Comparative study of the Security Analysis of IoT systems using attack trees algorithm

RAFALIA Najat[1], ETTAZI Haitam[1], ABOUCHABAKA Jaafar[1]
[1]Faculty of Sciences, University of Ibn Tofail, Kenitra, Morocco

Abstract. The Internet of Things (IoT) is a rapidly evolving environment that allows users to use and control a wide variety of connected objects. The 20 billion IoT devices that will be employed by 2020 are only the top of the iceberg. According to IDC, the overall amount of connected devices will rise to 41.6 billion over the next five years, producing over 80 Zettabytes of data by 2025 which will impact environment severely. These connected environments increase the attack surface of a system since; the risks are multiplied by the number of connected devices. These devices are responsible for more or less critical tasks, and can therefore be the target of users malicious, in this paper we present a methodology to evaluate the security of IoT systems. We propose a way to represent IoT systems, coupled with attack trees in order to assess the chances of success of an attack on a given system.

Index Terms— IoT Systems;Environment Factors; Security; Attack Trees modeling; Machine learning; AI.

1 Introduction

We often see the Internet of Things (IoT) as an upcoming revolution. We imagine a world in which information is accessible everywhere and human intervention on tedious tasks is rendered useless. In your everyday life, this means your home cleans itself and you don't need to stop worry about weekly shopping. Depending on your consumption habits, if you expect whether or not you have guests and taking into account your guests tastes and the occasion, your fridge will order the products. Your groceries will then be delivered, regardless of your presence at home, thanks to your connected lock. If you were to do anyway shopping, queues would be a thing of the past. You can simply enter the store, grab what you need and leave. Your account is automatically charged and you are notified, according to your preferences. Driving to work is not much of driving as we see it now: your car smart does most of the work. You can prepare your meetings from the day in the morning or finish your work while checking your children's homework on the way back. Being late because of traffic is no longer possible since it is regulated taking into account the main home-work travel patterns, extracted personal data of commuters. No longer slowed down by the garbage truck. The garbage is now automatically routed to the nearest sorting center and treated according to its nature. Hospitals are never crowded because a large majority of patients are treated at home, via a real-time link with health personnel. Your doctor has continuous, real-time access to metrics such as your fitness and performance, your caloric intake, daily activities, your schedule and quality of sleep, etc [1].

On the industry side, smart farming allows fields to be ploughed and harvested without human physical intervention, reducing costs, accidents and increasing efficiency. During

this time, humidity, air temperature, soil quality and real- time forecast data would be analysed to create a program optimal treatment and watering. Once the grain has been harvested, the stocks are distributed according to the needs of use and ready to be transmitted. Supply Chains being equipped with appropriate sensors, trackers and other devices, the goods are conveyed in complete safety, using the most optimal means. [1] Time monitoring real allows users to follow them at any time and ensure that the conditions of storage remain compatible with the constraints of the product.

IoT Security Issues IoT security is extremely important because everything a smart device uses can serve as an entry point for cybercriminals to access the network. Once the hackers gain access through a device, they can then move laterally throughout the organization, accessing high-value assets or conducting malicious activities, such as theft of data, IP addresses or information sensitive. In some attacks, such as a denial of service (DoS) attack, the[6] cybercriminals will take control of the device and use it to overwhelm the server traffic jams, preventing legitimate users from carrying out normal activity. Traditionally, organizations and consumers protected their devices through a range of security measures, such as anti-virus software and firewalls. However, these measures may not be suitable for the protection of IoT devices, as many cannot support the processing and storage requirements of these tools. As such, there is a need for organizations to develop a comprehensive Cyber security strategy that protects against a wide range of cyber-attacks on all devices, both terminal and network level [2].

Unfortunately, many IoT devices are not designed with security. In many cases, these devices lack the processing power and capabilities storage required to support installation of additional security on the device itself, which means that businesses and users cannot protect the terminal beyond existing security features. Instead, the businesses must rely on network security capabilities to prevent attacks, as well as to detect and remediate threats as they arise. Even devices that support the installation of additional security measures may not be compatible with existing cyber security toolset of the company [2]. Disparate operating systems and a variety of guaranteed hardware weave almost that the organization will not be able to protect all devices connected using the same tools, policies and procedures In addition, devices IoT, like traditional terminals, requires patches and updates to the operating system. The large number of connected devices makes it difficult to manage of this activity by organizations, particularly if the devices belong to employees [3].

Connected devices may not require solid password practices- a point that is compounded by the fact that many people underestimate the risk posed by non-traditional connected devices [4].

This work addresses security issues in distributed environments, and more especially in the Internet of Things. Our goal is to represent the systems IoT in an abstract but realistic way as well as the possible attacks against them. These representations are then used to perform statistical analyzes in order to calculate the chances of success of attacks on systems. Therefore, this work can be divided into three main aspects: the model of the system, the representation of the attacks and statistical analysis. This chapter reflects these three main tracks that make up this thesis. We start by giving an overview of the Internet of Things in the section 2.1, before addressing the security challenges it currently faces. The section 2.2 presents several ways to represent attacks [5].

Finally, section 2.3 present different ways to perform statistical analysis methods for rare events as well as similar approaches [4]. And finally, I'll present my attribution and

comparative study [5] where a framework wad dedicated to password and authentication security [6].

