SAFECARE

Integrated cyber-physical security for health services

Grant Agreement Number: 787002

State-of-the-art analysis and known vulnerabilities

Deliverable 3.2

Lead Author: ISEP

Contributors: APHM, CCS, EMAUG, ENC, SPF, ISEP, CNAM, KUL, ISMB, CSI, ASLTO5, AMC, MS, SEM, PEN, FMI, BEIA

Deliverable classification: PU

Co-funded by the Horizon 2020 Framework Programme of the European Union

Version Control Sheet

Title	*Sate of the art analysis and know vulnerabilities*
Prepared By	*Isabel Praça / Eva Maia*
Approved By	
Version Number	
Contact	

Revision History:

Version	Date	Summary of Changes	Initials	Changes Marked
V0.0		Initial draft of the structure	ISEP	
V1.0	08/02/2019	First complete version	All	
V2.0	15/02/2019	More information added	ISEP	
V2.1	20/02/2019	Review by David Lancelin	CCS	
V2.2	21/02/2019	Changes according to the review	ISEP	

The research leading to these results has received funding from the European Union's Horizon 2020 Research and Innovation Programme, under Grant Agreement no 787002.

1 Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

AAA	Authentication, Authorization, Accounting
ACS	Access Control System
ADMIT	Anomaly-based Data Mining for InTrusions
AES	Advanced Encryption Standard
AI	Artificial Intelligence
ANN	Artificial Neural Networks
ATP	Advanced Threat Protection
BCP	Business Continuity Plan
BSM	Basic Security Module
BYOD	Bring Your Own Device
CATO	Corporate Account Take Over
CBRNE	Chemical, Biological, Radiological, Nuclear, and Explosive
CEI	Critical Energy Infrastructures
CFATS	Chemical Facility Anti-Terrorism Standards
CI	Critical Infrastructure
CIP	Critical Infrastructure Protection
CJEU	Court of Justice of the European Union
COI	Chemicals of interest
CPS	Cyber-Physical Systems
CT	Computer tomography
DDoS	Distributed Denial of Service
DHS	Department of Homeland Security
DLP	Data Loss Prevention
DoS	Denial of Service
DRP	Disaster Recovery Plan
ECC	Elliptic curve cryptography
ECtHR	European Court on Human Rights
FDA	Food & Drug Administration
GA	Genetic Algorithm
GDPR	General Data Protection Regulation
GIDA	Game Theory Inspired Defense Architecture
HIDS	Host Intrusion Detection Systems
HIPS	Host Intrusion Prevention System

HSA	Hybrid Situation Awareness
IDS	Intrusion Detection System
IMD	Implantable Medical Devices
IoT	Internet of things
IPS	Intrusion Prevention System
IRS	Intrusion Response Systems
ISG	Information Security Governance
ISMICT	International Symposium on Medical Information and Communication Technology
ISO	International Standards Organization
KDD	Knowledge Discovery and Data
KVM	Keyboard, video, and mouse
MARS	Multivariate Adaptive Regression Splines
MBR	Master boot record
MDM	Mobile Device Management
MEDiSN	Medical Emergency Detection in Sensor Networks
NAT	Network Address Translation
NIDS	Network Intrusion Detection Systems
NIS	Network and information systems
OES	Operators of Essential Services
OS	Operation Systems
OT	Operation technology
PM	Physiological Monitors
PPS	Personnel into a Physical Security System
PSIM	Physical Security Information Management
RA	Reference architecture
RADIUS	Remote Authentication Dial-In User Service
RAID	Redundant Array of Inexpensive/Independent Disks
RBPS	Risk-based performance standards
RP	Relay Points
SaaS	Software as a service
SCADA	Supervisory control and data acquisition
SCP	Smart City Platforms
SDLC	Secure Development Life Cycle
SIEM	Security Information and Event Management

SVM	Support Vector Machine
UCB	Upper Confidence Bound
VCA	Video Content Analysis
VMS	Video Management System
WAF	Web Application Firewall

The SAFECARE Project

Over the last decade, the European Union has faced numerous threats that quickly increased in their magnitude, changing the lives, the habits and the fears of hundreds of millions of citizens. The sources of these threats have been heterogeneous, as well as weapons to impact the population. As Europeans, we know now that we must increase our awareness against these attacks that can strike the places we rely upon the most and destabilize our institutions remotely. Today, the lines between physical and cyber worlds are increasingly blurred. Nearly everything is connected to the Internet and if not, physical intrusion might rub out the barriers. Threats cannot be analyzed solely as physical or cyber, and therefore it is critical to develop an integrated approach in order to fight against such combination of threats. Health services are at the same time among the most critical infrastructures and the most vulnerable ones.

Health service providers are widely relying on information systems to optimize organization and costs, whereas ethics and privacy constraints severely restrict security controls and thus increase vulnerability. The aim of this proposal is to provide solutions that will improve physical and cyber security in a seamless and cost-effective way. It will promote new technologies and novel approaches to enhance threat prevention, threat detection, incident response and mitigation of impacts. The project will also participate in increasing the compliance between security tools and European regulations about ethics and privacy for health services. Finally, project pilots will take place in the hospitals of Marseille, Turin and Amsterdam, involving security and health practitioners, in order to simulate attack scenarios in near-real conditions. These pilot sites will serve as reference examples to disseminate the results and find customers across Europe.

Executive Summary

The challenge of SAFECARE is to bring together the most advanced technologies from the physical and cyber security spheres to achieve a global optimum for systemic security and for the management of combined cyber and physical threats and incidents, their interconnections and potential cascading effects. The project focuses on health service infrastructures and works towards the creation of a comprehensive protection system, which will cover threat prevention, detection, response and, in case of failure, mitigation of impacts across infrastructures, populations and environment.

Over a 36-month time frame, the SAFECARE Consortium will design, test, validate and demonstrate 13 innovative elements, developed in the Document of Actions, which will optimize the protection of critical infrastructures under operational conditions. These elements are interactive, cooperative and complementary, aiming at maximizing the potential use of each individual element. The consortium will also engage with leading hospitals, national public health agencies and security Stakeholders across Europe to ensure that SAFECARE's global solution is flexible, scalable and adaptable to the operational needs of various hospitals across Europe and meet the requirements of newly emerging technologies and standards.

To develop a useful and comprehensive protection system it is crucial study the security and crisis management solutions in health infrastructures already known. It is also very important understand the typical physical and cyber vulnerabilities in the healthcare sector. Thus, this deliverable (D3.2) will include the list of physical and cyber vulnerabilities, and how they might impact the likelihood of attacks and their effects and will also provide the state-of-the-art analysis about security controls in health infrastructures.

1 Introduction

Over the last decade cybercrime is the greatest threat to every sector in the world. Due to its critical and vulnerable infrastructure, the health sector is an easy target for hackers. Moreover, healthcare organizations are some of the entities we trust the most and that hold the most valuable information, so exploiting its vulnerabilities brings a huge potential for financial and political gain.

Hacking and malware (including ransomware) are the leading causes of health data breaches. These data breaches result in large financial losses, but also in loss of reputation and reduced patient safety. Some know data breaches are:

- 2017 WannaCry attack infected more than 300,000 computers across the world demanding that users pay bitcoin ransoms. Despite this attack was not specifically directed at healthcare organizations, other ransomware has specifically targeted the healthcare sector and, according to US media, the Presbyterian Medical Centre shut down for 10 days until it paid a $17,000 ransom.
- one UK health care trust suffered an unspecified cyber-attack which led to the shutdown of its IT systems and cancellation of almost all planned operations and outpatient appointments for four days.
- Medjack ("Medical Device Hijack") is another known attack that injects malware into unprotected medical devices to move laterally across the hospital network

The aim of project SAFECARE is to provide solutions that will improve physical and cyber security to prevent and detect attacks, to promote incident responses and mitigate the impacts.

1.1. Deliverable 3.2

The main objective of task 3.2 was analyze the state-of -the-art of physical and cyber security solutions and detail the most common physical and cyber vulnerabilities. To design new solutions that will improve the physical and cyber security in healthcare sector, it is very important to be aware of the physical and cyber vulnerabilities in health infrastructures, how they facilitate malicious actions, and how they may increase the likelihood and impact of human errors and system failures. It is also crucial to have a clear understanding of actual security strategies and controls implemented at health targeted infrastructures (medical devices, ICT and network infrastructures, e-health services, information systems, etc.).

This document, a first version of the Deliverable 3.2, presents a state-of-the-art analysis about security controls in health infrastructures. It is also described some typical physical and cyber vulnerabilities, and how they might impact the likelihood of attacks and their effects. The ethics, privacy and data protection legislations are also discussed.

2 Related Projects

In this section we will present the on-going projects approved by European Commission under the topic *CIP-01-2016-2017: Prevention, detection, response and mitigation of the combination of physical and cyber threats to the critical infrastructure of Europe*.

Project Name	Link to the website
Defending the European Energy Infrastructures (DEFENDER)	http://defender-project.eu/
Scalable multidimensionAl sitUation awaReness sOlution for protectiNg european ports (SAURON)	https://www.sauronproject.eu/
Strategic, Tactical, Operational Protection of water Infrastructure against cyber-physical Threats (STOP-IT)	https://stop-it-project.eu/
RESIlience enhancement and risk control platform for communication infraSTructure Operators (RESISTO)	http://www.resistoproject.eu/
Integrated Framework for Predictive and Collaborative Security of Financial Infrastructures (FINSEC)	https://www.finsec-project.eu/index.html

Table 1. Related Projects Overview

2.1 DEFENDER

Modern critical infrastructures are increasingly turning into distributed, complex cyber-physical systems that need proactive protection and fast restoration to mitigate physical or cyber incidents or attacks. Most importantly, combined cyber-physical attacks need to be considered, which are much more challenging, and are expected to become the most intrusive attack. This is particularly true for the Critical Energy Infrastructures (CEI). Defending the European Energy Infrastructures (DEFENDER[1]) is a project that will adapt, integrate, upscale, deploy and validate a number of different technologies and operational blueprints. The project has a vision to develop a new approach to safeguard existing and future European CEI operation over cyber-physical-social threats, based on novel protective concepts for lifecycle assessment, resilience and self-healing offering "security by design", and advanced intruder inspection and incident mitigation systems. Moreover, DEFENDER will create a culture of security, where trusted information exchange between trained employees and volunteers will complement cyber-physical protection, while preserving the privacy of the citizens involved.

To achieve its vision, DEFENDER will implement the four strategies:

[1] http://defender-project.eu/

- **Assess Risk**. This strategy gives to energy sector asset owners, utilities and service providers a thorough understanding of their current security posture, enabling them to continually assess evolving cyber/physical threats and vulnerabilities, their risks, and potential countermeasures.
- **Protective measures** to reduce risk by design. New protective (proactive) measures will be developed to reduce system risks, including vulnerabilities and emerging threats. These measures will be built into next-generation CEI and will help the electricity infrastructures stakeholders to offer CEI "defense in depth and by design" and offer components that are interoperable, extensible, and able to operate even in a degraded condition during a cyber incident.
- **Manage Incidents**. Managing incidents is critical, as physical disasters can be generalized, cyber assaults can be sophisticated, and at the end any system can become vulnerable to emerging threats as absolute security is not possible. When protective measures are not applied or fail to prevent an incident, detection, remediation, recovery, and restoration activities should minimize its impact and quickly return to normal operations.
- **Build a Culture of Security.** Post-incident analysis and forensics enable CEI stakeholders to learn from the incident. Integrated with reliability practices, risk management practices will be periodically reviewed and challenged to confirm that established security controls remain in place, while physical and cyber-security best practices should be disseminated at pan-European level.

To evaluate, validate and demonstrate how and to what extent the DEFENDER framework will enable and effective holistic cyber-physical security, DEFENDER will involve four real pilots in three different countries:

1. Bulk Energy Generation in Belgium;
2. Wind Farm Decentralized RES Generation in Italy;
3. Transmission System (TSO) power network in Slovenia;
4. Distribution System (DSO) power network C&I Prosumer in Italy.

2.2 SAURON

Scalable multidimensionAl sitUation awaReness sOlution for protectiNg european ports (SAURON[2]) project proposes a holistic situation awareness concept as an integrated, scalable and yet installation-specific solution for protecting EU ports and its surroundings. This solution combines the more advanced physical features with the newest techniques in prevention, detection and mitigation of cyber-threats, including the synthetic cyber space understanding using new visualization techniques (immersive interfaces, cyber 3D models and so on). In addition, a Hybrid Situation Awareness (HSA) application capable of determining the potential consequences of any threat will show the potential cascading effect of a detected threat in the two different domains (physical and cyber). On the other hand, through SAURON approach the public in the surroundings and the rescue/security teams will be able to be informed on any potential event/situation that could put in risk their integrity. Thus, SAURON proposes as main objective to ensure an adequate level of both physical and cyber protection for the EU ports and limiting,

[2] https://www.sauronproject.eu/

as far as possible, the detrimental effects for the society and citizens of a combined attack (physical & cyber) to an EU port.

The SAURON platform is composed by four main pillars:

- **PSA: A complete physical SA** system which includes novel features such as: dynamic location of resources and assets; location, management and monitoring of sensors, including cameras mounted on drones (under the conditions of and in compliance with all pertinent legal requirements at national and European level); security perimeter control; robust and secure tactical communication network and so on. This PSA system will be adapted to the EU ports characteristics, requirements and needs for protecting them against any kind of physical threat.
- **CSA: An advanced and scalable cyber SA** system capable of preventing and detecting threats and in case of a declared attack, capable of mitigating the effects of the infection/intrusion. This CSA system will include new visualization paradigms for the cyber space.
- **HSA: A Hybrid SA system receiving both physical and cyber alarms** on potential threats from the real world and the cyber space respectively. The HSA application and will show the potential consequences/effects of these threats in the other planes including cascading effects.
- **EPWS: An Emergency Population Warning System**, allowing local, regional, or national authorities to contact members of rescue/security teams and the public, also integrating Smart City Platforms (SCP) in order to warn them and draw their attention to an immediate hazard. This will encourage them to take a specific action in response to an emergency event or threat.

2.3 STOP-IT

Water critical infrastructures (CIs) are essential for human society, life and health and they can be endangered by physical and cyber threats with severe societal consequences. To address this, STOP-IT (Strategic, Tactical, Operational Protection of water Infrastructure against cyber-physical Threats[3]) assembles a team of major Water Utilities, industrial technology developers, high tech SMEs and top EU R&D providers, to find solutions to protect critical water infrastructure against physical and cyber threats.

The main aims of STOP-IT are:

- Identification of current and future water infrastructure risks;
- co-development of an all-hazards risk management framework for the phisical and cyber protection of critical water infrastructures.

The STOP-IT solutions are going to include mature technologies, such as public warning systems and smart locks that are improved by their combination and embedment. Novel technologies will also be included, for example, fault-tolerant control strategies for supervisory control and data acquisition (SCADA) integrated sensors, high-volume real-time sensor data protection via

[3] https://stop-it-project.eu/

blockchain schemes, irregular human detection using new computer vision methods and WiFi and efficient water contamination detection algorithms.

The solutions will be tested and demonstrated with a front-runner/follower approach, where four advanced front-runner utilities, Aigües de Barcelona (Spain), Berliner Wasserbetriebe (Germany), Mekorot (Israel) and Oslo VAV (Norway) are twinned with Hessenwasser (Germany), Bergen Kommune (Norway), Emasagra (Spain) and DeWatergroep (Belgium), to stimulate mutual learning, transfer and uptake of the STOP-IT solutions.

2.4 RESISTO

RESIlience enhancement and risk control platform for communication infraSTructure Operators (RESISTO[4]) is an innovative solution for Communication Infrastructure providing holistic (cyber or physical) situation awareness and enhanced resilience. RESISTO will help Communications Infrastructures Operators to take the best countermeasures and reactive actions exploiting the combined use of risk and resilience preparatory analyses, detection and reaction technologies, applications and processes in the physical and cyber domains.

RESISTO main objective is to improve risk control and resilience of modern Communication CIs, against a wide variety of cyber-physical threats, being those malicious attacks, natural disasters or even un-expected. It has 5 main functionalities:

- **Identification**: Define and maintain a knowledge base on physical and cyber security risks to systems, assets, data, and capabilities characterizing Telecommunication Cls.
- **Protection**: Develop and implement the appropriate safeguards to ensure delivery of CI services.
- **Detection**: Early and timely discover the occurrence of physical and cyber security events.
- **Reaction**: Orchestrate and implement effective response to a detected security event.
- **Mitigation**: Develop and implement the appropriate activities to mitigate the impacts of the threat and to restore as much as possible capabilities or services that were impaired due to a security event.

2.5 FINSEC

Integrated Framework for Predictive and Collaborative Security of Financial Infrastructures (FINSEC [5]) is a 3-years security innovation action project funded by the European Commission. It will develop, demonstrate and bring to market an integrated, intelligent, collaborative and predictive approach to the security of critical infrastructures in the financial sector. To this end, FINSEC will introduce, implement and validate a novel reference architecture for integrated physical and cyber security of critical infrastructures, which will enable handling of dynamic, advanced and asymmetric attacks, while at the same time boosting financial organizations' compliance to security standards and regulations. As a result, FINSEC will provide a blueprint for the next generation security systems for the critical infrastructures of the financial sector.

[4] http://www.resistoproject.eu/

[5] https://www.finsec-project.eu/index.html

FINSEC will provide a mature implementation of the reference architecture (RA), based on the enhancement and integration of novel solutions of the partners (eg., Anomaly Detection, AI CCTV Analytics, Risk Assessment Engines, Collaborative Risk Analysis & Management, Compliance), which will be bundled in a toolbox. The RA implementation and the toolbox will be validated through realistic pilots involving stakeholders in the identification, assessment and mitigation of threats. The five pilots involve high-impact scenarios including SWIFT network protection, buildings and ATM networks security, peer-to-peer payments network protection, risk assessment for insurance purposes and securing financial SMEs. The pilots will engage >=500 security & finance experts, while providing a representative coverage of the financial services industry (banking, capital management, insurance, card & P2P payments), which is a sound basis for FINSEC's broader impact. Towards maximum impact, FINSEC will establish an ecosystem of security solutions for the financial sector, which will be supported by the partners' dense network of sales, marketing, standardization and regulation channels worldwide. The key aspects of this project are:

- **Business**: FINSEC is a joint effort of prominent stakeholders in the financial sector and global leaders in physical & IT security, towards introducing a novel standards-based reference architecture (RA) for integrated (cyber & physical) security.
- **Technical**: FINSEC will provide a mature implementation of the RA, based on the enhancement and integration of novel solutions of the partners (e.g., Anomaly Detection, AI CCTV Analytics, Risk Assessment Engines, Collaborative Risk Analysis & Management, Compliance), which will be bundled in a toolbox.
- **Technological**: The RA implementation and the toolbox will be validated through realistic pilots involving stakeholders in the identification, assessment and mitigation of threats.
- **Pilots and Applications**: The five pilots involve high-impact scenarios including SWIFT network protection, buildings and ATM networks security, peer-to-peer payments network protection, risk assessment for insurance purposes and securing financial SMEs. The pilots will engage security & finance experts, while providing a representative coverage of the financial services industry (i.e., banking, capital management, insurance, card & P2P payments), which is a sound basis for FINSEC's broader impact.

3 Physical Security

In the last years, research on *Critical Infrastructure Protection* (CIP) has become one of the primary matters for the development of modern societies. Water, power, banking, transportation and communication systems are only a few examples of essential infrastructures to daily human activities and their protection is a concept relating to the preparedness and response to serious incidents that could threaten them.

Physical security is one of the most fundamental aspects of the protection. It concerns the use of physical controls for protecting premises, sites, facilities, buildings or other physical assets belonging to the critical sectors. The application of physical security is the process of using layers of physical protective measures to prevent unauthorized access or harm. This harm can involve terrorism, theft, destruction, sabotage, vandalism, espionage, and similar.

A crucial element which contributes to improve the protection of critical infrastructures seamlessly is technology. Thanks to fast technological progress, it is possible to build complex surveillance systems able to integrate heterogeneous sources which can monitor environments

potentially at risk. In this way, *resilience* may be accomplished, for example, through hardening the system by adding *redundancy* and *robustness*. However, for enhancing significantly the protection level, integration of different technologies is not enough, but a collaborative approach is essential. A strong protection calls for interoperability not only among ICT systems, but also among different operators, organizations, companies, and any other entity belonging to the public security sector. Nevertheless, the security designer must determine how best to combine elements like fences, barriers, sensors, procedures, security systems, and security personnel into a Physical Security System (PPS) that can achieve the protection objectives. For this reason, another important element is to conduct a systematic evaluation in which quantitative and/or qualitative techniques are used to predict overall system effectiveness, by identifying exploitable weaknesses in asset protection for a given threat[6].

3.1 Physical Security Solutions

Protection of Critical Infrastructures (CI). The Council of the European Union states "protection" means all activities aimed at ensuring the functionality, continuity and integrity of critical infrastructures in order to deter, mitigate and neutralize a threat, risk or vulnerability[7].

As reflected in this definition, the term protection is a broader concept in which three main aspects can be individuated: *safety*, *security*, and *emergency*. Safety involves the safeguard or protection against events or situations such as malfunctioning or faults of systems, accidents caused by people either intentionally or not. Instead, security refers to the safeguard or protection of people and assets against attacks, assaults, and damages carried out voluntarily by individual or organizations in order to harm. This includes civil disturbances, sabotage, theft of critical property or information, pilferage, extortion or other intentional attacks on assets by a human. Emergency refers to all those activities which have to be undertaken when safety and/or security fail and consequently require intervention of rescue teams such as first responders, civil protection, fire brigade, and so on. Thus, it regards the containment of attacks/hazard and minimization of damages.

In the case of chemical facilities, like those hosted into hospitals, the Department of Homeland Security (DHS) of the United States implements the *Chemical Facility Anti-Terrorism Standards (CFATS)* program to better protect those facilities from the threat of terrorist attack or exploitation. The CFATS program identifies and regulates high-risk chemical facilities to ensure they have security measures in place to reduce the risks associated with specific chemicals of interest (COI). Any entity that has a COI at certain quantities and concentrations, be it a university, hospital, community swimming pool, or even a brewery, is a chemical facility that is subject to CFATS and must submit a Top-Screen survey to give the Department basic information about the facility and the COI it possesses. Facilities at high-risk must submit a security plan and implement a variety of security measures—both physical and cyber in nature—to meet applicable risk-based performance standards (RBPS). Nevertheless, CFATS is a non-prescriptive program, meaning that

[6] **Drago, Annarita.** Methods and Techniques for Enhancing Physical Security of Critical Infrastructures. *PhD Thesis.* 2015

[7] **Council of the European Union.** Council directive 2008/114/ec on the identification and designation of european critical infrastructure (eci) and the assessment of the need to improve their protection. *Official Journal of the European Union.*

facilities are able to choose different security measures to meet the corresponding RBPS, which also best fit their individual facility concerns[8].

Physical security strategies. In order to face critical threats for physical security, the strategies that can be adopted are fundamentally three: *proactive* (stop the event before it occurs), *reactive* (act to limit the impact of the event, if the previous strategy fails), and *forensic* (get information to put the system back in operation). Obviously, from a security perspective, the best is to be as proactive as possible (stopping terrorists before they burst their bomb is preferable than finding the culprits), but it is very hard to meet this objective from a technological point of view. Using disparate technologies surely will help to have a reactive behavior to threats, while the proactive effect will be strictly limited to the motivations inciting an attacker. Generally, a proactive approach involves assessment methodologies, more or less detailed and complex, able to analyze or prove the protection levels of an infrastructure. Eventually, the forensic strategies can help to understand the dynamics of a successful attack in order to discover where and why the system failed. In addition, they allow to gather important information useful for facing future threats.

Physical security measures. The choice among the physical security measures to be adopted dep-ends greatly on what assets need to be protected, where they are located, and what threats, vulnerabilities, and risks pertain to them. Thus, applying an appropriate level of protection requires a specific understanding of environment under consideration as well as the threats to which it is exposed. In order to accomplish this, it is clear that an effective design have to be carried out. So, an effective design involves the use of multiple layers of interdependent systems and covers all the means and technologies for perimeter, external and internal protection such as barrier, lighting, different kinds of sensors, closed-circuit television, access control, and people. In this phase the choice of technological security systems and the adoption of architectures for integrating such systems play an important role, since they are effective means that contribute to increase resilience of a CI, providing early warning of threats and improving the response to eventual disasters. However, the activities tied to security design are very difficult considering the complexity and interconnectedness of current infrastructures, the lack of standards in physical security matters, the diversity of threats, and the different local regulations. Furthermore, the cost of physical security is not insignificant. Reaching an appropriate balance between adequate levels of protection and the cost of the systems enabling physical protection can be hard. Too little security leaves vulnerabilities in place, increasing risks. Too much security may mitigate threats and vulnerabilities and reduce risks but leads to unnecessary expenditures. Inefficient application of security controls (spending more than you need for a physical security service or product) may use scarce resources that otherwise would be available for additional protective measures[9].

Consequently, a trade-off between costs and effective protection based on their contributions to risk reduction is necessary. Translating strategic security objectives into wise choices is a

[8] **US Department of Homeland Security.** Critical Infrastructure: Cyber and Physical Security Essential Components in Securing Our Nation's High-Risk Chemicals. *American Infrastructure.* [Online] https://americaninfrastructuremag.com/critical-infrastructure-cyber-physical-security-essential-components-securing-nations-high-risk-chemicals/.

[9] Halibozek, Gerald L Kovacich and Edward P. *The manager's handbook for corporate security: establishing and managing a successful assets protection program.* s.l. : Butterworth-Heinemann.

challenging design problem both when designing a new physical security system and when upgrading to an existing system. In the context of infrastructure resilience and protection some considerations have been made for guaranteeing a certain risk level[10].

