribution rules in this weighted graph allow analysing consequences of failure of an edge. Between centrality has been used to investigate overload breakdowns in scale-free networks (Holme, 2002; Holme & Kim, 2002). Triggered events and random failure are juxtaposed in (Kim & Obah, 2007) concerning graphical features, such as critical path lengths and small world-ness index. A node capacity model is presented in (Motter, et al., 2002) to analyse terms, such as capacity considering that a component either works properly or fails. Topological models are mostly generic and can work in complex networks beyond the limits of a specific sector. Nonetheless, such models may result in a confused outcome for specific sectors (Hines, et al., 2010). Notwithstanding, they provide a good basis for more advanced models as shown in (Dey, et al., 2016). Several extensions work with maximum flow theory to model power grids attitude, as presented in (Fan, et al., 2016) and (Rinaldi, et al., 2001). Despite these models are generally applicable, they do not elaborate the underlined situation, analysing thoroughly and thus are error-prone concerning their prediction capability.

Stochastic models allow simulation that may be used to validate the model or make predictions and provide simulation on several possible events. (Lai, et al., 2019) estimate the robustness of asymmetric cyber-physical power systems against cyber-attacks taking into consideration the effects of computer malware spreading, power redistribution as well as overloading and the interrelations between the coupled networks, adopting a stochastic failure model to calculate the time interval between the initial cyber-attack and a given level of power loss counting on simulation results. Popular stochastic processes are Markov chains and branching processes often used for modelling as presented previously in attack graph representations as well. The model in (Zhang, et al., 2017) describes the failure dynamics of the entire network through a power flow model for the failure propagation combined with a stochastic model for the time between failures. It provides a simulation of the cascading procedure and investigates the systems robustness. Whenever detailed information about the system at hand is available, more accurate predictions based on more involved simulation models are possible. Such simulations are more evolved and less applicable for real time predictions.

Indicative dynamic models presenting the cascading process based on a linear programming approach (the OPA model) are introduced in (Carreras, et al., 2002; Mei, et al., 2009) enabling simulation of the patterns of cascading blackouts in power systems considering the dynamics and the potential mitigation actions. The COSMIC model (Song, et al., 2016) is a nonlinear dynamic model for cascading failures in power systems which describes many mechanisms by recursive computations of the corresponding differential equations. A dynamic probabilistic risk assessment is used in (Henneaux, et al., 2012) in order to indicate the coupling between events in cascading failure and the dynamic response of the grid to the perturbation which has been extended (Henneaux, et al., 2016).

Agent based models are created either from existing system dynamic models or discrete event models (Borshchev & Filippov, 2004). The N-ABLE model aims to identify physical effects of cyber-attacks in infrastructures (Kelic, et al., 2008). Additionally, agent based models are used to model risk responses in a complex society (Busby, et al., 2016) which in turn provide a basis to simulate responses of society to incidents in critical utilities (Busby, et al., 2016) and to analyse maintenance strategies in critical infrastructures (Kaegi, et al., 2009). Moreover, agent-based models can provide detailed information about the potential consequences of incidents for a given scenario and thus allows a very detailed "what-if"-analysis.

5.3 Data Visualization Techniques

Nowadays, the amount of data related with cyber-security events, as well as their interdependencies and correlations, are big and complex enough to handle in a timely manner. Therefore, the use of a visualization tool is considered necessary as a compulsory part of any cybersecurity solution. Specifically, cyber-security data visualization refers to creating charts, graphs and similar visual material from cyber-security data within a specific context. The data in visuals can be gathered from various sources. With the help of data visualisation methods, security officers or end users can have demonstrated and exposed: a) a comparative graph of malicious activity patterns; b) heat maps that illustrate the scope and severity of a security incident; and visualizing patterns that allows to detect possible additional attacks and from where such attacks can hit an infrastructure or organization. The threat-detection and proactive capabilities of CyberSANE will take advantage of a cyber-dashboard which is capable of visualizing and linking a CII network with regards to a detected threat. Such a thing is feasible by adopting one or more visualization techniques, where usually a graph representation of connections and relationships is used to depict and analyse a malicious event. Noel et al. (Noel, et al., 2016) developed a unified graph-based cyber-security model (i.e. CyGraph) capable of capturing security events, building a predictive model of possible attacks, and correlating events to known vulnerability paths. CyGraph also prioritizes exposed vulnerabilities, mapped to potential threats, in the context of mission-critical assets.

When analysing cyber-security data, the connections -between devices, events, locations, IPs, signatures, and so on- hold the key to uncovering anomalies, threats and vulnerabilities. Visualization solutions (e.g. Cambridge Intelligence[48], Qlik[49], Tableau[50], etc.) need to integrate new computational and theory-based algorithms with innovative interactive techniques and be compatible with advanced big data analytics frameworks, in order to act as a mediator of human-information discourse between humans and these information resources. Besides, visualization alone cannot effectively manage levels of details about the data or prioritize different information in the data - hence the need for analysis and interaction. The visual designs should be also based on human cognitive and perceptual principles. In addition to this, the challenges faced by a cyber-security professional are more complex ranging from fraud detection, network forensics, data privacy issues and data provenance problems (Jayasingh, et al., 2016). The response time is a desirable factor in cyber-security analytics that is also relevant from a visualisation point of view. Presenting results in a manner that is understandable by people without proven skills and experience in data science is highly relevant (Fan, et al., 2017). A reasonable degree of transparency, a key prerequisite to increase trust in data, is also relevant with respect to some of the issues identified. The densely connected nature of graph data can be complex to unravel. A few simple techniques, like filtering and layouts, can make insights easier to understand and communicate. In addition to this, time bars and macro-views of the datasets allows to investigate long-term patterns and trends (i.e. malware propagation, unusual login/access habits, unusual network traffic, etc.).

Finally, it is worth looking at two important dimensions to the collected and analysed data that are difficult to convey with standard node-link visualisation: time and location. By understanding the

[48] https://cambridge-intelligence.com/products/
[49] https://www.qlik.com/us/
[50] https://www.tableau.com/

physical location of an entity and the time of an event can unlock insights into patterns and trends, helping to uncover the origin of an attack, or predict what will happen in the future[51]. Using time bars in combination with maps, it is possible to filter the data by time and date and observe the evolution of an event as it occurs.

[51] https://www.informationisbeautiful.net/visualizations/worlds-biggest-data-breaches-hacks/

Chapter 6 Conclusions

This deliverable has thoroughly described both the mainstream and the latest advances in R&D on proactive detection and response methodologies that could be of potential use in CyberSANE. As aforementioned, these areas were the more critical for us and the main goal in this document was to provide new strategies and solutions that we could adopt or to use for enhancing our solutions.

Therefore, the key elements we study and analyse here are incident handling and response approaches along with a several solutions of digital chains of evidence that could be used for evidence correlation and event display purposes.

The process we followed was composed of different phases where we identified the different topics, analysed the improvement and studied how the tools of CyberSANE could benefit from this work. We performed this one area at a time in order to have all partners involved for each area and have a better understanding of the improvements and the impact that adopting improved algorithms or approaches would have for the whole hybrid net. This was necessary as we thought it was important to evaluate the impact of the different solutions we find in the project.