2 Material and Methods

2.1 Modeling techniques for the Internet of objects

The Internet of Things (IoT) can be defined as[3] a "comprehensive presence around us present in a variety of things or objects which, thanks to unique addressing patterns, are able to interact with each other and cooperate with their neighbors to achieve common goals. They also define different IoT paradigms, depending on the scientific community that considers it. But the children games that arise with the ubiquity of the IoT are the same for everyone and relate [11], among other things, to security and confidentiality [7].

This section deals with existing work related to these security issues, and in particular securing the representation of systems. We start by giving an overview of security oriented representation of IoT systems, subsection

Then, in subsection 2.1.2, we give an overview of the existing languages to represent IoT systems, whether they are security oriented or not [8].

2.1.1 IoT and security

IoT systems [art4] are composed of three successive layers: the layer of perception, the layer of transport and the application layer. The perception layer concerns the collection of information, the perception and object control. The transport layer is responsible for providing access to the first layer, and data transmission. Finally, the role of the application layer is to support business services and perform intelligent calculations and allocation resources. This vision of the IoT systems makes it possible to detail the functionalities and security requirements, and therefore to adapt the security solution, depending on the level device target. All the descriptive aspect of this work is based on the technical description system architecture. Indeed, by dividing the global architecture into layers, it becomes easier to understand and deal with vulnerabilities, depending on their level. On the other hand, detecting cross-layer vulnerabilities requires seeing the system as a whole and, in this case, finding the appropriate security measures could be more hard. Due to their interconnected nature, IoT systems are vulnerable to a wide spectrum of attacks. Indeed, it must be considered that a single unsecured device may affect the security of the overall system. A way to identify new surfaces attack using a visual representation has been proposed in [5]. In this work, IoT security is not seen as a binary concept but rather as a[8] spectrum of device vulnerability. They then propose a new visual grammar to describe IoT systems at a high abstract level. Their representation is as a three-level structure constructed as follows: the first layer is the high-level representation that uses visual grammar to provide a model of the IoT system. The second layer shows object profiles and logs the first and the third layer which provides system implementation details. The second layer maps the high level representation to the low level representation which gives us technical details about the connected objects that make up the system. To build the high-level model, they use the device characterizations of all objects that are part of the system. Then they mapped the device descriptions on one or more existing devices, and finally, they generate a representation of high level [9].

2.2 IoT language representations

In this subsection, we separate two types of representations: the one that has been created for security purposes and others.

2.2.1 IoT modeling

The goal is to implement a language using a visual representation based on UML notation[art5] powerful enough for professional needs and still being user friendly and easy to read. In this language, the basic element of an IoT system is called a "thing". The thing can be real or virtual. When things are arranged stored in the collections, they make up a "subsystem". Things can contain elements that can be inputs, outputs or components. All these concepts which include pose the language have a virtual representation inspired by class diagrams UML. Even though the virtual representation is not security oriented, as

[9] SOML (Security Oriented Modeling Language), we could take inspiration from it to create a visual representation of our language. Indeed, for the moment, to describe a system using SOML, it is necessary to know the syntax of the language. By providing an interface with visuals, we might as future work consider automatically generating system models from a visual interface. Another language, GroupeSens-L [1] proposes to represent the Activity Recognition of Band. The objective is to model the physical activity of a group using an EBNF modeling language. It makes it possible to model the activity of the group with its conditions and constraints vary the level of detail and generate a billing model [10].

We use a meta model to help developers build their IoT system through a model-driven development process. From of the meta-model, the developer builds his own model and can automatically generate only Java code, which is intended to be used as a starting point for development. Implementation of its IoT system. The meta model is built around "Human Object View". This means that it represents the interconnection between human connected objects and physical objects as well as how humans use them to extend their communication skills. This point of view shares our vision of communication between man and objects, but since it is not security oriented, it does not appear of verification, of the critical aspect of information exchanges, which is essentially our work. Another approach is the IRON [art5] language, which is an ECA (Event- Condition- Action) with formally defined semantics, intended to identify behaviors of an IoT system. The IRON language has a static part and a dynamic part. The static part concerns variable declarations. Variables are the connected devices which are described using an identifier and a name. The part static also defines the constraints that ensure the validity of the information exchanged. The the dynamic part concerns the ECA rules which describe the behavior of the devices. An Event-Condition-Action rule indicates how an action is performed during an event when the condition is met. Formal semantics is capable of assuring the reliability of the results of the execution. Even if this modeling language allows the execution of the IoT model to guarantee the correct functioning of the system, the errors are only considered from a logical point of view. Attacks against the system, as well as malicious users or actions are not taken into account [11].

2.2.2 IoT secure representations

The IOT Standard Specification and Description Language (SDL) art6 is a language executable intended to model and simulates the operation of IoT or other systems distributed environments. Using SDL, one can model an IoT system as a set of connected agents that communicate with each other. The SDL simulator allows checking the behavior

of the system before its deployment. The abstract machine the semantics of the language makes it possible to define the communication architecture of the system (communication elements with different levels of abstraction) and behavior complete system. But if SDL makes it possible to model the behaviors of the different actors in the system, whether malicious or desired, it does not address the attacks against the system. Its only concern is to detect account security issues maintained the normal operation of the system. This approach and our work have a close similarity. First of all, it uses a high-level model of the system and it allows you to perform simulations in order to find anomalies in its behavior. But our approach aims to simulate one or more external actors attacking the system and watch their progress. SDL does not take into account possible external attacks. The security issues that are considered here are those that may occur during the normal operation of the system. In this case, the model works like a realistic replica of the system and allows problems to be corrected before deploying the environment. IoTSec art7 is a UML-based language intended to encapsulate knowledge security studies to model IoT systems. The work uses the approach model-based system engineering. It is based on the idea that to solve problems security issues in IoT, we need to do security design as well as an analysis of vulnerabilities and threats before implementation. This extension sion UML/SysML redefines several UML diagrams to model the problems of security in IoT systems such as class diagram, sequence diagram, the state machine diagram,

... This work tends to respond to the security problems of the system by modeling them during its design. Models are not intended to be executed and display vulnerabilities and potential threats. These techniques cannot therefore not be used in our approach [12].