3.2 Technologies enabling physical security solutions

In the past decade, the security landscape has dramatically changed with the introduction of several new security technologies to deter, detect and react to more disparate attacks. Organizations are constantly introducing new technologies and upgrading existing ones in order to ensure the security of their most valuable assets such as people, infrastructure, and property. Typical systems include access control systems, CCTV systems, intrusion detection systems, *firefighting systems*, *Chemical, Biological, Radiological, Nuclear, and Explosive* (CBRNE) *sensors*, *content video analytics*, *intelligent sound detection*, *perimeter intruder detection*, and so on[11].

Redundancy and diversity of sensing technology is essential to build effective surveillance systems, but this increases the number of sensing devices and, consequently, of alarms to be managed. So, the integration of such security systems has been one of the primary requirements in the scenario of the physical security. However, regarding the information integration and management, the industry is still underdeveloped. In fact, potential capabilities of traditional systems are limited by their low abilities in data analysis and interpretation, resulting in an inadequate prevention and real-time reaction[10].

In practical applications, each monitoring system is managed by means of an ad-hoc software platform. The traditional surveillance solutions include, for example, *Video Management System* (VMS), *Access Control System* (ACS), etc. They provide an overview of the installed devices (with a related report of diagnostic, warning, and alert messages) and a set of basic functionalities (e.g. for data acquisition, control, configuration, and rules setting). In this way, each event is handled separately without an effective information sharing, resulting in a very fragmented approach to the physical security[12].

Furthermore, the separated use of multiple systems can even complicate the security management. For example, take a human operator at the control center: in case of attack he may be inundated of alert messages, coming from multiple separated interfaces, one for each management system of the single technology. Hence, industrial needs require supporting platforms capable of integrating monitoring components with data processing sub-systems, with also final consumers of produced warnings. A well-designed integrated security system allows the full control of a CI, unifying alarm signaling, management and control procedures, optimizing

[10] Drago, Annarita. Methods and Techniques for Enhancing Physical Security of Critical Infrastructures. *PhD Thesis.* 2015.

[11] Vittorini V. Flammini F., Pappalardo A. *Effective Surveillance for Homeland Security: Balancing Technology and Social Issues, chapter Challenges and Emerging Paradigms for Augmented Surveillance.* s.l. : Chapman and Hall/CRC, 2013. 9781439883242.

[12] Drago, Annarita. Methods and Techniques for Enhancing Physical Security of Critical Infrastructures. *PhD Thesis.* 2015.

the human resource necessary. The general architecture is composed by three fundamental components: field subsystems, communication network, and supervision and control system[13].[14]

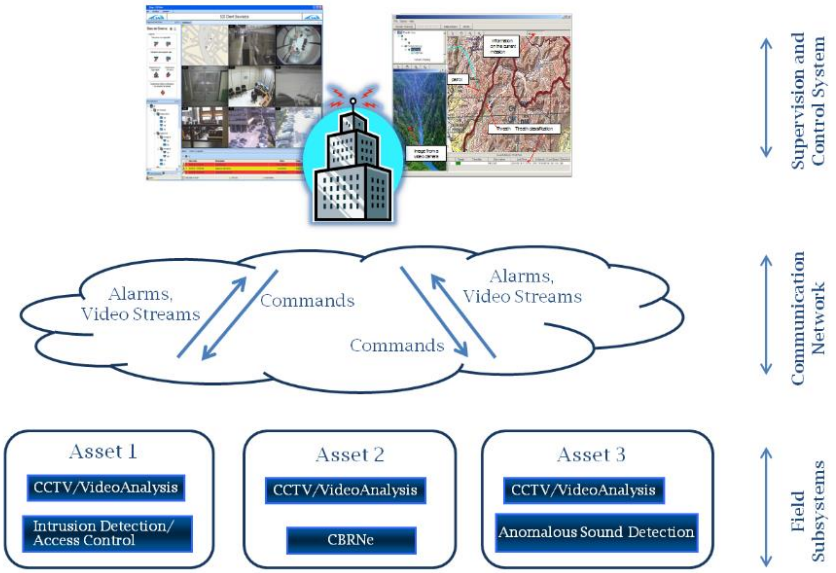

Figure 1 – General architecture of an integrated security system

The different subsystems are distributed within the infrastructure and are able to send alarms and video streams to the supervision and control system through the communication network. The supervision and control system is in charge of analyzing and possibly elaborating data in order to support the decision making. In addition, it can send commands to field subsystems still through the communication network. When the emergencies occur, one of the fundamental tasks is to get the right information to allocate the right resources. Quick collaboration between local, state, and, in some instances, federal agencies is critical to saving lives and critical assets. Therefore, security systems must be tightly integrated with policies, procedures, and protocols to empower decision makers to quickly make the proper decision.

3.2.1 Physical Security Systems

Investments in security monitoring are likely to increase. Even if the human observers theoretically offer the greatest security, they are not enough since it is necessary to take account the drawbacks of human inattention and limited senses. The ability to continuously monitor the environment, to detect abnormal conditions, and to capture information of interest, all in real-time, gives the opportunity to reduce inspection costs while providing for increased security to the public. Generally, security monitoring requires several sensor devices that are based on more or less sophisticated technologies, basically according to the application need. The strong need to have more and more intelligent surveillance systems has resulted in a new generation of sensors,

[13] F., Garzia. Security System design and integration. [aut. libro] F. Flammini. *Critical Infrastructure Security: Assessment, Prevention, Detection, Response.* s.l. : WitPress, 2012.

[14] Chong, James I. *Next generation multiagency fusion centers - people, process & technologies.* s.l. : VidSys, 2012.

smart sensors. The fundamental difference between a traditional sensor and a smart sensor is the latter's flexible communication and information processing capability. Each sensor has an on-board microprocessor that can be used for digital signal processing, self-identification, self-adaptation and self-diagnostics functions. Furthermore, all smart sensor platforms use wireless communication technology[15]. Actually, in the last years, the scientific community distinguishes between the concepts of "smart" and "intelligent", pointing out that the former is related to technological aspects while the latter to functional ones[16]. Thanks to the technological progress in the miniaturization techniques, the size of sensors has decreased over time as well as their costs. This allowed to build more complex systems for disparate applications able to implement effective protection strategies. Thus, modern surveillance systems integrate heterogeneous security systems equipped with smart sensors. In the following, the most relevant systems in security field are presented.

Examples of tools and capabilities to create safe and secure hospital environments for patients, staff and visitors include video surveillance, access control, intelligent controllers, electronic locks, motion sensors, panic buttons, threat level management, and visitor management[17].

3.2.1.1. Fences/Walls

Limiting physical access of unauthorized personal guarantee that only trusted and qualified staff is able to access areas contains critical infrastructures or sensitive data. Restricted areas should be surrounded by a fence. Entering or exiting restricted areas will be possible from controlled passage.

3.2.1.2. Guards

Despite all logical defences and security automation, security guard will always be part of the defence layer. Human guards can make decisions and conclusions on ongoing events and react accordingly in a fast and accurate way.

3.2.1.3. Building control

Building control systems control and automate the building systems such as elevators, air conditioning, water and electricity management, etc.

3.2.1.4. Intrusion detection and access control

Intrusion detection and access control belong to two diverse typologies of system but are closely connected between them. Intrusion Detection System (IDS) groups several devices able to detect unauthorized access of people into sensitive areas. It involves magnetic contacts, volumetric sensors, glass break detectors, etc. However, in order to differentiate the accesses unauthorized from the authorized ones, an Access Control System (ACS) is required. ACS is based on three main concepts: possess (e.g., a card), knowledge (e.g., a pin) and biometric feature (e.g., fingerprint). According to the required protection level or the permit level assigned, ACS can manage more combinations of entry to or exit from secured areas. ACS is constantly incorporating

[15] B. F. Spencer, Manuel E. Ruizs, and Narito Kurata. *Smart sensing technology: Opportunities and challenges*Journal of Structural Control and Health Monitoring, p. 349-368.

[16] Y., Yurish S. *Sensors: Smart vs intelligent.* Sensors and Transducers Journal.

[17] Prime Communications, Inc. The Top 3 Reasons Hospitals Need a Physical Security Plan. *Prime Communications.* [Online] https://www.primecominc.com/the-top-3-reasons-hospitals-need-a-physical-security-plan/.

improvements in communications and security technologies; nevertheless, each technology has a certain level of vulnerability to be considered. For this reason, hybrid approaches which combine technologies based on the three concepts above are preferred.

3.2.1.5. Video surveillance

Cameras are the most widespread devices in the surveillance field and their level of maturity is getting higher both in indoor and outdoor applications. Monitoring through video streams of a Closed-Circuit Television (CCTV) system allows a quick recognition of a situation in order to prevent or detect possible malevolent intents, as well as to conduct post incident analysis. Video surveillance is a field whose development keeps abreast of technology evolution. Indeed, cameras are equipped more and more with special features and often the CCTV system is combined with a Video Content Analysis (VCA) system. VCA is the capability of automatically analyzing video to detect and determine temporal and spatial events. This technical capability is used in a wide range of domains including transport, safety, security, health-care, retail, automotive, home automation and entertainment. Many different functionalities, more or less complex, can be implemented in VCA. Relating to the complexity of the algorithm, it can be hosted on the camera (using on board processing units) or on a dedicated server. Despite the continuous enhancement in this field, VCA still presents several limits. Their effectiveness may be reduced by multiple factors such as the difficulty of modeling complex behaviors (i.e. isolating individual people in crowds is hard)[18], the sensitivity to changes of lighting conditions, the presence of reflective surfaces in the scene, etc. In addition, VCA is often topic of debate for ethical issues (e.g. facial recognition is not allowed in all countries).

Video surveillance in hospital settings can help provide high-quality, affordable healthcare. For instance, it can be useful to centralize patient observation, lowering the cost of patient observers/caretakers; to remotely monitor crowded emergency departments (e.g., to count the number of people who enter and exit or to detect when a person crosses the threshold of a room); to monitor and deter drug diversion; to help prevent theft and hoarding of equipment; and to increase safety in public areas (e.g., parking lots and other public areas for break-ins and suspicious persons)[19].

3.2.1.6. Audio surveillance

An emergent security solution is the audio surveillance. By combining audio sensors with advanced algorithms, this kind of technological tool is able to recognize automatically abnormal or unexpected noises such as scream, glass breaks, explosions, and shots. This security system is particularly useful in situations of inadequate or absent visibility; in this case, the sound constitutes an essential information source for discriminating between suspicious events. In addition, this approach is especially advantageous if compared to other systems (e.g., VCA systems) since it is independent from lighting conditions and it has low computational needs. In contrast, their effectiveness decreases in areas where the noise is very high. Despite of this, adopting adaptive frameworks is possible to detect atypical situations under adverse conditions containing highly nonstationary background noise (e.g., see here).

[18] Rodriguez, Mikel & Laptev, Ivan & Sivic, Josef & Audibert, Jean-Yves. (2011). Density-aware person detection and tracking in crowds. Proceedings of the IEEE International Conference on Computer Vision. 2423-2430. 10.1109/ICCV.2011.6126526.

[19] Cisco Systems, Inc. Cisco Video Surveillance in Hospitals: Ten Ways to Save Money and Improve the Patient Experience . *Cisco.* [Online]

3.2.1.7. CBRNE sensors

CBRNE is the term for protective measures taken against Chemical, Biological, Radiological, Nuclear and Explosive attacks. CBRNE systems are a good security solution for environmental monitoring and are very specific technologies for particular threats. So, it is clear they constitute a powerful countermeasure against attacks where weapons of mass destruction are expected. This technology allows an effective identification of bombs, drugs, metallic and nonmetallic weapons and explosives at long distance. Actually, unlike radio-logical and explosive sensors, chemical ones have still some problems about the coverage range. For overcoming this restriction, often these tools dedicated to explosive detection are combined with the deployment of dog patrols. [20] Unfortunately, the cost of this technology is rather high so it is essential to balance the security needs with budgetary constraints. In practice, this limits the number of checkpoints for dangerous substances detection; thus, their locations must be evaluated accurately. Furthermore, the current solutions for CBRNE for people scanning are not directly suitable to all situation due to their excessive processing time (the mass-transit system is not compatible with the crowd flows)[21].

3.2.1.8. Physical Security Information Management (PSIM) systems

Given the proliferation of the variety of interconnected systems, the willingness to develop an "open" system architecture with the backdrop of interoperability, and driven from needs to include other value-added functionality, PSIM solutions have been developed. Born initially as a physical security integration enhancement, PSIM is rapidly evolving to encompass information management systems insomuch as it draws the attention of government agencies and businesses from a wide range of markets[22]. It is a software platform that collects, correlates and manages information from disparate security devices and information systems into one common situation picture in order to empower personnel to identify and proactively resolve situations. The key element is its ability to integrate different complex subsystems easily, as well as its interoperability with third-party applications and legacy security systems without being "locked-in" to any specific vendor. Many security benefits hail from adoption of PSIM solutions, like better situation awareness, decrease of reaction times to events, driven management of the procedural actions in case of crisis situations, support to post event analysis, etc. For this reason, they are assuming a strategic role for properly responding to any kind of emergency and are essential to respond and deal with the wide range of potential security risks. In detail, in order to provide a complete Situation Assessment and Situation Management this new generation of systems should fulfill five key capabilities [23] shown in Figure 2.

---

[20] Ansaldo STS. *Report on consolidation of functional results.* 2014.

[21] Report on consolidation of industrial results. *SECUR-ED Urban Transportation - European Demonstration (SECUR-ED).* [Online] 2014. http://www.secured.eu/wpcontent/uploads/2014/10/D46.5_Report_on_consolidation_Industrial.pdf .

[22] Howe, Ellen. Psim: An effective risk management tool. s.l. : VidSys, 2014.

[23] Roadnight, J. *Will physical security information management (psim) systems change the global security world?* s.l. : CornerStone.

Figure 2- Key capabilities of a PSIM system

1. Gathering: the system gathers data from a wide range of disparate devices and subsystems;

2. Analysis: the gathered data should be analyzed in order to recognize the situation and to give the right priority to a possible emergency;

3. Confirmation: the system shows the situation to the security operator in a clear and concise way enabling an accurate and quick confirmation of the appeared alarms;

4. Resolution: the PSIM system should clearly present to the security operator the steps of the procedure to carry out for managing the situation in real time;

5. Reporting: all activities should be recorded for supporting the post-event investigative analysis.

PSIM is analogous to SIEM (Security Information and Event Management) software. Basically, it does for physical security what SIEM does for cyber security, simplifying the surveillance activities while improving security and reducing time, cost and effort that physical security requires.

4 Cyber Security

Today, technology plays a crucial role in every industry, but healthcare is definitely one of the most important. Electronic healthcare technology has the potential to extend, save and enhance lives. This evolution brings a lot of benefits, but there are increasing concerns relating with the security of the healthcare data and devices. In the last few years healthcare was one of the industries most threatened by cybersecurity risks. The main reasons why healthcare is so vulnerable are[24]:

- Increasingly connected technology to provide multiple ways of connecting to medical devices, which are often easily accessible: a single device could provide a potential entry point to larger hospital networks, bypassing the firewalls;

[24] Cybersecurity in healthcare: A narrative review of trends, threats and ways forward. Coventry, Lynne et al. Maturitas, Volume 113, 48 - 52

- More devices being used in the wider healthcare setting increases vulnerability to breaches;
- Mobile consumer devices (e.g., smartphones) being widely adopted; making it difficult to protect health data from risks posed by general purpose devices;
- Lack of funding for cybersecurity.

It is known that there is no effective way to avoid cyber security breaches. However, it is important to guarantee that the risk is minimized. For that all the healthcare providers must be aware of cybersecurity trends and threats as they emerge to improve their cyber security solutions.

4.1 Cyber Security Solutions

In current times ICT systems deal with an enormous amount of sensitive data, such as: personal files, company records, government files, medical records, back amounts, etc. ICT systems are also often in control centers of critical infrastructure. These scenarios make ICT systems a desirable target for computer hackers that want to gain unauthorized access to the information stored and/or want to exploit control systems.

In order to ensure security in every ICT system, both in securing data and in preventing access control to unauthorized parties, it is important to follow the principles of security of information[25]:

- **Confidentiality** - Consists in protecting the information against individuals, entities or processes that are not allowed to access it.
- **Integrity** - Consists in preventing the information from being modified or corrupted by unauthorized means.
- **Availability** - Consists of maintaining the systems in operation, preventing the existence of disturbances. Services must be available whenever requested.
- **Authenticity** - Consists in ensuring that a party of interest or the origin / destination of a communication is who to claim to be. Each party is authentic to the other.
- **Non-repudiation** - Consists in ensuring that each party does not deny having signed or created information. Non-repudiation provides evidence that a party has performed a particular action.

Moreover, some simple and easy practices can be a crucial help to keep data safe and well-protected, for example regular and secure backups, which are essential to maintain resilience and be able to recover quickly in case of attack and keeping software up to date to ensure security patches are installed.

4.1.1 Security of ICT systems

As in physical protection, CI facilities may also be subject to Cyber-attacks, requiring additional Cyber-security measures. Most of them are meant to protect the CI against Information leakage/loss, prevent unauthorized access, and secure all communication in and out of the facilities:

[25] P. M. Jawandhiya, M. M. Ghonge, M. S. Ali, and P. J. S. Deshpande, "A Survey of Mobile Ad Hoc Network Attacks," vol. 2, no. 9, pp. 4063–4071, 2010.

- **Access rights management** – proper access rights management is important for both, the physical and the cybernetic domains. Access rights are usually implemented in all CI facilities, and allow for granular management of all authorized personnel, with ability to restrict access to only some parts of the CI facility or its information.
- **Monitoring sensors** – Most CI facilities are properly monitored both physically (for example via specific sensors that ensure the health state of critical equipment) and electronically (via Cyber security sensors and control systems), allowing efficient detection of malfunctions and possible security breaches as they happen (or even before that), thus preventing potential damage to the systems.
- **Command and Control Centers** – Most CI facilities use centralized control over all equipment (physical and virtual alike), to create a full picture regarding the health state of each facility or the entire CI altogether. The vast amount of information flowing to one centralized location, makes it easier to discover interconnections between different incidents, and manage them more efficiently.
- **Incident response teams** – Since all CI may be subject to all sorts of incidents (such as cyber security breaches, physical sabotage and technical malfunctions), many CI facilities maintain dedicated incident response teams to deal with each scenario. This enables a well-controlled response, by a crew, specifically trained to treat such incidents.
- **Regulation and standards** – Like all big and important facilities, the CI facilities use standards and regulations that dictate homogenous requirements regarding the facilities, the required security and protection measures, and standardized management for all similar facilities across all CI.

Table 2 describes some cyber-security measures used to mitigate the most common types of Cyber-attacks.

Domain	Objective	Measures
(D)DoS	To reduce impact of an attacks in denied service (protection measures)	Automatic mitigation Mail alert to permit an adaptation of the filtering Emergency process (internet physical disconnection)
Vulnerability attacks	To reduce the number of vulnerability (prevention measures)	Several level of filtering with highest level security increased monitoring of servers by an identified correspondent
Virus attacks	To reduce the number and the impact of incidents (protection and prevention measures)	Double decipherment Multi analysis by several independent motors Systematic installation of antivirus on computer station of each user
URL filtering	To reduce the number of attacks (prevention measures)	Several levels of filtering URL blacklist
Spams/phishing	To reduce the number and the impact of	Spam/ phishing by e-mail : cloud filtering ➔ e-mails are analyzed before going to the recipient

	incidents (protection and prevention measures)	Using a tool to analyze e-mail, sent URL to a sandbox and rewrite the URL before the user receive the e-mail. The URLs are blocked if they are dangerous Using the antivirus for a last analysis
Internal virus and data leak	To reduce the number of incidents (prevention measures)	Using proxy for analyzing each request
Maintainer access	To reduce the number and the impact of incidents (protection and prevention measures)	Bcoastion of computer administration
Internet flow	To reduce impact (protection measures)	Protection against saturation of the bandwidth by our users (level, waiting line…)
Protection against credential phishing	To reduce the number of incidents (prevention measures)	When identifiers of a person of the entity are detected into a form, the tool uses blocks the send of the form
Internet Access	To reduce the number and the impacts of incidents (protection and prevention measures)	Firewall chaining Filtering with independent computer case
WI-FI access	To reduce the number and the impacts of incidents (protection and prevention measures)	Using a VPN IPsec Using a tool of vulnerability detection Using a tool of AI in response of cyberattacks

Table 2. Some Security Measures per Domain

4.1.2 Cyber Security Premises and Solutions

Operation technology (OT) broadly refers to hardware and software dedicated to control and automate physical systems such as power plants, energy distribution, etc. Recently this field is undergoing a fast convergence with IT technologies thanks to the exponential growth in the usage of internet of things (IoT) connected devices in settings previously dominated by OT devices only. This phenomenon led to a stark increase in the attack surface, enabling cyberattacks against vital automation systems in critical infrastructure[26,27].

Within the OT domain, building automations systems are gaining a lot of traction and show the most pronounced IT/OT convergence. Healthcare facilities, due to their operation complexity, are generally required to adopt BASs for improving the efficiency, comfort and safety of the hospital buildings. However, due to the above reasons, this can also make them more vulnerable to

[26] Nimrod Stoler (2018). Anatomy of the Triton Malware Attack, https://www.cyberark.com/threat-research-blog/anatomy-triton-malware-attack/

[27] N. Falliere, L. O. Murchu, E. Chien (2011). W32.Stuxnet Dossier. Symantec Corporation.

cyberattacks[28] against BASs which attackers can exploit to either interrupting hospital operations or move laterally into the network to more sensitive areas such as server rooms.

To protect hospitals from such threats, several requirements must be met. In the next subsections some of these requirements are described and a list of available solutions is presented.

4.1.2.1 *Data Protection*

Data is the most valuable asset in any industry, and even more in health sector. Thus, it is crucial to guarantee the protection of data from corruption, compromise or loss. Below, we describe some measures that help in the data protection.

Encryption file systems

Disk Encryption technology refers to the conversion of clear-text code and information files stored on a hard-drive, into a bit-by-bit encrypted information, inaccessible to a third party without knowing the passcode or without holding the encryption key.

Disk encryption technology implements various encryption solutions, such as encryption software or dedicated encryption hardware, to encrypt every bit of data that goes on the disk.

The "FDE" or "Full Disk Encryption" term, may often suggest that the entire disk is encrypted, including the bootable operating system partitions. This full encryption often doesn't include the MBR (master boot record) to allow initial access to the hard-disk, to start the decryption process upon authorized access. Even so, there are some hardware-based FDE systems that can actually encrypt an entire boot HD, including the MBR.

Full Disk Encryption is used in all sorts of environments, organizational and private alike. The main idea behind this solution is to protect data-in-rest (statically stored on the HD) from unauthorized access, by physically accessing the computer or HD, in cases such as a stolen laptop or a PC, containing sensitive information, stolen from a corporation.

It is important to note that these solutions don't protect against unauthorized network-based access, or other common attacks such as malware or viruses.

Available solutions on the market

Most operating systems have some sort of built-in encryption abilities, such as the EFS and Bit-Locker for Microsoft Windows systems, FileVault for MAC users, and various encryption solutions for Linux, depending on the distribution and version of OS.

Third-party solutions can be divided into open-source and free solutions and commercial solutions.

Open Source:

- VeraCrypt (Successor to TrueCrypt)
- AxCrypt

[28] European Union Agency for Network and Information Security (ENISA), Cyber security and resilience for Smart Hospitals (2016).

- GNU Privacy Guard (GnuPG)

Commercial solutions:

- Symantec Drive Encryption
- Trend Micro Endpoint Encryption
- McAfee Complete Data Protection

Data Loss Prevention (DLP)

Data Loss Prevention (DLP) is a way of ensuring that corporate users will not, accidentally or intentionally, transfer or copy sensitive information outside the corporate network.

The term is also used to describe security software products used by network administrators and corporate security teams that allow control over what data is considered sensitive and what the corporate users can or cannot do with it.

The DLP products implement business rules and business security policies to classify and protect confidential and sensitive data. In essence, the software prevents accidental or malicious disclosure of information via most known electronic methods, such as: Corporate email, web sharing, cloud storage, external USB devices, physical printing of documents, and even screen capturing (print screen).

DLP solutions are very common in big and mid-sized companies, containing sensitive information, in virtually any industry. Whether it is customers' private data (like credit cards, SSN's, etc.) or corporate information such as financial information or intellectual property of the company, this kind of information, if leaked, can cause damage to the company and even expose the company to various law suits.

The DLP solutions are intended to prevent such information disclosure across the entire corporate network, using local end-point agents, network monitoring, implementing web proxies, etc.

Available solutions on the market

- Forcepoint (formerly known as Websense) Triton-APX DLP
- Intel Security (McAfee) – Total Protection for Data Loss Prevention
- CA – Data Protection (formerly known as CA DataMinder)
- Symantec DLP
- Trustwave DLP

Online storage and backups

Online or Cloud storage, is a data storage model in which all digital data is stored in a logical storage pools, while the physical infrastructure supporting it, can span across multiple servers in multiple server farms and even in different countries. The physical storage environment is typically owned and managed by a separate hosting company, which is responsible for the entire infrastructure and for keeping the data available and accessible. The host company can also offer various backup plans and implement various security measures to protect the data, such as secure access, data encryption, multiple backups, etc.

Cloud storage services can be accessed in various ways, depending on the service, and may include dedicated end-point applications, web-interface, and some additional interfaces, such as dedicated APIs.

Online and Cloud based storage and backups are becoming more and more popular, with almost every company putting at least part of its non-sensitive information in those storage solutions.

Some companies rely entirely on Cloud based storage, thus putting all their information and files online and maintain various safety measures to secure their information.

Private Cloud solutions are also very popular, allowing an end-user full access to his/her files from virtually anywhere, via a dedicated application or via web access.

