Artificial Intelligence in Protecting Smart Building's Cloud Service Infrastructure from Cyberattacks

Petri Vähäkainu, Martti Lehto, Antti Kariluoto and Anniina Ojalainen

Abstract Gathering and utilizing stored data is gaining popularity and has become a crucial component of smart building infrastructure. The data collected can be stored, for example, into private, public, or hybrid cloud service infrastructure or distributed service by utilizing data platforms. The stored data can be used when implementing services, such as building automation (BAS). Cloud services, IoT sensors, and data platforms can face several kinds of cybersecurity attack vectors such as adversarial, AI-based, DoS/DDoS, insider attacks. If a perpetrator can penetrate the defenses of a data platform, she can cause significant harm to the system. For example, the perpetrator can disrupt a building's automatic heating system or break the heating equipment by using a suitable attack vector for a data platform. This chapter focuses on examining possibilities to protect cloud storage or data platforms from incoming cyberattacks by using, for instance, artificial-intelligence-based tools or trained neural networks that can detect and prevent typical attack vectors.

Keywords artificial-intelligence-based applications ·artificial intelligence ·cloud service ·data platform ·attack vectors

1 Introduction

Artificial intelligence is a major buzzword nowadays and is considered as the new "oil" of the future with the potential for great societal impact. AI has been under research for many decades, and it was originally presented as a novel way to mimic the cognitive functions of the human brain. AI has the capacity to process vast amounts of data, it has far-reaching applications, and it has been used in armed forces, construction, education, healthcare, space exploration and transportation around the world. In the healthcare sector, AI has suc-

P. Vähäkainu (*) · M. Lehto · A. Kariluoto· A. Ojalainen
Faculty of Information Technology, University of Jyväskylä, Jyväskylä, Finland
e-mail: petri.vahakainu@jyu.fi

M. Lehto
e-mail: martti.lehto@jyu.fi

A. Kariluoto
e-mail: anjuedka@jyu.fi

A. Ojalainen
e-mail: anniina.m.t.ojalainen@jyu.fi

ceeded in providing accurate diagnoses to prevent skin cancer, treatment recommendations, and provided surgical aid. In the field of smart buildings, AI can assist in finding anomalies and providing future forecasting in order to reduce maintenance costs.

Artificial intelligence can be defined as a system that thinks and acts rationally thinks and acts in such a way as to mimic rational humanlike behavior (Deshpande, 2008). AI is a combination of information technology and physiological intelligence, which can be computationally used to reach goals defined. Intelligence is the ability to think through memory formation, pattern recognition, adaptive decision-making and experimental learning. Artificial Intelligence can make machines behave like humans and even surpass them in efficiency. (Lehto, 2015)

Artificial intelligence can be applied to a range of fields, such as healthcare, predictive building maintenance, military and cybersecurity. Cybersecurity provides the means to access the data and the data stored on them. Effective cybersecurity controls provide a cyberspace infrastructure, which is reliable and resilient. Lacking or absent controls lead to an insecure cyberspace. According to Bayuk, Healey, Rohmeyer, Sachs, Schmidt & Weiss (2012) cybersecurity applied to prevent, detect and recover from damage to confidentiality, integrity, and availability of information in cyberspace. In order to use all these factors, people, processes and technologies are utilized.

Smart buildings can be seen as a cyber-physical system (CPS) in which smart sensors automatically measure usage, functions, and variables describing the state of a building (Schmidt & Åhlund, 2018). Energy, electricity, and water consumption, inside temperature, humidity, and other relevant variables are examined and used to automatically adjust, for instance, the heating system of a smart building. A building can be considered smart, even if only some of these variables are measured.

Cyber-physical systems provide a way to gather relevant data through smart sensors. The data has to be stored privately, securely, and it has to be available. Cloud services enable data replication and strategic storage on multiple servers spread across various geographical locations (Shahapure & Jayarekha, 2015). Replication improves data availability, reliability, and ensures fault tolerance. Smart services, such as AI-assisted data-intensive automatic heating adjustment system, can be developed using the stored data. In order for such services to work, they need working business models and replication functions to operate globally.

Cybercriminals are constantly looking for new ways to exploit vulnerabilities. Data gathered and stored into cloud storage, or distributed data platforms need to be secured. Cybercriminals today are able to leverage sophisticated attack vectors, including artificial-intelligence-based attacks, in determining exploitable vulnerabilities. These days, cybercriminals can utilize even more sophisticated attack vectors such as artificial intelligence-based attacks in looking for vulnerabilities they can exploit. An insider threat can be a significant threat to the system. An insider threat causes one of the most if even the most significant threat to the system. A malicious insider familiar with an organization's security practices, data, and computer systems can circumvent security controls to gain access to the system and the data. This motivates the need to research novel ways to detect exploitable vulnerabilities and prevent these high-risk cyberattacks.

IoT devices can be used for the collection of data to be stored on cloud services. Artificial intelligence and machine learning can be used in optimizing data usage efficiency. This is why this chapter is organized as follows: Section 2 examines the basics of artificial intelligence and machine learning. Section 3 deals with the basics of cyberspace and

cybersecurity. Section 4 introduces smart buildings and services. Section 5 examines cloud services and data platforms. Section 6 presents common cloud vulnerabilities and attack vectors, including DoS and DDoS attacks, IoT based attacks, and insider threats. These attack vectors were selected because they came up repeatedly in scholarly sources. In addition, these vectors are both related to smart homes and cloud services. Section 7 discusses countering cloud cyberattacks and section 8 concludes the chapter.

2 Artificial intelligence and machine learning

Artificial intelligence (AI) and machine learning (ML) are disciplines with algorithms capable of learning representations from data. ML is a subset of AI (Nicholson). ML contains the research areas of deep learning (DL) and deep reinforcement learning (DRL). Relations of these fields can be seen as overlapping circles. AI refers to systems that simulate human behavior. ML refers to systems capable of adapting themselves based on the situation. DL refers to the actual size of structure of the ML model. This applies to DRL as well, but DRL is mainly known for how the ML model learns. Learning is based on an action-feedback loop.

Artificial Intelligence is used in cases which humans consider time-consuming or tedious, or when an AI model can be trained faster than programming an explicit solution. Tasks in which AI and ML algorithms have succeeded particularly well include image recognition, image classification, image generation, and natural language processing. They have also been used for social media monitoring, marketing, predictive health monitoring, robotics, fraud detection (appliedAI, 2019). Burnap and Williams (2015) used ML for hate speech detection from Twitter and Zhao, Zhong, Zhang, & Su (2016) used artificial neural networks (ANN) to predict building energy usage.

One common feature among these types of models is the need to train the algorithms that need to be trained (such as supervised learning) first before the actual use. Supervised, unsupervised, and DRL methods are used widely. Supervised methods refer to cases where there are pre-labeled data for the training of the algorithm. Unsupervised refers to cases in which the ML algorithm estimates these labels itself. DRL is a special case of training algorithms because it uses feedback in order to learn and that feedback can come from a human expert or from the surrounding system.

In general, algorithmic learning happens based on data inputs and the desired output. Therefore these can be abstracted into a functional representation: f(input) = output. In the case of neural networks (NN), which are extremely popular in ML and AI research, when the algorithm learns, the training changes hidden values based on the results of an activation function for each node of the neural network. Training continues until the model has reached a sufficient level of accuracy. Accuracy is calculated with the minimizing function, which calculates the differences of predictions of the ML algorithm and the given true values.

According to Ghahramani (2015), training these algorithms means that they learn models that represent part of the data or the behavior of the data. Another common feature is that AI solutions tend to be data-intensive systems. For example, using the NSL-KDD dataset Potluri & Dietrich (2016) trained their DNN model in parallel to accelerate the learning of different attack types. They did not have enough data for all attack types, which resulted in decreased classification performance on those attack types. In order to perform well, DL models tend to need lots of quality data and GPU time. Currently, there is a trend among researchers to find ways to lessen the number of data samples.

3 Cyber space and cyber security

The word Cyber comes from the Greek word κυβερεω (*kybereo*), which means to direct, guide, and control. Cyber refers to the digital world, which includes the surroundings and being present in our daily lives. In the year 1984, William Gibson's Neuromancer novel connected the words ”cyber” and ”space.” Defining cyberspace is still a challenging task. Cyberspace is described in United States Cyberspace Policy Review as an interdependent network of information technology infrastructures and includes the internet, telecommunications network, computer systems, and embedded processors and controllers in critical industries. Typical usage of the term also refers to the virtual environment of information and interactions between people. (Cyberspace policy review) Based on literacy review cybersecurity can be connected to cyberspace as follows: ”cybersecurity is the organization and collection of resources, processes, and structures used to protect cyberspace and cyberspace-enabled systems from occurrences that misalign de jure from de facto property rights” (Craigen, Diakun-Thibault & Purse, 2014).

There is no universally accepted definition of cybersecurity, but the term is broadly used in literacy. Even though there is no universal definition of cybersecurity, a description of the word should bind human and information system component together. Cybersecurity can be defined as a range of actions taken in defense against cyberattacks and their consequences and includes implementing the required countermeasures. Cybersecurity is built on the threat analysis of an organization or institution. The structure and elements of an organization's cybersecurity strategy and its implementation program is based on the estimated threats and risk analyses. In many cases, it becomes necessary to prepare several-targeted cybersecurity strategies and guidelines for an organization. (Lehto, 2015, 3 - 29.)