2.2.3 Attack representations

Model attacks, threats, vulnerabilities, etc. anything that can be a risk for the good the functioning of a system has always been a real concern in research. The objectives can be either to prevent it, to detect it or as is our case: estimate the probabilities for an attack to occur. There is a variety option with different possible applications. According to art8, we separate them into three main categories: representations, Attack Graph and others [13]

2.2.4 Tree-related representations

This subsection describes all attack representations that use an attack representation. Tree-like feeling. We start first with the attack trees, which are the way we have chosen to represent the ongoing attacks against the system that we are auditing. Indeed, attack trees have the advantage of being easy to understand thanks to their intuitive representation. It is also possible to modify the quantity information they contain, depending on our needs for analysis. Several alternatives, based on attack trees exist such as defense trees, threat trees which we detail in this subsection. They prolong the attack trees with information such as defense information or economics-index for further analysis. In this work we did not need to extend the attack tree because the information necessary for the analysis is contained in the system model. But it is possible as future work to consider conducting a different representation or deep analysis on the IoT systems. In this case, another tree representation can be used, which contain different types information's about attacks against the model.

2.2.5 Attack tree

The attack trees art9 are defined as a formal way of describing the security of a computer system and to be used to make security decisions. One of the advantages was then to be able to easily scale attack trees as has been developed and security issues have changed the purpose of an attack tree is to cover more or less, on the needs, all possible attacks we can

observe on the system. an example of attack tree defined by [11] exhaustively in figure 1.1, according to the needs, the possible attacks against a system. The root node represents the goal of the attack: it is a question here of how to open a safe. There are several ways to achieve this goal, such as crocheting the lock or cut the safe open. Annotations on tree nodes can differ according to need. Following the work of [11,12] propose a formal semantics for attack trees art11. In their work, they formalize the concepts introduced by Schneier and define writing rules to standardize the creation of attack trees. This can be useful, for example in threat analysis computer aided. Due to their simplicity, attack trees can be used in various environments and they allow users to get an overview overall attack. But because of this, it can be argued that attack trees do not have a sufficient level of technical specificity to enable them to represent attacks directly against cyber threats art12. But T. Tidwell et. Al. art13 found a way to overcome this problem. Indeed, they offer an attack visualization system using trees in which they are used to represent attacks Internet qualified as "multi-stage".

For our case study, we base ourselves on the attack tree proposed in the article to be able to structure intelligent approaches to dedicate to fortifying the faults that arise in nodes, and integrate frameworks for each layer of attack to fortify connected hardware, and defend the cloud server they are connected to. For the application we will attack the lowest node in the tree which concerns authentication administrator, the most exploited flaw of IOT machines, for this we are working on a password prediction and generation model secure and strong against tools of modern 'Brute force' attack, that represents the cause of more than 90% of infected devices, using the decision tree/regression algorithm tested on a database of 0.7 million reel passwords made public after the 2014 webhost hack.

2.2.6 What is IoT

Internet of Things, refers to the process of connecting physical objects Internet, everyday objects such as thermostats, medical devices, portable devices, smart devices, etc.