Another interesting aspect of the analysis is that the CyberSANE's threat taxonomy was used as the medium to enumerate the most typical tools, platforms and techniques within the proactive detection and response domain. Our cyber-security related survey continued across several research backgrounds of interest in anomaly-based detection methodologies, including all types of statistical analysis and machine learning, data mining and deep learning techniques, ultimately concluded with a couple of works in genetic algorithms area.

Finally, we investigated risk assessment methodologies, cascading effects and visualization environments in order to provide the CyberSANE platform with improved prevention and modelling capabilities.

All the above mentioned techniques are taken into consideration and will be scrutinised for their use and suitability in the upcoming tasks of WP5 for the implementation of automated and parameterized detection and response approaches for CIIs.

Chapter 7 List of Abbreviations

Abbreviation	Translation
AAM	Adaptive Assignment Manager
ADS	Anomaly Detection System
AHP	Analytic Hierarchy Process
ANN(s)	Artificial Neural Network(s)
APT(s)	Advanced Persistent Threat(s)
BCP	Business Continuity Plan
BD(s)	Block Diagram(s)
BM(s)	Boltzmann Machine(s)
BDMP	Boolean logic Driven Markov Processes
CB-IDPS	Cloud Based Intrusion Detection and Prevention Systems
CDN	Content Delivery Network
CERT	Computer Emergency Response Team
CI(s)	Critical Infrastructure(s)
CII(s)	Critical Information Infrastructure(s)
CNN(s)	Convolutional Neural Network(s)

CSIRT(s)	Computer Security Incident Response Team(s)
CTL	Computation Tree Logic
CVSS	Common Vulnerability Scoring System (CVSS)
DBN(s)	Deep Belief Network(s)
DDoS	Distributed Denial of Service
DL	Deep Learning
DNN(s)	Deep Neural Network(s)
DNS	Domain Name Service
DoS	Denial of Service
DSNSF	Digital Signature of Network Segment using Flow analysis
DTD	Document Type Description
ENISA	European Union Agency for Cybersecurity
FAHP	Fuzzy Analytic Hierarchy Process
FT(s)	Fault Tree(s)
GA(s)	Generic Algorithm(s)
GAN(s)	Generative Adversarial Network(s)
GOOSE	Generic Object Oriented Substation Event
ICS	Industrial Control Systems

ICT	Information and Communication Technology
IDIP	Intruder Detection and Isolation Protocol
IDS	Intrusion Detection System
IKS	Intrusion Kill Chain
IPS	Intrusion Prevention System
IR	Incident Response
MHN	Multiple Honeypot Solution
MISP	Malware Information Sharing Platform
ML	Machine Learning
MCUSUM	Multivariate CUmulative SUM
NADS	Network Anomaly Detection System
NFV	Network Function Virtualization
NIST	National Institute of Standards and Technology
NN(s)	Neural Network(s)
OTT	Open Threat Taxonomy
PAM	Partitioning Around Medoids
PCA	Principal Component Analysis
RA	Risk Assessment

RBM(s)	Restricted Boltzmann Machine(s)
RNN(s)	Recurrent Neural Network(s)
SCADA	Supervisory Control and Data Acquisition
SDN	Software Defined Networking
SMV	Sampled Measured Value
SoS	Systems-of-Systems
SPC	Statistical Process Control
STIX	Structured Threat Information Expression
SVM(s)	Support-Vector Machine(s)
SWaT	Secure Water Treatment
TAXII	Trusted Automated Exchange of Intelligence Information
TOCSR	Taxonomy of Operational Cyber Security Risks
TVA	Topological Analysis of Network Attack Vulnerability
VPN	Virtual Private Network
WSN	Wireless Sensor Networks

Chapter 8 Bibliography

Abdollah, M., Mas'ud, M., Yusof, R. & Selama, S., 2009. *Threshold verification using statistical approach for fast attack detection.* Kuala Lumpur, s.n.

Agrawal, S. & Agrawal, J., 2015. Survey on Anomaly Detection using Data Mining Techniques. *Procedia Computer Science,* Volume 60, pp. 708-713.

Al Ghazo, A., Ibrahim, M., Ren, H. & Kumar, R., 2019. A2G2V: Automatic Attack Graph Generation and Visualization and Its Applications to Computer and SCADA Networks. *IEEE Transactions on Systems, Man, and Cybernetics: Systems,* pp. 1-11.

Al Ghazo, A. & Kumar, R., 2019. *Identification of Critical-Attacks Set in an Attack-Graph.* New York City, IEEE, pp. 0716-0722.

Alberts, C. & Dorofee, A., 2002. *Managing Information Security Risks: The OCTAVE Approach.* Boston: Addison-Wesley Longman Publishing Co..

Alcaraz, C. et al., 2009. *Adaptive dispatching of incidences based on reputation for SCADA systems.* s.l., Springer, pp. 86-94.

Alcaraz, C. & Zeadally, S., 2015. Critical infrastructure protection: Requirements and challenges for the 21st century. *International Journal of Critical Infrastructure Protection,* Volume 8, pp. 53-66.

Alzaid, H., Foo, E. & Gonzalez Nieto, J., 2008. *Secure Data Aggregation in Wireless Sensor Network: a survey.* Wollongong, Australian Computer Society, pp. 93-105.

Amarbayasgalan, T., Jargalsaikhan, B. & Ryu, K., 2018. Unsupervised Novelty Detection Using Deep Autoencoders with Density Based Clustering. *Applied Sciences,* 8(9), p. 1468.

Artz, M., 2002. *NetSPA : a Network Security Planning Architecture,* Cambridge: Massachusetts Institute of Technology.

Ayodele, T. O., 2010. Types of machine learning algorithms. In: Y. Zhang, ed. *New advances in machine learning.* Rijeka: InTech, pp. 19-48.

Baldoni, R. & Montanari, L., 2016. *Italian National Cyber Security Framework.* Athens, The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp), pp. 168-174.

Barbara, D., Wu, N. & Jajodia, S., 2001. *Detecting novel network intrusions using bayes estimators.* Chicago, Society for Industrial and Applied Mathematics, pp. 1-17.

Barnum, S., 2014. *Standardizing Cyber Threat Intelligence Information with the Structured Threat Information eXpression (STIX™),* s.l.: MITRE.

Basicevic, I., Ocovaj, S. & Popovic, M., 2015. Use of Tsallis entropy in detection of SYN flood DoS attacks. *Security and Communication Networks,* 8(18), pp. 3634-3640.

Beggs, C. & Warren, M., 2009. *Safeguarding Australia from cyber-terrorism: A proposed cyber-terrorism SCADA risk framework for industry adoption.* Perth, School of Computer and Information Science, Edith Cowan University, Perth, Western Australia, p. 5.

Bereziński, P., Szpyrka, M., Jasiul, B. & Mazur, M., 2015. Network Anomaly Detection Using Parameterized Entropy. In: K. Saeed & V. Snášel, eds. *Computer Information Systems and Industrial Management. CISIM 2015. Lecture Notes in Computer Science.* s.l.:Springer, Berlin, Heidelberg, pp. 465-478.