Available solutions on the market

- Amazon S3 (Simple Storage Service)
- Microsoft Azure
- Google Cloud Storage
- Rackspace
- Dropbox for business

Data Redundant Array of Inexpensive/Independent Disks (RAID)

"RAID" is a data storage virtualization technology that implements multiple physical hard disks into one logical unit, in order to achieve data redundancy (thus higher fault tolerance), higher performance, or both.

RAID technology is divided into multiple RAID Levels, representing different data distribution strategies among the array of available hard disks on the RAID machine, as well as the minimum amount of physical hard disk required to implement each level. These levels are represented by numbers, and the most common RAID setups include the RAID 0, RAID 1 and RAID 5. RAID 0 represents a setup made for performance, higher overall disk space and no redundancy or fault tolerance. RAID 1 offers high tolerance at the expense of lower overall performance and lower overall disk space. RAID 5 offers both, a higher performance and fault tolerance, but requires at least one more hard disk to work (3 hard disks, instead of only 2 required on RAID 0 and 1).

Some other, less common RAID Levels include RAID 2,3,4,6,10,50 and 0+1, mainly representing different variations of the basic three (RAID 0,1,5).

RAID solutions are very common these days and are used by most enterprise level storage systems. The use of RAID in storage devices and servers, adds the ability to swap faulty hard disks, without the need for down time, and in most cases (except RAID 0) overcome disaster and data loss in case of a faulty disk.

RAID solutions are also common in home environment use, especially when small storage devices are used to store important data on multiple hard disks (such as important document or photos, etc.).

Available solutions on the market

RAID technology is available in almost every server and storage machine available in the enterprise market today and can be implemented by configuring the desired RAID level on each machine.

Data on transit Encryption

One of the security basics is keeping data integrity and confidentiality at all times. In many cases sensitive data must be transferred over unsecure networks (internet, 3rd party networks, etc.) and must be secured. This can be achieved by encrypting the data transferred over the network using common encryption protocols such as TLS, IPSEC, and VPN or in some cases by using built-in security mechanisms of protocols when available.

Encryption can be implemented in many ways. Traffic can be encrypted inside and outside the organization. Encryption of data can be achieved locally at the local operating system or by specific dedicated gateway.

An outside organization encryption example is VPN. It encrypts all traffic from client to server (usually a firewall) ensuring all passing data will be encrypted, even over public network like the internet.

An inside organization encryption is IPsec or TLS. They ensure all traffic is encrypted end-to-end and cannot be read by any unauthorized attacker in the middle.

4.1.2.2 *Network Monitoring*

The network can be used to execute the most frequent cyber-attacks. Network monitoring is a good way to understand the system and avoid attacks. It is very important to use tools that can help with this monitoring:

Intrusion Prevention System (IPS)

Attackers can activate manipulations on end-point operating systems, services, protocols and network communication. IPS can detect these manipulations and compare it to an attack-signatures database (which is updating continuously to new attacks signatures) and block the attack.

Another way IPS detects threats over the network is by examination of communications and search for anomalies such as port scanning (which is usually the first step performed before an attack can occur) or when a pre-defied policy or limit is approached, for example, the maximum number of sessions from one source IP, before consider that source as non-human (or malicious) behaviour.

In case IPS detects an attack, it will drop all the suspicious packets and prevent it from happening.

Most Firewall devices offer IPS capabilities inside it. IPS detects and prevents attack patterns to be transferred over the network. IPS can also be a dedicated appliance or as host-based application.

Available solutions on the market

- Dedicated Appliances:
- Cisco FirePower
- McAfee IPS
- IBM Security Network IPS.
- Feature in firewall:

- Fortinet
- Checkpoint

Intrusion Detection Systems

Intrusion detection systems (IDS) are software tools used as a form of protection against intrusions caused by malware and other malicious intrusions. IDS can be divided into two kinds. Host Intrusion Detection Systems (HIDS) detect host-level malicious intrusions by monitoring operating system interactions, file access, and reviewing data. And Network Intrusion Detection Systems (NIDS) detect malfunctions by monitoring network traffic[29]. Traditional cyber security software works by identifying malicious files that match against a database of malware signatures already known[30]. However, this approach has the disadvantage that, since it relies on parsing process signatures to detect attacks, new types of malicious intrusions cannot be detected, thus it is required that the signatures database is maintained constantly up to date when new malicious software is detected on hosts, or malicious activity is detected on the network. Due to these limitations there is interest in applying new and innovative techniques, such as Artificial Intelligence, in intrusion detection scenarios in order to obtain better results[31].

Intrusion detection system solutions are a key building block for a complete cybersecurity platform for modern healthcare facilities. These solutions are both open source and commercial.

Open Source:

- Suricata
- Zeek

Commercial solutions:

- Forescout SilentDefense
- Claroty
- CyberX
- Nozomi
- Orion Malware

Intrusion Response Systems

While IDS focus on detecting Cyber Security threats and attacks there is still the need to act in order to defend against or mitigate the effects of the cyber- threats and attacks. This creates the need for Intrusion Response Systems (IRS) that, in the event that an intrusive behavior is detected, choose the necessary action to take to prevent attacks and ensure the security of networks and computational systems [8]. IRS are categorized by the way they deal with

[29] A. Milenkoski, M. Vieira, S. Kounev, A. Avritzer, and B. D. Payne, "Evaluating Computer Intrusion Detection Systems: A Survey of Common Practices," *ACM Comput. Surv.*, vol. 48, no. 1, p. 12:1 - 12:41, 2015.

[30] N. Idika, "A Survey of Malware Detection Techniques," *Purdue Univ.*, p. 48, 2007.

[31] M. Egele, T. Scholte, E. Kirda, and C. Kruegel, "A survey on automated dynamic malware-analysis techniques and tools," *ACM Comput. Surv.*, vol. 44, no. 2, pp. 1–42, 2012.

responding to cyber- threats and attacks. The IRS can be passive if they only notify the system owner with data about the attack or active if the IRS takes actions to mitigate the effects of the attack. The level of automation may also be different depending in the IRS[32,33,34].

Levels of automation enable 3 different scenarios[35]. A notification only system, that notifies the systems administrator when an attack is identified with the necessary information for the administrator to take further defensive measures. A manual response system, where the system administrator is notified of attacks and has options to select immediate defensive measures that the system will perform. And automatic response systems, these are the most advanced systems, because they operate autonomously: when an attack is detected these systems select and preform the defense measures without human interaction. Within automatic response systems it is also relevant to consider if the response selection mechanism is static or adjust over time. The response time depends on whether the systems performs proactive measures when the detected threat is not completely confirmed. Finally, depending on the scenario the IRS might have to consider several systems with their own IRS, and perform cooperative actions in order to mitigate the cyber-threats in distributed environments.

Available solutions on the market

- SARA
- TBAIR
- SoSMART
- CSM
- EMERALD
- CITRA
- FLIPS
- Cymerius

Segmentation

Network segmentation is to divide a physical network to multiple logical sub networks called VLANs. VLANs configuration can be achieved with layer 2 (in the OSI model) devices like switches. Basically, VLANs are separated from one each other and communication are not allowed between them unless a routing device is allowing route between VLANs.

[32] T. Toth and C. Kruegel, "Evaluating the impact of automated intrusion response mechanisms," *Proc. - Annu. Comput. Secur. Appl. Conf. ACSAC*, vol. 2002–Janua, pp. 301–310, 2002.

[33] C. Carver, J. Hill, J. Surdu, and U. Pooch, "A Methodology for Using Intelligent Agents to provide Automated Intrusion Response," *Proc. IEEE Syst. Man, Cybern. Inf. Assur. Secur. Work.*, vol. 77843, no. July, pp. 110–116, 2000.

[34] D. J. Ragsdale, C. A. Carver, J. W. Humphries, and U. W. Pooch, "Adaptation techniques for intrusion detection and intrusion response systems," in *Systems, Man, and Cybernetics, 2000 IEEE International Conference on*, 2000, vol. 4, pp. 2344–2349.

[35] N. Stakhanova, S. Basu, and J. S. Wong, "A Taxonomy of Intrusion Response Systems," *Int. J. Inf. Comput. Secur.*, vol. 1, no. 1, pp. 169–184, 2007.

VLANs routing can be achieved with layer 3 (in the OSI model) devices such as routers or firewalls. The major advantage of the routing capabilities, beyond the ability to connect between different VLANs, is to apply set of rules (policy or access list) that will allow communication between two or more VLANs. Rules can allow specific IP accessing a specific resource on a specific port.

Segmentation is recommended to be used in almost every network. It is best practice to divide between assets and clients across the network. Configuring assets in dedicated VLANs and configuring all clients' computers in another VLANs. Communicating between these VLANs will be allowed through a Firewall device with a strict policy or a Router device with strict Access List. Access List or Policy will detail which device can access another and in which port or protocol it is allowed. The Firewall or Router will drop and ignore every other packet crossing the network that do not answer to policy or access list permissions.

Available solutions on the market

All major vendors providing VLAN capabilities in switching products such as Cisco, HP, Dell, Fortinet, Juniper and more.

Network Device Redundancy

Network devices are delivering services from the suppliers (servers) to their clients (end-computers). Failure in providing services due to technical issues or malicious activity would cause a reduction in productivity.

Network availability is crucial in critical infrastructure since no unavailability can be taken into consideration. This can be achieved by redundant all network devices to allow business continuity in case of a device failure.

There are 2 main redundant configurations: active-active and active-passive. Active-active describe at least 2 network devices that back up each other while they are both active and able to receive and process data. This configuration is not fully supported in all network devices. Active-passive describe at least 2 network devices that back up each other while only one of them is able to receive and process data while the other one is in standby mode. When the active device fails, the standby device becomes active and operational.

To provide network redundancy, two devices of each certain product are required. Firewall, for example, can be redundant by installing two firewall units, set them as a cluster (to consider them as one logical unit) and connect them correctly to the network. While one of the firewall devices are unavailable for any reason, the other one will come into action.

Other devices, like switches, can be setup in stack configuration. This configuration requires connecting the switches with special stuck cables which redundant both switches for power issues and hardware failure issues. In switch stack configuration both switches considered and one, they are both active-active (active-passive mode not supported) and the second switch is seen as an extension of the first one.

Network Address Translation (NAT)

NAT is a functionality that translates private IPs to public IPs. NAT can also translate IPv6 address to IPv4 address. NAT functionality usually is part of a firewall functionality since all traffic goes through it.

NAT help to protect networks by hiding private IPs from being exposed to the outside world. Without NAT, every communication will carry the real IP address of the device it was sent from as the source IP. This will allow attackers to know the inside devices IPs and try to communicate them directly. Knowledge of inside IP addresses can also imply of network size and configuration and may help attackers to achieve their goals.

Available solutions on the market

- Feature in firewalls: Fortinet, Checkpoint.
- Feature in Layer 3 Switches: Cisco, HP, Dell, Fortinet, Juniper etc.

Firewalls

Firewalls are the fundamental protection in every network. Logically, all communications flows through it, routed and examined by it.

By design, most firewalls are blocking all traffic unless specific policy set by the firewall administrator allows it. Such policy usually includes the source IP, destination IP and service or the port in use.

Because all communications flow through it, firewall can offer and deliver additional security mechanisms such as anti-virus, anti-malware, anti-bot, anti-spam, IPS, web filtering, application control, load balancing, VPN, etc.

Firewall can be deployed as physical appliance, virtual machine, cloud service or software.

To provide a defence layer over the network, firewall is deployed. Firewall allows conditioned routing between VLANs or separated physical networks while applying control, security and audit for the passing traffic.

Available solutions on the market

- Fortinet
- Checkpoint
- Juniper
- Sophos
- Cisco
- Palo-Alto

Web Application Firewall (WAF)

WAF examines layer 7 communication in HTTP and HTTPS protocols. They can read and understand the protocol and how it is used to communicate with the web server. By doing this they can prevent application attacks from taking place. Application attacks can be SQL injection, cross site scripting, CSRF and so on.

WAF checks for various vectors such as HTTP/S request headers (to determine the nature of the communication and it's propose), application attack signatures, abnormal behavior of the client, source IP reputation, etc.

WAF protect web server from application attacks. WAF is set in front of the web server and examines all incoming traffic. When malicious traffic is detected WAF drops it.

Available solutions on the market

- Imperva
- F5
- Incapsula (cloud)
- Rebase (Cloud)

Denial of Service (DoS)/Distributed Denial of Service (DDoS) prevention

DoS stand for Denial of Service which is a kind of attack that will cause a server to be unavailable by taking advantage of a legitimacy service provided it over and over again until it exhausted or fail. A web server for example is sending web pages stored on it to whomever requests it. DoS attack will request these webpages in a large scale simultaneously until the server reached its limit of concurrent requests and fails to answer new requests.

DDoS, stands for Distributed Denial of Service, and it is the same kind of attack but from multiple sources like botnet network which can contain thousands of devices communicating with the web server.

DoS\DDoS prevention solutions can protect servers and services from such attacks by examine all traffic to and from the protected resource. Every traffic that considered as part of an attack will be dropped and won't be transferred to the protected resources. DoS\DDoS mitigation solutions can observe large scale attacks depending on its configuration.

There are basically two forms of DoS\DDoS mitigation solutions: on premise and cloud-based. On premise solutions will be physical appliance logically installed in front of the firewall (to protect it). This form of solution has one major disadvantage – it cannot defend against DoS/DDoS volumetric attacks which is a kind of attack that consume bandwidth so no new legitimate traffic can pass and answered. Another form of DoS\DDoS solution is cloud-based. In this solution the traffic is distributed through data centers across multiple geographical locations (CDN network) so an attack will be dropped way before it reaches its destination.

Available solutions on the market

- Imperva Incapsula
- F5
- Arbor
- Akamai
- CloudFlare
- DoS arrest

Clustering

Computer clustering refers to various methods of duplicating existing infrastructure – such as critical servers (holding various services and applications) and configuring the duplicate servers (in a cluster) to take over the roles of the original servers in case of down-time or excessive load on the original server (the latter may also be considered as load-balancing). Clustering can also be used to increase performance, by using numerous separate computers to perform the same task simultaneously.

The cluster usually consists of two (or more) "nodes" – representing the physical (or virtual) machines running duplicate instances of the same operating system and applications required by the cluster.

In most cases, the nodes will use the exact same hardware and operating system on each node, and will be connected to each other through fast LAN connection, which allows full and constant replication between them.

Most clusters will use a single shared storage used by all nodes. This configuration allows each node to interpret the shared storage as its local drive, even though it is shared between all separate nodes.

Clustering technology has a wide range of applicability and deployment, varying from small business clusters to anything like super-computers.

Clustering is widely used to reduce down-time to minimum, and also allow continuous system maintenance while still providing service to the end-users.

Available solutions on the market

Clustering is an integral part of the Microsoft Windows server operating system and is available in other operating systems as well.

Load Balancers

Load balancers specialized in distributing workloads over multiple servers based on pre-defined policy. These allow to deliver high availability services and scalability capabilities.

There are multiple forms of load balancing: DNS-based, availability-based, network-based, etc. load balancers can be physical appliance, virtual machines or cloud-based services.

Some load balancer can provide multiple features such as SSL offloading – decrypt encryption before reaching destination servers.

Load balancers comes in various sizes and shapes. It can be implemented over a wide several of products and solutions. Main usage of load balancers is in front of web or application servers. Multiple web/application servers, delivering the same content, will be presented by a load balancer IP address. Every request for a web page will arrive to the load balancer which will route the traffic to the right server based on the policy pre-defined by the load balancer administrator. If HTTPS (SSL Encryption) protocol is applied, the load balancer can decrypted the information for the web/application servers and reduce their workload.

Available solutions on the market

- F5
- Incapsula
- Citrix
- Radware.

4.1.2.3 *Endpoint Monitoring*

In addition to the Network, also the endpoints should be protected, because they are another way to get into all the systems. Typically, the endpoints are connected to each other, so it is very important to monitor what is done on each device.

Endpoint Security

Endpoint Protection is known as the last line of defence for workstations, laptops and smartphones who are connected to organization networks. Various systems and software suites provide a wide range of solutions to address most common challenges. They can be separated to deal with a specific threat or unified as a platform against several ones.

The structure of the solution can be set as client-server model or software as a service (SaaS) model to enforce policies and run security modules. Anti-Malware, Host Intrusion Prevention System (HIPS), Data Loss Prevention (DLP), HDD encryption, Mobile Device Management (MDM), etc. are all examples for Endpoint Security solutions.

Endpoint security must be integrated by default in all devices regardless their usage or the user position. This is widely recommended as best practice.

This array of solutions almost declines unintentional security breaches and decrease lateral movement possibilities.

Available solutions on the market

- McAfee Endpoint Protection
- Trend Micro Endpoint Protection
- Symantec endpoint solutions.

Advanced Threat Protection (ATP)

The vast major of information security products are based on detection of known threats that were already discovered and learned with no different if the threat is a network pattern representing an attack, a virus \ malware \ ransomware file or a malicious website. The security mechanisms create a unique signature of the threat and that signature is used to identify the threat in the future. Attackers can almost easily evade detection just by slightly change the attack pattern (for example, such a change can be a single character added or subtract from the attack program code) and that changes the signature completely so the security mechanisms no longer recognize the new signatures until those are discovered and learned.

Information security product vendors was required to find a set of solutions that are not signature based and can adopt themselves to the current threats and detect new threats never seen before ("zero-day attacks"). In this line of solutions, we can.

ATP Solutions are integrated in many products because of their ability to detect new threats never seen before. ATP is implemented in mail relay solutions to detect threats over email traffic to content filtering products which examine passing traffic and searching for threats. ATP is also implemented in sandboxing solutions that examine files sent to it and detect threats suck as malicious code or behavior.

Available solutions on the market

- Fortinet FortiSandbox
- Checkpoint SandBlast
- Palo Alto WildFire

Application vulnerability scanning

As in code review bullet, mentioned earlier, Application vulnerability scanning refers to automatic scanning tools that help find code weakness, and security vulnerabilities, in applications – mainly referring to web-applications.

These tools usually address the well-known security issues, such as cross-site scripting, SQL injection, command execution, directory traversal and insecure server configuration.

Although effective, these tools are still used in conjunction with more traditional human code review practices, mainly to fill the gaps of less obvious vulnerabilities and reduce false positives from the tools.

Application vulnerability scanners are widely used by many development companies. They can greatly reduce the code reviewing process time, and reduce the time needed for human review – thus reducing cost and effort. These tools are used by companies from all sizes, and even by security companies offering code review services, as a helping tool to find additional issues in code.

Available solutions on the market

There are many solutions on the market, offering vulnerability scanning tools. Part of these tools are open source and free to download by anyone, and some are commercial tools, usually offered to organizations, that include support for their tool, and sometimes higher abilities for code review and app vulnerability scanning. Some of the available tools include:

- Acunetix WVS
- IBM AppScan
- HP – Fortify
- Trustwave App Scanner

OS vulnerability scanning

OS vulnerability scanning systems are assessing security weaknesses.

Unlike Antiviruses, OS vulnerability scanning are matching publicly published vulnerabilities to the operating system patches and current security-level. Commonly part of OS Patch Management or OS Vulnerability Management solutions.

The system checks specific attack vector preconditions and run tests on potential vulnerability for exploitation.

OS vulnerability scanning systems are typically used before and after OS patching cycle as verification and validation of the process, by user demand or at pre-defined times.

It is intended to use in large scale server farms, datacenters and organization with numerous endpoints, saving precious time in Vulnerability and Risk management objectives.

Available solutions on the market

Some systems are network-based scanners, available as appliance and others are host-based client or software. Among solutions:

- Tenable - Nessus Professional
- GFI – LanGuard
- Rapid 7 Nexpose

Patch management

Patch management systems are controlling all the patching process of updates for popular third-party software.

In addition to tighten security, these patches provide new features and improves software performance.

The system acquires, test, and install updates from relevant vendors to specific software suites. It acquires available updates as they're released, test their compatibility, automatically install and check that it was properly done.

Note that some updates require OS restart, commonly scheduled in advanced.

Patch management systems are constantly running in background but can be set to install updates only at predefined times.

It is intended to use in large scale organizations.

The system could overcome the demanding task of individually update numerous software types on every endpoint.

Available solutions on the market

Some systems are vendor-centric, others are cross-vendor solutions, and they can run in agent-based or agent-less mode. Among solutions:

- SolarWinds - Software Patch Management
- GFI – LanGuard
- IBM Tivoli Provisioning Manager

Operation Systems (OS) Patch management

OS patch management systems are controlling all the patching process of security and system updates.

In addition to tighten security, these patches provide new OS features and improves performance.

The system acquires, test, and install updates on correspondent servers or endpoints operating system. It acquires available updates as they're released, test their compatibility on present replicas, automatically install and check that it was properly done.

Note that some updates require OS restart, commonly scheduled in advanced.

OS patch management systems are constantly running in background but can be set to install updates only at predefined times.

It is intended to use in large scale server farms, datacenters and organization with numerous endpoints.

The system could overcome the demanding task of individually update every operation system.

Available solutions on the market

Some systems are platform-centric other are cross-platform solutions, most of them can run in agent-based or agent-less mode. Among solutions:

- GFI – LanGuard
- IBM Tivoli Provisioning Manager
- WSUS server (Microsoft environments only)
- Microsoft SCCM solutions (Microsoft environments only).

Content Filtering (DLP, Email Filtering, URL Filtering)

Content filtering is a term that describes various technologies exists to block and prevent access to a certain resource. These blocking abilities can be achieved by a deep inspection of passing traffic. A resource can be a website, email address, etc.

The most commonly use of content filtering is a URL filtering feature. URL filtering blocks access to a certain website which was pre-defined by the URL filtering administrator. Blocking policy can apply automatically over a group of websites answers to a certain category.

Another content filtering mechanism is email filtering which inspects all mail traffic and can apply block or allow policy base on various variables as email headers, email body content, attached file types, source domain and IP reputation, etc.

Another common content filtering mechanism is Data Loss Prevention (DLP) which is able to detect key words or file types that has some kind of valuable data to the organization such as personal identity information or datasheets contains sensitive financials data. On detection, it prevents the spreading of that file\data to be leaked outside the organization. DLP describes also set of solutions that logically prevent connecting personal devices to corporate computing systems such as USB, mobile devices, etc.

Content filtering can be a physical appliance, virtual machine, local software or a feature in a firewall.

With URL Filtering, an administrator can limit access to known malicious websites using the content filtering feature and preventing user's accidently surfing into malicious websites and download malicious content. Email filtering does the same in email communications. It prevents malicious content from being sent through emails.

DLP installed on network's gateway can detect if personal data (social ID number or credit cards numbers), or business-critical data (Excel datasheets, for example), are sent outside the corporate's network and blocks it from being leaked.

Available solutions on the market

- Zscaler
- Blue Coat
- Websense
- Fortinet
- Checkpoint
- Symantec
- Digital Guardian
- ForcePoint.

4.1.2.4 *Authentication and Access Control*

To guarantee the security of the platform is very important to ensure who has access to the data. For this, it is very important implement an authentication system (only authorized people can access), but also define roles and permissions. In the following, we describe some useful tools to help in the implementation of these requirements.

Authentication, Authorization, Accounting (AAA)

Authentication, Authorization and Accounting (AAA), refers to a commonly used framework, which allows control over access to digital resources and allows policies enforcement and Usage auditing. These three components are considered an important combination to effectively and securely manage network assets.

Authentication – is the first process inline, allowing user identification, usually by having the user enter valid credentials (such as username and password) before being granted access to the network. The authentication process is usually based on an identifier, unique to each user. If the user credentials match the ones stored in the system, the user is granted access. Otherwise the user access will be denied.

Authorization – refers to the authenticated user rights to perform certain tasks or issue commands inside the interface. In essence, authorization is the process of enforcing policies and determining what types of access rights the user has in the system/network.

Accounting – is the final part of the AAA framework, referring to the auditing and logging options available to monitor user activity inside the system. These auditing and logging options, may include usage statistics, overall load on the system, access to sensitive information, login hours, etc.

These three main services are often provided by a dedicated AAA server, via a program that performs these functions.

The AAA framework is implemented by the all known RADIUS (Remote Authentication Dial-In User Service) network protocol, and also by its newer counterpart the Diameter protocol.

Similar technique is implemented in almost every application and service today that requires identity-based access to all sorts of information or services.

Information right assignment

Information rights assignment refers to the process of managing sensitive information access and protecting it from unauthorized access.

Most organizations have some sort of hierarchy and different departments, responsible for various aspects of the organization (for example: Management, Sales department, Finance department, HR department, etc.).

In most cases, the different departments only deal with specific data, related to their department, which means that they don't necessarily need access to data related to other departments.

These conditions allow a granular information right assignment, in which the users from one department can't access (read or make changes) files of other departments.

Information right assignment can be performed in various ways, with NTFS permissions being one of the most popular, allowing granular control over access rights to files and folders based on LDAP users and groups.

NTFS rights management is implemented in almost every company, running some sort of centralized storage.

Some other implementations are available for Linux based and OSX systems that also divide access rights to data, according to users and groups in the system environment.

Password Vault

Access to almost every service today, requires some sort of authentication that involves (at least) a username and password.

This is especially true for privileged access to various corporate services and infrastructures, such as servers, network devices, and security appliances.

The large number of personal and corporate passwords in use by each user, combined with the struggle to remember them all, may lead to a use of weak and/or repetitive passwords for many different services, personal and organizational alike.

This kind of behavior is very common among all users and may lead to major security issues that may eventually allow an unauthorized access to organizational assets and sensitive information and cause serious damage to an organization.

The Password Vaults, also known as Password Managers, were created to address these exact issues.

Password Managers store all passwords, locally or online, in a secure and encrypted manner, and protect all of them, using one very strong master password, selected by the user. Many password managers also include various ways of generating strong and pseudo-random passwords automatically, thus reducing the need to think of new passwords manually.