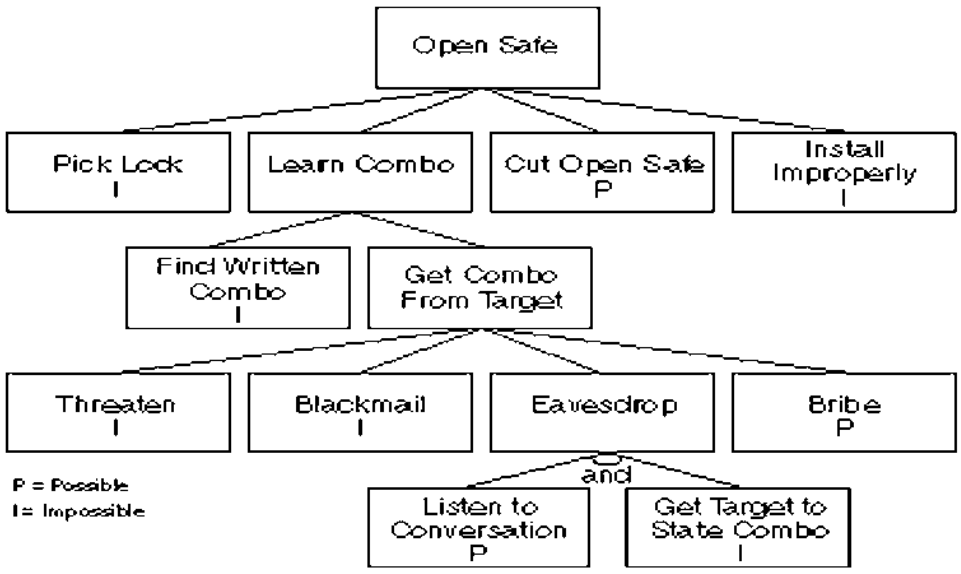

Fig. 1. Example of attack tree by Bruce Schneier

2.2.7 How does it work?

The Internet of Things refers to physical devices that send and receive data over wireless networks, with limited human manipulation. This technology is based on the integration of an IT infrastructure with all kinds of objects. For example, connected air conditioning (the adjective "connected" often refers to the Internet of Things) receives Geo-localization data transmitted from your connected car when you get home from work or your mobile phone, and adjusts the temperature of your home before you arrive. No intervention on your part is necessary, and the result is much better than if you had manually adjusted the thermostat. A classic IoT system, such as the connected home, works by sending, the constant reception and analysis of data according to a feedback loop. Depends- Due to the IoT technology used, the analyzes can be performed by human beings. hands or by an artificial intelligence dedicated to automatic learning (IA/AA), in near real time or over a long period. Let's go back to the example of the connected home. To consider the optimal time to trigger the air conditioning before you return your IoT system can connect to the API Google Maps, and thus obtain real-time traffic modeling in your area. Data relating to your usual journeys that your car has collected over a long period of time. period can be used. The IoT data collected by the air conditioning of each customer can be grouped into customer clusters and analyzed by energy distributors to optimize their services at scale.

2.2.8 IoT in the business word

The public masses, however, remain wary of the Internet of Things, as their experiences with technologies such as smart watches are affected by privacy and security issues intrinsic to always-on connectivity. Companies must remain vigilant in this aspect for all types of IoT projects they are considering, especially if the project is intended for the general public. IoT solutions business models improve existing business operating models and to create new connection channels with customers and partners. However, some problems also arise. Of course, the amount of data generated by a range of connected devices can become very broad (we often speak of of Big Data). Integrating Big Data into existing systems and analyzing it can turn out to be a heavy task. IT security can be positioned as a critical pillar when developing IoT systems. Even so, the benefits that the Internet of Things provides to businesses largely cover the efforts necessary for its implementation and, at present, companies in all fields of activity are already using it, with success.

2.2.9 IoT and edge computing

The public masses, however, remain wary of the Internet of Things, as their experiences with technologies such as smart watches are affected by privacy and security issues intrinsic to always Edge computing makes it possible to reinforce the computing power at the periphery of a network IoT, to reduce communication latency between IoT devices and networks central computers to which these devices are connected. The ability of devices to harness computing power is becoming extremely interesting when you need to quickly analyze real-time data. If the simple reason for see sending or receiving data marked the advent of the IoT, it is the ability send, receive and analyze data using IoT applications which turn out to be essential for the future. In the cloud computing model, computing resources and services are generally grouped and centralized in large data centers. These data centers are accessible by objects connected via the network. This model reduces the costs and share resources more effectively. However, in order for an infrastructure IoT to be effective, computing power must be reinforced closer to the real location of the physical device. Edge computing distributes computing resources at the edge, while other resources are centralized in the cloud. This specific location of resources formats provides quick access to actionable in- sights based on data varying over time. The coordination of a fleet

of autonomous vehicles that transport and carry containers equipped with connected tracking devices is a somewhat extreme, but there are many other practical applications, such as improving health care through point-of- care data analytics. Let's take as an example RFID tags in the transport sector: communication between the device RFID and the reader is always unidirectional. The RFID device cannot receive updates, and a central computer network would not be able to send back data to the RFID device. This constraint limits the tracking of the container upon arrival in some places and does not allow continuous monitoring. If the device can synchronize with IoT sensors installed in the vehicle carrying the container, all data can be managed in a central computer network. To realize this scenario, each physical IoT device should have a large computing power, especially if the company uses complex machines such as than autonomous vehicles. IoT devices will no longer be just devices data exchange systematically waiting to receive instructions from a data center centralized data via a Wi-Fi connection. They will be able to process the data and make informed decisions independently. The deployment of power computing at the edge of the network than in an centralized data center is called edge computing Let's take a final example: a construction company brings in a machine equipped of Bluetooth on a construction site. This machine sends data via the worker's smartphone, allowing the company to track their tasks and localization; If 10 people work around this device all day, their smartphones will continuously interrogate the central server to determine its location. Such activity may overload the computer system. However, an app Mobile IoT can use a smartphone as a small, low-power server energy and unnecessarily reduce the need for a central server.

2.2.10 Security in IOT

The Internet of Things (IoT) has changed the way we interact with the world that surrounds us. Many devices switch from offline to online mode, connecting with each other through the Internet, providing more features for users. Despite the increase in user quality of life provided by IoT devices, it is also necessary to establish trust in the privacy and security of final users. With this level of connectivity, the amount of data exchanged between devices is also increasing, leading to malicious activities. One of the main problems is the lack of regulation in the IOT industry, especially between different manufacturers. There are no formal safety regulations and manufacturers may choose not to install security mechanisms. Therefore, it is necessary to promote the adoption of security measures. One way to achieve this is through device and IoT system certification. In recent years, IoT certifications have emerged. Meanwhile, the European Union passed the Cyber Security Act to unify and regulate the certifications of security in member states

IoT security is extremely important because any smart device can serve as an entry point for cybercriminals to gain access to the network. Once the opponents have access via a device, they can then move laterally throughout the organization, accessing high-value assets or carrying out malicious activities, such as the theft of data, IP addresses or sensitive information. In some attacks, such as a Denial of Service (DoS) attack, cybercriminals take control of the device and use it to overwhelm servers with traffic web, preventing legitimate users from carrying out normal activity. Traditionally, organizations and consumers protected their devices through a range of security measures, such as anti-virus software and firewalls. However, these measures may not be suitable for the protection of IoT devices, because many cannot support processing and storage requirements of these tools. As such, there is a need for organizations to develop a comprehensive cyber security strategy that protects against a wide range of cyber-attacks on all devices, both at terminal and network level.