Berman, D. S., Buczak, A. L., Chavis, J. S. & Corbett, C. L., 2019. A Survey of Deep Learning Methods for Cyber Security. *Information,* 10(4), p. 122.

Bhatt, P., Yano, E. T. & Gustavsson, P., 2014. *Towards a Framework to Detect Multi-Stage Advanced Persistent Threats Attacks.* Oxford, IEEE, pp. 390-395.

Bhuyan, M., Bhattacharyya, D. & Kalita, J., 2014. Network Anomaly Detection: Methods, Systems and Tools. *IEEE Communications Surveys & Tutorials,* 16(1), pp. 303-336.

Bigham, J. et al., 2004. *Dynamic Trust Management of Semi-Automated Complex Systems.* Austin, IIIS: International Institute of Informatics and Systemics.

Bilge, L., Kirda, E., Kruegel, C. & Balduzzi, M., 2011. *EXPOSURE: Finding Malicious Domains Using Passive DNS Analysis.* San Diego, NDSS, pp. 1-17.

Bloomfield, R. et al., 2010. Stochastic Modelling of the Effects of Interdependencies between Critical Infrastructure. In: E. Rome & R. Bloomfield, eds. *Critical Information Infrastructures Security. CRITIS 2009. Lecture Notes in Computer Science.* Bonn: Springer, Berlin, Heidelberg, pp. 201-212.

Borshchev, A. & Filippov, A., 2004. *From System Dynamics and Discrete Event to Practical Agent Based Modeling: Reasons, Techniques, Tools.* Oxford, s.n.

Bottazzia, G. & Mea, G., 2017. *Cybercrime-Funded Terrorism and the Threats Posed by Future Technologies: Appealing Economics and Targets.* Dublin, IOS Press, p. 109.

Bottazzi, G., Italiano, G. & Rutigliano, G., 2017. *An Operational Framework for Incident Handling.* Venice, s.n., pp. 126-135.

Bou-Harb, E., Debbabi, M. & Assi, C., 2013. *A Statistical Approach for Fingerprinting Probing Activities.* Regensburg, IEEE, pp. 21-30.

Brenner, J., 2013. Eyes wide shut: The growing threat of cyber attacks on industrial control systems. *Bulletin of the Atomic Scientists,* 69(5), pp. 15-20.

Brugman, J., Khan, M., Kasera, S. & Parvania, M., 2019. *Cloud Based Intrusion Detection and Prevention System for Industrial Control Systems Using Software Defined Networking.* San Antonio, IEEE, pp. 98-104.

Buczak, A. L. & Guven, E., 2016. A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection. *IEEE Communications Surveys & Tutorials,* 18(2), pp. 1153-1176.

Busby, J., Gouglidis, A., Rass, S. & König, S., 2016. *Modelling security risk in critical utilities: The system at risk as a three player game and agent society.* Budapest, IEEE, pp. 1758-1763.

Busby, J., Onggo, B. & Liu, Y., 2016. Agent-based computational modelling of social risk responses. *European Journal of Operational Research,* 251(3), pp. 1029-1042.

Byres, E., Franz, M. & Miller, D., 2004. *The Use of Attack Trees in Assessing Vulnerabilities in SCADA Systems.* Lisbon, Citeseer, pp. 3-10.

Cadzow, S. et al., 2015. *Supply Chain Integrity: An overview of the ICT supply chain risks and challenges, and vision for the way forward,* s.l.: European Union Agency for Cybersecurity (ENISA).

Cardenas, A. et al., 2011. *Attacks against process control systems: risk assessment, detection, and response.* New York, Association for Computing Machinery, p. 355–366.

Carreras, B., Lynch, V., Dobson, I. & Newman, D., 2002. Critical points and transitions in an electric power transmission model for cascading failure blackouts. *Chaos: An interdisciplinary journal of nonlinear science,* 12(4), pp. 985-994.

Cebula, J., Popeck, M. & Young, L., 2014. *A Taxonomy of Operational Cyber Security Risks Version 2 (No. CMU/SEI-2014-TN-006),* Pittsburgh : Carnegie Mellon University - Software Engineering Institute.

Cheminod, M., Durante, L. & Valenzano, A., 2013. Review of Security Issues in Industrial Networks. *IEEE Transactions on Industrial Informatics,* 9(1), p. 277–293.

Cherdantseva, Y. et al., 2016. A review of cyber security risk assessment methods for SCADA systems. *Computers & Security,* Volume 56, pp. 1-27.

Chizoba, O. & Kyari, B., 2020. Ensemble classifiers for detection of advanced persistent threats. *Global Journal of Engineering and Technology Advances,* 2(2), pp. 1-10.

Chochliouros, I., Spiliopoulou, A. & Chochliouros, S., 2009. Methods for Dependability and Security Analysis of Large Networks. In: M. Pagani, ed. *Encyclopedia of Multimedia Technology and Networking.* Milan: IGI Global, pp. 921-929.

Cichonski, P., Millar, T., Grance, T. & Scarfone, K., 2012. *Computer security incident handling guide.* s.l.:NIST Special Publication 800(61).

Cimatti, A., Clarke, E., Giunchiglia, F. & Rover, M., 1999. *NuSMV: A New Symbolic Model Verifier.* Trento, Springer-Verlag, pp. 495-499.

Cisar, P. & Cisar, S., 2007. *EWMA Statistic in Adaptive Threshold Algorithm.* Budapest, IEEE, pp. 51-54.

Cohen, M., 2008. PyFlag–An advanced network forensic framework. *Digital Investigation,* Volume 5, pp. S112-S120.

Connolly, J., Davidson, M., Richard, M. & Skorupka, C., 2014. *The Trusted Automated eXchange of Indicator Information (TAXII™),* s.l.: MITRE.

Cowsalya, T. & Mugunthan, S., 2015. Hadoop architecture and fault tolerance based hadoop clusters in geographically distributed data center. *ARPN Journal of Engineering and Applied Sciences,* 10(7), pp. 2818-2821.

Cruz, T. et al., 2016. A Cybersecurity Detection Framework for Supervisory Control and Data Acquisition Systems. *IEEE Transactions on Industrial Informatics,* 12(6), pp. 2236-2246.

Daley, R., Millar, . T. & Osorno, M., 2011. *Operationalizing the coordinated incident handling model.* Waltham, IEEE, pp. 287-294.

Daryabar, F. et al., 2012. *Towards Secure Model for SCADA Systems.* Kuala Lumpur, IEEE, pp. 60-64.

de Abreu, S., Kendzierskyj, S. & Jahankhani, H., 2020. Attack Vectors and Advanced Persistent Threats. In: H. Jahankhani, S. Kendzierskyj, N. Chelvachandran & J. Ibarra, eds. *Cyber Defence in the Age of AI, Smart Societies and Augmented Humanity. Advanced Sciences and Technologies for Security Applications.* s.l.:Springer, pp. 267-288.

Debar, H., Becker, M. & Siboni, D., 1992. *A neural network component for an intrusion detection system.* s.l.:IEEE.