European commission defined cybersecurity in the Cybers Security Strategy of the European Union as the safeguards and actions that can be used to protect the cyber domain both in the civilian and military fields, from those threats that are associated with its interdependent networks and information infrastructure or that may harm them. Cybersecurity strives to preserve the availability and integrity of the networks, infrastructure, and the confidentiality of the information contained therein. (EUR-Lex)

Original Martin C. Libicki's model of cyberspace consisted of a three-layer model: semantic, syntactic, and physical. We created and enhanced our unique Libicki's model by adding the cognitive and the service layer into the original Libicki's model for it to better describe the cyber environment concerning a smart building concept discussed in this chapter.

The physical layer is the first layer, which consists of physical components of an information network. The physical layer includes all the equipment necessary to send, receive, store, and interact with and through cyberspace. The hardware and devices concerned are e.g., cables, routers, switches, transmitters, receivers, computers, and hard disks. The layer acts as a bridge between the physical layer and the syntactic layer.

The syntactic layer uses protocols and software to send, receive, store, format, and present gathered data through the physical layer. The syntactic layer divides into sub-layers by using, for example, OSI-model (Open System Interconnection). The syntactic layer is responsible for interaction between the devices connected to the network.

The semantic layer contains all the information and gathered datasets from smart building's IoT sensors and stores them into data storage, such as data warehouses, located on cloud services (data platforms). All the data stored needs to be secure, and current

information security goals should be followed. Those information security goals being: confidentiality, integrity, availability of information, authenticity, accountability and non-repudiation, and reliability. (BS ISO/IEC 27002, 2013)

The service layer includes digital smart services that implement data gathered from smart building's IoT sensors. Digital smart services can be, for example, smart lock, automatic heating adjusting system, snowplowing service, or digital caretaker. The service layer also includes information security and data management services.

The cognitive layer's meaning is to provide an environment to understand visualized and analyzed information. The information considered is beneficial for decision-makers who build and maintain smart buildings. On the cognitive layer, information gathered is being analyzed to form a contextual understanding of information for a decision-maker.

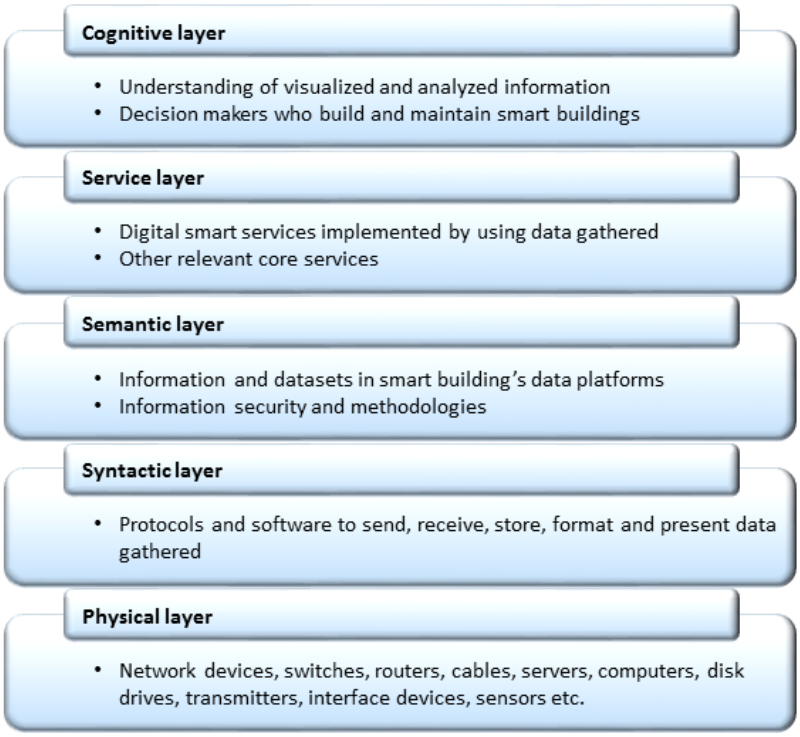

Fig. 1 Libicki's model of Cyberspace (modified by authors)

4 Smart buildings and services

There are several types of smart buildings, such as smart homes, smart airports, smart hospitals, smart factories. For example, Alam, Reaz, & Ali (2014) define smart home as a home that has sensors and appliances, which communicate with each other and the smart meter that continuously pushes and receives information to and from the smart grid. The Smart grid forms from the union of information and communication technologies with the traditional power grid (Iyer & Agrawal, 2010). This information transfer is intended to minimize power consumption.

Authors define smart buildings loosely as buildings that have devices for energy consumption optimization of the structure while using sensors to gather data of the building conditions and actuators to maintain building conditions at an acceptable level for all inhabitants using some guiding method that could be perceived intelligent, such as AI. Devices and applications can be the Internet of Things (IoT) devices and operate under many

different protocols, such as BTLE, PaaS & IaaS, ZigBee, SFP. These buildings are meant to protect both inhabitants and IoT devices against elements of nature. Smart buildings can include a multitude of sub-systems, such as smart homes, and altering energy sources and energy source combinations since some of them might be energy producers and consumers simultaneously. For example, Nagpal, Basu, & Staino (2018) suggested a concept of cooperative energy consumption optimization for use with buildings. They showed that building automation systems (BAS) could be used together for a cluster of buildings with the results leading to up to 15 percent reduction of energy consumption. On a similar note, Wang, Lee & Yuen (2018) trained an ensemble model for dynamic short-term cooling load forecasting of a building.

Smart buildings are also cyber-physical (CP) systems that combine both the physical aspects of the building and cyber (virtual) aspects of the cloud-based solution. Physical attributes include building, sensors, and actuators (Legatiuk et al., 2018). Building functions as the frame of the system, providing a place to integrate sensors and actuators while protecting these devices against weather conditions and manipulation. Sensors gather data from surroundings, which include the building itself. Actuators are devices that given a command they produce an action, which alters device settings and causes some change in the structure eventually. In the CP system, cyber refers to making decisions in the cloud. It can use knowledge of previous measurements and calculation results as well as the most current measurements from the sensors. It calculates new commands for the actuators.

Smart buildings and their sub-systems, such as structural health monitoring (SHM) systems and IoT -based devices, should be made to follow similar guidelines as CP systems since this could improve the gathering and utilization of data. According to Abate, Budde, Cauchi, Hoque, & Stoelinga (2018), high reliability, availability, maintainability, and safety – standards are necessary. Jiang (2018), although having a focus on smart factories, suggests an 8C architecture or guideline for CP systems as an improvement for design known as 5C architecture, those 8 Cs being connection, conversion, cyber, cognition, configuration, coalition, customer, and content. The last three Cs are providing more room for mass production and customization.

For collecting data, Legatiuk et al. (2018) recommend that these kinds of systems should be recorded mathematically. Sensors of the buildings or target structures ought to be modeled mathematically together with corresponding measured data. These should then be further formulated to cover also sensor groups and groups of sensor groups. This approach comes with the benefit of being general but also mathematically precise. On the downside, this method does not provide suitable information for every use case. However, Wang & Srinivasan (2017), note that not all AI approaches need a high level of structural information about the building. Abate et al. (2018) suggest using fault management trees (FMT) with smart building maintenance since they are dynamic event trees easing the decomposing of fault modes of the system.

Unused data has little value. Services that utilize IoT, wearable devices, portions of Big Data, and AI -based systems to ease the handling of the above-mentioned Big Data to provide continuous, traceable, and preemptive services for customers can be called smart services. These services are often novel.

On the one hand, with smart buildings and smart homes, it is well advised to consider the physical aspects of possible services. For example, when considering smart services for smart building energy usage control, Byun & Park (2011) brought forth three main issues with smart services at the time: services were centralized which can lead to performance

issues, fixed rule-based control is not necessarily capable of handling complex situations, and, lastly, physical parts have different lifetimes and sensor nodes tend to die out. Their proposed solution was a self-adapting intelligent system that had been designed to have distributed devices, which could alter their functioning based on measurements from the building, environment, and users. With a decentralized adaptive control system, it was possible to handle also the problem of dying batteries. On the other hand, transforming data into information can be vital for services. For example, Dao, Pongpaichet, Jalali, Kim, Jain & Zettsu (2014) proposed a system based on cloud computing called EvIM that comprised both gathering data and using it for different services. The proposal made it possible to unite CP systems together with users and alleviated some of the rule-based rigidness mentioned above while providing real-time event handling.

Smart buildings bring forth yet another challenge, and that is the preservation of privacy since there are different kinds of intelligent buildings, and some might be controlled together. Therefore, smart services should be such that they do not expose user(s) or user data to third parties without user consent. Also, according to Sta (2017), systems should be made capable of handling imperfect information when dealing with Big Data. Digital twin should be used with smart services, as well, since according to Qi, Tao, Zuo & Zhao (2018) digital twin is versatile and combines both physical aspects and virtual aspects with connections between them. Utilization of digital twins could lead to a simulation of various situations and eventually to better smart services. Lim, Kim, Kim, Heo, Kim, & Maglio (2018) remind that for service to succeed it needs to bring value to the system.