2.2.11 IoT security risks

Unfortunately, many IoT devices are not designed with security. In many cases, these devices lack the processing power and capabilities storage required to sup- port installation of additional security on the device itself, which means that businesses and users cannot protect the terminal beyond existing security features. Instead, the businesses must rely on network security capabilities to prevent attacks, as well as to detect and remediate threats as they arise. Even devices that support the installation of security measures additional may not be compatible with existing cyber security toolset of the company. Disparate operating systems and a variety of guaranteed hardware weave almost that the organization will not be able to protect all devices connected using the same tools, policies and procedures. Additionally, IoT devices, like traditional terminals, require patching and operating system updates. The large number of connected devices makes difficult for organizations to manage this activity, especially if the devices belong to employees. Finally, connected devices may not require solid password practices, a point that is compounded by the fact that many people underestimate the risk posed by non-traditional connected devices Since there is no single security tool capable of providing uniform and complete across all connected devices, IoT security requires mix elements from both the endpoint security policy and the policy cloud security. The following features can help keep all devices safe connected and are considered a necessity for all modern organizations:

•*Prevention*: Next Generations Antivirus (NGAV) uses advanced technologies, such as AI and machine learning, to identify new and emerging threats by examining more things, such as file hashes, URLs, and IP addresses.

•*Detection*: Endpoint detection and response (EDR) End- point Detection and Response (EDR) is a solution that provides continuous complete and global visibility on what is happening on the terminals in real time. Companies must look for solutions that offer advanced detection and investigation capabilities threat and response, including research and investigation of incident data tooth, alert triage, suspicious activity validation, threat hunting and detection and containment of malicious activity.

•*Manage*: Threat Hunt Managed Threat research is conducted by elite teams that learn lessons reports of incidents that have already occurred, aggregate data from crowd-sourcing and provide guidance on how best to respond when a malicious activity is detected.

•*Threat Intelligence Integration*: To stay one step ahead of attackers, organizations need to understand respond to threats as they evolve. Sophisticated adversaries and Advanced Persistent threats (APT) can move quickly and stealthily, and the security teams need up-to-date information to ensure that defenses are set automatically and precisely.

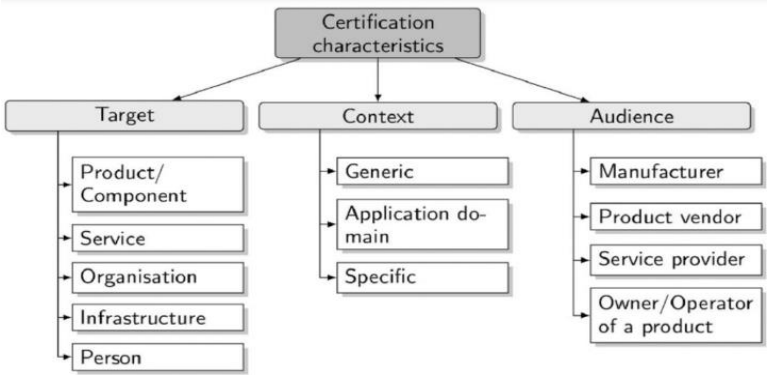

Fig. 2. Certification characteristic by Bruce Schneier

3 Material and Methods

3.1 Comparative Analyses

After this analysis of the previous work we saw, that represents the level where the security of IOT is currently at, and the various threats and vulnerabilities alongside the different techniques and approaches used on different levels of a particular connected system, we can gather a wide and intensive picture of the layers of security present in the IOT environment, therefore have a work-plan model representing all aspects of the architecture and the potential gateways to the system.

As we know by now, there are 3 layers to IOT systems: Cloud layer, communication layer and perception layer.

3.1.1 Cloud layer

Attribute-based access control (ABAC) is an access control method where subject requests to perform operations on objects are granted or denied based on assigned attributes of the subject, assigned attributes of the object, environment conditions, and a set of policies that are specified in terms of those attributes and conditions. ABAC has three main functional points, which is as follows:

- The PEP or Policy Enforcement Point: is a component that serves for protecting apps data in which ABAC is applied. The PEP inspects the request and describes the user's attributes to the Policy Decision Point PDP.
- The PDP or Policy Decision Point: is the component that makes the determination of whether a user's request is authorized or not by evaluating incoming requests against policies it has been configured with. The PDP returns a Permit / Deny decision. The PDP may also use PIPs to retrieve missing metadata.
- The PIP or Policy Information Point: serves as the re- trivial source of attributes and bridges the PDP to external sources of attributes e.g. LDAP or databases.