DePaul University, 2002. *A Framework for Incident Response (Draft),* Chicago: Information Security Team.

Dey, P. et al., 2016. Impact of Topology on the Propagation of Cascading Failure in Power Grid. *IEEE Transactions on Smart Grid,* 7(4), pp. 1970- 978.

Di Mattia, F., Galeone, P., De Simoni, M. & Ghelfi, E., 2019. *A Survey on GANs for Anomaly Detection,* s.l.: arXiv.

Djordjevic, I. et al., 2002. Model Based Risk Management of Security Critical Systems. In: C. Brebbia, ed. *WIT Transactions on Modelling and Simulation.* Southampton: WIT Press.

Dua, D. & Graff, C., 2019. *KDD Cup 1999 Data Data Set.* [Online] Available at: [http://archive.ics.uci.edu/ml

Durumeric, Z., Bailey, M. & Halderman, J., 2014. *An Internet-Wide View of Internet-Wide Scanning.* San Diego, UNISEX Association, pp. 65-78.

Ekelhart, A., Fenz, S. & Neubauer, T., 2009. *Automated Risk and Utility Management.* Las Vegas, IEEE, pp. 393-398.

Ektefa, M., Memar, S., Sidi, F. & Affendey, L., 2010. *Intrusion detection using data mining techniques.* Shah Alam, IEEE, pp. 200-203.

ENISA, 2016. *Guidelines for SMEs on the security of personal data,* s.l.: European Union Agency for Cybersecurity (ENISA).

ENISA, 2020. *Inventory of Risk Management / Risk Assessment Method.* [Online] Available at: https://www.enisa.europa.eu/topics/threat-risk-management/risk-management/current-risk/risk-management-inventory/rm-ra-methods
[Accessed 2 6 2020].

ENISA, 2020. *Proactive detection – Measures and information sources,* s.l.: European Union Agency for Cybersecurity (ENISA).

Ericsson, G., 2009. Information Security for Electric Power Utilities (EPUs)—CIGRÉ Developments on Frameworks, Risk Assessment, and Technology. *IEEE Transactions on Power Delivery,* 24(3), pp. 1174 - 1181.

Ester, M., Kriegel, H., Sander, J. & Xu, X., 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. *KDD-96 Proceedings,* 96(34), pp. 226-231.

Eusgeld, I., Nan, C. & Dietz, S., 2011. “System-of-systems” approach for interdependent critical infrastructures. *Reliability Engineering & System Safety,* 96(6), pp. 679-686.

Ezell, B., 2007. Infrastructure vulnerability assessment model (I-VAM). *Risk Analysis: An International Journal,* 27(3), pp. 571-583.

Fan, W., Huang, S. & Mei, S., 2016. Invulnerability of power grids based on maximum flow theory. *Physica A: Statistical Mechanics and its Applications,* Volume 462, pp. 977-985.

Fan, X. et al., 2017. A personal visual analytics on smartphone usage data. *Journal of Visual Languages & Computing,* Volume 41, pp. 111-120.

Farwell, J. & Rohozinski, R., 2011. Stuxnet and the Future of Cyber War. *Survival,* 53(1), pp. 23-40.

Ferrag, M., Maglaras, L., Moschoyiannis, S. & Janicke, H., 2020. Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study. *Journal of Information Security and Applications,* Volume 50, p. 102419.

Filippini, R. & Silva, A., 2011. A modeling language for the resilience assessment of networked systems of systems. In: C. Berenguer, A. Grall & C. Soares, eds. *Advances in Safety, Reliability and Risk Management.* s.l.:CRC Press, p. 401.

Fiore, U., Palmieri, F., Castiglione, A. & De Santis, A., 2013. Network anomaly detection with the restricted Boltzmann machine. *Neurocomputing,* Volume 122, pp. 13-23.

Foglietta, C. et al., 2019. From Detecting Cyber-Attacks to Mitigating Risk Within a Hybrid Environment. *IEEE Systems Journal,* 13(1), pp. 424 - 435.

Francia III, G., Thornton, D. & Dawson, J., 2012. *Security Best Practices and Risk Assessment of SCADA and Industrial Control Systems.* Las Vegas, The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp), p. 1.

Friedman, J. & Rafsky, L., 1979. Multivariate Generalizations of the Wald-Wolfowitz and Smirnov Two-Sample Tests. *The Annals of Statistics,* 7(4), pp. 697-717.

Garfinkel, S., 2010. Digital forensics research: The next 10 years. *Digital Investigation,* Volume 7, pp. S64-S73.

Gertman, D., Folkers, R. & Roberts, J., 2006. *Scenario-based approach to risk analysis in support of cyber security.* Albuquerque, 5th International Topical Meeting on Nuclear Plant Instrumentation Controls, and Human Machine Interface Technology (NPIC and HMIT 2006), pp. 542-547.

Giannopoulos, G., Filippini, R. & Schimmer, M., 2012. *Risk assessment methodologies for Critical Infrastructure Protection. Part I: A state of the art,* s.l.: JRC Technical Notes.

Goh, J., Adepu, S., Tan, M. & Lee, Z., 2017. *Anomaly Detection in Cyber Physical Systems Using Recurrent Neural Networks.* Singapore, IEEE, pp. 140-145.

Gong, D. et al., 2019. *Memorizing Normality to Detect Anomaly: Memory-Augmented Deep Autoencoder for Unsupervised Anomaly Detection.* Seoul, Computer Vision Foundation, pp. 1705-1714.

Goodfellow, I. et al., 2014. Generative Adversarial Nets. *Advances in Neural Information Processing Systems,* Volume 27, pp. 2672-2680.

Gorzelak, K. et al., 2011. *Proactive detection of network security incidents,* s.l.: European Union Agency for Cybersecurity (ENISA).

Govindarasu, M., Hann, A. & Sauer, P., 2012. *Cyber-Physical Systems Security for the Smart Grid,* s.l.: Power Systems Engineering Research Center (PSERC).

Grudziecki, T. et al., 2012. *Proactive detection of security incidents II - Honeypots,* s.l.: European Union Agency for Cybersecurity (ENISA).

Guo, H., Zheng, C., lu, H. & Fernando, T., 2017. A critical review of cascading failure analysis and modeling of power system. *Renewable and Sustainable Energy Reviews,* Volume 80, pp. 9-22.

Haimes, Y. & Horowitz, B., 2004. Adaptive Two-Player Hierarchical Holographic Modeling Game for Counterterrorism Intelligence Analysis. *Journal of Homeland Security and Emergency Management,* 1(3), p. 121.

Hamamoto, A. et al., 2018. Network Anomaly Detection System using Genetic Algorithm and Fuzzy Logic. *Expert Systems with Applications,* Volume 92, pp. 390-402.

Harris, B., 2018. *Generative Adversarial Networks and Cybersecurity: Part 1.* [Online] Available at: https://securityintelligence.com/generative-adversarial-networks-and-cybersecurity-part-1/
[Accessed 7 8 2020].

Harris, B., 2018. *Generative Adversarial Networks and Cybersecurity: Part 2.* [Online] Available at: https://securityintelligence.com/generative-adversarial-networks-and-cybersecurity-part-2/
[Accessed 7 8 2020].