5 Cloud services and data platforms

Cloud computing can be defined in various ways in the literature. It can be referred to as a way to store and access data over the Internet instead of one's computer's storage media. Through cloud computing, a user with a pay-as-you-go pricing business model can rent computing power, database storage, applications, and other relevant IT resources. According to Karthikeyan & Thangavel (2018), Cloud computing can be thought of as a computing paradigm providing dynamically scalable infrastructure for application, data, and file storage. A large pool of various systems is connected in public and private networks to form the basis of the paradigm.

The concept of cloud computing can be traced back to the 1950s, the era of mainframe computers, which were accessible via thin terminal clients. The development towards nowadays cloud computing started in the 1980s with cluster computing. Cluster computing was followed by grid computing, focusing on solving significant problems with parallel computing. Grid computing lead to utility computing in the 1990s offering computing resources (clusters) as virtual platforms for computing with a metered service. Clusters are usually distributed locally using the same hardware and operating system, which provides the possibility to use them as a supercomputer by using the pay-per-use approach. In 2001 software as a service (SaaS) concept was introduced focusing on network-based subscription of applications. Figure 2 illustrates commonly agreed SPI (SaaS, PaaS, IaaS) framework of three primary services provided through the cloud.

Public cloud service is the most widely used service delivery model in cloud computing currently available. Public clouds can be owned, operated, and managed by third parties, such as government institutions, businesses, academic institutions, or a combination of them (Castro-Leon & Harmon, 2016). Public clouds are highly scalable, they provide large

capacity, and shared resources require minimal IT investments and decrease operating costs in the long run by using a pay-as-you-use –model (Usman, Bawazir & Kabir, 2014). All customers of public cloud providers share the same pool of security protections without a possibility to affect it. Major public providers in the market are Amazon Web Services (AWS), Microsoft, and Google.

Unlike public cloud services, private cloud services are intended for a single enterprise. Private clouds provide better controls and data security, which public cloud services are lacking. The private cloud divides into two categories, as follows: 1. On-Premise (internal) Private Cloud, and 2. Externally (External) Hosted Private Cloud. Internal clouds provide standardized process and protection, but size and scalability are limited and operated within one's own data center. Internal cloud fits for applications requiring control and configuring capabilities of the infrastructure and security. External clouds are externally operated with a cloud provider, which produces an exclusive cloud environment ensuring a high level of privacy. External clouds fit for companies, which require a highly secure cloud service not sharing of physical resources. (Karthikeyan & Thangavel, 2018)

Hybrid clouds combine private and public cloud services. Hybrid cloud services increase flexibility as hybrid cloud providers can use third-party cloud provider services in full or partially depending on the need. The hybrid cloud enhances the capabilities of a private cloud, providing a possibility to use public cloud services when e.g., the computing power of a private cloud is not enough. (Karthikeyan & Thangavel, 2018) To eliminate security risks, an enterprise can use private cloud services to host sensitive and critical workloads and use 3rd party, public cloud provider services to host less-critical tasks i.e. testing and improving new services. The hybrid cloud reduces initial investment costs when developing services by using a pay-per-go model without a need to make a substantial investment beforehand.

In SPI SaaS model software is licensed on a subscription basis or by pay-per-use model. The cloud provider provides the hardware infrastructure and software applications, and applications are run, for example, via web portals. A single instance of the service runs on the cloud concerned, and multiple users can access it. Customers do not need to worry about investment in infrastructure, licensing, and maintenance of the software or environment scalability issues, as they are the provider's tasks. Security, customization, and components can be issues on SaaS –layer as customers cannot control them. Service on SaaS can be CRM, email, virtual desktop, communication, or games (Goel & Sharma, 2014).

In SPI PaaS model, the cloud service provider provides software and product development tools on its cloud infrastructure. The provider's task is to offer system resources such as network, server, storage, operating systems, databases, development tools, and other relevant resources to customers. The customer can design, implement, and deploy his/her applications into the cloud service and run them there. The client must keep his/her deployed software updated to confirm security. The disadvantage of PaaS is the mandatory use of the service provider's API. Service on PaaS can be e.g., execution runtime, database, web server, and development tools (Goel & Sharma, 2014).

In SPI IaaS model the cloud service provider controls and provides the infrastructure required to run customer's developed and deployed applications. IaaS layer offers storage and computing capabilities as a service. IaaS model also provides flexibility in the means of security as customers can also affect it. The customer needs to make sure the software deployed is up to date, configured, and appropriately integrated. Service on IaaS –layer can be virtual machines, servers, storage, load balancer, or network (Goel & Sharma, 2014).

There is a need to introduce an additional Data-as-a-Service (DaaS) service layer into SPI–framework. DaaS provides a new architecture model in which, for example, private clouds can be located inside a public cloud service. DaaS is a service, which provides means and capabilities to transform raw stored data into meaningful assets e.g., smart service development and/or analysis from various data sources, such as databases, data warehouses, data lakes, filesystems, applications, data science platforms, applications, and BI tools. DaaS provides functions such as collection, integration, enrichment, curation, contextualization, aggregating, and analysis of the data (Randall, 2016).

Instead of copying or moving all the data from data sources into a data warehouse or a monolithic data lake, DaaS services can be implemented between them to gather the data required. Data lakes and data warehouses (DW) are centralized storage repositories that can store a significant amount of data. Repositories are different as data lakes can store data in native/raw (both structured and unstructured) format and DW handles only structured and cleansed data. Both repositories can be used as a source for DaaS and in conjunction to complement each other. DaaS can decrease redundancy and cut costs by placing relevant data into one location, providing data usage and modification for many users through one convenient service. Regardless of data location, structure, and size, DaaS enables users to examine, classify, and analyze the data. Users can use analytics tools they favor the most, such as Python, QlikSense, and R.

One way to use DaaS service is through Amazon (AWS) or Azure cloud services, which offer Data-as-a-Service functionalities in conjunction with open source Dremio DaaS platform solution. Cloud platforms generally provide various kinds of solutions and services for computing, security, AI, and storing, managing, and analyzing the data. Amazon object-oriented Simple Storage Service (S3) or Azure cloud service can be connected to Dremio DaaS service to discover and explore, curate, share, and analyze the data. The Dremio service includes data catalog, which provides a way to find and explore real and virtual datasets, which are automatically updated when new data source is added and when datasets evolve. Dremio also supports SQL syntax for advanced transformations, learning about the data and various kinds of transformations recommendations. Dremio can be deployed on-premises or in a public cloud service. (Dremio, 2019)

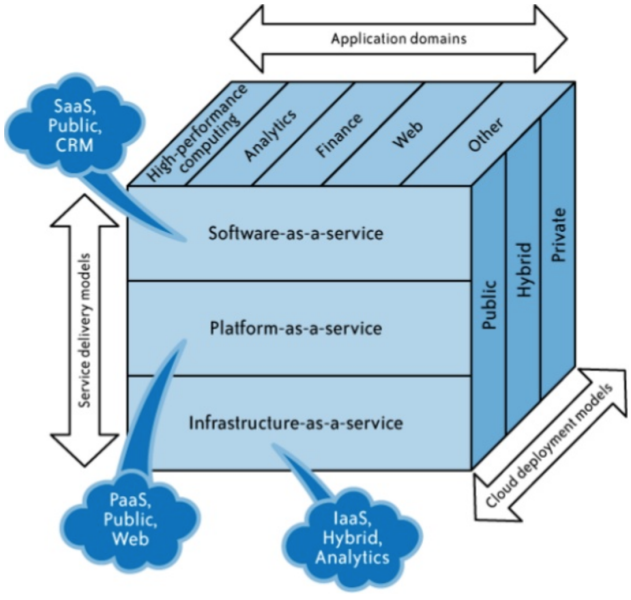

Fig. 2 SPI service model (Mather, Kamaraswamy, Latif, 2009)

6 Cloud service security

6.1 Situational awareness in the cyber world

Digitalization is taking significant steps ahead continuously. Due to digital transformation advances, organizations are accelerating the migration of data to the cloud services creating an enormous increase in attack surface and numerous amounts of novel types of risks for organizations to manage. At the same time, cyberattackers are using more sophisticated attack methods to penetrate an organization's defenses that more and more located on the cloud service. Generally, organizations react after the cyber incident has already occurred. To prepare for cyber-attacks in advance, organizations should assess their cyber risk profile beforehand, fix current problems and proactively manage the defense. Situational awareness is the key to surviving in the cyber world.

Situational awareness is a crucial asset for an organization, as without it, organizations cannot build functioning cybersecurity resilience. Organizations would need to clarify what are potential threats, what kind of harm could they provide, and what do they mean to the organization. Organizations may feel familiar with attacks they confronted in the past years, but they may still lack the ability to deal with current more sophisticated, advanced, and emerging attack methods. Updating situational awareness concerning these kinds of emerging threats should be in high priority.