Some solutions also offer various ways of automatically filling in the passwords, on various authentication platforms, thus eliminating the need to know the current passwords for each service the user uses.

Password Vaults are commonly used for private use. But can also be implemented in organizations for various use cases. Some password vault solutions are enterprise oriented (such as the Cyber Ark's Enterprise Password Vault) and allow integration with LDAP services and other custom authentication methods specific to corporate use.

Implementation of such password managers and vaults in a corporate environment, can significantly contribute to the overall security of the authentication process.

Available solutions on the market

Password Vault (or Password Managers) can be divided into two main groups: Private use oriented, and Enterprise oriented. Though both can be implemented in an enterprise environment, the enterprise-oriented solutions, usually offer additional management options, such as multi user management, and full (or partial) integration with LDAP services, such as Active Directory users. Some of the enterprise level solutions available are:

- Cyber Ark – Enterprise Password Vault
- Micro Focus – NetIQ - Privileged Account Manager
- DELL - Privileged Password Manager
- IBM - Security Privileged Identity Manager
- Manage Engine – Password Manager Pro

Digital Vault

Digital Vaults represent a modern technology solution, comparable to the more traditional physical vaults, in use by many companies to store sensitive and critical information or objects. Although, many companies today still store physical copies of sensitive documents inside physical vaults, the growing need to access this information or documents on a day-to-day basis, by numerous people across the entire organization, requires a technological solution that comes in a shape of "Digital Vault".

Most Digital Vault solutions contains a list of key security features, essential for securing sensitive information across the organization.

The vaults solution should include a granular control over access rights of users across the organization, only allowing access to authorized personnel, and then managing specific access rights to different safes or "digital deposit boxes" inside the vault itself.

The access control should also incorporate abilities to control the allowed time intervals, in which each user is allowed to access the sensitive info and also control the geographical and network locations, from which the access would be allowed.

Additional abilities should include: Full auditing of access to the information, including user information, access timing and type of action performed on the documents (read, change, delete, copy, move, etc.). Also, it should have the ability to control access to the vault/safe using multiple authorization levels, for example a security officer, or a manager, that should explicitly authorize a user to access some sensitive info.

Digital Vaults are usually designed to protect its data using various methods, such as encryption, and also implement network security measures, independent of the security measurements implemented in the rest of the organizational network. This makes it easier to protect the sensitive information and documents, without the need to reconstruct the entire network around the vault.

The use of digital vaults is very common among corporations dealing with sensitive information, especially if it needs to be transferred securely to third parties. Digital vault solutions often allow a secured client-based access to the vault by external users, thus allowing co-operating companies to securely share sensitive information, without actually allowing access to the company's servers or network, except the digital vault and safe needed to be shared.

Available solutions on the market

- Cyber-Ark's digital vault solution is considered one of the major players in the digital vault market today.
- Covertix is yet another major player in the data protection business, offering similar vault based solution called SmartCipher Enterprise.

Access/change auditing

Almost every organization has some sort of sensitive information that needs to be protected, and accessed only by authorized personnel on "need to know" basis.

Access and change auditing, refers to various monitoring and auditing mechanisms, allowing a full record of all actions performed on this kind of information.

The auditing process usually includes full mapping of the sensitive information in the organization, and a definition of when this information can be accessed and by whom.

The second step includes an implementation of an auditing system, configured to log all actions performed upon the mapped sensitive data (for example: digital documents or various data-bases). These actions may include actions such as: Read, edit, delete, move, copy or changing permissions on various files or databases.

These systems will usually also record additional information, such as the time of action, the user that initiated it, and the details of the computer the user used to perform those actions (such as its IP address, username, operating system, etc.).

Most systems also allow integration with centralized monitoring software, such as SIEM solutions. This integration also allows configuration of warnings in live mode, meaning unauthorized access can be monitored and detected as it happens. Allowing quick response by the IT and security crews.

Various auditing are used in the majority of organizations, allowing varying levels of auditing and control over the sensitive information of the company.

Available solutions on the market

Various companies offer different monitoring solutions, allowing different aspects of access monitoring. Some of the known brands include:

- Imperva
- Fortinet
- Guardium
- Netwrix

Network Access Control (NAC)

NAC describes a solution meant to identify endpoint devices and computers before they access the network.

Connecting to a network in the traditional way (using network cable) or using wireless, eventually leads to the company core network switch, so just the connection itself potentially grants some level of access to network or data. VLAN segmentation can limit the level of access available from a specific port but VLAN cannot decide to whom it serves data to.

To answer this security issue, NAC is in use. Every network card has a unique address called MAC. While connected a device to a port in the switch it can learn its MAC address and remember it. MAC solutions can basically remember MAC address and allow or block access to them base of pre-set policy. Some solutions can set the VLAN of specific port based of the MAC address connected to it. Another way of identification can be in form of checking the endpoint operating system version, a specific registry value present, certificate, anti-virus software installed (and updated) or other pre-defined baseline value.

Unknown endpoints can be leaded to dedicate VLAN that has no valuable resource in it until administrator will manually set the correct VLAN to it.

Available solutions on the market

- ForeScout
- Portnox
- Cisco.

4.1.2.5 *Software Development*

During the software development it is important to guarantee not only the quality of the tool but also its security. Features and deadlines are always at the top of the development list, and security is often left behind. It is very important to ensure safe development processes and implement some best practices, such as the following.

Code review

Code review refers to the process of a systematic examination of computer programming source code. The main purpose of code review is to find overlooked mistakes in code that occurred during the initial development phase.

A well performed code review, can have a significant value in making the software more robust, more secure and better performing in general.

Code review can often help find and remove common vulnerabilities such as memory leaks, buffer overflow and more, thus improving software security.

Code review practices can be roughly divided into two categories: formal code review and lightweight code review.

Code review is an integral part of a SDLC (Secure Development Life Cycle) incorporated in almost every development team today.

It is crucial in developing secure software products that will protect users and service providers alike and will prevent potential damage or information leakage to malicious users that can exploit potential vulnerabilities in the code.

Available solutions on the market

Many security companies offer various services – such as secure code review, allowing the development team to outsource the intensive work of code review to security experts, while concentrating on code development for the software.

There are also various automatic tools offered by companies like "Checkmarx" – allowing automated process of code review and quick detection of common code security issues. These tools still require a human being inspecting the results due to possible false positives and false negatives generated by the tool.

Input validation

Input validation refers to secure coding methods, implemented in application development. The various methods verify that the input from the end-user (or the service communicating with the application) is not malformed and can't damage the program or make it operate in an unexpected way (for example make it crash or expose sensitive information to unauthorized user).

User input in a program can never be trusted by the developer and all input should always be checked for correctness, meaningfulness and implement security checks to prevent possible attacks on the program/application.

Some examples of validation techniques include: Data type validation (integers, strings etc.), range and constraint validation, code and cross-reference validation and structured validation. Each technique can be implemented depending on the security requirements of the system.

Input validation is used in all kinds of applications and programs, regardless of their usage profile. It is also considered best practice and an inherent part of every secure development cycle.

Available solutions on the market

Input validation checking is done on the programming language level and can include:

- Data type validation.
- Range and constraint validation.
- Code and Cross-reference validation.
- Structured validation.

4.1.2.6 *IoT Sensors for Health*

Some of the cyber security solutions known in health services are:

- **CodeBlue** is a combined hardware and software platform for medical sensor networks. This framework provides protocols for device discovery and publish/subscribe multihop routing, as well as a simple query interface that is tailored for medical monitoring.. The moves are equipped with elliptic curve cryptography (ECC) and Tinysec to enable the symmetric encryption. However the framework is not HIPAA compliant.[36]
- **MEDiSN (**Medical Emergency Detection in Sensor Networks**)** is a wireless sensor network for monitoring patients' physiological data in hospitals and during disaster events. MEDiSN comprises Physiological Monitors (PMs), which are custom-built, patient-worn motes that sample, encrypt, and sign physiological data and Relay Points (RPs) that self-organize into a multi-hop wireless backbone for carrying physiological data. Moreover, MEDiSN includes a back-end server that persistently stores medical data and presents them to authenticated GUI clients. [37]
- **Alarm-Net** has been implemented as a network of MICAz sensors, stargate gateways, iPAQ PDAs, and PCs. It depends on WSN for environmental monitoring, BSN to monitor the patient, a database to store the information, PDA for a caregiver, a network gateway. On the security side the system was enabled with a built-in cryptosystem for sensors, enhanced with an encryption algorithm based on Advanced Encryption Standard (AES), and by using a remote password, the mechanism can perform authentication. Still, Alarm-Net lacks HIPAA compliance because integrity check pseudonymization is not considered.[38]

4.2 Artificial Intelligence and Cyber Security

Artificial Intelligence (AI) is an emerging field in the area of Computer Science. AI techniques are creating machine capabilities and the automation of certain tasks that until recently were not possible. One of the advantages of recent AI techniques is their versatility, these techniques are usually easy to modify and apply in different contexts and problems while retaining the ability to achieve good results when applied properly. This versatility makes AI techniques suitable for several areas of application. Currently AI is applied in several areas such as healthcare, energy markets, finance, and more.

Cybersecurity is one of the areas with high potential for the application of AI techniques. Computer systems are often victims of cyber-attacks by ill intended parties. As already referred, to combat and protect computer systems against these attacks, cyber security experts make use of tools like Intrusion Detection Systems (IDS) and Intrusion Response Systems (IRS) to aid in the protection of the systems.

---

[36] Victor Shnayder, Bor-rong Chen, Konrad Lorincz, Thaddeus R. F. Fulford Jones, and Matt Welsh. 2005. Sensor networks for medical care. In Proceedings of the 3rd international conference on Embedded networked sensor systems (SenSys '05). ACM, New York, NY, USA, 314-314. DOI: https://doi.org/10.1145/1098918.1098979

[37] Ko, Jeonggil & Hyun Lim, Jong & Chen, Yin & Musvaloiu-E, Rvazvan & Terzis, Andreas & M. Masson, Gerald & Gao, Tia & Destler, Walt & Selavo, Leo & Dutton, Richard. (2010). BMEDiSN: Medical emergency detection in sensor networks. ACM Trans. Embedded Comput. Syst.. 10. 10.1145/1814539.1814550.

[38] Wood, A & Virone, G & Doan, Tam & Cao, Q & Selavo, Leo & Wu, Y & Fang, L & He, Zhijun & Lin, S & Stankovic, J. (2008). ALARM-NET: Wireless Sensor Networks for Assisted-Living and Residential Monitoring.

4.2.1 Artificial Intelligence techniques

Artificial intelligence techniques are often subdivided into three major sub-groups, namely: Supervised Learning, Unsupervised Learning and Reinforcement Learning. Each of these techniques brings a set of different advantage and disadvantages, and in can have distinct applications in the field of cyber-security.

4.2.1.1 *Supervised Learning*

Supervised Algorithms are ideal for scenarios where the attacks are already known. These algorithms work by learning the classifications of attacks in training datasets and are then capable of detecting such attacks. These algorithms present advantages when it comes to identifying attacks with detail, however they fall short when a computer system is being attacked by an attack that is still unknown.

Artificial Neural Networks

An Artificial Neural Network (ANN) [12] is a processing unit inspired by the neuron connections in the human brain, neural networks acquire knowledge through a learning process from the surrounding environment, and the knowledge is stored on the weights used in the interconnections between the neurons.

A neural network is characterized by two basic aspects, its architecture that is related to the type, number of processing units and the way the neurons are connected; and the learning rules used to adjust the weights of the network and the information used by the network. Multilayer Perceptron (MLP) is the architecture most used in neural networks in the application of pattern recognition problems both in general scope and in the detection area of anomalies, as shown in Figure 3.

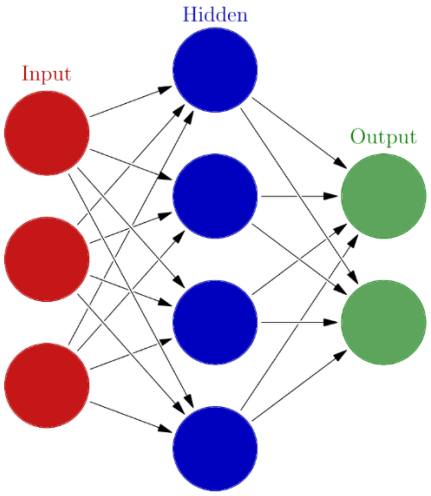

Figure 3 - Multilayer Perceptron

An MLP network is divided into three classes an input layer that receives the information to be processed, an output layer where the processed data is found, and one or more hidden layers between these two. Processing units play a very simple role. Each input terminal of a neuron receives a value. The received values are weighted and combined by a mathematical activation

function. The output of the function is the response of the neuron to the next neuron. Neural networks may have backpropagation or feedback connections. These connections allow a neuron to receive input from a neuron in the same or layer or following layers. Networks without backpropagation connections are the most commonly used and named by feedforward[39].

Support Vector Machine

Support Vector Machines[40] are a supervised learning technique that has the ability to solve classification and regression problems. This type of algorithm focuses on the search for a hyperplane of dimension n that best divides a dataset into two classes. The support vectors are the points that are near the hyperplane, Figure 4. These are considered the critical elements of the dataset because they are responsible for determining the position of the hyperplane and if removed, the position of the hyperplane would change.

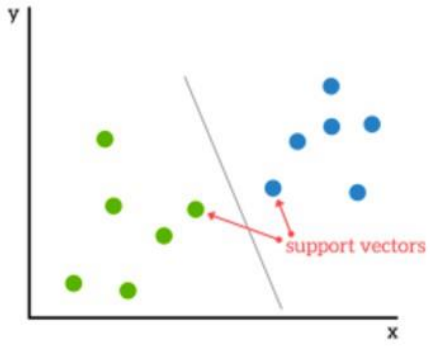

Figure 4 - Two-dimensional plane with support vectors and hyperplane

SVM are effective in sorting linearly separable data or data with an approximately linear distribution. However, there are many cases where it is not possible to divide training data in a hyperplane. When this type of problem appears SVMs preform a mapping of the training set from its original space to a new higher dimension space, called feature space[41]. To calculate the scalar products between objects mapped in the new space, functions called Kernels are used. The usefulness of kernels lies in the simplicity of their calculation and in their ability to represent abstract spaces. Some of the kernels most used are the polynomials, the radial base function and the sigmoid function. Each of them has parameters that can be defined to obtain ideal results[42].

C4.5

C4.5 is an algorithm, that given an already classified dataset (training set) constructs a classifier in the form of a decision trees. These decision trees are a very effective supervised learning method. Its purpose is to partition a dataset into groups of homogeneous data for the forecast of

[39] M. O. João Gama, André Ponce de Leon Carvalho, Katti Faceli, Ana Carolina Lorena, *Extração de Conhecimento de Dados*. 2015.

[40] Kdnuggets, "support-vector-machines-simple-explanation @ www.kdnuggets.com." .

[41] E. Osuna and J. Platt, "Support vector machines," 1998.

[42] M. O. João Gama, André Ponce de Leon Carvalho, Katti Faceli, Ana Carolina Lorena, *Extração de Conhecimento de Dados*. 2015.

a given variable. It receives as input a sorted dataset and returns a tree that resembles a diagram where each node of the end of the tree (leaf) is a class and each internal node represents a test. Each leaf represents the decision to belong in a class, that verified by the path from the root, through the test nodes to the ends. Shannon–Hartley theorem is the basis of algorithms that use decision trees. Shannon's entropy is the most applied in that it defines the amount of information provided by an event. Given a distribution probability $P = (p_1, p_2, p_3)$ and a sample S, then the amount of information of this distribution, called Entropy of P, is given by expression [43] (1):

$$Entropy(P) = \sum_{i=1}^{n} p_i \times \log(p_i) \qquad (1)$$

C4.5 applies the information gain, using the entropy as an impurity measure to generate the decision tree. To determine how good a test condition is, one must compare the degree of entropy of the parent node with the degree of entropy in child nodes. The attribute that generates the greatest difference is chosen as the test condition.

Random Forest

Random Forest is an algorithm that, as the name implies, creates a forest with a certain number of decision trees. This is a set method whose purpose is to create a strong learner through a group of less complex learners. Each classifier, in this case, each decision tree is a "week learner", while the joining of all decision trees is considered a "strong learner"[44]. The Random Forest algorithm used decision tree stuttered as $\{h(x, \Theta k), k = 1, \ldots,\}$ where $\{\Theta k\}$ are identically distributed independent random vectors and x is a pattern of entry[45]. In the training process, the Random Forest algorithm creates multiple CART decision trees[46], each of which is trained with a data sample taken from the training dataset, and only looks at a random sub dataset of the input variables to determine a division for each node of the tree. In the classification, each tree of the Random Forest algorithm represents a vote for the most popular class of the input variable x. The result of the algorithm is the class with the majority of the votes given by decision trees.

Naive Bayes

The Naive Bayes is an algorithm based on Bayes' Theorem, that aims to simplify the classification problem by making the assumption of features independence within each class. Assumptions of independence are usually not correct but perform well in the Bayesian learning scheme. The relationships between the dependent events can be described by means of the Bayes' Theorem, as shown in expression (2):

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(B \cap A)}{P(B)} \qquad (2)$$

[43] B. Hssina, A. Merbouha, H. Ezzikouri, and M. Erritali, "A comparative study of decision tree ID3 and C4.5," *Int. J. Adv. Comput. Sci. Appl.*, no. 2, pp. 13–19, 2014.

[44] D. Benyamin, "A Gentle Introduction to Random Forests, Ensembles, and Performance Metrics in a Commercial System." 2012.

[45] L. Breiman, "Random Forests," pp. 1–33, 2001.

[46] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*. 1984.

The notation $P(A|B)$ can be read as the probability of event A, knowing that event B happened. It is also known as conditional probability since the probability of A is conditioned by event B[47].

K Nearest Neighbor

The algorithm K Nearest Neighbor varies by the number of k neighbors defined. The simplest of its variations is the nearest neighbor algorithm (1NN) where $k = 1$. In this algorithm each object represents a point in the space defined by the attributes. By selecting a metric in this space, it is possible to calculate the distances between these two points. The most used metric is the Euclidean distance, given by (3), where X_i and X_j are two objects represented by vectors in space Rd, with x_i^l and x_j^l being elements of these vectors, which correspond to with the attributes of the coordinate l.

$$d(X_i, X_j) = \sqrt{\sum_{i=1}^{d} (x_i^l - x_j^l)^2} \qquad (3)$$

This variation of the K Nearest Neighbor algorithm where $k = 1$ is quite simple. In the training phase, the algorithm memorizes the classified data samples of the training set. In the classification phase, the unclassified samples, that is, whose class is not known, are analyzed by the algorithm. The distance between the vector of attribute values of the new unclassified sample and each already classified sample in memory is calculated. The label of the class associated with the nearest training set sample is the selected as output to the current classification[48].

4.2.1.2 *Unsupervised Learning*

Unsupervised Learning Algorithms contrast with Supervised Learning Algorithms because by their nature they learn without a classified dataset. Looking specifically in cyber-security, these algorithms are ideal for outlier detection, this means, identifying data or patterns of behavior in a computer system or computer network that seems to be out of place. With outlier detection techniques it is possible to identify behaviors that indicate the presence of a cyber-attack, even if the attack is new and operates in an unknown way. However Unsupervised Learning Algorithms have the disadvantage that when an attack is found the information provided is not has detailed has with Supervised Algorithms, this is because the algorithm only detects the presence of an attack but not which type of attack. This behavior is often called as one class classification.

Autoencoder

Autoencoder [49] is an unsupervised learning algorithm designed as a subtype of neural network. The idea around this algorithm is the process of compression and decompression of data. The neural network is structed as a sequence of neuron layers, where the number of neurons gets

[47] A. Hajek, "What Conditional Probability Could Not Be," *Synthese*, vol. 137, no. 3, pp. 273–323, 2003.

[48] M. O. João Gama, André Ponce de Leon Carvalho, Katti Faceli, Ana Carolina Lorena, *Extração de Conhecimento de Dados*. 2015.

[49] J. Chen, S. Sathe, C. Aggarwal, and D. Turaga, "Outlier Detection with Autoencoder Ensembles," in *Proceedings of the 2017 SIAM International Conference on Data Mining*, pp. 90–98.

smaller in each layer to a minimum number and then increases again to the original number of neurons. The objective with this algorithm is to obtain an output is a reconstruction of the initial data and a reconstruction error. The autoencoder network can be used to detect outliers, by learning only with normal data. After the learning process, when a new sample is processed, the reconstruction error is expected to be high if the sample represents an anomaly. To classify the data as normal or anomaly the autoencoder only needs a threshold value so that when the reconstruction error is above the threshold, the sample is classified as an anomaly.

One-Class K-Means

K-Means[50] is one of the most widely used clustering algorithms. One-Class K-Means works by initially learning with a set of training data. The algorithm calculates the distances between the characteristics of all samples in the training data and creates clusters with samples that have small distances between their characteristics. To apply this algorithm in an anomaly detection scenario, the learning process must use a data set containing only normal data, then, in the classification process, the algorithm calculates the distance to the nearest cluster. Once again, a threshold value is defined, if this distance calculated is greater than the defined threshold, the sample is classified as an anomaly.

One-Class Nearest Neighbor

One-Class Nearest Neighbor is an adaptation of the original K-Nearest Neighbor supervised algorithm[51], it is used in anomaly detection in a way similar to the One-Class K-Means algorithm, but instead of using clusters, this algorithm uses distances between Neighbors (near points). The One-Class Nearest Neighbor algorithm begins by learning the distances in a training set containing only normal data. These distances are used to identify a maximum distance and set a limit. Then, in the classification process, the algorithm begins by calculating the distance of the sample to the first point closest to the training set. If this distance is greater than the defined maximum threshold, the sample will be classified as an anomaly.

Isolation Forest

Isolation Forest[52] is a learning algorithm used for anomaly detection that makes use data structures called trees, such as binary trees. Each tree is created by partitioning samples recursively, randomly selecting an attribute and dividing value between the maximum and minimum values of the selected attribute. Each partition phase represents a node in the tree, and the number of partitions needed to isolate a point is equivalent to the length of the path from the root node to a leaf node. The length of the path to isolate a normal point is greater than the path

[50] K. A. Yoon, O. S. Kwon, and D. H. Bae, "An Approach to Outlier Detection of Software Measurement Data using the K-means Clustering Method," in *First International Symposium on Empirical Software Engineering and Measurement (ESEM 2007)*, 2007, pp. 443–445.

[51] F. Angiulli and C. Pizzuti, "Fast Outlier Detection in High Dimensional Spaces," *Princ. Data Min. Knowl. Discov.*, pp. 15–27, 2002.

[52] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation-Based Anomaly Detection," *ACM Trans. Knowl. Discov. Data*, vol. 6, no. 1, p. 3:1--3:39, 2012.

length of an anomaly point. Therefore, this algorithm is able to calculate the average path length of a set of trees and uses this value to determine the score of the anomaly.

Reinforcement Learning

Reinforcement learning algorithms are quite different from both Supervised and Unsupervised learning algorithms manly because their learning process rests on a mathematical model of the problem and not on datasets. Reinforcement learning technics learning by an iterative process of observing the world, preforming an action, and evaluating the results. This learning process make these technics good candidates for scenarios of uncertainty and rapid change, such as the process of selecting the ideal course of action to combat cyber threats.

Problems where Reinforcement Learning Algorithms (MDP) are applied are represented as Markov decision processes. MDP[53] is a non-deterministic search problem, that makes a representation of the environment of the world where an agent can be used to make act, and learn with its decisions in the face of a stochastic problem.

An MDP makes a representation of the world in a finite list of states and a finite list of actions. At any given time, the world is represented by one of the possible states. Performing an action results in what is called a transition. A transition can leave the world in the same state or lead it to a different state. The state to which the transition is made is obtained by following a probabilistic distribution that relates the current state with the action taken, and when the transition is completed a reward, or loss, corresponding to that transition is obtained. An MDP can be formally defined as follows:

- S - Finite number of states,
- s_0 – Initial state,
- A – Finite number of actions,
- $T(s, a, s')$ – Possible transitions from a state s to a state s' carrying out an action a,
- $\delta: S \times A \rightarrow dist(S)$ – Function that given a state s and an action a, possible to take in state s, returns a probabilistic distribution of which possible states the transition can be made to and the associated probability for each
- $R(t)$ – Reward or loss associated with a transaction t.

Q-Learning

Q-learning[54] is a commonly used RLA that aims to compute the optimal value function. Q-learning follows the steps of: observing the environment; selecting an action preform; preforming the action; observing the results in the environment, and when a reward is received for a given state s_n and an action a_n, the formula (4) is used of determine the value of that action.

$$\begin{aligned}\hat{Q}(s_n, a_n) &:= (1-\alpha_n)\hat{Q}(s_n, a_n) + \alpha_n[r_n + \gamma \max_{a'} \hat{Q}(s_{n+1}, a')] \\ &= \hat{Q}(s_n, a_n) + \alpha_n[r_n + \gamma \max_{a'} \hat{Q}(s_{n+1}, a') - \hat{Q}(s_n, a_n)]\end{aligned} \quad (4)$$

[53] T. Brázdil, K. Chatterjee, V. Forejt, and A. Kučera, "Trading performance for stability in Markov decision processes," *J. Comput. Syst. Sci.*, vol. 84, pp. 144–170, Mar. 2017.

[54] N. Shimkin, "Reinforcement Learning – Basic Algorithms." 2011.

The formula takes into consideration the following parts:

- old value: $Q(s_n, a_n)$
- learning rate: α_n
- reward received: r_n
- discount factor: γ
- estimated future value: $\max Q(s_{n+1}, a)$

The learning rate and the discount factor are variables that can be chosen by the user, and this choice should take into consideration the specific learning problem where they are being applied. The values learnt by the Q-Learning algorithm will, over time, influencing the action selection process so that the best actions are selected more often than not.