Harrou, F. et al., 2015. Improved principal component analysis for anomaly detection: Application to an emergency department. *Computers & Industrial Engineering,* Volume 88, pp. 63-77.

Hasan, K., Shetty, S. & Ullah, S., 2019. *Artificial Intelligence Empowered Cyber Threat Detection and Protection for Power Utilities.* Los Angeles, IEEE, pp. 354-359.

Hasan, S. & Foliente, G., 2015. Modeling infrastructure system interdependencies and socioeconomic impacts of failure in extreme events: emerging R&D challenges. *Natural Hazards,* 78(3), p. 2143–2168.

Hashemi, M. & Zarif, M., 2020. A novel two-stage distributed structure for reactive power control. *Engineering Science and Technology, an International Journal,* 23(1), pp. 168-188.

Henneaux, P., Labeau, P. & Maun, J., 2012. A level-1 probabilistic risk assessment to blackout hazard in transmission power systems. *Reliability Engineering & System Safety,* Volume 102, pp. 41-52.

Henneaux, P., Labeau, P., Maun, J. & Haarla, L., 2016. A Two-Level Probabilistic Risk Assessment of Cascading Outages. *IEEE Transactions on Power Systems,* 31(3), pp. 2393-2403.

Hines, P., Cotilla-Sanchez, E. & Blumsack, S., 2010. Do topological models provide good information about electricity infrastructure vulnerability?. *Chaos: An Interdisciplinary Journal of Nonlinear Science,* 20(3), p. 033122.

Hinton, G., Sejnowski, T. & Ackley, D., 1984. *Boltzmann Machines: Constraint Satisfaction Networks that Learn,* Pittsburgh: PA: Carnegie-Mellon University, Department of Computer Science.

Hogarth, R., 1987. *Judgement and Choice: The Psychology of Decision.* 2nd edn. ed. Chichester: John Wiley & Sons.

Holme, P., 2002. Edge overload breakdown in evolving networks. *Physical Review E,* 66(3), p. 036119 .

Holme, P. & Kim, B., 2002. Vertex overload breakdown in evolving networks. *Physical Review E,* 65(6), p. 066109.

Homer, J., Ou, X. & Schmidt, D., 2009. *A Sound and Practical Approach to Quantifying Security Risk in Enterprise Networks,* s.l.: Kansas State University Technical Report.

Homer, J. et al., 2013. Aggregating vulnerability metrics in enterprise networks using attack graphs. *Journal of Computer Security,* 21(4), pp. 561-597.

Hong, J., Liu, C. & Govindarasu, M., 2014. Integrated Anomaly Detection for Cyber Security of the Substations. *IEEE Transactions on Smart Grid,* 5(4), pp. 1643-1653.

Huang, X. et al., 2013. The robustness of interdependent clustered networks. *EPL (Europhysics Letters),* 101(1), p. 18002.

Hutchins, E. M., Cloppert, M. J. & Amin, R. M., 2011. Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains. In: J. Ryan, ed. *Leading Issues in Information Warfare & Security Research.* Reading: Academic Publishing international Limited, p. 80.

Igure, V., Laughter, S. & Williams, R., 2006. Security issues in SCADA networks. *Computers and Security,* 25(7), pp. 498-506.

Imamverdiyev, Y. N. & Abdullayeva, F. J., 2020. Deep Learning in Cybersecurity: Challenges and Approaches. *International Journal of Cyber Warfare and Terrorism (IJCWT),* 10(2), pp. 82-105.

ISO/IEC JTC 1/SC 27, 2016. *ISO/IEC 27035-1:2016 Information technology — Security techniques — Information security incident management — Part 1: Principles of incident management,* s.l.: International Organization for Standardization (ISO).

ISO/IEC JTC 1/SC 27, 2018. *ISO/IEC 27000:2018 Information technology — Security techniques — Information security management systems — Overview and vocabulary,* s.l.: International Organization for Standardization (ISO).

ISO/TC 262, 2018. *ISO 31000:2018 Risk management - Guidelines,* s.l.: International Organization for Standardization (ISO).

Jajodia, S., Noel, S. & O'Berry, B., 2005. Topological Analysis of Network Attack Vulnerability. In: V. Kumar, J. Srivastava & A. Lazarevic, eds. *Managing Cyber Threats.* Boston: Springer, pp. 247-266.

Jayasingh, B., Patra, M. & Mahesh, D., 2016. *Security issues and challenges of big data analytics and visualization.* Noida, IEEE, pp. 204-208.

Jha, S., Sheyner, O. & Wing, J., 2002. *Two formal analyses of attack graphs.* Cape Breton, IEEE, pp. 49-63.

Jha, S., Sheyner, O. & Wing, J., 49-63. *Two formal analyses of attack graphs.* Cape Breton, IEEE.

Johnson, P., Lagerström, R. & Ekstedt, M., 2018. *A Meta Language for Threat Modeling and Attack Simulations.* Hamburg, Association for Computing Machinery, pp. 1-8.

Kaegi, M., Mock, R. & Kröger, W., 2009. Analyzing maintenance strategies by agent-based simulations: A feasibility study. *Reliability Engineering & System Safety,* 94(9), pp. 1416-1421.

Kailath, T., 1967. The Divergence and Bhattacharyya Distance Measures in Signal Selection. *IEEE Transactions on Communication Technology,* 15(1), pp. 52-60.

Kalogeraki, E., Apostolou, D., Polemi, N. & Papastergiou, S., 2018. Knowledge management methodology for identifying threats in maritime/logistics supply chains. *Knowledge Management Research & Practice,* 16(4), pp. 508-524.

Kalogeraki, E., Papastergiou, S., Mouratidis, H. & Polemi, N., 2018. A Novel Risk Assessment Methodology for SCADA Maritime Logistics Environments. *Applied Sciences,* 8(9), p. 1477.

Kalogeraki, E., Papastergiou, S., Panayiotopoulos, T. & Polemi, N., 2018. Modeling SCADA Attacks. In: X. Yang , A. Nagar & A. Joshi, eds. *Smart Trends in Systems, Security and Sustainability. Lecture Notes in Networks and Systems.* Singapore: Springer, pp. 47-55.

Kelic, A., Warren, D. & Phillips, L., 2008. *Cyber and Physical Infrastructure Interdependencies,* Albuquerque: Sandia National Laboratories.

Kendall, M., 1948. *Rank correlation methods.* s.l.:s.n.

Kerr, O., 2005. *Digital Evidence and the New Criminal Procedure.* s.l.:Columbia Law Review.

Kim, C. & Obah, O., 2007. Vulnerability Assessment of Power Grid Using Graph Topological Indices. *International Journal of Emerging Electric Power Systems,* 8(6).

Kim, J., Kim, J., Thu, H. & Kim, H., 2016. *Long Short Term Memory Recurrent Neural Network Classifier for Intrusion Detection.* Jeju, IEEE, pp. 1-5.

Knapp, E. & Langill, J., 2019. *Industrial Network Security: Securing critical infrastructure networks for smart grid, SCADA, and other Industrial Control Systems.* 2nd edn. ed. s.l.:Syngress.