The proposed architecture for the authentication of things in IoT-Clouds is a combination of ABAC Technology and PKI Infrastructure (Figure 3). It forces a smart object in IoT to pass a double check authentication system to ensure that the data is collected from the correct IoT object and not from a fake one. For the users that need to get access the IoT data stored in the cloud, we used another multi-authentication factor which is the mobile two factors authentication MPTFA.

3.1.2 Communication layer

There are various cryptographic techniques [11] that can be used to ensure high communication performance and contribute in an efficient and secure searching for shared data over the smart devices of IoT. Here we focus on the various threats while sharing data among smart devices of IoT. Implicit threats generated by the system itself because of malicious functioning and explicit threats that are generated by the unauthorized users using the devices are one more issue.

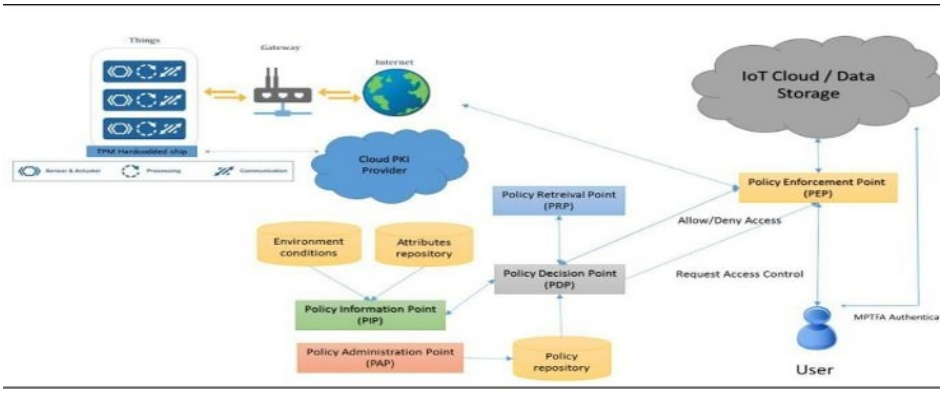

Fig. 3. Proposed Architecture

Cryptographic Technique	*Description*
Secret Key Encryption	By using secret key the user will send and receive secured data. Devices using secure communication principals.
Public Key Encryption	It is a two key mechanism, public key and a secret key. Public key can be used before sending data and secret key is used for decrypting the data.
Searchable Secret Key Encryption	It uses secret key by using trapdoor for authorized user devices only.
One Way Hash Algorithms	It is used for integrity check with hash functions i.e. if any data is modified between sender and receiver.
Digital Signature	Public and secret key are operated by the authorized users with digital signatures.

Fig. 4. Cryptographic Techniques Description

3.1.3 Perception layer

Authentication plays a key role in protecting our systems and databases. But when the process is bypassed by a hacker or an attacker, the whole system becomes vulnerable so, it's always crucial to secure data. The following is the types of authentication attacks:

- Bypass attacks: The attacker's first intention is to bypass the security and authentication is no exception. If a hacker bypasses the authentication, it is not his credit; it is either the software system's failure to impose access policies or a poorly designed authentication system. Let's have an example, a code written by a programmer is capable of performing authentication using users' credentials may enforce strict password input, but may fail to block a blank password, which may endanger our data. In another case, the code developers place the protected and unprotected files in a same folder by mistake. This makes the site more vulnerable for attacks by displaying the code easily. These Bypass attacks are usually tried in case of customized authentication systems not the robust industry standard ones where tunneling is more difficult.

- Session eavesdropping: An attacker can use packet- capturing tools that intercept sessions and decipher credentials while authentication is in progress. Most probably, websites that use simple HTML forms of authentication with no SSL will have simple text passwords that are easy to track. The same case may be used in alternate form

where hacker can use session takeover attack. Here, the valid token received after authentication is captured and used to personify the victim.

• Server side authentication attacks: there are also challenges when a web server or LDAP server is performing authentication. The attacker would try to steal user's credentials like username and passwords and login to the main authentication server and steal the entire data in database. Usually, the server side attacks involve the injection of scripts or installation of a Trojan to open the ports. Attacker now uses those ports to run commands on server to fetch required information. Today, modern websites use SQL based authentication that internally use SQL back-end for validation. Insecure web requests or running malicious scripts can let attacker to steal your SQL data.

• Brute force attacks: We have a single level security even today which is our passwords. The attackers don't just guess our password as we do in case of cracking our friend's password. They use scripts or customized software that contains every possible combination against authentication system. The combinations may include all the words in English, very commonly used IDs and passwords with numeric values and special characters and at last huge databases of credentials used by various people in the past. Cracking password by trying every possible combination till one is accepted is a risky job, but it has become easier with improved computing power and network bandwidths. And when it comes to IOT the risk is even greater due to the low security and regulations around the IOT environment, most connected objects operate with factory default authentication that can be easily found online through the official website of the product.

• More than 90% of all compromised devices are hacked through this particular vulnerability, today a lot of pre- programmed scripts exists to carry specific brute-force attack depending on the devices targeted. In this sense we propose a framework to strengthen and fortify this front, by integrating a smart algorithm that test the strength of a given password and generating solid usable password to further protect the device. We used a regression- based model tested on a database of more than 0.7 million real password made open source after the 2014 webhost hack. By integrating our framework we actively protect against the number one exploit used against inter- connected environments, responsible for over 90% of all IOT related security fails.