Threat, vulnerability, risk, and asset form an intertwined entity in the cyber world (Lehto, 2015, 3 - 29.). According to Threat Analysis Group (2010), the asset can mean people, tangible or intangible valued property and information, such as databases, software code, and information system records. The asset is a resource that has to be protected. The threat is a hazardous cyber event that can exploit the vulnerability, accidentally or intentionally to obtain, damage, or destroy an asset. Vulnerability is a weakness or gap in the security of the system that can be exploited by threats to get unauthorized access to an asset. Vulnerabilities can be divided into human actions, processes, or technologies according to where they exist. Risk is the potential of the expected damage, loss or destruction of an asset, and it can be seen as the intersection of assets, threats, and vulnerabilities. It can be assessed from the viewpoint of its economic consequences or loss of loss at face value (Lehto, 2015, 3 - 29).

According to ENISA (2018, p. 125) Threat Landscape Report 2018, an attack vector is a path or means by which a threat agent, for example, hacker or cracker, can gain access to a computer or network server, abuse weaknesses to achieve a specific outcome. Attack vectors include viruses, e-mail attachments, WWW-pages, chat rooms, and deception. Cybercriminals continuously seek new attack vectors they can utilize in attacking e.g., cloud service infrastructure. There exist various attack vectors, which threaten cloud services, but some of the common ones are AI/machine learning-based attacks, DoS/DDoS attacks, insider threats, IoT attacks.

6.2 Utilizing artificial intelligence in cyber-attacks

Artificial intelligence can be used when executing targeted attacks. An attacker can teach and utilize AI algorithms to recognize persons who are the most suitable target victims and provide them with malware. A perpetrator can also use AI to gain information from the target security solution through the perpetrator's reconnaissance actions on the target network. Attacks towards IoT devices are substantially growing, and due to underestimation

of the situation, IoT devices generally lack necessary security measures and use relatively weak default device credentials opening a way to malware penetration. An attacker targeting IoT devices could use AI, for example, to generate credentials, find new vulnerabilities, learn the standard processes and behavior, distribute algorithms across all the nodes of a botnet for collective learning. (Kubovič, Košinár & Jánošík, 2018)

Artificial intelligence can be taught to find a way for new vulnerabilities by fuzzing, in which an attacker provides the algorithm with invalid, unexpected, or random input data. AI can also be a powerful technology to find the most effective way to attack. An attacker can abstract and combine attack techniques to identify the most effective ways of attacking. In the case of detection by the defender, the attacker needs to rerun the algorithm to follow a new learning path. Artificial intelligence can be utilized by the perpetrator in protecting himself by detecting intruders and suspicious nodes in their networks. The perpetrator can utilize artificial intelligence to spread disinformation, generate phishing emails and high-quality spam, and choose the best target, misuse a defender's AI model solution as a black box. He might use the same configuration in identifying what kind of traffic can pass through the defenses, and so on. (Kubovič, Košinár & Jánošík, 2018)

A perpetrator can also utilize adversarial examples when attacking machine learning models used, e.g., in cloud services, such as convolutional (CNN) neural networks or deep neural networks (DNNs), which can be used in implementing smart building services. Adversarial examples can be malicious inputs to DNNs providing erroneous model outputs while appearing to be unmodified in human eyes. This incident knocks out the classifier. (Papernot, McDaniel, Goodfellow, Jha, Celik & Swami, 2017). Adversarial input attacks are a threat to CNNs as instead of generalizing well and learning high-level representation (less prone to noise), they easily learn superficial dataset regularity (Bursztein, 2018). Defending against adversarial attacks is difficult because the theoretical model of adversarial example crafting process is hard to construct. In theory, machine-learning models would be needed to defend against them to produce the right outputs for every possible input. In practice, ML models may only work on a relatively small number of potential data available that they face; models may block one type of an attack, but leave vulnerabilities open for the perpetrator to exploit. (Goodfellow, Papernot, Huang, Duan, Abbeel& Clark, 2017).

6.1 Common attack vectors

6.1.1 DoS– and DDoS –attacks

A Denial-of-Service (DoS) is a malicious attempt in which the perpetrator tries to disrupt data traffic to targeted service with limited bandwidth, machine or network resource by overloading the resource with a flood of traffic intending to make the service low or make it temporarily or entirely unusable (Gillespie, 2016). The DoS attack can be described as a traffic jam hitting ordinary traffic on a highway by preventing its normal flow. A similar situation can happen when an overwhelming amount of people are in the midst of booking for concert tickets or buying discounted products at the same time when discounts are announced.

There are various DoS attacks, but popular ones are buffer overflow, ICMP flood (ping of death) and TCP SYN flood attacks. In a buffer overflow, attack the perpetrator sends more traffic than the target system can handle. In an ICMP attack, the attacker sends a huge amount of spoofed large-sized ICMP echo requests to target host enforcing each computer on the target network to ping instead of just the target computer attempting to switch it

offline or keep it busy. TCP SYN flood uses a three-way handshake trying to make a connection with an invalid return address without completing the handshake. (Elleithy, Blagovic, Cheng & Sideleau, 2006)

A Distributed-Denial-of-Service (DDos) is similar to DoS attack, but it takes advantage of several compromised computers when carrying out an attack. Exploited 'cluster' of machines may consist of ordinary hacked computers or IoT 'bots' or 'zombies' devices and commanded by an attacker. Using bots provides means for attackers to use various IP addresses from different areas of the world at the same time, making it more complicated for service providers to defend themselves from incoming attacks as blocking one IP address will not make much of a difference. (Gillespie, 2016) Detecting the location of an attack is a challenging task as the attack system can be randomly distributed. Distributed DDoS attacks are used for the reason it is challenging to be detected; it is also efficient and cheap to execute due to hacked zombie computers. Usually, DDoS attacks focus to do damage to a victim for personal reasons, gain material benefits or popularity. Through DDoS attacks cloud services could get jammed, and the data become inaccessible.

6.1.2 IoT based attacks

IoT sensors can gather data from the real world by measuring the physical quantity and converting it into a signal and eventually data for the digital domain. These days, there are estimated to exist more than 50 billion sensors connected via IoT. According to HP security research (2014), 80 % had privacy concerns and bad passwords, 70 % had lacked encryption, 60 % had vulnerabilities in UI and insecure updates. According to Gartner, there are more than 6 billion relatively vulnerable IoT devices on the Globe. Therefore, IoT devices provide an excellent attack surface for cybercriminals targeting cloud services. A vast amount of IoT devices still use default login credentials, which makes penetrating the defenses an easy task. DDoS attacks also pose a significant risk on IoT devices, which became true in 2016 in the form of Mirai botnet malware reaching up to 1 Tbps of traffic through hundreds of thousands of compromised IoT devices. (Wani, 2018) In IoT poisoning attacks, utilizing adversaries can cause remarkable risks on sensors when manipulating the training data by altering the sensor's measurements. Poisoning attacks greatly decrease performance causing misclassification or other kind of bad behavior. Through poisoning attacks, backdoors and neural Trojans are sneaked in. To prevent this kind of incident, collecting poisonous data and training an arbitrary supervised learning model could work as a defense strategy. (Baracaldo, Chen, Ludwig & Safavi, 2018)

IoT devices' vulnerabilities can be divided into system hardware or system software-based vulnerabilities. System hardware is vulnerable to exposure since the devices are often left unattended. Through exposure, the attacker might steal the device, extract sensitive data, modify the device's programming, or replace the original device with a malicious one. (Padmavathi, Shanmugapriya et al., 2009). In this case, the data in the cloud service might get tampered, and therefore the reliability is decreased.

IoT devices' software vulnerabilities are linked to application software, control software, and operating systems. Through system software, the attackers can execute, for example, access attacks or privacy attacks. In an access attack, an unauthorized person gains access to networks or devices in which they have no permission to enter. (Abomhara & Køien, 2014). Attack can also affect the cloud services, to which the devices are connected. According to Abomhara & Køien (2014), the access attack can be targeted at the physical machine or to IP-connected devices.

Privacy attacks are directed, for example, to data mining, cyber espionage, and tracking. IoT devices' information can be highly sensitive since the collected data can be, in some cases linked to one's home or workplace. Through IoT device hacking an attacker might be able to tell when there are people inside the building, what are they doing at the moment, and so forth.

6.1.3 Insider threats

Insider threats are substantial and increasing problem causing significant risk to organizations. Bonderud (2018) claims that one in the four attacks start inside corporate networks. Insider threat is a current or former employee, contractor, or business partner who has or had authorized access to an organization's network, system, or data, and intentionally exceeded or misused access which negatively affected to the confidentiality, integrity, or availability of the organization's information system (Costa, 2017). Cloud service can be targeted by an insider threat conducted by, for example, rogue administrator, an employee utilizing cloud weaknesses for unauthorized access, or an insider who uses cloud service resources to execute attacks against an organization's IT infrastructure. Motivations for conducting attacks can be e.g., financial aspect, theft of sensitive information, intellectual property, or fraud.

According to Ca Technologies report (2018), accidental or unintentional insider threat causes the most considerable risk (51 %) to the organization, and malicious or deliberate risk is the second largest risk (47 %). Regular employees pose the most significant security threat of 56 % to the organization, privileged IT users/admins 55 %, contractors/service providers, or temporary workers 42 %. The most vulnerable data is confidential business information such as financials, customer data, or employee data. Cybercriminals are highly interested in the organization's databases (50 %), fileservers (46 %), cloud applications (39 %) and cloud infrastructure (36 %). According to the report, up to 90 % of companies surveyed, felt vulnerable to insider threats.