Roth-Erev

Roth-Erev[55,56] is a RLA developed by Roth and Ever, and is based on two fundamental principles of psychology that provide a formal definition on how learning is done by humans and animals. The principles are the following:

- The Law of Effect – "… choices that have led to good outcomes in the past are more likely to be repeated in the future [30]"
- Power Law of Practice – "… learning curves tend to be steep initially, and then flatter [30]"

The main objective of Roth-Erev is to simulate learning strategies used in games. This is done by trying the most options possible in the beginning of the learning processes similar to how an inexpedient player tries different possibilities when playing a new game and in the later phase of the learning giving lower weights to the learning done in the beginning of the learning processes. These two parameters are called "experimentation" and "forgetting" respectively, these approach makes the algorithm capable of learning a lot of possibilities of action, while being capable of rejecting the bad ones.

Multi-Armed Bandit algorithms

Multi-armed bandit algorithms[57] is a family of RLA that have as a base the idea of trying to find a, slot machine also known as "one-armed bandit" or simply "arm", that may have a biased reward probability distribution picked a priori, looking for the best arm is called the exploratory phase, and using that information to make the biggest profit possible, this is the exploitation phase. The aim on these algorithms is to try the different options until enough confidence is built on what option is the best. These algorithms are usually called Upper Confidence Bound (UCB) algorithms.

- UCB1 combines the exploratory phase and an exploitation phase, and the algorithm choses one of those two actions in each iteration depending on the rewards received. The algorithm also has a concept of a regret function that is used to try to find the loss

[55] A. Roth and I. Erev, "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term," Jan. 1995.

[56] I. Erev and A. E. Roth, "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria," *Am. Econ. Rev.*, vol. 88, no. 4, pp. 848–881, 1998.

[57] G. Burtini, J. Loeppky, and R. Lawrence, "A Survey of Online Experiment Design with the Stochastic Multi-Armed Bandit," pp. 1–49, 2015.

correspondent with each arm. The arm with the lowest value in the regret function is considered the best option.

- UCB2 is an improvement on UBC1 that manages to shorten the exploratory phase time and consequentially achieve better results.

Adversarial Bandits

Adversarial Bandits algorithms are a sub family of the Multi-armed bandit algorithms, however, instead of fixed distributions, adversarial bandits follow the idea that an "Adversary" is changing the rewards distributions in each time step. This make the process have three steps: the adversary picks the reward probabilities, the agent selects an arm, the agent them processes the received reward. When working with Adversarial Bandits problems it is important to keep in mind that the adversary may have access to more information that the agent.

- Exponential-weight algorithm for exploration and exploitation (Exp3) is an algorithm for Adversarial Bandits problems. The algorithm uses a parameter called egalitarianism, $\gamma \in [0, 1]$, this parameter is used for the exploration. The objective with this parameter is to determine the amount of time $(1 - \gamma)$, in which the algorithm is doing a weighted exploration/exploitation. The weighted exploration/exploitation is based on the current estimated reward, and the rewards received from the weighted exploration/exploitation are immediately used to update the correspondent arm's weight with formula (3) where *i* indicates the arm, and *Pi* represents the received reward for the arm, and formula (4) is used to calculate the current probability for each arm.

$$w_{i,t} = w_{i,t-1} \cdot e^{\gamma \cdot \frac{\rho_i}{p_{i,t} \cdot K}} \tag{3}$$

$$p_{i,t} = (1 - \gamma)\frac{w_{i,t}}{\sum_{j=1}^{K} w_{j,t}} + \gamma \cdot \frac{1}{K}. \tag{4}$$

- Exponential-weight algorithm for exploration and exploitation with expert advice (Exp4) extends the Exp3 by adding a set of weight vectors that substitute the previously used weight values. Each vector represents a different context that is similar to "expert" information for the specific probability distribution in use. The previous formulas (3) and (4) are replaced with (5) and (6) respectively, where *j* indicates the index of the current context.

$$w_{j,t} = w_{j,t-1} \cdot e^{\gamma \cdot \frac{\rho_j \cdot \xi_{j,t}}{p_{j,t} \cdot K}} \tag{5}$$

$$p_{i,t} = (1-\gamma)\sum_{i=1}^{N}\frac{w_{j,t}\xi_{j,t}(j)}{\sum_{k=1}^{K} w_{k,t}} + \gamma \cdot \frac{1}{K} \quad (6)$$

4.2.2 Datasets

To train and apply these methodologies, and to develop new algorithms it is necessary to have valid and consistent data. This data is necessary to test and evaluate the algorithms used with the possibility of comparing results between other algorithms and authors. This brings the need for datasets that are specifically designed for this purpose.

4.2.2.1 *Darpa 1998*

DARPA's Intrusion Detection Evaluation program 1998 provides a large sample of attacks in computational systems[58,59]. The create this dataset data from a TCPDUMP and the audit data from Basic Security Module (BSM) was collected on a network that simulated communication traffic from an Air Force local network. The dataset created contains seven weeks of training data and two weeks of test data, with 38 types of network attacks and several scenarios of realistic intrusions performed as normal background data.

4.2.2.2 *KDD 99*

In 1999 a dataset released in the third international competition of the Knowledge Discovery and Data Mining (KDD) conference, called KDD99[60]. The objective of this competition was to construct a predictive model of intrusion detection in a network. This data set includes a wide variety of intrusions, simulated in a military-grade network. The records are separated into two sub datasets, one with classified data (training dataset) and one without classifications (test dataset). The samples were separated in into five categories including normal data that represents no malicious activity and four kinds of simulated intrusions: Denial of service (DoS), Remote to Local (R2L), User to Root (U2R) and Probe.

4.2.2.3 *NSL-KDD*

NSL-KDD is a dataset created by Tavallaee et al. (2009)[61] as an improvement of the original KDD 99 dataset. The creation of the NSL-KDD dataset aimed at solving the problems of the KDD 99 dataset. NSL-KDD is divided into two subsets of data, one to train the algorithms containing 125,973 observations and another to test the algorithms containing 22,544 observations. Each of these subsets has 43 attributes. The NLS-KDD dataset differs from KDD 99 in the following ways:

- It does not include redundant samples in either the training and test datasets

[58] Y. Liao and V. R. Vemuri, "Using Text Categorization Techniques for Intrusion Detection," in *Proceedings of the 11th USENIX Security Symposium*, 2002, pp. 51–59.

[59] L. P. Richard *et al.*, "DARPA Intrusion Detection Data Sets." .

[60] UCI Machine Learning Repository, "KDD Cup 1999 Data." p. 92697, 2015.

[61] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," *IEEE Symp. Comput. Intell. Secur. Def. Appl. CISDA 2009*, no. Cisda, pp. 1–6, 2009.

- The sample distribution of the dataset was changed to reflect the difficulty to classify some samples. Firstly, the samples were divided into difficulty levels based on the number of machine learning algorithms that can correctly classify the records. Then, samples were randomly selected of each difficulty level in a fraction that is inversely proportional to the fraction of distinct records of the KDD data set. In this way, the results of the classification of the different methods of machine learning, vary in a larger interval, what makes it more efficient, to evaluate and compare these methods;
- The number of records in the training and test datasets was adjusted. This means the evaluation of results from different research papers will be more consistent and comparable.

4.2.2.4 *ISCX*

The ISCX dataset[62] developed at the Canadian Institute for Cybersecurity. This dataset is based on the concept of profiles that contain detailed descriptions of intrusions and abstract distributions for low-level network entities, applications, protocols and services. Real network communication interactions were analyzed to create profiles for agents that generate real traffic to HTTP, SMTP, SSH, IMAP, POP3 and FTP protocols. A set of guidelines was created to delineate a valid dataset that establishes the basis for profiling. These guidelines are vital to the effectiveness of the dataset in terms of realism, total capture, integrity and malicious activity [37]. ISCX contains 7 days of network traffic captured with normal activity and 4 of these days are mixed with malicious activity form one type of attack each. These attacks were: local network infiltration, HTTP Denial of Service, Distributed Denial of Service using an IRC Botnet and SSH Brute Force.

4.2.2.5 *CICIDS2017*

CICIDS2017[63] is a recently developed Network based intrusion detection dataset. Samples in this data set are stored as in a real world like scenario, similar to PCAPs. One of the objectives when creating this dataset was to obtain realistic background traffic. To achieve this the authors proposed a B-Profile system that generates human traffic representing the normal activity in the dataset. This generated traffic corresponds to 25 users and includes the HTTP, HTTPS, FTP, SSH and email protocols. In total this dataset contains activity corresponding to 5 days, and includes the attacks: Brute Force FTP, Brute Force SSH, DoS, Heartbleed, Web Attack, Infiltration, Botnet and DDoS; with the following distribution:

- Monday - Benign.
- Tuesday - BForce,SFTP and SSH.
- Wednesday - DoS and Hearbleed Attacks, slowloris, Slowhttptest, Hulk and GoldenEye
- Thursday - Web and Infiltration Attacks, Web BForce, XSS and Sql Injection, Infiltration Dropbox Download and Cool disk.

[62] A. Shiravi, H. Shiravi, M. Tavallaee, and A. A. Ghorbani, "Toward developing a systematic approach to generate benchmark datasets for intrusion detection," *Comput. Secur.*, vol. 31, no. 3, pp. 357–374, 2012.

[63] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization," no. Cic, pp. 108–116, 2018.

- Friday - DDoS LOIT, Botnet ARES, PortScans (sS, sT, sF, sX, sN, sP, sV, sU, sO, sA, sW, sR, sL and B)

4.2.3 Practical applications

4.2.3.1 *J. Meira, R. Andrade, I. Praça, J. Carneiro, G. Marreiros*

In this work[64] the authors explore the usage of artificial intelligent algorithms in the context of unknown cyber-attack detection. This research was conducted with the publicly available NSL-KDD and ISCX datasets. The algorithms applied are based on one class classification: Autoencoder, Neural Network, K-Means, Nearest Neighbor and Isolation Forest. The algorithms were also paired with different preprocessing techniques to see if they would improve the results. The results showed that all algorithms had acceptable results in finding the unknown attacks. The performance of the algorithms was slightly better in the ISCX dataset. The authors conclude that these techniques show promising results and should continue to be studied and improved.

4.2.3.2 *E. Balkanli, J. Alves and A. N. Zincir-Heywood*

The work of these authors[65] explores the performance of two Network Intrusion Detection Systems and two Supervised Learning algorithms applied in the context of DDoS attacks Detection. The objective of this work is to compare traditional security methodologies, namely Bro[66] and Corsaro[67], with artificial intelligent ones, namely CART Decision Tree and Naive Bayes classifiers, and find if there is an advantage to artificial intelligent methods. The experimentation was conducted utilizing a publicly available dataset with registers containing information that represents Denial of Service attacks. This work showed that the best results were achieved with the Corsaro IDS and the Decision Tree classifier, the authors conclude that a combination of these two approaches would yield great results.

4.2.3.3 *M. Yousefi-Azar, V. Varadharajan, L. Hamey and U. Tupakula*

This work[68] aims at creating better pre-processing data methodologies for cyber security scenarios by applying an Auto-Encoder algorithm with the objective of creating a more effective representation of the feature set. The experimentation was conducted with several classifier algorithms using two publicly available datasets: Microsoft Malware Classification Challenge[69] was used as a malware dataset and the already mentioned NSL-KDD was used as network

[64] J. Meira, R. Andrade, I. Praça, J. Carneiro, and G. Marreiros, "Comparative Results with Unsupervised Techniques in Cyber Attack Novelty Detection BT - Ambient Intelligence – Software and Applications –, 9th International Symposium on Ambient Intelligence," 2019, pp. 103–112.

[65] E. Balkanli, "Supervised Learning to Detect DDoS Attacks," *2014 IEEE Symp. Comput. Intell. Cyber Secur.*, pp. 1–8, 2014.

[66] "Bro." [Online]. Available: https://www.bro.org/.

[67] "Corsaro." [Online]. Available: http://www.caida.org/tools/measurement/corsaro/.

[68] M. Yousefi-Azar, V. Varadharajan, L. Hamey, and U. Tupakula, "Autoencoder-based feature learning for cyber security applications," *2017 Int. Jt. Conf. Neural Networks*, pp. 3854–3861, 2017.

[69] R. Ronen, M. Radu, C. Feuerstein, E. Yom-Tov, and M. Ahmadi, "Microsoft Malware Classification Challenge," *CoRR*, vol. abs/1802.10135, 2018.

anomaly datasets. The classifiers used were: Naive Bayes, K-NN, SVM and Xgboost; only the Xgboost showed a slight loss in accuracy while the other showed better preforming results. The authors conclude that an approach like this brings benefits specially when it comes to combine classification for malware and network-based anomalies.

4.2.3.4 *R. Ganesan, Su. Jajodia, A. Shah and H. Cam*

This work incorporates Artificial intelligence techniques in the context of cyber security in a novel way. Currently security analysts are required to analyze reposts generated by traditional security tools like SNORT[70] and Bro[71]. The authors point to problems caused the task of scheduling all the analysis, due to different specialization of each analyst, time taken by each analysis, preferred shift hours and days-off. With this problem in mind the goal of this work was to create a dynamic scheduling system, by applying Reinforcement Learning Algorithms, that are capable of deciding when and which Cyber Security expert should make an analysis of which security reports. Thus, adequately combining the security analysts with the required task and leaving other security analysts free for other necessary tasks. The concept of risk was also taken into consideration in this work, risk being defined as the percentage of alerts that were not analyzed. The authors propose a stochastic optimization model using Reinforcement Learning and their results showed it to be a viable option to apply in this problem.

4.2.3.5 *S. Roy, C. Ellis, Q. Wu, D. Dasgupta, V. Shandilya and S. Shiva*

Game theory is a sub field of Artificial intelligence that focuses on understanding and applying the same strategies that produce the ideal results in games. This work [72] proposes an architecture for cyber security that considers malicious activities and parties as the attacker and the system administrator as a defender in a game. The proposed architecture is called Game Theory Inspired Defense Architecture (GIDA), this is a holistic approach, meaning GIDA preforms all the activities in all parts necessaries in cyber security. This work is presented as a stepping stone to implement a complete system and combine with other learning methodologies.

4.2.3.6 *C. Symons and J. Beaver*

Semi-supervised learning is the name given to learning algorithms that incorporate behaviors from both supervised and unsupervised learning methodologies. These algorithms are capable of using labeled and unlabeled data in the learning process. This work[73] proposes the application of these algorithms in network intrusion detection scenarios. The authors experimented with the Kyoto2006+ dataset [74]. The results presented in in this work present great performance for this

[70] "Snort." [Online]. Available: https://www.snort.org/.

[71] "Bro." [Online]. Available: https://www.bro.org/.

[72] S. Roy, "Game Theory for Cyber Security," 2010.

[73] C. T. Symons, O. Ridge, and J. M. Beaver, "Nonparametric Semi-Supervised Learning for Network Intrusion Detection : Combining Performance Improvements with Realistic In-Situ Training Categories and Subject Descriptors," pp. 49–58.

[74] J. Song, H. Takakura, Y. Okabe, M. Eto, D. Inoue, and K. Nakao, "Statistical Analysis of Honeypot Data and Building of Kyoto 2006+ Dataset for NIDS Evaluation," in *Proceedings of the First Workshop on Building Analysis Datasets and Gathering Experience Returns for Security*, 2011, pp. 29–36.

task and shows evidence that it could be an effective method to apply in real world cyber security applications.

4.2.3.7 *R. Lippmann and R. Cunningham*

One of the problems in cyber-attack detection is the high false-alarm rates that occur in some systems, and improving those results is a major necessity. This work[75] explores exactly that point by utilizing a combination of keyword selection and neural networks. This work was developed with the DARPA 1998 dataset. The proposed methodology used a LNKnet pattern classification to detect and add new patterns of keywords. The authors conclude that with a simple keyword detection system, and the addition of new keywords it became possible to improve the results obtained with the DARPA dataset.

4.2.3.8 *A. Bivens, S. Rasheda, P. Chandrika and S. Boleslaw*

This work[76] focuses on network-based cyber intrusion detection. The authors propose a modular system that applies Self-Organizing Maps and Neural Network techniques to analyses the data. The data is obtained by preforming a tpcdump on the network machines and thus obtaining the traffic data in windows of time. With this proposed architecture both the Neural Network and Self-Organizing Maps start by preforming a learning phase, that is feed by live tcpdump data and tpcdump data from the DARPA dataset. The results show that with this approach the Neural Network had low performance on detecting groups of attacks, but at good results at identifying single types of attacks.

4.2.3.9 *J. Luo, and S. Bridges*

This work[77] focuses on the usage of association rules and frequency episodes integrated with fuzzy logic, applied within the context of intrusion detection. The authors propose the concept of fuzzy frequency episodes and an algorithm targeted at mining these episodes, that is an algorithm that is capable of applying fuzzy logic and association rules to temporal data. The combination of these techniques makes this approach able to identify if a sample corresponds to a pattern of normal activity or an anomaly, that may represent a cyberattack. The authors conclude that this methodology showed itself to be a viable option for intrusion detection problems.

4.2.3.10 *C. Kruegel, D. Mutz, W. Robertson and F. Valeur*

Being able to detect new and previously unknown attacks is a major goal in cyber security research. This work[78] aims towards finding viable techniques to apply in this context. The authors identify problems in result aggregation, when more than one classifier is used, and in the fact that

[75] R. P. Lippmann and R. K. Cunningham, "Improving intrusion detection performance using keyword selection and neural networks," vol. 34, pp. 597–603, 2000.

[76] A. BIVENS, S. RASHEDA, P. CHANDRIKA, and S. BOLESLAW, "NETWORK-BASED INTRUSION DETECTION USING NEURAL NETWORKS," vol. 12, pp. 579–584, 2002.

[77] J. Luo and S. Bridges, "Mining fuzzy association rules and fuzzy frequency episodes for intrusion detection," *Int. J. Intell. Syst.*, vol. 15, pp. 687–703, 2000.

[78] C. Kruegel, D. Mutz, W. Robertson, and F. Valeur, "Bayesian event classification for intrusion detection," in *19th Annual Computer Security Applications Conference, 2003. Proceedings.*, 2003, pp. 14–23.

unusual behavior patterns might be classified as outliers, and there for attacks, even when they represent normal benign activity. To combat these problems a Bayesian network model is proposed. By applying the Bayesian network, it was possibly to aggregate multiple results with higher performance. The results show how this approach was able to achieve far less false positives.

4.2.3.11 *K. Sequeira and M. Zaki,*

This work[79] explores the problems preforming intrusion detection in real-time scenarios. The authors targeted the specific problem of detecting if the user in a terminal based computer system was real or an intruder. Aside for the traditional password protection the authors propose their Intrusion Detection system, called Anomaly-based Data Mining for InTrusions (ADMIT). The proposed system worked by first learning from a record of the real user interaction with the machine. The system was them able of classifying if future interactions were taken with the real or false user. In the test preformed the system showed acceptable results. The authors identify the need to reduce the required training and improve the overall system in future work.

4.2.3.12 *S. Mukkamala, A. Sung and A. Abraham*

Combining several artificial intelligence algorithms is one of the possible ways to improve upon the results of single technique approaches. This work[80] explores the performance of three algorithms in a cyber intrusion detection scenario, namely: performance of Artificial Neural Networks (ANNs), Support Vector Machines (SVMs) and Multivariate Adaptive Regression Splines (MARS); and compares it to a combination of all the three. This experiment is conducted using the DARPA 1998 dataset. The results show that the combination approached achieves greater results in all the scenarios tested by the authors.

4.2.3.13 *W. Lee, S. Stolfo and K. Mok*

In this work[81] the authors explore the concept of adaptively building Intrusion Detection models. The proposed is to preform data collecting in networks and host sessions in order to extract a feature set that can be analyzed by data mining techniques such as: classification, meta-learning, association rules, and frequent episodes. The authors conducted an experiment to evaluate the performance of the proposed methodology, using the DARPA 1998 intrusion detection dataset. The experimental results show that the data mining techniques applied produce reliable anomaly detection models. The authors conclude by expressing the need to continue to work on these techniques, especially making then suitable for real-time intrusion detection systems.

---

[79] K. Sequeira and M. Zaki, "ADMIT: Anomaly-based data mining for intrusions," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2002, pp. 386–395.

[80] S. Mukkamala, A. H. Sung, and A. Abraham, "Intrusion detection using an ensemble of intelligent paradigms," *J. Netw. Comput. Appl.*, vol. 28, no. 2, pp. 167–182, 2005.

[81] W. Lee, S. J. Stolfo, and K. W. Mok, "A data mining framework for building intrusion detection models," in *Proceedings of the 1999 IEEE Symposium on Security and Privacy (Cat. No.99CB36344)*, 1999, pp. 120–132.

4.2.3.14 *N. Amor, S. Benferhat, and Z. Elouedi*

Naive Bayes and Decision Trees are two artificial intelligence techniques commonly used in cyber-attack detection scenarios. In this work[82] a comparative analysis is preformed between these two techniques in a cyber-attack detection problem. The experiment was conducted using the KDD99 dataset as a source for intrusion data. The authors considered three scenarios: classifying all attacks, classifying samples as one of five possible groups and anomaly detection, that is, classifying if the sample as normal data or an attack. The results were similar in all three scenarios the results very similar. Both algorithms made slightly above 90% correct classifications. The authors conclude that the algorithms have the potential to be applied together in order to obtain even better results.

4.2.3.15 *T. Shon and J. Moon*

This work[83] focuses on the problem of detecting unknown security attacks. The authors propose a new anomaly detection algorithm to apply in these scenarios. This algorithm is called Enhanced Support Vector Machine (SVM), which is an improvement upon the SVM algorithm, in order to obtain its performance for this context. This proposed methodology follows four steps, them being: the use of a Self-Organized Feature Map to create the profile of normal data, packet filtering to reject incomplete network traffic, the usage of a Genetic Algorithm (GA) to extract optimized information and a preprocessing step that considers temporal relationships in the data preprocessing. The experimental work was preformed using the DARPA intrusion dataset. The results show that this approach was capable of achieving good results and with a low false positive rate, which is an important factor in real world intrusion detection systems.

4.3 Medical Devices Cyber Security

In recent years the state of the art in cybersecurity for medical devices is catching up with other domains [84,85]. This follows medical device reaching a tipping point with software-driven functionality and increasing connectivity of devices through the Internet or over networks[86]. Before, cybersecurity for networked medical devices has been usually "bolted on" by manufacturers at the end of the design cycle, rather than integrated as a key factor of the product

[82] N. Ben Amor, S. Benferhat, and Z. Elouedi, *Naive Bayes vs decision trees in intrusion detection systems*, vol. 1. 2004.

[83] T. Shon and J. Moon, "A hybrid machine learning approach to network anomaly detection," *Inf. Sci. (Ny).*, vol. 177, no. 18, pp. 3799–3821, 2007.

[84] *The Evolving State of Medical Device Cybersecurity.* Schwartz, Suzanne, et al. 2018. 2, s.l. : AAMI, 2018, Biomedical Instrumentation & Technology, Vol. 52, pp. 103-111.

[85] *Think Like a Hacker: Insights on the Latest Attack Vectors (and Security Controls) for Medical Device Applications.* Khera, Mandeep. 2017. 2, s.l. : SAGE, 2017, Journal of Diabetes Science and Technology, Vol. 11, pp. 207-212.

[86] *A brief chronology of medical device security.* Burns, A. J., Johnson, M. and Honeyman, Peter. 2016. 10, s.l. : ACM, 2016, Communications of the ACM, Vol. 59, pp. 66-72.

development and value creation process[87]. Medical devices got challenged by basic cybersecurity hygiene that must be addressed during early engineering and design [88,89].

4.3.1 Expert guidance from academia and industry

To get medical devices cybersecurity state to a proper level, experts from academia and industry put together guidance and regulatory bodies started to define their expectations. This state of the art focuses on the technical measures and not the organizational measures proposed by these guidances. Of course, in practice organizational and non-technical measures are very important, including a number of aspects like cybersecurity organization, security and risk management frameworks and policies, and concepts of Secure Development Lifecycle (SDLC) and Security by Design[90].

ENISA covers security for smart hospitals and presents a number of security good practices for technical security measures. The scope explicitly includes networked medical devices[91]: mobile devices (e.g. glucose measuring devices), wearable external devices (e.g. portable insulin pumps, wireless temperature counters), implantable devices (e.g. cardiac pacemakers), stationary devices (e.g. computer tomography (CT) scanners, life support machines, chemotherapy dispensing stations), and supportive devices (e.g. assistive robots).

Similarly, Haigh and Landwehr present a building code that provides a basis for reducing the risk that software used to operate medical devices is vulnerable to malicious attacks. The code consists of elements organized in 10 categories[92]. A platform approach and reference architecture, specifying requirements for security mechanisms and functionality, is another contribution to secure medical devices[93].

4.3.2 FDA pre- and post-market guidance

The FDA (Food & Drug Administration) introduced guidance to medical device manufacturers. The FDA guidance on Premarket Management on Cybersecurity in Medical Devices[94] identified general principles to be applied together with a number of explicit requirements for cybersecurity functions reflecting priorities on addressing a number of urgent issues, e.g. hard-

[87] *A Value Blueprint Approach to Cybersecurity in Networked Medical Devices.* Tanev, George, Tzolov, Peyo and Apiafi, Rollins. 2015. 6, June 2015, Technology Innovation Management Review, Vol. 5.

[88] Fu, Kevin. 2016. On the Technical Debt of Medical Device Security. *Fronteers of Engineering: Reports on Leading-Edge Engineering from the 2015 Symposium.* s.l. : The National Academies Press, 2016.

[89] *Cybersecurity Concerns and Medical Devices Lessons From a Pacemaker Advisory.* Kramer, Daniel B. and Fu, Kevin. 2017. 21, s.l. : JAMA, 2017, Vol. 318, pp. 2077-2078.