3.2 Proposed Approach

We are going to discuss the authentication types supported by the Azure IoT Hub Device Provisioning Service and Azure IoT Hub. There are other authentication methods out there, but these are the ones we have found to be the most widely used.

• X.509 certificates are a type of digital identity that is standardized in IETF RFC 5280. If you have the time and inclination, I recommend reading the RFC to learn about what makes X.509 certificates useful in IoT scenarios. There are several ways certificates can be authenticated:

- Thumbprint: A hex string uniquely identifying a cert generated by running a thumbprint algorithm on the cert.
- CA authentication based on a full chain: Ensuring the certificate chain was signed by a trusted signer somewhere in the cert.
- Many customers rely on external vendors for certificates.
- Management comes at a price, adding to the overall solution cost.

– Lifecycle management can be a challenge due to the logistical complexities involved.

Trusted Platform Module (TPM): TPM can refer to a standard for securely storing keys used to authenticate the platform, or it can refer to the I/O interface used to interact with the modules implementing the standard. TPMs can exist as discrete hardware, integrated hardware, firmware- based modules, or software-based modules. Some of the key differences between TPMs and symmetric keys (discussed below) are that:

Cons

- TPMs are difficult to use in general if you're not familiar with them.
- Difficult to develop for without either a physical TPM or a quality emulator.
- May require board re-design to include in hardware.
- You can't roll the EK without essentially destroying the identity of the chip and giving it a new one. It's like if you had a clone, your clone would have the same physical characteristics as you but they are ultimately a different person. Although the physical chip stays the same, it has a new identity in your IoT solution.

Symmetric key:

A symmetric key is known to both the device and the service, and the key is used to both encrypt and decrypt messages sent between parties. Azure IoT supports SAS token based symmetric key connections. The best way to protect symmetric keys is via a hardware security module.

Cons

- Less secure than X.509 certificates or TPM because the same key is shared between device and cloud, which means the key needs protecting in two places. For certificates, TPM, and PKI in general the challenge is all about proving possession of the key without ever revealing the private portion of the key.
- Easy to have bad security practices. Folks using sym- metric keys tend to hardcode the keys in the clear (unencrypted) on devices, leaving the keys vulnerable. It's possible to mitigate some risk by securely storing the symmetric key on the device, but in general, folks using symmetric keys aren't necessarily following best practices around key storage. It's not impossible, just uncommon.
- Shared symmetric key: Using the same symmetric key in all your devices. A very common strategy which is not recommended at all and makes a security malpractice regarding IoT authentication.
- Really, don't use the same symmetric key in all devices. The risks far outweigh the benefit of easy implementation. It would be security malpractice to suggest that shared symmetric key is a serious solution for IoT authentication.
- Very vulnerable to attack.
- Anyone can impersonate your devices if they get a hold of your key.
- Likely to lose control of devices if you rely on shared symmetric key. It can also be actively used on botnets.
- After analyzing the different protocols used for securing authentication, we can say that the permanent issue faced is the adoption of the security process and protocols in

order to protect your system, but the vast majority of IOT users are not well equipped to face and act to secure their devices, so we propose a framework that can be adopted to secure the authentication process without the need for user intervention.

- Our model uses a regression algorithm trained on a database of 0.7 million password categorized from into three clusters: weak identified with 0, strong with 1 and very strong with 2.
- We start by importing our data and cleaning it alongside correcting false and empty data, before applying the models.

```
data = pd.read_csv(r"C:\Users\elkhe\Desktop\PFE\data.csv",',',error_bad_lines=False)
data
```

	password	strength
0	kzde5577	1
1	kino3434	1
2	visi7k1yr	1
3	megzy123	1
4	lamborghin1	1
...	...	...
669635	10redtux10	1
669636	infrared1	1
669637	184520socram	1
669638	marken22a	1
669639	fxx4pw4g	1

669640 rows × 2 columns

```
data.isnull().sum()
```

```
password    1
strength    0
dtype: int64
```

```
data.dropna(inplace = True)
```

Fig. 5. Data Set used

Then import and train the regression model on our database with the help of the command '.fit' from "sklearn" Python package: Our model has a 98% accuracy, a very high score that fits our needs. It's the best we got from all the other machine learning models we tested.

```
[18]: X_train,X_test ,y_train,y_test = train_test_split(x,y,test_size = 0.20,random_state = 42)

[19]: xg = xgb.XGBClassifier()

[20]: xg.fit(X_train,y_train)

[20]: XGBClassifier(base_score=0.5, booster='gbtree', colsample_bylevel=1,
              colsample_bynode=1, colsample_bytree=1, gamma=0, gpu_id=-1,
              importance_type='gain', interaction_constraints='',
              learning_rate=0.300000012, max_delta_step=0, max_depth=6,
              min_child_weight=1, missing=nan, monotone_constraints='()',
              n_estimators=100, n_jobs=0, num_parallel_tree=1,
              objective='multi:softprob', random_state=0, reg_alpha=0,
              reg_lambda=1, scale_pos_weight=None, subsample=1,
              tree_method='exact', validate_parameters=1, verbosity=None)

[21]: xg.score(X_test,y_test)

[21]: 0.9865524759572307

[23]: y_pred=xg.predict(X_test)
```

Fig. 6. Data Splitting and Training

```
xg.score(X_test,y_test)
```

```
0.9865524759572307
```

Fig. 7. Model accuracy

We train the model in the same way as the previous one: we test its accuracy and compare it with the other models: accuracy=0.87

For our next model, we used k-fold cross-validation with random-forest classifier on 4 folds.