Cloud services can be targeted by an insider threat conducted by e.g., rogue administrator, an employee utilizing cloud weaknesses for unauthorized access or an insider who uses cloud service resources to execute attacks against an organization's IT infrastructure. Motivations for conducting attacks can be, for example, financial aspect, theft of sensitive information, intellectual property, or fraud. Shaw, Ruby, and Jerrold (1998) identified a coherent cluster of risk factors characteristic of a vulnerable subgroup of critical information technology insiders. The factors that reduce inhibitions against potentially damaging acts are negative personal and social experiences, reduced loyalty towards the organization, personal and professional frustration, and ethical "flexibility," feeling of entitlement, anger, and lack of empathy. Also, stressors like family problems, substance abuse, disappointments at work, and threatened layoffs may trigger insider attacks. (Shaw, Ruby, and Jerrold, 1998)

7 Countering cloud cyberattacks

7.1 Encrypting the data

Cloud services are becoming more and more popular due to the organization’s interest in deploying applications and store their data into cloud service platforms. Cloud services are also gaining attention among smart building administrators. Cloud services provide many

kinds of benefits, but one of the biggest worries is confidentiality. Organizations have to be sure that the cloud service provider has stored their data securely, and proper encryption methods have been used. Cryptographic algorithms provide a means to secure the data concerned, but they also limit the functionality of the cloud storage. According to Gupta, Ghakraborty & Rajput (2015) two main categories of encryptions, symmetric (e.g., DES, AES, 3-DES, RC6, IDEA, Blowfish) and asymmetric (e.g., RSA, ECC, Elgamal), are being used in cryptography to achieve confidentiality, integrity, availability, and authentication. While using symmetric algorithms, encryption and decryption use the same algorithm and the same key to encipher and decipher the message. Symmetrical algorithms are useful to ensure confidentiality, integrity, and availability, but not authenticity. Asymmetric algorithms use two keys, one is a private key, which only recipient knows and the other is a public key, which everyone knows. Both of the keys can be used to encrypt and decrypt the message. Asymmetric algorithms provide better key sharing than symmetric algorithms, but they are slower than symmetric algorithms. Asymmetric encryption is slower than symmetric one due to the longer key lengths used and complexity of the encryption algorithms used. Conventionally known cloud service providers, e.g., Google cloud service platform, use AES128 and AES256, Amazon AWS AES128, AES192, and AES256 symmetrical algorithms and asymmetric RSA and Elliptic Curve Cryptography algorithms.

Encryption keys should be kept on a separate server on a storage block. Especially sensitive data needs to be encrypted after it is collected or created and uploaded to the cloud service data storage or an organization's private cloud service after the encryption process. The process mentioned may also bring out issues as if the data has been uploaded to the cloud service is encrypted and then later downloaded onto another media, it does not already have the decryption key providing useless encrypted data. (Business.com, 2018) Using homomorphic encryption could circumvent this issue by allowing data to be sent the cloud service to be analyzed without having to decrypt it first. Using homomorphic encryption provides only users needed to be able to analyze the data leaving cloud service providers no chance to know what kind of information is contained on the data. (Machmeier & Kunzke, 2019)

7.2 IoT based attacks

IoT devices typically are low powered, they have low storage, low computing resources, and they have been massively deployed and connected to each other. Partly due to a lack of resources, they are vulnerable to many kinds of cyber threats. Encryption provides an effective countermeasure, and nowadays, encryption is becoming a more and more crucial part of IoT sensor devices in various environments that formerly did not require it.

Ordinary cryptography methods, such as AES encryption and SHA-hashing, RSA signing is widely used in systems, which have enough processing power and memory. They are not fit for IoT sensor networks providing considerably less capability. Elliptic curve cryptography has been successfully applied on sensor nodes though. Therefore, lightweight cryptography methods are being developed and standardized to provide suitable means for IoT sensors with fewer resources. An adversary attack poses a real threat to an IoT sensor and sensor nodes by eavesdropping and modifying the data. Hence, be able to provide a secure routing protocol to ensure authentication, availability, and integrity is vital. Handful of lightweight cryptography protocols and primitives has been standardized as the ISO/ICE 29121 standard and primitives have been included in IPSec and TLS.

According to Buchanan, Li, and Asif (2018), the disadvantage of lightweight cryptography is less secured than conventional ones due to limited resources on sensors. Lightweight cryptography implementations are usually bound to use short key sizes, which increase the risk for key-related attacks. Sometimes read-only (masking) technology is used to permanently burn keys into IoT device chips to decrease key space consumed. When considering lightweight cryptography, IoT device clock, memory, storing internal and key states should be evaluated.

Using proper authentication and data encryption alone is not enough for ensuring data security. According to Chang adversary attacks can be injected into sensor nodes through compromised nodes. Intrusion detection systems (IDS) can be used to monitor suspicious and anomalous patterns of activity, which are different compared to ordinary and expected behavior. It is widely assumed that an intruder has significantly different behavioral patterns than legitimate users usually have in the network. Rule-based IDSs can be used to detect known patterns of intrusions, and anomaly-based IDSs can be used to detect new or unknown intrusions. Anomaly-based IDSs provide notably higher false alarm rates compared to rule-based IDSs.

Focusing on proper authentication and encryption and using intrusion detection systems can secure IoT devices. To prevent any incidents, collecting poisonous data, and training an arbitrary supervised learning model could work as a defense strategy. (Baracaldo, Chen, Ludwig & Safavi, 2018). IoT devices can be secured by focusing on confidentiality, integrity, authentication, accountability, auditing, and privacy. (Abomhara & Køien, 2014). Overall, the benefits of IoT devices are exceeding the downsides.

7.3 AI based tools in countering cloud cyberattacks

7.3.1 Insider attacks

Existing data protection techniques can be effective against insider attacks if implemented carefully and in the right way. Current technologies to prevent insider threats are Data Loss Prevention (DLP), encryption, identity, and access management solutions. In detecting active insider threats, organizations can utilize, for example, intrusion detection and prevention (IDS) services, log management, Security Information and Event Management System (SIEM) platforms, User Activity Monitoring (UAM), Privileged Access Management (PAM), DLP.

The monitoring of sensitive assets can be utilized in order to prevent and restrict insider threats that organizations are facing. According to Ca Technologies report (2018), 78% of organizations inventory and monitor all or most of their key assets, and more than 93% of them monitor access to sensitive data. Due to the increase in insider threat volume, organizations have begun to utilize User Behavior Analytics (UBA) tools and solutions to detect, classify, and alert anomalous behavior. Finding insiders who cause the highest risk is a crucial part of threat prevention. Organizations can monitor their behavior and work patterns, such as hostility towards colleagues, missing work, an excessive amount of work outside ordinary working hours, declined performance. In addition to UBA monitoring, comprehensive data access, movement analysis, and security analytics can be utilized.

Various solutions can be used to tackle insider threat issues, such as Darktrace Vectra Cognito. Darktrace uses the Enterprise Immune System technology (EIS) utilizing machine learning algorithms and mathematical principles to detect anomalies. EIS can adapt and automatically learn user, device, or an information network behavior to identify behaviors

reflecting threats, such as an insider threat. Darktrace uses mathematical approaches, such as Bayesian estimation to produce behavioral models for individual people and devices they use to detect unusual behavior and reveal possible insider attack. Darktrace (2018) Vectra Cognito works in a bit similar way as Darktrace, and it continuously learns from an organization's network activity. Cognito uses data science, supervised and unsupervised machine learning, and behavioral analytics to reveal attack behaviors and attacks such as an insider attack. Vectra Cognito can monitor and detect suspicious access to critical assets, policy violations related to, for example, cloud service usage, or another means of moving data. (Vectra Cognito, 2019)

To lower the risk of insider attacks towards cloud services, an organization should avoid management errors. According to Shaw, Ruby, and Jerrold (1998), organizations should understand the personality and motivation of the at-risk employee. Clear and standardized rules about the use of company information systems should be created. Also, the consequences of misuse should be made clear, and rule violations should be enforced.

7.3.2 DoS/DDoS attacks

DoS and DDoS attacks are ones of the most frequent, causing significant damage, and they impact cloud service performance. These kinds of attacks can be tricky to detect and block as the attack traffic can be easily tangled with legitimate traffic causing it to be challenging to trace. Especially application layer (Layer 7) DoS attacks can be hard to detect as the traffic appears to be like regular traffic with complete Transmission Control Protocol (TCP) connections and following protocol rules. Therefore, these attacks can target applications, which bypass the firewall (Ballal, Prasad, Rajappa & Khader, 2018).