Kral, P., 2012. *Incident Handler's Handbook,* s.l.: SANS Technology Institute.

Kravchik, M. & Shabtai, A., 2018. *Detecting Cyber Attacks in Industrial Control Systems Using Convolutional Neural Networks.* Toronto, Association for Computing Machinery, pp. 72-83.

Kriaa, S., Bouissou, M. & Piètre-Cambacédès, L., 2012. *Modeling the Stuxnet attack with BDMP: Towards more formal risk assessments.* Cork, IEEE, pp. 1-8.

Kumar, D., Khan, A., Nayyar, H. & Gupta, V., 2020. *Cyber Risk Assessment Model for Critical Information Infrastructure.* Mathura, IEEE, pp. 292-297.

Kumar, M., Hanumanthappa, M. & Kumar, T., 2011. Forensic Analysis Using Intrusion Detection System. *International Journal of Computer Technology and Applications,* 2(3), pp. 612-618.

Kumar, R. et al., 2017. *End-to-End Network Delay Guarantees for Real-Time Systems Using SDN.* Paris, IEEE, pp. 231-242.

Kuntze, N. & Rudolph, C., 2011. *Secure Digital Chains of Evidence.* Oakland, IEEE.

Lai, R., Qiu, X. & Wu, J., 2019. Robustness of Asymmetric Cyber-Physical Power Systems Against Cyber Attacks. *IEEE Access,* Volume 7, pp. 61342-61352.

Lanzrath, M., Suhrke, M. & Hirsch, H., 2020. HPEM-Based Risk Assessment of Substations Enabled for the Smart Grid. *IEEE Transactions on Electromagnetic Compatibility,* 62(1), pp. 173 - 185.

Launius, S., 2018. *Evaluation of Comprehensive Taxonomies for Information Technology Threats,* s.l.: The SANS Institute.

Lee, W. & Stolfo, S., 2000. A framework for constructing features and models for intrusion detection systems. *ACM transactions on Information and system security (TiSSEC),* 3(4), pp. 227-261.

Line, M. B., 2013. *A Case Study: Preparing for the Smart Grids - Identifying Current Practice for Information Security Incident Management in the Power Industry.* Nuremberg, IEEE, pp. 26-32.

Liu, L. et al., 2018. Detecting and Preventing Cyber Insider Threats: A Survey. *IEEE Communications Surveys & Tutorials,* 20(2), pp. 1397-1417.

Long, M., Wu, C. & Hung, J., 2005. Denial of service attacks on network-based control systems: impact and mitigation. *IEEE Transactions on Industrial Informatics,* 1(2), pp. 85-96.

Maj, M., Reijers, R. & Stikvoort, D., 2010. *Good Practice Guide for Incident Management,* s.l.: European Union Agency for Cybersecurity (ENISA).

Mandia, K., 2001. *Incident Response: Investigating Computer Crime.* 1st edn. ed. s.l.:McGraw-Hill Professional.

Marinos, L., 2016. *ENISA Threat Taxonomy: A tool for structuring threat information,* Heraklion: European Union Agency for Cybersecurity (ENISA).

Markovic-Petrovic, J. & Stojanovic, M., 2014. An Improved Risk Assessment Method for SCADA Information Security. *Elektronika ir Elektrotechnika,* 20(7), pp. 69-72.

Markovic-Petrovic, J., Stojanovic, M. & Rakas, S., 2019. A Fuzzy AHP Approach for Security Risk Assessment in SCADA Networks. *Advances in Electrical and Computer Engineering,* 19(3), pp. 69-75.

Mattioli, R. & Moulinos, K., 2015. *Analysis of ICS-SCADA Cyber Security Maturity Levels,* s.l.: European Union Agency for Network and Information Security.

McGee, S., Frittman, J., Ahn, S. J. & Murray, S., 2016. Implications of cascading effects for the hyogo framework. *International Journal of Disaster Resilience in the Built Environment,* 7(2), pp. 144-157.

McQueen, M., Boyer, W., Flynn, M. & Beitel, G., 2006. *Quantitative Cyber Risk Reduction Estimation Methodology for a Small SCADA Control System.* Kauia, IEEE, pp. 26-226.

Meira, J. et al., 2018. *Comparative Results with Unsupervised Techniques in Cyber Attack Novelty Detection.* Toledo, Springer, pp. 103-112.

Mei, S. et al., 2009. An Improved OPA Model and Blackout Risk Assessment. *IEEE Transactions on Power Systems,* 24(2), pp. 814-823.

Mell, P., Scarfone, K. & Romanosky, S., 2007. *A Complete Guide to the Common Vulnerability Scoring System Version 2.0,* Gaithersburg: National Institute of Standards and Technology (NIST).

Meng, C., 2015. *Research on Dynamic and Static Risk Assessment for Power Information System.* Shanghai, East China University of Science And Technology.

Mitropoulos, S., Patsos, D. & Douligeris, C., 2006. On Incident Handling and Response: A state-of-the-art approach. *Computers & Security,* 25(5), pp. 351-370.

Monni, C., Pezzè, M. & Prisco, G., 2019. *An RBM Anomaly Detector for the Cloud.* Xi'an, IEEE.

Moore, C., 2016. *Detecting Ransomware with Honeypot Techniques.* Amman, IEEE, pp. 77-81.

Motter, A., De Moura, A., Lai, Y. & Dasgupta, P., 2002. Topology of the conceptual network of language. *Physical Review E,* 65(6), p. 065102.

Naik, N. et al., 2019. *Lockout-Tagout Ransomware: A Detection Method for Ransomware using Fuzzy Hashing and Clustering.* Xiamen, IEEE, pp. 641-648.

National Academy Press, 1983. *Risk Assessment in the Federal Government: Managing the Process.* Washington: The National Academies Press.

Ndatinya, V. et al., 2015. Network forensics analysis using Wireshark. *International Journal of Security and Networks,* 10(2), pp. 91-106.

NIST, 2014. *Framework for Improving Critical Infrastructure Cybersecurity,* s.l.: National Institute of Standards and Technology (NIST).

NIST, 2020. *NIST CSRC Taxonomy.* [Online] Available at: https://csrc.nist.gov/topics [Accessed 31 07 2020].

Noel, S. et al., 2016. CyGraph: Graph-Based Analytics and Visualization for Cybersecurity. *Handbook of Statistics,* Volume 35, pp. 117-167.

Noto, K., Brodley, C. & Slonim, D., 2012. FRaC: a feature-modeling approach for semi-supervised and unsupervised anomaly detection. *Data mining and knowledge discovery,* 25(1), pp. 109-133.

Nychis, G. et al., 2008. *An Empirical Evaluation of Entropy-Based Traffic Anomaly Detection.* Vouliagmeni, Association for Computing Machinery, pp. 151-156.

O'Hagan, A. et al., 2006. *Uncertain Judgements: Eliciting Experts' Probabilities.* Chichester: John Wiley & Sons.

Oliva, G., Panzieri, S. & Setola, R., 2012. Modeling and simulation of critical infrastructures. In: *WIT Transactions on State-of-the-art in Science and Engineering.* s.l.:WIT Press.