We train our model using a 'For' loop for each fold, the same way we did with previous models, and test its accuracy across all folds:

Our model has its best accuracy reaching the forth fold with a score of 94 still less than the regression model we tested.

```
from sklearn.ensemble import RandomForestClassifier
```

Fig. 8. Random Forest Library

```
xg.score(X_test,Y_test)
```

```
0.8726037522058726
```

Fig. 9. Random Forest Score

```
from sklearn.model_selection import KFold
```

```
folds = KFold(n_splits=4)
folds.get_n_splits(X)
```

Fig. 10. Cross Validation

```
for train_index, test_index in folds.split(X):
    X_train, X_test, y_train, y_test = X.iloc[train_index], X.iloc[test_index], y.iloc[train_index], y.iloc[test_index]
    model = RandomForestClassifier()
    model.fit(X_train, y_train)
    y_pred = model.predict(X_test)
    fold+=1
    print(f"Accurac in fold {fold}:", accuracy_score(y_pred, y_test))
```

```
accuracy in fold 1: 0.9173228146688084
accuracy in fold 2: 0.9242815846600084
accuracy in fold 3: 0.9200227146612084
accuracy in fold 4: 0.9417179146620314
```

Fig. 11. Splitting Data and Model Training

After testing all the models, we can positively judge the regression model as best fit for our work, with the highest accuracy score of 98%, and can be used on any database of classified passwords.

By applying our framework, we can successfully protect different devices along with the environment connected to the device, and actively reduce the risk that roams around IOT industry and secure the security breach responsible for over 90% of compromised IOT devices.

There's always a risk of authentication failure even with the solution we propose, as we saw earlier different types of authentication attacks, especially when data is not well encrypted or sensitive documents being stored regularly, but it still represent a higher breach level and low maintenance for the attack to be successful, and it only represent around 6% of known successful attacks.

4 Conclusion and Perspectives

There are a number of challenges to securing IoT devices and providing end-to-end security in an IoT environment. Because the idea of networking devices and other objects is relatively new, security has not always been considered a top priority during the product design phase. In addition, because IoT is an emerging market, many product designers and manufacturers are more interested in getting their products to market quickly than in taking the necessary steps to build in security from the start.

One of the main issues cited with IoT security is the use of hard-coded or default passwords, which can lead to security breaches. Even if passwords are changed, they are often not strong enough to prevent infiltration.

Organizations must learn to view security a common problem, from manufacturer to service provider and end user. Manufacturers and furnaces Service providers must prioritize the security and privacy of their products, and also provide encryption and authorization by default, for example. But the responsibility does not stop there; end users should take care to take their own precautions, including changing their passwords, installing patches when available and using security software.

In future work, it is very interesting to manage and reduce the amount of IoT devices which will help in saving the planet from a pile of trash, so it is important to overcome those issues by optimizing the use and efficiency of those devices.

References

1. Atzoria. L,Ieraauthor vitae, g. morabito "the internet of things: A survey, (2010)
2. Burch, J.R., Clarke, E.M., McMillan, K.L., Dill, D.L.: Sequential circuit verification using symbolic model checking. In: 27th ACM/IEEE Design Automation Conference. pp. 46–51. IEEE (1990)
3. Cacciagrano, D.R., Culmone, R.: Formal semantics of an iot-specific language. In 2018 32nd International Conference on Advanced Information Net- working and Applications Workshops (WAINA). pp. 579–584. IEEE (2018)
4. Caltagirone, S., Pendergast, A., Betz, C.: The diamond model of intrusion analysis. Tech. rep., Center For Cyber Intelligence Analysis and Threat Research Hanover Md (2013)

5. Jacobs, K.: Session details: Book review of K. Jacobs (ed.), Information technology standards and standardization: A global perspective, European Journal of Information Systems (2001)
6. Jing, Q., Vasilakos, A.V., Wan, J., Lu, J., Qiu, D.: Security of the internet of things: perspectives and challenges. Wireless Networks 20(8), 2481–2501 (2014)
7. Liaropoulos, A.: 13th european conference on cyber warfare and security eccws-2014 the university of piraeus piraeus, greece (2014)
8. Rodriguez-Mota, A., Escamilla-Ambrosio, P.J., Happa, J., Nurse, J.R.: Towards iot cybersecurity modeling: From malware analysis data to iot system representation. In: 2016 8th IEEE Latin-American Conference on Communications (LATINCOM). pp. 1–6. IEEE (2016)
9. Ruijters, E., Stoelinga, M.: Fault tree analysis: A survey of the state- of-the-art in modeling, analysis and tools. Computer science review 15, 29–62 (2015)
10. Watson, H.A., et al.: Launch control safety study. Bell labs (1961)
11. Bruce Schneier, Security Analysis of IoT Systems using Attack Trees, (1990)

    Chahid, Ismail and Abderrahim Marzouk. "A Secure IoT Data Integration in Cloud Storage Systems using ABAC Access Control Policy." International Journal of Advanced Engineering Research and Science 4 (2017): 34-37.
12. Sjouke Mauw and Martijn Oostdijk. "Foundations of Attack Trees". In: Information Security and Cryptology - ICISC 2005. Ed. by Dong Ho Won and Seungjoo Kim. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 186–198. isbn: 978-3-540-33355-5, (2005)