Often security experts who deal with these kinds of issues are busy, so additional means to deal with these attacks are needed. There exist various tools to treat the problem, such as the PatternEx AI2 platform, that can predict incoming cyber-attacks, such as DoS or DDoS. AI2 uses three different unsupervised machine-learning methods and clusters data into patterns showing the top abnormal events to security analysts for further analysis to confirm attacks are real attacks. In the following phase, the platform builds a supervised model for the next set of data, which enables further active learning. This process will eventually improve the attack detection rate of the algorithms requiring less security analyst time. Currently, AI2 is able to detect up to 85 % of attacks while false positives are reduced by factor 5. (Conner-Simons, 2016)

Classical DDoS defense tools take advantage of rate-limiting and manual signature creation in mitigating cyber-attacks. Rate limiting tends to produce a significant amount of false positives while providing effective means in mitigating attacks. Manual signatures created can be then utilized to prevent or decrease the amount of false-positive results. Identifying the attack traffic is time-consuming as it requires human security analysts to analyze the attack vector, and it can be only done when the attack is already started. Hence, time to mitigation increases resulting in ineffective defense strategy. (Radware, 2018)

Radware Defense Pro offers means to prevent, protect and mitigate DDoS and IoT botnet attacks, such as fast-moving, high volume, encrypted or short-duration attacks, and IoT attacks, such as Mirai, Pulse, Burst, DNS, TLS/SSL, PDoS and Ransom Denial-of-Service (RDoS). Defense Pro provides behavioral mitigation capabilities to circumvent the manual signature creation and rate-limiting problems. It uses automatic machine-learning algorithms to create signatures and adapt defenses in changing attack-vector environment. Defense Pro can learn real-time behavior of legitimate traffic and to quickly detect an attack

when rate and rate-invariant parameters indicate an anomaly compared to legitimate traffic. Defense Pro offers negative and positive protection models and rate limiting, ensuring zero time to mitigation with scarce human cybersecurity professional intervention. (Radware, 2019)

Reblaze offers DoS/DDoS protection solution that provides defense from DDoS botnet assaults until single malformed-packet DoS attempts. The Reblaze solution protection mechanism if effective against various forms of DoS/DDoS attack vectors, such as amplification and reflection attacks, application-layer vulnerabilities, malicious inputs, protocol exploits, volumetric flooding, resource depletion, and exhaustion. Reblaze provides DoS/DDoS protection towards attacks on ISO/OSI layers 3 (network), 4 (transport), and 7 (application) and blocks attacks in the cloud service. Full protection from DDoS attacks is not common to many so-called “DDoS solutions,” but layers 3 and 4 are protected more comprehensively. Reblaze can run natively on Google Cloud Platform and integrate with Cloud Armor service. It augments Cloud Armor’s capabilities and uses Machine Learning in self-learning and adapting when there will be changes in the cyber threat environment. The learning process is automated and constantly adapting. It provides pattern recognition and behavioral analysis to detect early-stage attacks generating a small amount of false-positive results. (Reblaze, 2019)

7.4 Utilizing AI and ML based methods in combating cyberattacks

It is challenging to protect against insider threats, DoS/DDoS attacks, and adversarial attacks. Due to the increased amount of Big Data, AI, and ML methods are needed to combat these threats. Insider threats are challenging for AI, since not necessarily all malicious influence on the user can be prevented. According to Le, Khanchi, Zincir-Heywood & Heywood (2018), insecure habits of the user can be used as adversarial examples for AI-based IDS. Therefore, human experts shouldn't be allowed to decide what data they label when using supervised learning on AI models. Gavai, Sricharan, Gunning, Hanley, Singhal & Rolleston (2015) compared supervised and unsupervised ML models while investigating the usage of employees' social and web usage data, such as email frequency and machine access patterns, as possible features for the detection and prevention of insider threats. They found their unsupervised model to exceed their supervised model by a few percent. Zhang, Zheng, Wen, Xu, Wang, Yu & Meng (2018) studied a way to classify possible insider threats based on user behavior logs. They used long-short term memory NN (LSTM), which is used typically for sequences, in order to find anomalies from role-based user log data.

DoS/DDoS attacks are dangerous attacks because they can be hard to detect in the early stages of the attack, malicious packages can hide between legitimate traffic, attacks can inconvenience the target server, and the attacker can hide among zombie computers, which might be IoT devices. AI and ML techniques are needed for the automated detection of DOS/DDOS attacks. According to Diro & Chilamkurti (2018), the interconnectivity of smart cities is a potentially tempting playground for attackers. Rangaraju, Sriramoju, Sarma (2018) list in their article several ML techniques, such as Naïve Bayes (NB), Support Vector Machines (SVM) and genetic algorithms, that are used for detection and prevention of cyberattacks. NB as well as SVM are techniques based on probability while the genetic algorithm is an umbrella term for algorithms that are inspired by evolutionary theorem. Rathore & Park (2018) introduced their fog-computing framework against distributed attacks that used an extreme learning machine (ELM) for faster generalization. Instead, Han, Yang, Sun, Huang, & Su (2018) proposed a defensive framework that focused on the

detection of DDOS attacks both on data plane and on control plane while collaboratively distributing attack load on multiple defense applications. The NN model that the authors used was a stacked combination of autoencoder and softmax-classifier, which could detect a multitude of DDOS-attack types.

Adversarial attacks can be defended against with AI in some cases. Both CNN's and DNN's are commonly known to be susceptible to this cyber-attack type. Since no system is perfect, it is best to assume that all AI systems are vulnerable to adversarial attacks.

Adversarial training means modifying legitimate inputs to make AI classifier to learn to be more robust (Tramèr, Kurakin, Papernot, Goodfellow, Boneh, & McDaniel, 2018). In other words, the use of various counts of modified inputs used together with unmodified inputs in the training stage helps NNs to compact against false (modified) inputs by broadening their "understanding," where understanding refers to what the inputs would mean to a human. The creation of modified inputs is typically done using two NNs of which the first one tries to produce falsified inputs, and the second one attempts to classify inputs as true or false. Many ways to alter input data exists. Ganin, Ustinova, Ajakan, Germain, Larochelle, Laviolette, Marchand & Lempitsky (2016) suggested training classifiers with either labeled or unlabeled training data, which come from a different distribution than the intended data but have the same features. This is also known as domain adaptation, hence the name domain-adversarial neural networks (DANN). However, with this method, features need to exist in both domains and remain the same. Samangouei, Kabkab, & Chellappa (2018) proposed a new structure to protect NN used for classification called Defense-GAN. It finds similar input as given from its database and uses that as input for the classifying NN model. This has the benefit of protecting against adversarial attacks geared towards the classifier; however, it seems likely that the AI system would have to have a comprehensive and specific input domain to function properly as part of the CP system.

Defensive distillation has also been used in attempts to train NN models to resist adversarial attacks. Goldblum, Fowl, Feizi, & Goldstein (2019) presented adversarial robust distillation (ARD) which can help smaller NNs to lean robustness of a bigger model. According to Papernot & McDaniel (2016), defensive distillation is done by first training an NN model, where the last layer is a softmax-layer, with a labeled dataset. Then this model predicts new probability values from the training set. Using original input as input and output from the first NN, a second NN model (has softmax-layer) can be taught. However, there is a trick; a term called "temperature" in the softmax-layers. If the temperature is greater than 1, the probabilities get distributed more uniformly, meaning that for each class these probabilities are very nearly the same, when the temperature goes to infinity. If the temperature is 1, softmax-function will output probabilities closer to one that corresponds with most likely class labels. Both NNs are to be trained with the same temperature values that are greater than one, but after training of the second NN model is done, its temperature is set to one. According to a short article written by Papernot et al. (2016), defensive distillation can work against adversarial attacks. However, according to Carlini, & Wagner (2016), defensive distillation does not necessarily work against carefully constructed adversarial attacks.

Adversarial noise removal refers to techniques that can help reduce the effect of input noise typical to adversarial attacks. Gu & Rigazio (2015) tried both adding extra noise to image inputs and removing noise with the usage of autoencoders. They found that autoencoders work well for noise reduction but using them together with the original AI model leaves the compound model still vulnerable to even smaller adversarial noise.

According to Liang, Li, Su, Li, Shi, & Wang (2019), when inputs are images, varying de-noising techniques should be used based on the image space, since over and under de-noising should be avoided. Possible input transformation techniques can be bit-depth reduction, compression, total variance minimization, image quilting (Guo, Rana, Cissé, & van der Maaten, 2018), scalar quantization, and smoothing spatial filtering (Liang et al., 2019). Guo et al. (2018) managed to defend against 90 percent of black-box attacks, while Liang et al. (2019) managed to get a high 94 percent F1-score.

Sometimes if one model does not work, its performance may be possible to increase with an ensemble. Ensembles typically refer to a NN that takes outputs of other NNs or ML models as its input and calculates new output for the entire system. The beauty of an ensemble is that it can produce more acceptable results compared to a single ML -model. For example, Jia, Huang, Liu, & Ma (2017) devised an ensemble classifier model for the detection of DDOS attacks. This model consisted of Bagging, k-Nearest Neighbors (k-NN), and Random Forest, which were each trained and tested with cross-validation. Jia et al. (2017) reported that their model reached similar classification results as the Random Forest method on its beating Bagging and K-NN with a significantly higher true negative score. Other kinds of ensembles exist. Sengupta (2017) proposed a model that uses several NN models as defenders against adversarial attacks. This would make attacking a black-box model challenging, since defenders could alternate and therefore obscure decision boundaries. Tramèr et al. (2018) used an ensemble of different adversarial attack models to train a defensive NN. They showed that, in some cases, learned robustness from some attack could be transferred and it can be used similarly against other attacks.

8 Conclusion

Artificial Intelligence in protecting smart building's cloud service infrastructure represents a potential research area as the importance of cybersecurity in cloud services is growing. This chapter presented a general overview of cyberspace, artificial intelligence, cloud services, smart buildings, and typical attack vectors a perpetrator can utilize when attacking towards cloud services.