Olsson, J. & Boldt, M., 2009. Computer forensic timeline visualization tool. *Digital Investigation,* Volume 6, pp. S78-S87.

Oluwasegun, A. & Aminat, A., 2019. *Mitigating Advanced Persistent Threats Using A Combined Static-Rule And Machine Learning-Based Technique.* Abuja, IEEE, pp. 1-6.

Ou, X., Govindavajhala, S. & Appel, A., 2005. *MulVAL: A Logic-based Network Security Analyzer.* Baltimore, USENIX, pp. 113-128.

Ou, X. & Singhal, A., 2011. *Quantitative Security Risk Assessment of Enterprise Networks.* 1st edn. ed. New York: Springer.

Ouyang, M., 2014. Review on modeling and simulation of interdependent critical infrastructure systems. *Reliability Engineering & System Safety,* Volume 121, pp. 43-60.

Panda, A. & Bower, A., 2020. Cyber security and the disaster resilience framework. *International Journal of Disaster Resilience in the Built Environment.*

Papastergiou, S. & Polemi , D., 2017. Securing Maritime Logistics and Supply Chain: The Medusa and MITIGATE approaches. *Maritime Interdiction Operations Journal,* 14(1), pp. 42-48.

Papastergiou, S., Polemi, D. & Karantjias, A., 2015. *CYSM: An Innovative Physical/Cyber Security Management System for Ports.* Los Angeles, Springer, Cham, pp. 219-230.

Papastergiou, S. & Polemi, N., 2018. *MITIGATE: A Dynamic Supply Chain Cyber Risk Assessment Methodology.* London, Springer, Singapore, pp. 1-9.

Pederson, P., Dudenhoeffer, D., Hartley, S. & Permann, M., 2006. *Critical infrastructure interdependency modeling. A Survey of US and International Research,* s.l.: Idaho National Laboratory (INL).

Peng, H., Wu, P., Zhu, J. & Zhang, J., 2011. *Helix: Unsupervised Grammar Induction for Structured Activity Recognition.* Vancouver, IEEE, pp. 1194-1199.

Permann, M. & Rohde, K., 2005. *Cyber Assessment Methods For SCADA Security.* Nashville, Idaho National Laboratory (INL).

Poolsappasit, N., Dewri, R. & Ray, I., 2012. Dynamic Security Risk Management Using Bayesian Attack Graphs. *IEEE Transactions on Dependable and Secure Computing,* 9(1), pp. 61 - 74.

Proença, M. J., Coppelmans, C., Bottoli, M. & de Souza Mendes, L., 2006. Baseline to help with network management. In: J. Ascenso, L. Vasiu, C. Belo & M. Saramago, eds. *e-Business and Telecommunication Networks.* Dordrecht: Springer, pp. 158-166.

Radford, B., Apolonio, L., Trias, A. & Simpson, J., 2018. *Network Traffic Anomaly Detection Using Recurrent Neural Networks.* Springfield, arXiv.

Raghavan, S., 2013. Digital forensic research: current state of the art. *CSI Transactions on ICT,* 1(1), pp. 91-114.

Ralston, P., Graham, J. & Hieb, J., 2007. Cyber security risk assessment for SCADA and DCS networks. *ISA Transactions,* 46(4), pp. 583-594.

Reith, M., Carr, C. & Gunsch, G., 2002. An examination of digital forensic models. *International Journal of Digital Evidence,* 1(3), pp. 1-12.

Richter, J., Kuntze, N. & Rudolph, C., 2010. *Security Digital Evidence.* Oakland, IEEE, pp. 119-130.

Rinaldi, S., Peerenboom, J. & Kelly, T., 2001. Identifying, understanding, and analyzing critical infrastructure interdependencies. *IEEE Control Systems Magazine,* 21(6), pp. 11-25.

Ross, S. M., 2014. *Introduction to Probability Models.* 11th edn. ed. s.l.:Academic Press.

Rousseeuw, P. & Hubert, M., 2018. Anomaly detection by robust statistics. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery,* 8(2), p. 1236.

Roux, M., 2018. A Comparative Study of Divisive and Agglomerative Hierarchical Clustering Algorithms. *Journal of Classification,* 35(2), pp. 345-366.

Rubin-Delanchy, P., Lawson, D. & Heard, N., 2016. Anomaly detection for cyber security applications. In: A. Niall & N. Heard, eds. *Dynamic Networks and Cyber-Security.* s.l.:World Scientific, pp. 137-156.

Ruefle, R. et al., 2014. Computer Security Incident Response Team Development and Evolution. *IEEE Security & Privacy,* 12(5), pp. 12-26.

Rybnicek, M., Tjoa, S. & Poisel, R., 2014. Simulation-Based Cyber-Attack Assessment of Critical Infrastructures. In: J. Barjis & R. Pergl, eds. *Enterprise and Organizational Modeling and Simulation. EOMAS 2014. Lecture Notes in Business Information Processing.* Berlin: Springer, pp. 135-150.

Schnackengerg, D. et al., 2001. *Cooperative Intrusion Traceback and Response Architecture (CITRA).* Anaheim, IEEE, pp. 56-68.

Schneier, B., 1999. Attack Trees. In: *Dr. Dobb's Journal.* New Orleans: Counterpane Internet Security, pp. 21-29.

Schreck, T., 2018. *IT Security Incident Response: Current State, Emerging Problems, and New Approaches,* Erlangen: Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU).

Shafer, G., 1976. *A Mathematical Theory of Evidence.* Princeton: Princeton University Press.

Shannon, C., 2001. A mathematical theory of communication. *ACM SIGMOBILE mobile computing and communications review,* 5(1), pp. 3-55.

Sharma, M., Sheet, D. & Biswas, P., 2016. *Abnormality Detecting Deep Belief Network.* Bikaner, Association for Computing Machinery, pp. 1-6.

Shen, Z., Miao, C., Tao, X. & Gay, R., 2004. *Goal oriented modeling for intelligent software agents.* Beijing, IEEE, pp. 540-543.

Shin, D., Qian, D. & Zhang, J., 2014. Cascading effects in interdependent networks. *IEEE Network,* 28(4), pp. 82-87.

Shiravi, A., Shiravi, H., Tavallaee, M. & Ghorba, A., 2012. Toward developing a systematic approach to generate benchmark datasets for intrusion detection. *Computers & Security,* 31(3), pp. 357-374.

Singhal, A. & Ou, X., 2017. Security Risk Analysis of Enterprise Networks Using Probabilistic Attack Graphs. In: *Network Security Metrics.* s.l.:Springer, pp. 53-73.

Singh, S. et al., 2019. A comprehensive study on APT attacks and countermeasures for future networks and communications: challenges and solutions. *The Journal of Supercomputing,* 75(8), pp. 4543-4574.

Sinha, A., Vyas, R., Subramanian, V. & Vyas, O., 2020. Critical Infrastructure Security: Cyber-Physical Attack Prevention, Detection, and Countermeasures. In: *Quantum Cryptography and the Future of Cyber Security.* s.l.:IGI Global, pp. 134-162.