[90] Philips. 2018. *Position paper Philips and cybersecurity.* 2018.

[91] ENISA. 2016. *Smart Hospitals: Security and Resilience for Smart Health Service and Infrastructures.* 2016.

[92] Haigh, Tom and Landwehr, Carl. 2015. *Building Code for Medical Device Software Security.* s.l. : IEEE, 2015.

[93] *A Reference Architecture for Secure Medical Devices.* Harp, Steven, Carpenter, Todd and Hatcliff, John. 2018. 5, s.l. : AAMI, 2018, Biomedical Instrumentation & Technology, Vol. 52, pp. 357-365.

[94] FDA. 2014. *Content of Premarket Submissions for Management of Cybersecurity in Medical Devices – Final Guidance.* 2014.

coded password use. The new draft version takes this to a new level and expands on the general principles and risk assessment, recommends the application of NIST Cybersecurity Framework, and describes the specific design features and cybersecurity design controls it believes should be included in the design of a trustworthy device[95]. The design controls are grouped under for "Identify and Protect Device Assets and Functionality" and "Detect, Respond, Recover: Design Expectations". The FDA also published its guidance for Postmarket Management of Cybersecurity in Medical Devices to address and manage cybersecurity for devices after being on the market[96]. The guidance recommends some more design controls, i.e. technical security device features.

4.3.3 Standards

Standards and standards-based approaches are gaining traction for medical device security[97]. The FDA builds its guidance on the NIST Cybersecurity and lists a number of recognized standards, which it tracks in its FDA Recognized Consensus Standards Database[98]. Particularly, NIST SP 800-53[99] provides a baseline of many security and privacy controls that medical device may adopt where applicable. ISO 80001-2-2 also lists technical security controls that may be used to address cybersecurity in medical devices[100].

4.3.4 Security design controls and technical measures

The medical device security state of the art is moving towards getting on par with other domains, yet in some cases the medical setting presents a unique setting that sets it apart and puts itself at the state of the art frontier. Below we touch on these aspects along the design controls and technical security measures from abovementioned sources for guidance. In perspective of the sheer width of the frontier we focus on the controls that relate most to the Safecare priorities with potential for analytics enhancements in the next section.

Access control. Modern medical devices support fine-grained access control and authorization mechanisms, e.g. role-based access control, to restrict access to data and device functions to

[95] FDA. 2018. *Content of Premarket Submissions for Management of Cybersecurity in Medical Devices – Draft Guidance.* 2018.

[96] FDA. 2016. *Postmarket Management of Cybersecurity in Medical Devices - Final Guidance.* 2016.

[97] *Standards for Medical Device Cybersecurity in 2018.* Yuan, Sean, Fernando, Anura and Klonoff, David C. 2018. 4, s.l. : SAGE Publications, 2018, Journal of Diabetes Science and Technology, Vol. 12, pp. 743-746.

[98] FDA. 2018. *Content of Premarket Submissions for Management of Cybersecurity in Medical Devices – Draft Guidance.* 2018.

[99] NIST. 2014. *Security Controls and Assessment Procedures for Federal Information Systems and Organizations.* 2014. NIST SP800-53 (Rev.4).

[100] *Cybersecurity and medical devices: Are the ISO/IEC 80001-2-2 technical controls up to the challenge?* 2017. s.l. : Elsevier, 2017, Computer Standards & Interfaces, Vol. 56, pp. 134-143.

people and systems that need it[101]. Adaptive access[102] and break-the-glass[103] concepts support specific healthcare workflows.

Authentication. Single sign-on and multi-factor authentication are concepts are introduced to a wider variety of medical devices. Hospitals employ solutions like Imprivata to strike a balance between workflow (usability, efficiency) and security. Risk-based authentication is introduced to support this.[104]

In parallel, devices are factory-provisioned with unique and cryptographically strong identities, digital certificates and keys for authentication, encryption, etc. purposes. [105]

Code integrity. State of the art medical devices employ code / firmware signing. Device verify these signatures to enforce only authorized code will be execute at installation / upgrade time or at run-time, i.e. leveraging application whitelisting[101] as supported by modern operating systems.

Secure boot (based on a hardware root of trust) complements code signing by ensuring that all code running on a medical device can be trusted including boot loader, operating system and application code.

Hardware root of trust. Integration of secure elements, e.g. TPM on PC, is good practice in state of the art medical devices to provide the ability to store cryptographic keys securely. The hardware root of trust contributes to establishing a Trusted Computing Base (TCB)[106].

Encryption. Encryption of data in-transit is the norm today[107,108], e.g. realized through protocols like TLS, etc. Encryption at-rest follows supported by a wide variety of encryption technologies and ideally backed by a secure element. Encryption of removable media has a number of solutions, ranging from Bitlocker-to-go to IHE XMD Media Encryption, each with its set of limitations. End-to-end data encryption and data-centric encryption are concepts to gain maturity and broader adoption.

Security updates and patches. The state of the art for security patches and updates for medical devices includes their capability to securely and meaninglessly update in the field[103]. This is supported with strategies and infrastructure to reduce the time needed to deploy updates to

---

[101] ENISA. 2016. *Smart Hospitals: Security and Resilience for Smart Health Service and Infrastructures.* 2016.

[102] *Access Control Schemes for Implantable Medical Devices: A Survey.* Wu, Longfei, et al. 2017. 5, s.l. : IEEE, 2017, IEEE Internet of Things Journal, Vol. 4, pp. 1272-1283.

[103] *Lightweight break-glass access control system for healthcare Internet-of-Things.* Yang, Yang, Liu, Ximeng and Deng, Robert H. 2018. 8, s.l. : IEEE, 2018, IEEE Transactions on Industrial Informatics, Vol. 14, pp. 3610-3617.

[104] *e-SAFE: Secure, Efficient and Forensics-Enabled Access to Implantable Medical Devices.* Chi, Haotian, et al. 2018. Beijing : IEEE, 2018. IEEE Conference on Communications and Network Security.

[105] ENISA. 2016. *Smart Hospitals: Security and Resilience for Smart Health Service and Infrastructures.* 2016.

[106] Haigh, Tom and Landwehr, Carl. 2015. *Building Code for Medical Device Software Security.* s.l. : IEEE, 2015.

[107] ENISA. 2016. *Smart Hospitals: Security and Resilience for Smart Health Service and Infrastructures.* 2016.

[108] Philips. 2018. *Position paper Philips and cybersecurity.* 2018.

devices and keep their versions current, thereby shortening the time window at which devices are vulnerable.

Host firewalls. Medical devices may employ firewalls themselves in form of host firewalls[103]. This contributes to attack surface minimization, defense-in-depth, etc. concepts. Host firewalls may also provide mitigations against certain vulnerabilities pending deployment of a security patch.

Malware protection. Eligible medical devices employ anti-virus and anti-malware software[103]. To guarantee undisturbed operation these detect malware based on signatures of known malware. Medical device vendors will validate updates before these are deployed on devices in the field.

Logging. Medical devices log security relevant events to enable auditing. Log events may be stored on the device as well as remote, e.g. by leveraging standards like syslog[109], IHE ATNA, DICOM Part 15 Annex A5[110], FHIR AuditEvent. A challenge faced by existing and new medical devices is to systematically log anything that is potentially security relevant in form of structured data.

To meet HIPAA and GDPR privacy requirements de-identification (anonymization) principles are applied as part of the logging mechanism.

Remote management and monitoring. Medical devices may have remote management capabilities[103]. This enables authorized service engineers to provide support and manage security, e.g. apply security mitigations, inspect log files, apply updates, etc.

Remote monitoring can support the security management of medical devices in the field. An up-to-date view on the state of a device, e.g. software versions and patch levels or logged security events and alerts, can help to prioritize (remote) maintenance and support as well as risk management and security roadmapping. For these purposes remote management capabilities include the ability to upload relevant security data to a central infrastructure of e.g. the vendor.

4.3.5 Security analytics

Security analytics, e.g. leveraging artificial intelligence / machine learning / data science for the security purposes, has the potential to enable or improve (aforementioned) security functions including detection, monitoring and risk management[111]. This emerging field has been introduced in network infrastructure level, e.g. intrusion detection, and for enterprise security operations, but is now also entering the sphere of medical devices.

The state of the art of security analytics for medical devices is very much subject to research to reach maturity sufficient for broad deployment. Yet, its application is stimulated and expected. For example, the FDA states in its post-market guidance[112]:

> "Incorporation of Threat Detection Capabilities

[109] IETF. 2009. *The Syslog Protocol.* 2009. RFC 5424.

[110] DICOM. 2018. *Security and System Management Profiles.* s.l. : NEMA, 2018. PS3.15 2018e.

[111] Zorz, Zeljka. 2018. *The benefits and limitations of AI in cybersecurity.* s.l. : Helpnetsecurity, 12 20, 2018.

[112] FDA. 2016. *Postmarket Management of Cybersecurity in Medical Devices - Final Guidance.* 2016.

> Medical devices may not be capable of detecting threat activity and may be reliant on network monitoring. Manufacturers should consider the incorporation of design features that establish or enhance the ability of the device to detect and produce forensically sound postmarket evidence capture in the event of an attack. This information may assist the manufacturer in assessing and remediating identified risks."

Similarly, ENISA recommends to implement monitoring and intrusion detection mechanisms as part of state of the art measures[113].

Implantable Medical Devices (IMD) is a class of medical devices for which the concept of anomaly detection has been explored [114,115]. One approach is to monitor the IMD externally, particularly the radio-frequency wireless communications, for anomalies [116,117]. IMDs are a special class because they typically have a small and well-defined functional scope, perform life-critical functions, have restricted and infrequent external communication (typically for maintenance and management), and are (very) resource constraint.

Anomaly and intrusion detection on other classes of medical devices raises similar questions. Logical candidate areas are malware detection beyond current pattern / signature-based approaches, host firewall enhancements with network anomaly detection capabilities, etc. Empirical studies have to determine the effectiveness of such methods and how to optimally leverage them as a security control. The same applies to translation of these concepts from network- / host-level to application-level. Since false positives may disturb the medical function of the device, detection is likely the primary function and prevention a secondary derived function. This can be particularly powerful in combination with remote monitoring where the combined data may be used for security intelligence and risk management purposes. The combined availability of heterogeneous logs may enable a multi-analysis approach to study complex security events[118]. However, while promising scientific results are lacking until now. Chaundry et.al present a middleware approach to postmarket surveillance of devices to provide the operational details (excluding the private data of patient) of the devices to the manufacturers and give device manufacturers the means to closely monitor the functioning of devices, upgrade

[113] ENISA. 2016. *Smart Hospitals: Security and Resilience for Smart Health Service and Infrastructures.* 2016.

[114] *Ideas and Challenges for Securing Wireless Implantable Medical Devices: A Review.* Zheng, G., et al. 2017. 3, s.l. : IEEE, 2017, IEEE Sensors Journal, Vol. 17, pp. 562-576.

[115] *Access Control Schemes for Implantable Medical Devices: A Survey.* Wu, Longfei, et al. 2017. 5, s.l. : IEEE, 2017, IEEE Internet of Things Journal, Vol. 4, pp. 1272-1283.

[116] *MedMon: securing medical devices through wireless monitoring and anomaly detection.* Zhang, M., Raghunathan, A. and Jha, N. K. 2013. 6, s.l. : IEEE, 2013, IEEE transactions on biomedical circuits and systems, Vol. 7.

[117] *Syndrome: Spectral analysis for anomaly detection on medical IoT and embedded devices.* Sehatbakhsh, Nader, et al. 2018. Washington : IEEE, 2018. IEEE International Symposium on Hardware Oriented Security and Trust (Host). pp. 1-8.

[118] *HuMa: A multi-layer framework for threat analysis in a heterogeneous log environment.* Navarro, Julio, et al. 2017. s.l. : Springer, 2017. Foundations and Practice of Security. pp. 144-159.

devices, patch security vulnerabilities and monitor device performance thereby enhancing health care outcomes[119].

At macro scale analysis has been done such as the prevalence of security risks within the clinical setting based on publicly available databases maintained by the Food and Drug Administration (FDA) to evaluate recalls and adverse events related to security and privacy risks of medical devices[120].

4.4 Policies, Procedures and Awareness

4.4.1 Data classification

Data classification refers to the process of organizing and categorizing data according to various properties, such as its sensitivity, the department it relates to, etc.

The main categories should refer to: Data types, data locations, data access levels and data protection levels implemented, with reference to compliance regulations.

A well performed data classification can outline the most important parts of data across the entire organization and allow an effective process of managing this data and controlling all access to it.

The data classification process is an intensive mapping process that requires full cooperation between the data owners (for example: department managers) and the corresponding IT and security teams in the company.

After classifying all sensitive and important data in the organization, the IT and Security departments can implement various security and management measures, allowing to limit the access to sensitive data to allowed personnel only, and apply additional protection layers such as backups and redundancy to information marked as critical to the organization.

Usage

Almost every organization deals with some sort of data classification. Some organizations are subject to various regulations, requiring them to manage their data in a certain way, other organizations manage their data just to maintain order, thus allowing them to focus their protection upon the most sensitive data.

4.4.2 Password strength

Almost every service accessed today, requires authentication using a combination of (at least) a username and password. A successful attempt to guess a password of a certain user, can easily grant access to the system to an unauthorized user.

Password strength refers to the effectiveness of a password in resisting guessing and brute-force/dictionary attacks. The strength is usually measured by the password's length, complexity and unpredictability. By integrating all previously mentioned aspects of the password, it is

[119] *POStCODE Middleware for Post-Market Surveillance of Medical Devices for Cyber Security in Medical and Healthcare Sector in Australia.* Chaudhry, Junaid, et al. 2018. s.l. : IEEE, 2018. 12th International Symposium on Medical Information and Communication Technology (ISMICT).

[120] *Security and Privacy Qualities of Medical Devices: An Analysis of FDA Postmarket Surveillance.* Kramer, Daniel B., et al. 2012. 7, 2012, PLoS One, Vol. 7.

possible to estimate how many trials an attacker would need, on average, to guess the password correctly.

Brute-force and dictionary attacks can be implemented, to generate a large number of access attempts, using thousands of password variations per second. The longer and stronger the password, the harder it would be to get it using these automatic methods.

But password strength is not only about its length and complexity. Additional security measures must be implemented to properly protect the login stage. These security measures may include various techniques, such as a limit on the number of wrong password attempts in a certain period of time, a time-out between each attempt to login, and even an entire user block after a certain number of times.

Another important aspect of password strength is the password creation process. Passwords are created either automatically (using randomizing algorithms and special programs) or more commonly by a person. Unlike the random algorithms, humans select their passwords based on various patterns which may help them remember the password more easily. These patterns, if known to the attacker, can help greatly at discovering the password (for example: a pet name, a birthday date, the user's name, etc.).

Usage

Many organizational IT and Security departments implement various password policies. These policies are usually configured to bind the password creation and management to various rules. For example: limit the minimum password length, define what characters must be used in each password, define rules about password repetition, limit the password lifetime, etc.

4.4.3 User education – security awareness

User education refers to the formal process of educating and raising the awareness level of employees towards computer security.

There is almost no organization today that doesn't use computers and information technologies of some sort. This means, all users in all organizations are subject to various Cyber threats that they need to be aware of.

The vast majority of most companies' employees are not tech-oriented and most of them don't know the potential risks of using the web or corporate network and end-point computers.

This low level of security awareness is usually easily exploited by hackers to infiltrate the company's systems, gain access to sensitive data or even cause damage to the company. The infiltration is executed using various techniques, such as Social Engineering, Phishing Email, Malicious programs and attachments, etc.

Most of the attacks can be spotted or prevented altogether by a security-aware user, who will not open suspicious links, or download attachments from an unfamiliar person. Well educated and security aware users will also report potential risks to the relevant person in the organization, in case they suspect they were attacked.

A well implemented user awareness and education program is usually implemented on all levels of the organization, making it a second nature to all employees. This effort to improve employee security awareness usually originates from the management and from the IT and information security teams, responsible to the company's security.

Usage

Security awareness programs are implemented in various ways in many organizations today, regardless of their main business. Companies dealing with highly sensitive information are more likely to implement broad security awareness programs to their employees, thus reducing risk or potential security risks originating from human error.

Available solutions on the market

Many security consulting companies offer various education and awareness packs and dedicated courses. These packs may include various materials, such as posters, slogans, games, and articles regarding security awareness.

5 Physical and Cyber Security Vulnerabilities

Combating today's attack attempts that are steadily getting more advanced and complex is a significant challenge to protecting people's security with only security checks to prevent hazardous objects from entering airports, port facilities, power plants, and other critical facilities and police patrols. What is attracting attention amid such circumstances as a new approach to combating cyber-attacks is cyber- physical security. This aims to block off cyber-attacks by driving the latest ICT by combining information from cyber space, which includes cyber-attack information, log information, and SNS and other Web information, information from the physical world, such as biometric data, behavioral analysis data, and GPS data. For example, a terrorist's cyber space terrorism plans or access to confidential information by a cyber-attack on a critical infrastructure establishment can be detected in advance, and through physical monitoring, including facial recognition, behavior detection, and drone patrolling at the target facility, attacks can be blocked off or promptly dealt with in case an emergency arises. NEC promotes comprehensive, unified responses for cyber-physical security, with the world's top-class physical security and cyber security as two critical components of an overarching whole. NEC boasts integrated security control and implementation solutions, NEC Cyber Security Platform, Security Operation Center (SOC), and world's number one biometric (facial and fingerprint) authentication technologies, and the world's first behavior detection and analysis technologies. Not only do these options work alone as competitive edges in their respective fields, but they can be combined to respond with their collective strengths. NEC will continue its pursuit to block cyber-attacks as well as quickly implement post- attack actions in case of a contingency[121].

Definition of vulnerability. Vulnerability is the quality or state of being exposed to the possibility of being attacked or harmed, either physically or emotionally 122. The International Organization for Standardization defines vulnerability as: "A weakness of an asset or group of assets that can be exploited by one or more threats where an asset is anything that has value to the organization, its business operations and their continuity, including information resources that support the organization's mission"123. While, ENISA definition of vulnerability "The existence of a weakness, design, or implementation error that can lead to an unexpected, undesirable event compromising the security of the computer system, network, application, or protocol involved."[124]

5.1 Vulnerabilities of Critical Infrastructures

Critical Infrastructures (Power Grid, water management, national transport, healthcare) are important resources which malfunction, destruction or even a partial unavailability could create havoc in a country. Nowadays CI are managed through Cyber-Physical Systems, CPS is an integration of physical processes (sensors and actuators) with computation and communication.

[121] NEC. Commercial facilities as targets: New threats to critical infrastructure. [Online] https://www.nec.com/en/global/about/vision/report/pdf/SocialValueCreationReport_en_Vol.1.pdf.

[122] https://en.oxforddictionaries.com/definition/vulnerability

[123] ISO/IEC, "Information technology -- Security techniques-Information security risk management" ISO/IEC FIDIS 27005:2008

[124]ENISA Glossary – (accessed in November 2018) from: https://www.enisa.europa.eu/topics/threat-risk-management/risk-management/current-risk/risk-management-inventory/glossary#G52

For example, Building, Energy and Facilities Management Systems (BMS) are used to integrate and command HVAC, lighting, CCTV, access controls and perimeter security.

Effective protection demands the availability of proper methodologies and tools to evaluate the vulnerability of the assets, and the ability of the adopted protection systems to meet its objectives[125].

Vulnerability is the manifestation of the inherent states of the system (e.g. physical, technical, organizational, cultural) that can be exploited by an adversary to harm or damage the system[126].

Given their importance, CIs must be protected by different type of threats:

- Adversarial threats, which pose malicious intentions from individuals, groups organizations or state nations.
    - Hacktivists: Sabotage systems to make a political or social statement. For example, in 2014 Boston Children's Hospital was involved in case regarding a teenage girl taken into state custody. According to someone in hacktivist group Anonymous the state custody was an infringement of the girl's right, therefore he decided to punish the hospital with a distributed denial of service (DDos) attack, flooding the hospital's servers with traffic to bring them down.
    - Cybercriminals: Range from lone actors to large crime organizations. Goal is to steal identities and money. In May 2017 UK hospitals and General Practitioners were hit by an attack freezing systems and encrypting files. When employees tried to access the computers, they were presented with a demand for $300 in bitcoin, a classic ransomware tactic.
    - Nation States: Seeks secrets or intellectual property to help their host nation gain strategic advantage. During early 2000 IRAN was target of very complex and advanced attack aiming to delay nuclear program (Stuxnet). Attackers used vulnerabilities in windows computers and in SCADA systems to cause damages to centrifuges used to separate nuclear material.
    - Insiders: Employee, contractor, supplier or business, partner with system access, steals information or sabotages systems. In court papers filed in a U.S. federal court in New York, prosecutors alleged that Orlando Jemmott, a former emergency department clerk at Kings County Hospital in Brooklyn (owned by of NYC Health + Hospitals) from December 2014 to June 2017, "obtained individually identifiable health information relating ... and disclosed [it] to another person, under false pretenses and with the intent to sell, transfer and use said individually identifiable health information for personal gain
- Accidental threats caused accidentally or through legitimate CI components. In August 2016 Bon Secours Health System announced that patient names, health insurer names and patient ID numbers, Social Security numbers, and some clinical information were exposed over the Internet for a period of almost two months. The incident was caused by the actions of one of its business associates. The partner, a reimbursement optimization firm, inadvertently changed network security settings on its servers when performing maintenance between

[125] Drago, Annarita. Methods and Techniques for Enhancing Physical Security of Critical Infrastructures. *PhD Thesis.* 2015.

[126] *Risk modeling, assessment, and management.* Haimes, Yacov Y. s.l. : Wiley-Interscience, Wiley series in system engineering and management. 0-471-48048-7.

April 18 and April 21. The change meant the data stored on the server could be accessed via the internet by unauthorized individuals.

- Environmental which include natural disasters (floods, earthquakes), human caused disasters (fire and explosions) and failing of supporting infrastructures (power outage or communication loss).

The following table lists the healthcare CPS resources that could be attacked and the impacts of the compromise:

System	Impact of compromise
Management system	Lockout genuine users from system
Lighting	Deactivation of lights may cause safety issues Flickering of lights could cause health issues Increased situational awareness for criminals by activating light remotely Reduced situational awareness for guards/CCTV operator by deactivating light remotely
Access Control	Remote release of secure doors resulting in unauthorized access Deactivation of door release to inconvenience users/force use of green break glass Deactivation of authorised users Addition of unauthorised users Erasure of access logs to cover criminal activity
HVAC	Deactivation of cooling to cause plant/ICT equipment to overheat/shutdown/malfunction Activation of heating to cause plant/ICT equipment to overheat/shutdown/malfunction Deactivation of cooling/heating making normal working/living difficult
Lifts	Denial of service Override lift control
Building Information Modelling (BIM)	Criminals have a greater awareness of where key systems are located and how they are connected and powered
Building /Perimeter Intruder Detection System	Deactivation of system allowing unauthorized access Creating false alarms for distraction Erasure of events records to hide criminal activity

Fire Detection	Cause panic and disruption activating alarm or risk lives by deactivating it
CCTV	Increased situational awareness for intruder to be able to see guard locations and blind spots Ability to turn on off cameras Ability for intruder to erase footage
Communication	Mobile communication jamming Cut off wired communication
Plumbing	Sabotage medical gas distribution Disrupt water management
Monitoring and therapeutic equipment	Gaining unauthorized access to patient's data Putting patient's life in danger
Imaging Equipment	Gaining unauthorized access to patient's data Putting patient's life in danger
High interventionist areas	Putting patient's life in danger
Nurse call system for patients	Disrupt the service
Implantable Medical Devices	Putting patient's life in danger Making unavailable the device, failure to deliver the expected therapies
Wearable Devices	Gaining unauthorized access to patient's data

Table 3. Healthcare CPS Resources

5.1 Examples of cyber security attacks

Studying the attacks performed in recent years, it is easy to determine the most commonly used and effective threats. In the following sections some examples of such cyber-attacks are described.

5.1.1 Advanced Persistent Threat (APT)

APT consists in undercover and permanent computer hacking routines to gain access to an organization's resource. When the fake person reaches access, they often stay unrevealed for a significant lapse of time whereas quietly stealing data, committing fraud, or jeopardizing its reputation.

Although phishing email remains a significant attack mechanism for cybercriminals to mislead employees into downloading malware, there is an increased appetite toward social media platforms. The bulk of the social media scams were manually shared and are lucrative for cybercriminals because of their quick propagation, due to the fact that people are more likely to click on something posted by a friend.

Corporate Account Take Over (CATO)

In a CATO attack, cyber criminals impersonate the business identity and send fraudulent transactions to accounts controlled by them. Institutions under-equipped with limited security controls and safeguards are especially vulnerable to a CATO attack. Impacts stem from this form of cyber-crime could be substantial, and likelihood of recognition for these thefts is still low.

5.1.2 CryptoLocker

CryptoLocker is a type of malware that damages by encrypting data, avoiding access to the data on the infected computers. Once the computer is bugged, cyber criminals request the victims for a payment so their data can be decrypted and recovered. Usually the payment is demanded within three days since the attack through a third-party payment method (i.e. MoneyPak, Bitcoin). Obviously, there is no guarantee that payment brings the promised decryption key.

This malware is typically spread through malicious attachments included in phishing emails and is capable to encrypt files within shared network servers and file shares, USB devices, removable hard disks, and even some cloud storage drives. Furthermore, if one computer on a network becomes infected, mapped network drives might also turn infected.

5.1.3 Distributed Denial of Service

Not only DDoS attacks fill up networks with a massive amount of connection requests, turning them off to deal with legitimate user requests, but DDoS attacks are often used as a smokescreen or camouflage for other types of network intrusions as well, because whereas response team focus on DDoS mitigation, attackers have a greater chance of secretly overtake firewalls to tackle data and financial theft.