Cloud services are becoming even more popular these days due to the organization's interest in deploying applications and store data into cloud services. Cloud services can provide many benefits, but they also pose risks in security issues. Organizations need to be sure that robust encryption methods are used in storing and transferring data. In the smart home context, data can be gathered through IoT sensors, which are commonly known as vulnerable towards cyberattacks. Ordinary cryptographic encryption methods cannot always be used due to lack of processing power, storage space, and computing resources of IoT sensors. Lightweight cryptography methods can be applied, but they are less secure than conventional methods. Even proper authentication and data encryption alone is not enough for ensuring data security allowing perpetrators to use e.g., adversarial attacks when attacking the cloud service. In countering cyber-attacks, proper AI- and ML-models can be utilized.

Various solutions, which can be utilized in countering cyberattacks towards common attack vectors, exist. The solutions presented in this paper are Darktrace Vectra Cognito, PatternEx AI2 platform, Radware Defence Pro, and Reblaze. Vectra Cognito utilizing ML algorithms and mathematical principles in detecting anomalies can be used in tackling e.g., with insider threat issues. The solution produces behavioral models for individual people

and devices they use to detect unusual behavior and reveal possible insider attack. PatternEx AI2 platform can be used in predicting incoming cyberattacks, such as DoS and DDoS. AI2 utilizes unsupervised ML methods to present abnormal events to security specialists for further analysis to confirm attacks are real and produce a supervised model for further active learning. AI2 is estimated to reach 85% of detection accuracy. Radware Defence Pro offer means to detect and mitigate DDoS and IoT botnet attacks using ML algorithms to create signatures in real-time and adapt defenses. Defense Pro can separate anomalous attack traffic from legitimate one and to provide protection models ensuring zero time to mitigation. Reblaze's solution is to provide protection against DDoS botnet assaults until single malformed-packet DoS attempts. Reblaze's specialty is to provide more comprehensive protection than competing solutions against incoming DoS/DDoS attacks enabling ISO/OSI layers 3 (network) and 4 (transport) protection blocking incoming attacks towards cloud service. The solution integrates natively on public cloud service providers, such as Google Cloud Platform utilizing ML algorithms in self-learning and adapting under the continuous change in the cyber threat environment. The solution is able to detect early-stage attacks by using behavioral analysis and pattern recognition generating a small amount of false-positive results.

The results indicate that artificial intelligence can be used to prevent cyberattacks with some reservations. The architecture of chosen defensive AI model, defensive plan of the CP system, and how the model has been trained, determines how well artificial intelligence can combat attack vectors. Architecture and defensive plan of the CP system can help alleviate attacks on the CP system and the AI model, while under DDOS attack, for example, the system might start new defensive programs to mitigate load caused by the attack. When training models, utilizing different data manipulation schemes, such as adversarial training and defensive distillation, is important but not guaranteed to work perfectly. Classifiers trained with adversarial examples are more robust than classifiers trained with regular data, and this robustness can be transferred to other models. It was asserted that ensemble models could improve artificial intelligence performance, and it was found that ensemble training could do that as well. Training of models remains to be time-consuming.

Smart homes, smart building maintenance, and cloud services will get more popular over the years. Artificial intelligence is a huge part of that change. Artificial intelligence in protecting cloud services will gain more popularity in the future since current trends indicate that the number of cyber-attacks is increasing. For safety, various cyber threats towards cloud services should be researched thoroughly, and the adversarial attacks require further research.

References

Abate, A, Budde, C E, Cauchi, N, Hoque, K A, &Stoelinga, M. (2018). Assessment of Maintenance Policies for Smart Buildings: Application of Formal Methods to Fault Maintenance Trees. European Conference of the Prognostics and Health Management Society (2018).

Abomhara, M., & Køien, G. M. (2014, May). Security and privacy in the Internet of Things: Current status and open issues. In Privacy and Security in Mobile Systems (PRISMS), 2014 International Conference on (pp. 1-8). IEEE.

Alam, M, R, Reaz, M B I, Ali, M A M. (2012). "A review of smart homes—Past present and future", IEEE Trans. Syst. Man Cybern. C Appl. Rev.), vol. 42, no. 6, pp. 1190-1203, Nov. 2012.s

appliedAI. Accessed August 5th 2019https://appliedai.com/use-cases/1

Ballal S, Prasad L S, Rajappa M & Khader A. (2018). Bumper to bumper: detecting and mitigating DoS and DDoS attacks on the cloud. SecurityIntelligence. Accessed 16.8.2019 https://securityintelligence.com/bumper-to-bumper-detecting-and-mitigating-dos-and-ddos-attacks-on-the-cloud-part-1.

Baracaldo N, Chen B, Ludwig H &Safavi A. (2018). Detecting poisoning attacks on machine learning in IoT environments. IEEE international congress on internet of things (ICIOT), San Fransisco, CA, USA.

Bayuk J L, Healey J, Rohmeyer P, Sachs M H, Schmidt J, Weiss J. (2012). Cyber security policy guidebook, first edition. Wiley & Sons Inc, USA.

Bonderud D. (2018). Breaking bad behavior: can AI combat insider threats? SecurityIntelligence. Accessed 9.8.2019 https://securityintelligence.com/breaking-bad-behavior-can-ai-combat-insider-threats.

BS ISO/IEC 27002. (2013). Information technology – security techniques – code of practice for information security management.The British Standards Institution. BSI Standards Limited, Swizerland.

Buchanan W J, Li S & Asif R. (2018). Lightweight cryptography methods. Journal of cyber security technology, Vol. 1(3-4), 187-201.

Business.com. (2018) Cloud encryption: using data encryption in the cloud. Accessed 4.8.2019 https://www.business.com/articles/cloud-data-encryption.

Burnap, P. & Williams, M., L. (2015). Cyber Hate Speech on Twitter: An Application of Machine Classification and Statistical Modeling for Policy and Decision Making https://doi.org/10.1002/poi3.85 . Accessed August 7th 2019.

Ca technologies. (2018). Insider threat report. Cybersecurity insiders. Accessed 12.8.2019 https://www.ca.com/content/dam/ca/us/files/ebook/insider-threat-report.pdf

Castro-Leon E, Harmon R. (2016). Cloud as a service: understanding the service innovation ecosystem. Apress, USA.

Carlini N, Wagner D. (2016). Defensive Distillation is Not Robust to Adversarial Examples. ArXiv: 1607.04311v1 [cs.CR] 14 Jul 2016

Chang Z. Wireless and internet of things (IoT) security. Department of Mathematical Information Technology. University of Jyväskylä, Finland. Accessed 9.8.2019 users.jyu.fi/~timoh/TIES327/Wireless.pdf.

Costa D. (2017). CERT definition of 'insider threat'. Software engineering institute, Cargenie Mellon, University. Accessed 9.8.2019 https://insights.sei.cmu.edu/insider-threat/2017/03/cert-definition-of-insider-threat---updated.html.

Craigen D, Diakun-Thibault N, Purse R. (2014). Security in cyberspace. Targeting nations, infrastructures, individuals. Bloomsbury publishing, New York.

Cyberspace policy review. Assuring a trusted and resilient information and communications Infrastructure.https://www.energy.gov/sites/prod/files/cioprod/documents/Cyberspace_Policy_Review_final.pdf

Dao M-S, Pongpaichet S, Jalali L, Kim K, Jain R, Zettsu K. (2014). A Real-Time Complex Event Dixcovery Platform for Cyper-Physical-Social Systems. ICMR 2014, April 1-4, Glasgow, UK.

Deshpande N. (2009). Artificial intelligence. Technical publications. University of Pune, India.

Diro A A, Chilamkurti N. (2018). Distributed attack detection scheme using deep learning approach for Internet of Things. Future Generation Computer Systems, volume 82, May 2018, 761-768.

Darktrace. (2018). Darktrace enterprise – detects and classifies cyber-threats across your entire enterprise. Darktrace. Accessed 12.8.2019 https://www.darktrace.com/en/products.

Dremio. (2019). Enabling Data-as-a-Service for AWS and R.

Elleithy K, Blagovic D, Cheng W & Sideleau P. (2006). Denial of service attack techniques: analysis, implementation and comparison. Journal of Systemics, Cybernetics and Informatics. 3. 66-71.

ENISA. (2018). ENISA threat landscape report 2018 – 15 top cyberthreats and trends. European Union agency for network and information security.

EUR-Lex. (2013). Access to European Union law. Joint communication of the European parliament, the council, the European economic and social committee and the committee of the regions. Cyber Security strategy of the European Union: an open, safe and secure cyberspace. Document number 52013JC0001.

Ganin Y, Ustinova E, Ajakan H, Germain P, Larochelle H, Laviolette F, Marchand M, Lempitsky V. (2016). Domain-Adversarial Training of Neural Networks. Journal of Machine Learning Research 17, 2016, 1-35.

Gavai G, Sricharan K, Gunning D, Hanley J, Singhal M, Rolleston R. (2015). Supervised and Unsupervised methods to detect Insider Threat from Enterprise Social and Online Activity Data. JoWUA, Volume 6, number 4. DOI:10.22667/JOWUA.2015.12.31.047

Ghahramani, Z. (2015). Probabilistic machine learning and artificial intelligence. Nature volume 521, pages 452–459 (28 May 2015).