Smithson, M., 1989. *Ignorance and Uncertainty: Emerging Paradigms.* New York: Springer-Verlag.

Sommer, R. & Paxson, V., 2010. *Outside the Closed World: On Using Machine Learning for Network Intrusion Detection.* Berkeley/Oakland, IEEE, pp. 305-316.

Song, J., Cotilla-Sanchez, E., Ghanavati, G. & Hines, P., 2016. Dynamic Modeling of Cascading Failure in Power Systems. *IEEE Transactions on Power Systems,* 31(3), pp. 2085-2095.

Steinke, J. et al., 2015. Improving Cybersecurity Incident Response Team Effectiveness Using Teams-Based Research. *IEEE Security & Privacy,* 13(4), pp. 20-29.

Stouffer, K. et al., 2015. *Guide to Industrial Control Systems (ICS) Security,* Gaithersburg: NIST special publication SP800-82 Rev.2.

Tang, L., Jing, K., He, J. & Stanley, H., 2016. Complex interdependent supply chain networks: Cascading failure and robustness. *Physica A: Statistical Mechanics and its Applications,* Volume 443, pp. 58-69.

Tankard, C., 2011. Advanced Persistent threats and how to monitor and deter them. *Network Security,* Volume 8, pp. 16-19.

Tantawy, A., Erradi, A. & Abdelwahed, S., 2019. *A Modified Layer of Protection Analysis for Cyber-Physical Systems Security.* Rome, IEEE, pp. 94-101.

Tavallaee, M., Bagheri, E., Lu, W. & Ghorbani, A., 2009. *A detailed analysis of the KDD CUP 99 data set.* Ottawa, IEEE, pp. 1-6.

Ten,, C.-W., Manimaran, G. & Liu, C.-C., 2010. Cybersecurity for Critical Infrastructures: Attack and Defense Modeling. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans,* 40(4), pp. 853 - 865.

Ten, C., Hong, J. & Liu, C., 2011. Anomaly Detection for Cybersecurity of the Substations. *IEEE Transactions on Smart Grid,* 2(4), pp. 865-873.

Theocharidou, M. & Giannopoulos, G., 2015. *Risk assessment methodologies for critical infrastructure protection. Part II: A new approach. tech report EUR 27332 EN,* Luxembourg: Luxembourg Publications Office.

Tianfield, H., 2017. Data mining based cyber-attack detection. *System Simulation Technology,* 13(2), p. 15.

Tian, W. et al., 2020. Prospect Theoretic Study of Honeypot Defense Against Advanced Persistent Threats in Power Grid. *IEEE Access,* Volume 8, pp. 64075-64085.

Titchener, M., 1998. *Deterministic computation of complexity, information and entropy.* San Diego, IEEE, p. 326.

Tøndel, I., Line, M. & Jaatun, M., 2014. Information security incident management: Current practice as reported in the literature. *Computers & Security,* Volume 45, pp. 42-57.

Tuor, A. et al., 2017. *Deep Learning for Unsupervised Insider Threat Detection in Structured Cybersecurity Data Streams.* San Francisco, AAAI Publications.

Usmani, K., Mohapatra, A. & Prakash, N., 2013. An Improved Framework for Incident Handling. *Information Security Journal: A Global Perspective,* 22(1), pp. 1-9.

Veeramachaneni, K. et al., 2016. *AI^2: Training a Big Data Machine to Defend.* New York, IEEE, pp. 49-54.

Verendel, V., 2009. *Quantified Security is a Weak Hypothesis: A Critical Survey of Results and Assumptions.* Oxford, Association for Computing Machinery, pp. 37-50.

Wagner, C., Dulaunoy, A. & Wagener, G., 2016. *MISP: The Design and Implementation of a Collaborative Threat Intelligence Sharing Platform.* New York, ACM, pp. 49-56.

Wang, W. & Chen, G., 2008. Universal robustness characteristic of weighted networks against cascading failure. *Physical Review E,* 77(2), p. 026101.

Wei, D., Luo, X. & Zhang, B., 2012. Analysis of cascading failure in complex power networks under the load local preferential redistribution rule. *Physica A: Statistical Mechanics and its Applications,* 391(8), pp. 2771-2777.

Werlinger, R., Muldner, K., Hawkey, K. & Beznosov, K., 2010. Preparation, detection, and analysis: the diagnostic work of IT security incident response. *Information Management & Computer Security,* 18(1), pp. 26-42.

Wu, Y., 2013. SCADA system information security technology. *Autom. Panor,* Volume 2, p. 98–100.

Wyss, G. & Durán, F., 2001. *OBEST: The Object-Based Event Scenario Tree Methodology,* Albuquerque: Sandia National Laboratories.

Xiaolin, C., Xiaobin, T., Yong, Z. & Hongsheng, X., 2008. *A Markov Game Theory-Based Risk Assessment Model for Network Information System.* Hubei, IEEE, pp. 1057-1061.

Yang, L., Cao, X. & Geng, X., 2019. A novel intelligent assessment method for SCADA information security risk based on causality analysis. *Cluster Computing,* 22(3), p. 5491–5503.

Yan, R., Zheng, Q. & Peng, W., 2008. *Multi-scale entropy and renyi cross entropy based traffic anomaly detection.* Guangzhou, IEEE, pp. 554-558.

Yan, X. & Zhang, J., 2013. Early detection of cyber security threats using structured behavior modeling. *ACM Transactions on Information and System Security 5,* Volume 5.

Yavanoglu, O. & Aydos, M., 2017. *A review on cyber security datasets for machine learning algorithms.* Boston, IEEE, pp. 2186-2193.

Ye, N., Emran, S., Li, X. & Chen, Q., 2001. *Statistical process control for computer intrusion detection.* Anaheim, IEEE, pp. 3-14.

Yu, T. et al., 2017. *PSI: Precise Security Instrumentation for Enterprise Networks..* San Diego, Internet Society.

Zadeh, L. A., 1965. Fuzzy sets. *Information and Control,* 8(3), pp. 338-353.

Zambon, E., Etalle, S., Wieringa, R. J. & Hartel, P., 2011. Model-based qualitative risk assessment for availability of IT infrastructures. *Software & Systems Modeling,* 10(4), pp. 553-580.

Zhang, S., Ou, X., Singhal, A. & Homer, J., 2011. *An empirical study of a vulnerability metric aggregation method.* Las Vegas, s.n.

Zhang, X., Zhan, C. & Chi, K., 2017. Modeling the Dynamics of Cascading Failures in Power Systems. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems,* 7(2), pp. 192-204.

Zio, E., 2018. The future of risk assessment. *Reliability Engineering & System Safety,* Volume 177, pp. 176-190.

Zio, E. & Golea, L., 2012. Identifying groups of critical edges in a realistic electrical network by multi-objective genetic algorithms. *Reliability Engineering & System Safety,* Volume 99, pp. 172-177.

Zio, E. & Sansavini, G., 2011. Modeling Interdependent Network Systems for Identifying Cascade-Safe Operating Margins. *IEEE Transactions on Reliability,* 60(1), pp. 94-101.