5.1.4 Insider and Internal Threats

Any stakeholder can compromise company's security if access has been granted to systems and/or sensitive information. Employee, contractor, supplier, or partner has the chance to harm company assets and prestige, both intentionally or unintentionally, even more with the pervasive trends of Bring Your Own Device (BYOD), cloud-based systems and use of personal USB storage devices.

5.1.5 Physical factors

In a recent attack on Santander Bank in the UK, came into a branch and installed a KVM (keyboard, video, and mouse) switch, a device that enables one computer to remotely control many others by manipulating their keyboards, mice, and video screens.

BYOD truly has consequences. These devices are now part of the firm's ecosystem with no control over them. This is the rationale behind the corporation's consideration about real cost-savings of BYOD trend in opposition of an undermined corporate security. The required extension of corporate security means to protect personally owned devices belongs to a sensitive territory, so mandatory controls as a consequence of employment can be foreseen.

5.1.6 Supply Chain Infiltration

Cybercriminals are continually devising new ways to infiltrate organizations, pretending to be supplier employees to install infected equipment or hardware able to tamper transactions via mobile networks. Trusted providers of software and hardware have been targeted in recent years by cyber criminals seeking to gain physical and technical access to institutions.

5.1.7 Trojan-based attacks

Trojans addressing institutions have become one of the most widespread threats on the internet today. According Symantec report, nearly 95 percent of the raided organizations belong to the financial sector.

5.2 Best practices for Medical and Healthcare domain

The following table summarizes the best practices for medical and healthcare domain.

Practise	Description
Risk assessment and mitigation	Evaluation of the risks associated with healthcare operation based on established methodology like EBIOS Risk Management
Password security	Password must be strong and changed regularly
Software policy	Prevent unauthorized software installation Remove unnecessary software and browser plugins Keep software updated
Access Control	Every actor must be able to access only the information and resources that are necessary for its legitimate purpose Restrict access to physical ports on work machines Restrict access to dubious websites Restrict access to dubious websites Remove or disable unnecessary accounts t
User education and training	Regular training helps remind users about good security and privacy practices, as well as ensuring they are ready to respond appropriately when something goes wrong
Security incident management and response	In an ongoing security event, it is important to manage and orchestrate all involved factors in order to quickly eliminate the threat, minimize impact, gain full recovery and continue all routine operations.
Logging and Auditing	Logs are highly important since they can point for events that need administrator's attention. Such events can be hardware failure, security breach, data leak, malicious activity
Business Continuity Plan (BCP) and Disaster Recovery Plan (DRP)	BCP, or Business continuity plan, referring to all actions need to be taken in order to allow all business operations during and after a disaster. Disaster can be nature hazards such as fire or flood but it can also be data theft, data corruption or data deletion. DRP, or Disaster Recovery Plan, is a detailed technical plan describing all methods and actions that will allow specific business operation to function after disaster occurred. In data theft, corruption or lost example, it is DRP best practice to

	manage consist data backup operation that will allow data restore in case of need
Data Security	Encrypt all the data for maximum security
Documentation	Security and privacy policies need to be fully documented to be compliant with legislation and to be used for mitigating the impact of the disaster
Physical access and network design	Servers should be located in locked rooms. For critical areas you might even want to consider installing cameras. Choose operating systems designed with enterprise-level security. Linux is a great choice for workstations and servers because it provides superb access control. Install firewall software to protect your network from external internet threats; better still, invest in a commercial-grade hardware firewall. Wireless networks should employ the latest encryption standards and if they can't, consider upgrading to something current so you get the manufacturer's security updates.

Table 4. Best practices for Medical and Healthcare domain

6 Ontologies and Impact Propagation Models

With the development of the Internet and the opportunities it provides for business and organizations, there is more and more dependency on Internet-based technologies for critical operation. This led to the mix of network infrastructure, physical hardware, software, and human operations and increased the number and the variety of cyber vulnerabilities. It is more than ever necessary to implement reactive and proactive security approaches which are adapted to their current organization context. This is even more applicable to OES (Operators of Essential Services) like hospitals and industrial organizations whose technological infrastructure is built around CPS (Cyber-Physical Systems). To limit the damage and reduce the cost of recovery, it is essential to provide solutions able to rapidly recognize, analyze, and respond to an incident and to anticipate the future cyber-attacks will limit the damage and lower the cost of recovery. This explains the plethora of research works dealing with information system security in general and more particularly, these last years, with cyber-physical security.

It this section, we present a survey on cyber-physical risk modeling and impact propagation.

6.1 Cyber-physical risk and vulnerability models

We distinguish process-oriented approaches and model-oriented approaches. Within model-oriented approaches, we'll make a special focus on ontology-based models.

These approaches rely on models aiming to allow predicting and analyzing how vulnerabilities can be exploited by an attacker. Several formalisms and representations to model attacks and their dependencies are proposed in literature.

6.1.1 Process-oriented approaches

One of the methods used by organizations to face increasing cyber-attacks is security governance. These approaches take a lifecycle perspective and consider security at a high level of risk assessment. They rely on security requirements and provide descriptions on activities that should be performed and the flow of their execution. In Nicho M.[127], the author presented an Information Security Governance process (ISG). This model help selecting relevant IS security and governance frameworks (technical as well as non-technical) and mapping relevant IT controls upon ISG frameworks and standards. In such systems, it is essential to provide the system with validation and compliance functions or mechanisms to ensure the fulfillment of security policies and requirements. In Whitman *et al.*[128] authors define processes, providing governance approach to system resilience, as well as handling and recovering from incidents. In the adoption of a governance approach, we should consider the three levels of organizational structure to ensure the protection of resources including the medical devices and associated networked technologies. It is necessary to guarantee compliance with regulation, policy development, and business process at the strategic level. At the tactical level, proactive

[127] Nicho, M. (2018). A process model for implementing information systems security governance. *Information & Computer Security*, *26*(1), 10-38.

[128] Whitman M, Mattord H. Management of Information Security.3rd ed. Boston, Mass: Course Technology, Cengage Learning; 2010.

approaches to risk management, auditing, education, and contingency planning are needed. Finally, at the day-to-day operational level, everyday practices such as implementing technical controls and using suitable processes and workflows can ensure that mitigation is effective.

Once the secure business process has been established, process compliance could be an effective way to ensure security. Authors in Fellmann *et al.*[129] present a survey on business process compliance. Moreover, in Hashmi *et al.*[130], authors distinguish three process compliance strategies: design time, runtime, and auditing. Regarding rules-based process compliance, authors of Yip *et al.*[131] suggest modeling compliance knowledge using semantic web rules (SWRL) and ontology (OWL). Based on an existing work, the authors defined a semantic framework for intelligent compliance management including domain ontology. They illustrated their proposal on an information security domain fragment and defined compliance requirements as semantic web rules.

6.1.2 Ontology-based cyber and physical vulnerabilities, risks, and attacks management

Unlike process-oriented approaches, model-based approaches rely on a formal or semi-formal representation of the assets and their related risks. The underlying models could be graphical or non-graphical. In this section, our goal is to investigate existing works that has contributed to the modeling of physical and/or cyber vulnerabilities, risks and attacks in cyber and/or physical systems. The efforts dedicated to model security-related concepts have produced significant contributions. Indeed, existing works provide ontologies that model the human health risks[132] or for the use of IoT in the field of medicine[133]. On the other hand, there are some works that have proposed ontology-based approaches for modeling vulnerabilities and risks in Critical Infrastructure (CI) and in Cyber and Physical Systems (CPS). Among these contributions, few works have dealt with the representation of security concepts in the context of healthcare information systems in which, the CPS are part of the technological infrastructure. Based on the modeled security breach in cyber-physical systems, existing contributions can be classified to three major categories: risk & threat modeling, vulnerability modeling, and attacks & incident modeling.

6.1.2.1 *Cyber-physical risk and threat modeling*

McKone *et al.* and Abinaya *et al.* provide ontology-based representation of human health risks, for health risk assessment respectively. Other contributions, not necessarily related to the healthcare field, exist. We can mention the technical report proposed by Cebula *et al.*[134]. In this report, the authors proposed a rich taxonomy of operational cyber risk that attempts to identify

[129] Fellmann, M., & Zasada, A. (2014). State-of-the-art of business process compliance approaches.

[130] Hashmi, M., Governatori, G., Lam, H. P., & Wynn, M. T. (2018). Are we done with business process compliance: state of the art and challenges ahead. *Knowledge and Information Systems*, 1-55.

[131] Yip, F., Wong, A. K. Y., Parameswaran, N., & Ray, P. (2007, October). Rules and ontology in compliance management. In *11th IEEE International Enterprise Distributed Object Computing Conference (EDOC 2007)* (pp. 435-435). IEEE.

[132] McKone, T. E., & Feng, L. (2015). Building a human health risk assessment ontology (RsO): A proposed framework. *Risk analysis*, *35*(11), 2087-2101.

[133] Abinaya, Kumar, V., Swathika. "Ontology based public healthcare system in Internet of Things (IoT)." *Procedia Computer Science* 50 (2015): 99-102.

[134] James Cebula and Lisa Young. A taxonomy of Operational Cyber Security Risks. Technical note CMU/SEI-2010-TN-028. December 2010

and organize the sources of cyber risks. In Trucco *et al.*[135], the authors propose an ontology of hazards and threats that could affect a critical infrastructure. This ontology covers four sectors: energy, transport, water and telecommunication. Other works are more general since they propose model for risk representation[136].

6.1.2.2 *Cyber-physical vulnerability modeling*

Our review of the literature has identified the work of (Syed et al. 2016) where the UCO ontology is described (Unified Cyber security Ontology). This ontology unifies most commonly used cyber security standards. A more general ontology, called VDO ontology is presented in (Booth et al., 2016). It is promoted by the NIST institute and it allows characterizing vulnerabilities. On the other hand, we can cite commonly known knowledge bases such as CVE[137] and CWE[138]. These knowledge bases are structured as light ontologies.

6.1.2.3 *Cyber-physical attack and incident modeling*

Existing works propose ontology-based frameworks that model the different attacks and/or the relationship between cyber and physical systems. Authors in Cheh *et al.*[139] propose ontology for cyber-physical systems and attack steps that model the physical consequences of actions. In Yanambaka Venkata *et al.*[140], the authors propose an ontology-driven framework that captures the relationship between cyber and physical systems to semantically reason about the impact of cyber-attacks on the physical systems. Neuman and co-authors in Neuman *et al.*[141] discuss architectural approaches for decomposing the smart grid into protection domain. They also suggest approaches to model and mitigate the impact of threats. Camera *et al.* propose taxonomy of adverse events related to medical devices. In Miller[142] a taxonomy for classifying security incident that focuses on the cross domain and impact oriented analysis is presented. In Meng *et al.*[143] a detection model for events occurring in cyber physical systems is presented. In this same

---

[135] Trucco, P., Petrenj, B., Bouchon, S., & Mauro, C. D. (2016). Ontology-based approach to disruption scenario generation for critical infrastructure systems. International Journal of Critical Infrastructures, 12(3), 248-272.

[136] Vivek Agrawal: A Comparative Study on Information Security Risk Analysis Methods. JCP 12(1): 57-67 (2017)

[137] https://cve.mitre.org/

[138] https://cwe.mitre.org/

[139] Cheh, C., Keefe, K., Feddersen, B., Chen, B., Temple, W. G., & Sanders, W. H. (2017, November). Developing Models for Physical Attacks in Cyber-Physical Systems. In *Proceedings of the 2017 Workshop on Cyber-Physical Systems Security and PrivaCy* (pp. 49-55). ACM.

[140] Yanambaka Venkata R., Kamongi, P., Kavi, K. (2018). An Ontology-Driven Framework for Security and Resiliency in Cyber Physical Systems.

[141] Clifford Neuman, Kymie Tan: Mediating cyber and physical threat propagation in secure smart grid architectures. SmartGridComm 2011: 238-243

[142] Classifying and Cataloging Cyber-Security Incidents Within Cyber-Physical Systems. PHD thesis. https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=5344&context=etd. 2014.

[143] Meng Ma, Ling Liu, Yangxin Lin, Disheng Pan, Ping Wang: Event Description and Detection in Cyber-Physical Systems: An Ontology-Based Language and Approach. ICPADS 2017: 1-8

model events are described. Finally, the CAPEC[144] knowledge base reports attack patterns in the context of cyber security.

6.2 Impact propagation models

We have classified the approaches into structure-based and behavior-based approaches.

6.2.1 Structure-based approaches

Several formalisms have been proposed to model the concepts and their interdependencies. Both attacker and defender points of view are considered through the concepts of attack and impact. In Schneier[145] *attack trees* formalism is presented. It is a graphical formalism to structure, model and analyze the potential attacks on an asset. It is based on an incremental refinement of a security threat to represent an attacker's goal. Attack-defense trees [146] extend attack trees with defensive measures, also called countermeasures. It also allows modeling interaction between attackers and defenders.

Impact dependency graphs used by the author in Jakobson[147] propose a model of a cyber-attack based on an extended conceptual graph of impact dependencies. It aims to model a multi-level information structure containing assets, services, and their interdependencies. It defines an approach on how to infer plausible future cyber security situations to fulfill cyberspace situational awareness goals. The objective of these awareness goals is to ensure the resilience of systems. Following the same objective, author in Lei[148] propose to use impact graphs to illustrate how a mission can be affected by a compromised software asset step by step.

6.2.2 6.2.2 Behavior-based approaches

In Ten et al., 2008, the authors used Stochastic Petri nets to model the behavior of cyber-physical systems and to analyze the effect of intrusion detection and to assess vulnerabilities in SCADA systems. More recently, authors in Szpyrka & Jasiul[149] defined, on the basis of colored Petri nets, a new class called propagation nets. This class provides a formal model for risk propagation. The proposed method allows for model relations between nodes forming the network structure. In Genge *et al.*[150] the authors propose a behavioral assessment of physical processes. The proposed methodology is inspired by system dynamics research and the sensitivity analysis. The evaluation method calculates the covariance of observed variables before and after performing individual

[144] https://capec.mitre.org/index.html

[145] Schneier, B. (1999). Attack trees. Dr. Dobb's journal, 24(12), 21-29.

[146] Kordy, B., Mauw, S., Radomirović, S., & Schweitzer, P. (2014). Attack–defense trees. Journal of Logic and Computation, 24(1), 55-87.

[147] Jakobson, G. (2013, June). Mission-centricity in cyber security: Architecting cyber-attack resilient missions. In *Cyber Conflict (CyCon), 2013 5th International Conference on* (pp. 1-18). IEEE.

[148] Cyber situational awareness and mission-centric resilient cyber defense. In Computer Science and Network Technology (ICCSNT), 2015 4th International Conference on (Vol. 1, pp. 1218-1225). IEEE.

[149] Szpyrka, M., & Jasiul, B. (2017). Evaluation of Cyber Security and Modelling of Risk Propagation with Petri Nets. *Symmetry*, *9*(3), 32.

[150] Genge, B., Kiss, I., & Haller, P. (2015). A system dynamics approach for assessing the impact of cyber attacks on critical infrastructures. *International Journal of Critical Infrastructure Protection*, *10*, 3-17.

attacks against control variables. One main feature of this method is its applicability in situations where the physical process is not available.

An important aspect that is not always considered in cyber-physical systems is the impact of human interaction. In Fan *et al.*[151], authors presented a platform and associated methodology to effectively generate accident scenarios by modeling human machine interaction errors using model-level fault injection, followed by simulation to produce dynamic evolution of accident scenarios. Results show that human mode confusion triggered by false displays may lead to severe accidents.

6.2.3 Risk mitigation in Cyber-Physical Systems

Mitigation is related to how to recover a safe state in safety-critical systems. It is related to the resilience of the system. While risk assessment goal is situational awareness and diagnostics, resilience is taking one step forward while taking quick actions to maintain critical system functionality via remedial action schemes[152]. In a survey on remedial methods presented in Combita *et al.*[153], the authors identify different trends on automatic attack detection and response such as preventive and reactive responses. Preventive response takes place when vulnerabilities in a CPS have been identified. As a consequence, the system structure can be modified in order to increase the system resiliency to attacks. Most of the existing related works propose solutions in the power systems field. Nevertheless, there are some recent contributions in healthcare cyber-physical system domain. In Rao *et al.*[154], authors present a framework for threat detection and mitigation during deployment of medical devices. The solution is demonstrated through a smart-connected-pacemaker scenario. The proposed framework relies on a multimodal approach where modes are abstractions of functionalities ordered in a decreasing order of functional criticality. Mode 0 would, for example, contain functions needed to keep the device working or, in other words, keep the patient alive.

To conclude on this part of the state of the art, risk management is a complex problem with a variety of facets. A lot of work has been done for Cyber-Physical Systems. However, healthcare domain has specific needs and requirements that need to be elicited and analyzed to finally define a suitable solution. These are some of the challenges of the Safecare project.

---

[151] Fan, C. F., Chan, C. C., Yu, H. Y., & Yih, S. (2018). A simulation platform for human-machine interaction safety analysis of cyber-physical systems. *International Journal of Industrial Ergonomics*, *68*, 89-100.

[152] Arghandeh, R., Von Meier, A., Mehrmanesh, L., & Mili, L. (2016). On the definition of cyber-physical resilience in power systems. *Renewable and Sustainable Energy Reviews*, *58*, 1060-1069.

[153] Cómbita, L. F., Giraldo, J., Cárdenas, A. A., & Quijano, N. (2015, October). Response and reconfiguration of cyber-physical control systems: A survey. In *Automatic Control (CCAC), 2015 IEEE 2nd Colombian Conference on* (pp. 1-6). IEEE.

[154] Rao, A., Carreón, N., Lysecky, R., & Rozenblit, J. (2018). Probabilistic Threat

7 Crisis Management

7.1 Critical infrastructure

Critical infrastructures are defined by the EU as "…is an asset or system which is essential for the maintenance of vital societal functions. The damage to a critical infrastructure, its destruction or disruption by natural disasters, terrorism, criminal activity or malicious behavior, may have a significant negative impact for the security of the EU and the well-being of its citizens."

The criticality of these infrastructures results, among other things, from the high complexity and interdependence of physical and cyber physical systems, resulting in high vulnerability to disruptions and threats[155,156]. The risk potential is particularly high for the healthcare sector and especially for hospitals, as vulnerable groups of people are particularly affected. Accordingly, many scenarios are conceivable that represent threat situations and can lead to a crisis or catastrophe[157]. Measures to prevent or manage such crises at the time they occur are part of crisis management.

7.2 Crisis management

Starting from the natural course of a crisis, various approaches to systematizing can be found in the literature. For example, cycle models distinguish between various impacts, e.g. prevention, planning, acute reaction, recovery and learning[158,159,160]. However, with regard to this phase it must be critically questioned to what extent preparation is possible if, on the one hand, there is an almost unlimited number of conceivable scenarios and on the other hand the origins of the crisis are not controllable (e.g. terrorism)[151]. Moreover, traditional response processes and structures are no longer adequate due to the complexity of crises. Since the SafeCare project also deals with complex threat situations with high dynamics and cascading effects, a temporal classification is used on the first level of systematization and the various effects only on the second level.

7.2.1 Relevant aspects in the context of crisis management

Regardless of the type of crisis or catastrophe (e.g. Hurricane Katrina, 2005 or Fukushima, 2011), overarching aspects can be identified. These include communication, coordination and

[155] Boin, A. The new world of crises and crisis management: Implications for policymaking and research. *Review of Policy Research.* 2009, 26(4), pp. 367-377.

[156] Lagadec, P. A new cosmology of risks and crises: Time for a radical shift in paradigm and practice. *Review of Policy Research.* 2009, 26(4), pp. 473–486.

[157] Boin, A., & McConnell, A. Preparing for Critical Infrastructure Breakdowns: The Limits of Crisis Management and the Need for Resilience. *Journal of Contingencies and Crisis Management.* 2007, 15(1), pp. 50–59.

[158] Fink, S. *Crisis Management: Planning for the Inevitable.* Lincoln, NE : iUniverse, 2002.

[159] Regster, M. & Larkin, J. *Risk Issues and Crisis Management: A Casebook of Best Practice.* London : Kogan Page, 2002. Vol. 2nd Edition.

[160] Curtin, T., Hayman, D. & Husein, N. *Managing Crisis: A Practical Guide.* Houndmills, Basingstoke, Hampshire : Palgrave Macmillan, 2005.

cooperation within and between participating organizations and institutions, contingency planning and communication with the population[161,162].

Furthermore, there are specific aspects of crisis management for the time phases[163,164].

7.2.1.1 *Before an event*

The focus here is on preparing for possible threat situations in order to minimize unavoidable negative effects. To this end, contingency plans are being developed to minimize damage to people and the environment[165]. Following FEMA (2006), preparation means a continuous process in which the actors involved jointly identify threats, vulnerabilities and available resources. An important aspect of the preparation is to establish trusting relationships between the individual actors, media representatives, external stakeholders and experts[166,167]. Following Boin and 't Hart[162], the quality of communication, coordination, and cooperation between, above, and beyond the emergency forces is decisive for the quality of crisis management.

The development of contingency plans is also an important aspect of crisis management in the run-up to a crisis. Based on these, exercises or simulations can be held[168]. As these are only prepared very generally to prepare for all eventualities, there is a lack of recommendations for action and behaviour for one-off and unforeseen events[164]. It is also important to involve external actors in the planning process, such as private critical infrastructure operators, local businesses and municipalities[164]. In this way, partnerships can also be established at the same time in order to be able to react as a community in the event of a crisis.

---

[161] Boin, A., & McConnell, A. Preparing for Critical Infrastructure Breakdowns: The Limits of Crisis Management and the Need for Resilience. *Journal of Contingencies and Crisis Management.* 2007, 15(1), pp. 50–59.

[162] Gheytanchi, A., Joseph, L., Gierlach, E., Kimpara, S., Housley, J., Franco, Z. E., & Beutler, L. E. The dirty dozen: twelve failures of the hurricane katrina response and how psychology can help. *The American Psychologist.* 2007, 62(2), pp. 118–130.

[163] Kahan, J. H., Allen, A. C. & George, J. K. An operational framework for resilience. *Journal of Homeland Security and Emergency Management.* 2009, 6(1).

[164] Wyche, K., Pfefferbaum, R., Pfefferbaum, B. & Norris, F. Exploring community resilience in workforce communities of first responders serving Katrina survivors. *American Journal of Othopsychiatry.* 2011, 81(1), pp. 18-30.

[165] Baird, M. E. *The " Phases " of Emergency Management.* 2010.

[166] Boin, A. & ´t Hart, P. Organising for Effective Emergency Management: Lessons from Research. *Australian Journal of Public Administration.* 2010, 69(4), pp. 357–371.

[167] Seeger, M. W. Best Practices in Crisis Communication: An Expert Panel Process. *Journal of Applied Communication Research.* 2006, 34(3), pp. 232–244.

[168] Boin, A., & McConnell, A. Preparing for Critical Infrastructure Breakdowns: The Limits of Crisis Management and the Need for Resilience. *Journal of Contingencies and Crisis Management.* 2007, 15(1), pp. 50–59.

An important role is also played by the joint exercises of all actors involved in an emergency[169,170]. Under realistic conditions, however, such exercises are very costly and time-consuming. Therefore computer-based simulations are increasingly being used[171].

7.2.1.2 *During an event*

- Intra- and interorganizational communication, coordination and cooperation

Good communication is essential for the adequate and timely provision and dissemination of information. This applies to both intra- and interorganizational communication[172,173,174]. The goal of all actors must be the development of a "common operational picture"[166] or "shared situational awareness"[175]. Only in this way the necessary measures can be effectively coordinated[176] and the best possible decisions be made[177,178,171].

- Decision making

In the context of communication, coordination and cooperation, managers must also be considered as decision-makers. These are confronted with conflicting, diverse and dynamic requirements during a crisis situation. More specifically their work is characterized by

[169] Boin, A., & McConnell, A. Preparing for Critical Infrastructure Breakdowns: The Limits of Crisis Management and the Need for Resilience. *Journal of Contingencies and Crisis Management.* 2007, 15(1), pp. 50–59.

[170] Boin, A. The new world of crises and crisis management: Implications for policymaking and research. *Review of Policy Research.* 2009, 26(4), pp. 367-377.

[171] Quillinan, T. B., Bazier, F., Aldewereld, H., Dignum, F., Dignum, V., Penserini, L. & Wijngaards, N. Developing Agent-based Organizational Models for Crisis. *Proceeding of 8th International Conference on Autonomous Agents and Multiagent Systems .* 2009.

[172] Ansell, C., Boin, A. & Keller, A. Managing transboundary crises: identifying the building blocks of an effective response system. *Journal of Contingencies and Crisis Management.* 2010, 18(4), pp. 195-207.

[173] Okros, A., Verdun, J. & Chouinard, P. *The meta-organization: Research and conceptual landscape.* Ottawa, Canada : Defence Research and Development Canada and the Centre for Security Sciences, 2011.

[174] Wyche, K., Pfefferbaum, R., Pfefferbaum, B. & Norris, F. Exploring community resilience in workforce communities of first responders serving Katrina survivors. *American Journal of Othopsychiatry.* 2011, 81(1), pp. 18-30.

[175] Janssen, M., Lee, J., Bharosa, N. & Cresswell, A. Advances in multi-agency disaster management: Key elements in disaster research. *Information Systems Frontiers.* 2009, 12(1), pp. 1-7.

[176] Comfort, K., Ko, A. & Zagorecki, A. Coordination in rapidly evolving disaster response systems: the role of information. *American Behavioral Scientist.* 2004, 48(3), pp. 295–325.

[177] Palttala, P., Boano, C. Lund, R. & Vos, M. Communication Gaps in Disaster Management: Perceptions by Experts from Governmental and Non-Governmental Organizations. *Journal of Contingencies and Crisis Management.* 2012, 20(1), pp. 2-12.