Gillespie A. (2016). Cybercrime – Key issues and debates. Routledge, New York.

Goel N, Sharma T. (2014). Cloud computing – SPI framework, deployment models, challenges. International journal of emerging technology and advanced engineering. International conference on advanced deployments in engineering and technology, India.

Goldblum M, Fowl L, Feizi S, Goldstein T. (2019). Adversarially Robust Distillation. ArXiv:1905.09747v1 [cs.LG] 23 May 2019.

Goodfellow I, Papernot N, Huang S, Duan R, Abbeel P & Clark J. (2017). Attacking machine learning with adversarial examples. OpenAI. Accessed 3.8.2019 https://openai.com/blog/adversarial-example-research

Gu S, Rigazio L. (2015). Towards Deep Neural Network Architectures Robust to Adversarial Examples. ArXiv:1412.5068v4 [cs.LG] 9 Apr 2015.

Guo C, Rana M, Cissé M, van der Maaten L. (2018). Countering Adversarial Images Using Input Transformations. ArXiv:1711.00117v3 [cs.CV] 25 Jan 2018.

Gupta D, Ghakraborty P S & Rajput P. (2015). Cloud security using encryption techniques. International journal of advances research in computer science and software engineering, 5(2), SRM University, India.

Han B, Yang X, Sun Z, Huang J, Su J. (2018). OverWatch: A Cross-Plane DDOS Attack Defense Framework with Collaborative Intelligence in SDN. Hindawi, Security and Communication Networks, Volume 2018. https://doi.org/10.1155/2018/9649643.

HP security research. (2014). Internet of things research study. Accessed 7.8.2019 http://d-russia.ru/wp-content/uploads/2015/10/4AA5-4759ENW.pdf

Iyer, G, Agrawal, P. (2010). Smart power grids, 42nd Southeastern Symposium on System Theory (SSST), IEEE (2010), pp. 152-155.

Jia B, Huang X, Liu R, Ma Y. (2017). A DDOS Attack Detection Method Based on Hybrid Heterogenous Multiclassifier Ensemble Learning. Hindawi, Journal of Electrical and Computer Engineering, Volume 2017, https://doi.org/10.1155/2017/4975343.

Jiang, J-R. (2018). An improved cyber-physical systems architecture for Industry 4.0 smart factories. Advances in Mechanical Engineering, 2018, Vol. 10(6), 1-15.

Karthikeyan P, Thangavel M. (2018). Applications of security, mobile, analytic and cloud (SMAC) technologies for effective information processing and management. A volume in the advances in computer and electrical engineering (ACEE) book series. IGI Global, USA.

Kubovič O, Košinár P &Jánošík J. (2018). Can artificial intelligence power future malware? ESET white paper.

Le D C, Khanchi S, Zincir-Heywood A N, Heywood M I. (2018). Benchmarking Evolutionary Computation Approaches to Insider Threat Detection. Association for Computing Machinery. https://doi.org/10.1145/3205455.3205612.

Legatiuk D, Smarsly K. (2018). An abstract approach towards modeling intelligent structural system. 9th EWSHM, UK. CC-BY-NC license 4.0.

Lehto M. (2008). Phoenomena in the cyber world. Cyber security: analytics, technology and automation. Springer, Berlin.

Lehto M. (2015). Phenomena in the cyber world. Cyber security: analytics, technology and automation. Springer, Berlin.

Liang B, Li H, Su M, Li X, Shi W, Wang X. (2019). Detecting Adversarial Image Examples in Deep Neural Networks with Adaptive Noise Reduction. ArXiv:1705.08378v5 [cs.CR] 9 Jan 2019.

Libicki MC. (2007). Conquest in cyberspace – national security and information warfare. Cambridge University press, New York

Lim, C, Kim, K-H, Kim, M-J, Heo, J-Y, Kim, K-J, & Maglio, P P. (2018). From data to value: A nine-factor framework for data -based value creation in information-intensive services. International Journal of Information Management 39, 2018, 121-135.

Machmeier C &Kunzke F. (2019). How safeguarding sensitive data could lead to smarter AI. Sap News Center. Accessed 4.8.2019 https://news.sap.com/2019/01/homomorphic-encryption-safeguarding-sensitive-data-smarter-ai.

Mather T, Kamaraswamy S, Latif S. (2009). Cloud security and privacy – an enterprise perspective on risks and compliances, O'Reilly Media Inc., USA.

Nagpal, H, Basu, B, & Staino, A. (2018). Economic Model Predictive Control of Building Energy Systems in Cooperative Optimization Framework. ICC, January 4-6, 2018, IIT Kanpur, India.

Nicholson, C. Skymind. Accessed August 5 2019 https://skymind.ai/wiki/ai-vs-machine-learning-vs-deep-learning.

Qi, Q, Tao, F, Zuo, Y, Zhao, D. (2018). Digital Twin Service towards Smart Manufacturing. Procedia CIRP 72, 2018, 237-242.

Padmavathi & Shanmugapriya et al., (2009). A survey of attacks, security mechanisms and challenges in wireless sensor networks, arXiv preprint arXiv:0909.0576.

Papernot N, McDaniel P. (2016). On the Effectiveness of Defensive Distillation. ArXiv:1607.05113v1 [cs.CR] 18 Jul 2016

Papernot N, McDaniel P, Goodfellow I, Jha S, Celic Z B, Swami A. (2017). Practical black-box attacks against machine learning. Proceedings of the 2017 ACM Asia conference on computer and communications security, Abu Dhabi, UAE.

Potluri S, Diedrich C. (2016). Accelerated Deep Neural Networks for Enhanced Intrusion Detection System. 2016 IEEE 21st ETFA. 10.1109/ETFA.2016.7733515

Radware. (2018). Machine-learning automation to ensure zero time to mitigation. Accessed 16.8.2019 https://www.radware.com/pleaseregister.aspx?returnurl=732862c3-5149-4806-b060-ba20d2bca6eb

Randal, L. (2016). What is data as a service? The 3 key dimensions. BDQ big data quarterly. http://www.dbta.com/BigDataQuarterly/Articles/What-is-Data-as-a-Service-The-3-Key-Dimensions-114568.aspx

Rangaraju N K, Sriramoju S B, Sarma S. (2018). A Study on Machine Learning Techniques Towards the Detection of Distributed Denial of Service Attacks. International Journal of Pure and Applied Mathematics, Volume 120, No. 6 2018, 7407-7423.

Reblaze. (2019). Comprehensive DDoS protection DoS/DDoS datasheet – web application & API security. Accessed 22.8.2019 https://www.reblaze.com/wp-content/uploads/2019/05/Reblaze-DDoS-Datasheet.pdf

Schmidt M, Åhlund C. (2018). Samrt buildings as Cyber-Physical Systems: Data-driven predictive control strategies for energy efficiency. Renewable and Sustainable Energy Reviews, Volume 90, 742-756. https://doi.org/10.1016/j.rser.2018.04.013

Sengupta S. (2017). Moving Target Defense: A Symbiotic Framework for AI & Security. Proc. of the 16th International Conference on Autonomous Agents and Multiagent Systems.

Shahapure N H, Jayarekha P. (2015). Replication: a technique for scalability in cloud computing. International journal of computer applications (0975 – 8887), 122(5).

Sta, H B. (2017). Quality and the efficiency of data in “Smart-Cities”. Future Gener. Comput. Syst., 0167-739X, 74 (2017), pp. 409-416.

Shaw, E., Ruby, K. & Post, J. (1998). The Insider Threat to Information Systems. Security Awareness Bulletin, 2(98).

Threat Analysis Group (Tag). (2010). Threat, vulnerability, risk – commonly mixed up terms. Accessed 31.7.2019 https://www.threatanalysis.com/2010/05/03/threat-vulnerability-risk-commonly-mixed-up-terms.

Tramèr F, Kurakin A, Papernot N, Goodfellow I, Boneh D, & McDaniel P. (2018). Ensemble Adversarial Training: Attacks and Defenses. ICLR 2018.

Usman S H, Bawazir M A, Kabir A M. (2014). Cloud computing: a strategy to improve the economy of Islamic societies. International journal of computer trends and technology (IJCTT), 9(7).

Vectra cognito. (2019). Vectra security that thinks. Artificial intelligence powered automated threat hunting and network self-defense. Accessed 12.8.2019 https://www.beotech.rs/wp-content/uploads/2019/02/Vectra-Cognito-DataSheet.pdf.

Wang, Z, Srinivasan, R S. (2017). A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models. Renewable and Sustainable Energy Reviews 75, 3027, 796-808.

Wani S Y. (2018). Internet of things (IoT) security and vulnerability. Research proposal. DOI:10.13140/RG.2.2.29633.40801.

Zhang D, Zheng Y, Wen Y, Xu Y, Wang J, Yu Y, Meng D. (2018). Role-based Log Analysis Applying Deep Learning for Insider Threat De tection. Association for Computing Machinery, SecArch’18. https://doi.org/10.1145/3267494.3267495.

Zhao D, Zhong M, Zhang X, Su X. (2016). Energy consumption predicting model of VRV (Variable refrigerant volume) system in office buildings based on data mining. Energy 2016, 102, 287-97.