[178] Comfort, L. K. & Haase, T. W. Communication, coherence and collective action: the impact of hurricane Katrina on communications infrastructure. *Public Works Management Policy.* 2006, 10(4), pp. 328-343.

changes in urgency, scope, impact, type of suitable emergency forces and varying information and communication requirements of the emergency forces[179,180,181].

- Crisis communication with the population

Crisis communication with the general population is another focal point, as it can have a decisive influence on the behaviour of the population affected by the crisis[182,183]. This is particularly important when an immediate response is required, such as in the event of an evacuation or sheltering in place[184]. Whether the desired behaviour is also shown by the population depends on the content and design of the concrete warning message. In this respect, established good practice and guidelines can be used[185,178.186,179].

- Event-related factors

Further influencing factors on the concrete implementation of crisis management result from general conditions. These include, for example, the speed at which the event occurs, the time of year or the weather conditions[187]. If a breakdown of other critical infrastructures occurred, such as the power grid, it will not only have a direct influence on the behaviour of the general population and the emergency services but can also trigger or intensify cascading effects[188].

[179] Hofinger, G., Zinke, R. & Künzer, L. Psychological Requirements for Crisis and Emergency Decision-Support Systems for Public TransportControl Centers. *Proceedings of the 8th International ISCRAM Conference - Lisbon, Portugal, May 2011.* 2011, pp. 1-5.

[180] Janssen, M., Lee, J., Bharosa, N. & Cresswell, A. Advances in multi-agency disaster management: Key elements in disaster research. *Information Systems Frontiers.* 2009, 12(1), pp. 1-7.

[181] Paton, D. & Flin, R. Disaster stress: an emergency management perspective. *Disaster Prevention and Management.* 1999, 8(4), pp. 261–267.

[182] Seeger, M. W. Best Practices in Crisis Communication: An Expert Panel Process. *Journal of Applied Communication Research.* 2006, 34(3), pp. 232–244.

[183] Prevention, U.S. Department of Health and Human Servics - Centers for Disease Control and. *Crisis and Emergency Risk Communication: 2014 Edition.* 2014.

[184] Chandan, S., Saha, S. & Barrett, C. Modeling the Interaction between Emergency Communications and Behavior in the Aftermath of a Disaster. *Social Computing, Behavioral-Cultural Modeling and Prediction.* 2013.

[185] Palttala, P., Boano, C. Lund, R. & Vos, M. Communication Gaps in Disaster Management: Perceptions by Experts from Governmental and Non-Governmental Organizations. *Journal of Contingencies and Crisis Management.* 2012, 20(1), pp. 2-12.

[186] Sellnow, T. L. & Seeger, M. W. *Theorizing crisis communication. Foundations of communication theory.* Chichester, West Sussex : Wiley-Blackwell, 2013.

[187] Arvidsson, B. *Development of a method for studying cascading effects between critical infrastructures (CascEff project).* 2015.

[188] Petermann, T., Bradke, H., Lüllmann, A., Poetzsch, M. & Riehm, U. *What Happens During a Blackout.* 2011.

7.2.1.3 *After an event*

- Lesson learned

As explained above, no crisis is like another, so a systematic and critical reflection in the aftermath represents a unique opportunity to learn from past experiences and implement the findings in mitigation actions to improve both prevention and preparedness[189,190,191].

7.3 Potential societal impacts

Within the critical infrastructure, health care facilities and especially hospitals are of particular importance. The fulfilment of their health care mandate is vital for the citizens and therefore has the highest priority. At the same time, however, this special position makes them a target for physical and cyber physical attacks, which results in the special need for protection of such facilities[192].

[189] Baird, M. E. *The " Phases " of Emergency Management.* 2010.

[190] Gheytanchi, A., Joseph, L., Gierlach, E., Kimpara, S., Housley, J., Franco, Z. E., & Beutler, L. E. The dirty dozen: twelve failures of the hurricane katrina response and how psychology can help. *The American Psychologist.* 2007, 62(2), pp. 118–130.

[191] Robinson, L. Proceedings of the Workshop on Preparing for and Responding to Disasters in North America. *Homeland Security Affairs.* 2006, 1.

[192] McDaniels, T., Chang, S., Peterson, K., Mikawoz, J. & Reed, D. Empirical framework for characterizing infrastructure failure interdependencies. *American Society of Civil Engineers.* 2007, 13(3), pp. 175-184.

8 Ethics, Privacy and Data Protection

SAFECARE project aims to address physical and cyber aspects of security in order to globally improve resilience of health services. In that regard, an important category of critical assets deserving particular protection is medical devices, whose vulnerabilities are to be regarded as crucial factors able to increase the likelihood of cyber-attack events. Taking these risks into account, there was a need to address it specifically in distinct regulations.

This Section outlines the state of art concerning privacy and data protection (§8.1), critical infrastructures (§8.2) and medical devices (§8.3) legislations. In the first sub-section, the General Data Protection Regulation will be addressed focussing on the measures related to health data. In the second sub-section, the directive on network and information systems (NIS directive) will be examined, taking into account the national implementation measures. In the third section, attention will be drawn on the key actual and upcoming legislative acts as well as their main points of interest for the purposes of SAFECARE.[193] Due to the limited volume of this deliverable, a detailed attention could not be drawn on the peculiarities in each national system.

8.1 Privacy and Data Protection Regulation

The key point of legislation in the field of data protection law is represented by the General Data Protection Regulation (GDPR).[194] Besides data protection aspects, attention has also to be drawn to privacy matters and the case law of the European Court on Human Rights (ECtHR). Data protection is often referred to as the means to protect and control the data collected, whereas privacy would emphasise the importance of an area of someone's life that has to remain private.[195] The ECtHR consistently refers to the fact that "the protection of personal data, not least medical data, is of fundamental importance to a person's enjoyment of his or her right to respect for private and family life".[196]

8.1.1 Scope of the GDPR: material / territorial / personal

The GDPR defines personal data as "any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly […]".[197] By processing personal data, the GDPR refers to "any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation

[193] A more complete overview concerning the legal and ethical framework concerning Privacy, data protection and confidentiality is provided in D3.9 – 'Analysis of ethics, privacy and confidentiality constraints'.

[194] Regulation (EU) 2016/679 of the European Parliament and Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC. (which became directly applicable in all EU Member States as of 25th of May 2018).

[195] N. Purtova, *Property Rights in Personal Data, A European perspective* (Wolters Kluwer 2011) 226-227

[196] Z v. Finland, 25/02/1997, § 95; M.S. v. Sweden, 27/08/1997, § 41; L.L. v. France, 10/10/2006, § 44; I v. Finland 17/07/2008, § 38; Biriuk v. Lithuania, 25/11/2008, § 39; Armoniene v. Lithuania, 25/11/2008, § 40; Szuluk v. The United Kingdom, 02/06/2009, § 47; L.H. v. Latvia, 29/04/2014, § 56.

[197] Art. 4(1) GDPR.

or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction".[198]

The GDPR applies to controllers and processors.[199] In a synthetic way, the controller is defined as a "natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data".[200] A processor is defined as "a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller.[201]

The GDPR applies when "processing of personal data in the context of the activities of an establishment of a controller or a processor in the Union, regardless of whether the processing takes place in the Union or not",[202] and when "processing of personal data of data subjects who are in the Union by a controller or processor not established in the Union, where the processing activities are related to offering of goods and services to data subjects in EU; monitoring of behaviour of data subjects in the EU".[203] Concerning the question of the activities, the Court of Justice of the European Union (CJEU) ruled that: "that provision requires the processing of personal data in question to be carried out not 'by' the establishment concerned itself but only 'in the context of the activities' of the establishment'.[204] This means that the criterion is a large one. Even if the processing of the data is not realized by a healthcare provider and is carried out outside the European Union, the GDPR may apply.[205]

8.1.2 Stronger guarantees for health data

The guarantees for individuals are stronger when health data is at stake. Art. 9 GDPR states that the processing of such personal information is prohibited unless an exception applies. Herein below are outlined the main exceptions foreseen by Art. 9 GDPR:

- Concerning the **consent**: it has to be explicit (Art. 9(a) GDPR). On the other hand, consent is not required if it is impossible to retrieve it from a data subject who is physically or legally incapable of giving consent and when there is a need to protect vital interests of the data subject (Art. 9(c) GDPR). Furthermore, explicit consent is not required if the data is made public by the data subject (Art. 9(e) GDPR).
- Particular field of **employment / social security / social protection law**: "Processing is necessary for the purposes of carrying out the obligations and exercising specific rights of the controller or of the data subject […] in so far as it is authorised by Union or Member State law or a collective agreement" (Art. 9(b) GDPR).

[198] Art. 4(2) GDPR.
[199] See L. Feiler, N. Forgó, M. Weigl, *The EU General Data protection Regulation (GDPR): A commentary* (German Law Publishers 2018) 14.
[200] Art. 4(7) GDPR.
[201] Art. 4(8) GDPR.
[202] Art. 3(1) GDPR.
[203] Art. 3 (2) GDPR.
[204] CJEU, Verein für Konsumenteninformation v Amazon EU Sàrl, 28 July 2016, C-191/15, § 78.
[205] In the same sense, see L. Feiler, N. Forgó, M. Weigl, *The EU General Data protection Regulation (GDPR): A commentary* (German Law Publishers 2018) 16.

- Processing is necessary for the purposes of **preventive or occupational medicine**, for the **assessment of the working capacity** of the employee, **medical diagnosis**, the **provision of health or social care** or **treatment** or the **management of health or social care systems** and services on the basis of Union or Member State law or pursuant to contract with a health professional and subject to the conditions and safeguards referred to in paragraph 3 (h); with a need to be 'subject to the obligation of professional secrecy',[206] as defined within the relevant national legislations.
- processing is necessary for **reasons of public interest in the area of public health** (Art. 9(i) GDPR).

8.2 Protection of Critical Infrastructures

The main piece of European legislation in this field is the NIS directive which aims at enhancing cyber security of critical infrastructures.[207] The directive applies to network and information systems, including electronic communications networks or interconnected devices performing automatic processing of digital data.[208] These measures are transposed by national parliaments in order to be effective.[209]

The definition of network and information systems refers to electronic communications network as defined by Directive 2002/21/EC. This network is defined as follows: 'transmission systems and, where applicable, switching or routing equipment and other resources which permit the conveyance of signals by wire, by radio, by optical or by other electromagnetic means, including satellite networks, fixed (circuit – and packet-switched, including Internet) and mobile terrestrial networks, electricity cable systems, to the extent that they are used for the purpose of transmitting signals, networks used for radio and television broadcasting, and cable television networks, irrespective of the type of information conveyed'.[210] Interconnected devices relate to the internet of medical things. Security risks are for example related to 'data loss or unauthorized access to data'.[211] It applies to the healthcare sector, as it is considered by the directive an 'operator of essential services'.[212] Nevertheless, it has to be checked in the legislation of each Member State whether they consider the hospitals to fall under this category or not.

The three main aims of this directive are: (1) To ensure that the Member States are prepared to face cyber-security threats through a Computer Security Incident Response Team (CSIRT)[213] and

[206] Art. 9(3) GDPR.

[207] Directive (EU) 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union. Hereafter referred to as the NIS directive.

[208] Art. 4(1) NIS Directive.

[209] A regulation of the European Union produces immediately its effect in all Member States; a directive has to be transposed by each member state through their own parliament in order to be enforceable.

[210] Art 2(a) of Directive 2002/21/EC of the European Parliament and of the Council of 7 March 2002 on a common regulatory framework for electronic communications networks and services (Framework Directive).

[211] C. Storr, P. Storr, Internet of Things: Right to Data from a European Perspective, in M. Corrales, M. Fenwick, N. Forgó, *New Technologies, Big Data and the Law* (Springer 2017) 87.

[212] [212] Art 4 (4) NIS Directive and Annex II, point 5.

[213] Art. 12 NIS Directive.

a competent national network and information systems (NIS) authority; (2) Cooperation between Member States in this field; (3) Developing a culture of security in the concerned sectors. Attention is drawn on the 'requirement to notify security incidents' to these aforementioned authorities.[214]

Concerning the extent and degree of security mechanisms, the directive is not precise. It refers to the state of the art and the risk posed.[215] The state of the art refers to the current good practice in the field but this concept is not well defined.[216]

- **France**

Security of the information systems are defined by the French legislator as the ability to resist, at a certain confidence level, to actions compromising the availability, authenticity, integrity or confidentiality of the stored data, communicated or processed, and the related services offered by or accessed through these information systems.[217] Fines relating to the non-compliance with these provisions range from 50.000 to 100.000€.[218]

The French legislator listed 23 rules concerning the security for 'operator of essential services'.[219] These rules will have to be complied with from the moment when the 'operator of essential services' –e.g. hospitals – are categorized as such by a Prime Minister's decree ("arrêté du premier ministre"). This classification will be notified to the concerned parties.[220]

The first measure that will have to be complied with concerns the treatment of the alerts concerning information security systems.

These security rules will only be listed in this document while the specific focus will be put on Rule 22.[221] It concerns the treatment of the alerts. The legislative provision foresees that the hospital has to hold a service responsible for the exchange of information with the national

[214] J. César, J. Debussche, 'Novel EU Legal requirements in Big Data security, Big Data – Big Security Headaches', 8 J. Intell. Prop. Info. Tech. & Elec. Com. L. 79 (2017) 82; see also articles 14(3) and 16(3) of the NIS directive.

[215] Art 14(1) NIS Directive.

[216] Trend micro study finds 'state of the art' GDPR rule confuses businesses. Dow Jones Institutional News, 6th November 2017 (Available on ProQuest).

[217] Art. 1. LOI n° 2018-133 du 26 février 2018 portant diverses dispositions d'adaptation au droit de l'Union européenne dans le domaine de la sécurité ; Art 5; Annex 1 Décret n° 2018-384 du 23 mai 2018 relatif à la sécurité des réseaux et systèmes d'information des opérateurs de services essentiels et des fournisseurs de service numérique, PRESTATAIRES DE SOINS DE SANTÉ → Service concourant aux activités de prévention, de diagnostic ou de soins; PRESTATAIRES FOURNISSANT UN SERVICE D'AIDE MÉDICALE D'URGENCE → Réception et régulation des appels Service mobile d'urgence et réanimation.

[218] Ibid, Art. 15.

[219] Art. 4(4) NIS Directive, Annex 1 Arrêté du 14 septembre 2018 fixant les règles de sécurité et les délais mentionnés à l'article 10 du décret n° 2018-384 du 23 mai 2018 relatif à la sécurité des réseaux et systèmes d'information des opérateurs de services essentiels et des fournisseurs de service numérique

[220] Art. 3 Décret n° 2018-384 du 23 mai 2018 relatif à la sécurité des réseaux et systèmes d'information des opérateurs de services essentiels et des fournisseurs de service numérique.

[221] Arrêté du 14 septembre 2018 fixant les règles de sécurité et les délais mentionnés à l'article 10 du décret n° 2018-384 du 23 mai 2018 relatif à la sécurité des réseaux et systèmes d'information des opérateurs de services essentiels et des fournisseurs de service numérique, Annex 1, Rule 22.

agency involved with the security of information systems and take the appropriate measures. The hospital will also have to share information related to this service (name, phone and email contacts).

The other security rules are as follows: 1- Risk analysis; 2- Security policy; 3- Security homologation; 4- Indicators; 5- Security audits; 6- Cartography; 7- Configuration; 8- Partitioning; 9- Remote access; 10- Filtering; 11- Administrator accounts; 12- Administration information's systems; 13- Identification; 14- Authentication; 15- Access rights; 16- Procedure relative to security conditions; 17- Physical and environmental security; 18- Detection; 19- Logging; 20- Correlation and analysis of the logs; 21- Incident's response; 22- Treatment of the alerts; 23- Crisis behavior.

- **The Netherlands**

The Dutch legislator did not classify the healthcare sector as an operator of essential services.[222] In the preparatory work, it was assessed that there are already measures existing to guarantee confidentiality and integrity for the healthcare sector, such as the obligation to disclose calamities, i.e. an event related to the quality of the healthcare service impacting on a serious way the patient.[223] The parliamentary works also referred to the obligation to disclose data leaks, respecting normalisation norms issued by the NEN[224] and the control by the health and youth care inspectorate (*Inspectie Gezondheidszorg en Jeugd*) as well as the Dutch Data Protection authority (*Autoriteit Persoonsgegevens*).[225]

- **Italy**

In order to transpose the NIS Directive into national legislation, the Italian Council of Minister approved the Legislative Decree No. 61/2018[226]. Art. 4(2) thereof sets the criteria for the identification of operators of essential services.

Annex II of Legislative Decree No. 61/2018 identifies sectors and sub-sectors for the operators of essential services. Concerning the healthcare sector, 'health care providers' (in the meaning of

[222] TK Memorie van toelichting Cybersecurity NIB Richtlijn 'Vooralsnog wordt niet voorzien dat Nederland zorgaanbieders aanwijst als AED', p. 25. https://www.rijksoverheid.nl/documenten/kamerstukken/2018/02/15/tk-memorie-van-toelichting-cybersecurity-nib-richtlijn ; https://www.security.nl/posting/568604/Cybersecuritywet+ziet+ziekenhuizen+niet+als+essenti%C3%ABle+dienst

[223] For a definition of 'calamiteit', see Art. 1, Wet van 7 oktober 2015, houdende regels ter bevordering van de kwaliteit van zorg en de behandeling van klachten en geschillen in de zorg (Wet kwaliteit, klachten en geschillen zorg)

[224] NEN 7510 Medische informatica - Informatiebeveiliging in de zorg, D.1 Management systeem, D.2. Beheersmaatregelen; NEN 7512 Medische informatica - Informatiebeveiliging in de zorg - Vertrouwensbasis voor gegevensuitwisseling; NEN 7513 Medische informatica - Logging - Vastleggen van acties op elektronische patiëntdossiers. Available on https://www.nen.nl

[225] Kamerstukken 34 883 nr 3 (2017-2018) - Regels ter implementatie van richtlijn (EU) 2016/1148 (Cybersecuritywet) – pp. 12-13.

[226] Decreto Legislativo 18 maggio 2018, n. 65 Attuazione della direttiva (UE) 2016/1148 del Parlamento europeo e del Consiglio, del 6 luglio 2016, recante misure per un livello comune elevato di sicurezza delle reti e dei sistemi informativi nell'Unione. (18G00092) (GU Serie Generale n.132 del 09-06-2018).

Art. 3(1)(h) of Legislative Decree 38/2014[227]) are indeed recognised as operator of essential services.

Finally, Art. 4(1) of the Legislative Decree foresees that NIS competent authorities have to identify and list all the essential operators residing in the national territory. The essential operators providing healthcare assistance in Italy have to be identified by the national Ministry of Health ('Ministero della Salute').[228]

8.3 Medical Devices Regulations

Under the current European framework, a single clear set of rules and standards for cybersecurity does not exist in one single place.[229] Legal provisions concerning cybersecurity and medical devices in Europe come indeed from a fragmented legal framework. On the one hand, (a) privacy and data protection regulations (illustrated above) foresee specific provisions on security and confidentiality that are applicable also to medical devices and (b) critical infrastructure regulations require the adoption of cybersecurity measures that have impacts also on medical devices;[230] on the other hand, the EU legislation concerning medical devices foresees only small relevant references concerning cybersecurity. The subsections below illustrate European laws concerning medical devices, with a brief focus on cybersecurity related provisions.

8.3.1 The Actual Framework concerning Medical Devices

The actual European framework concerning medical devices is composed by three Directives: 1) Council Directive 90/385/EEC on Active Implantable Medical Devices[231] ('AIMDD'), Council Directive 93/42/EEC on Medical Devices ('MDD')[232] and Council Directive 98/79/EC on In Vitro Diagnostic Medical Devices ('IVDMD').[233]

The three Directives set a common European framework, that has been implemented by Member States by the means of national legislations. The approach proposed by MDD aimed to harmonise the EU framework to the essential requirements of medical devices thus leaving to Member States the opportunity to translate the European legal requirements into national ones.

The Directives have been further specified by standards, and relevant Bodies provided guidance through different documentation. Standards are and have been elaborated by the European

[227] For the sake of completeness is reported the definition thereof. See Decreto Legislativo 4 marzo 2014, n. 38, Attuazione della direttiva 2011/24/UE concernente l'applicazione dei diritti dei pazienti relativi all'assistenza sanitaria transfrontaliera, nonche' della direttiva 2012/52/UE, comportante misure destinate ad agevolare il riconoscimento delle ricette mediche emesse in un altro stato membro.) (GU n.67 del 21-3-2014, where 'Health care providers' are defined by Art. 3(1)(h) as "any natural or legal person or any other entity legally providing health care in the territory of a Member State of the European Union".

[228] The list of such essential operators – due to be issued by 9th November 2018 – was not made public yet.

[229] See E. Vollebregt, EU cybersecurity requirements under current and future medical devices regulation, Q1 conference, 25 July 2016, see: https://www.slideshare.net/ErikVollebregt/eu-cybersecurity-requirements-under-current-and-future-medical-devices-regulation.

[230] See Art. 4(b) and (c) of the NIS Directive.

[231] Council Directive of 20 June 1990 on the approximation of the laws of the Member States relating to active implantable medical devices (90/385/EEC).

[232] Council Directive 93/42/EEC of 14 June 1993 concerning medical devices.

[233] Directive 98/79/EC of the European Parliament and of the Council of 27 October 1998 on in vitro diagnostic medical devices.

Standards Group (CEN) and many are subsequently adopted or incorporated into international standards by the ISO (International Standards Organization); guidance documents (e.g. 'MEDDEV' documents) are issued by the European Medical Device Expert Group of the European Commission.

8.3.2 New EU Regulations on Medical Devices

The European framework concerning Medical Devices has been recently reformed. The Medical Devices Directives were in need of modernization as they seemed no longer reflecting the scientific and technological progress.[234] The three current Directives have been replaced by two new Regulations: Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices ('MDR')[235] – incorporating active implantable medical devices – and Regulation (EU) 2017/746[236] of the European Parliament and of the Council of 5 April 2017 on in vitro diagnostic medical devices ('IVDR'). These two regulations have been adopted on 5 April 2017, entered into force on 25 May 2017.

Differently from the previous framework, these Regulations will not need to be transposed by Member States with national legislation. They will be directly applicable and enforceable in all EU Member States as of spring 2020 for the Regulation on medical devices and as of spring 2022 for the Regulation on in vitro diagnostic medical devices.

8.3.3 Cybersecurity in the Medical Devices Framework

Cybersecurity for medical devices has not been explicitly addressed within the medical devices framework. To this end, the word 'cybersecurity' is not present in MDD, AIMDD, IVDMD or MDR, IVDR. Nonetheless, the current and future framework contain rules that may concern medical devices under the cybersecurity perspective.

One of the key elements of the MDD is that it requires the adoption of a general risk management approach by manufacturers. In the practice field, this has implied the use of EN ISO 19471:2012[237] on the application of risk management to medical devices. The new framework follows the path set by the precedent one and it contains rules emphasizing the requirement of such risk management approach.[238]

Also, it is worth to note that MDD has led to the application of standards such as IEC 62304 on medical device software – software life cycle processes,[239] which contains security requirements and typical cybersecurity points of software requirements. Now, Annex I, Requirement n. 17.2 of MDR, sets out that "for devices that incorporate software or for software that are devices in

[234] See G. Verhenneman, Medical Devices Regulations sound the legal future for heart valves and sticking plasters, 06 June 2018, in CiTiP Blog, available at: https://www.law.kuleuven.be/citip/blog/medical-devices-regulations-sound-the-legal-future-for-heart-valves-and-sticking-plasters/, last accessed: 04/12/2018.

[235] Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC.

[236] Regulation (EU) 2017/746 of the European Parliament and of the Council of 5 April 2017 on in vitro diagnostic medical devices and repealing Directive 98/79/EC and Commission Decision 2010/227/EU.

[237] ISO 14971:2007, Medical devices -- Application of risk management to medical devices.

[238] E.g. Annex I, requirement 1 foresees that "manufacturers shall establish, implement, document and maintain a risk management system".

[239] IEC 62304, Medical device software – software life cycle processes.

themselves, the software shall be developed and manufactured in accordance with the state of the art taking into account the principles of development life cycle, risk management, including information security, verification and validation".

Finally, it has to be noted that the new framework contains new provisions (in particular, MDR, Annex I) concerning the manufacturing and design of medical devices. These provisions are worth to be reported herein as they establish baseline criteria for manufacturers and designers for the cybersecurity of medical devices from the perspective of their cybersecurity within an IT environment. These are reported below:

- Annex I, Requirement n. 11.2(d) of MDR, which requires that devices shall be designated and manufactured in such a way as to remove or reduce as far as possible "the risks associated with the possible negative interaction between software and the IT environment within which it operates and interacts".
- Annex I, Requirement n. 17.4 of MDR, which states that "manufacturers shall set out minimum requirements concerning hardware, IT networks characteristics and IT security measures, including protection against unauthorised access, necessary to run the software as intended".

9 Conclusion

After the execution of the task 3.2, it can be said that the state-of-the-art of physical and cyber security solutions as well as the most common physical and cyber vulnerabilities have been studied. Using the knowhow acquired during this task, it will be possible to implement appropriate disruptive solutions that will improve physical and cyber security in healthcare infrastructures. All the collected information must be present at all phases of execution of the SAFECARE project.

In the final version of this report (Deliverable 3.3), it will be very useful to add more cyber security solutions that are actually used in health services. The solutions described in this report have been more focused on low-level solutions, so it is important to add more high-level solutions that are used in health sector.

It is important to note that new vulnerabilities are emerging every day, so it is important to update this report with the new findings during project execution to guarantee that the common vulnerabilities are known.