University of Piraeus

School of Information and Communication Technologies

Department of Digital Systems

Postgraduate Program of Studies

M.Sc. Digital Systems Security

Master Thesis

Internet of Things, Smart Homes and Privacy: Cyber Security and Personal Data Protection in a Fragile Environment

Supervisor Professor: S. Gritzalis

Athanasios Karampogias k.thanasis14@gmail.com MTE1911

Abstract

From the beginning of the "4th industrial revolution" and the rapid evolution of technology over the past decades, modern ICT's play a critical role in everyday life of governments, organizations and billions of individuals all over the world. The enhancement of the large capabilities the technological growth has brought in almost all sectors (governance, financial, communication, transportation, accommodation etc.) comes along with its risks, as most aspects of personal life that depend on technology. How "invasive" modern services can be to an individual's private life is a matter of deep consideration, along with how the protection of one's private/sensitive information could be secured; both from a technological and a regulatory/law perspective. These issues and, in general, security and privacy nowadays, are a major concern in developing/maintaining and expanding current and future ICTs, along with the policies applied to the collection and management of personal data by organizations, enterprises, public administration, supervising bodies and authorities.

The scope of this essay is to examine matters of security and privacy in the field of Internet of Things and especially in Smart Homes. The massive collection of personal data and their excessive misuse, creates a number of emerging needs that need to be dealt by the community; not only with secure and applicable engineering solutions that protect privacy in these environments, but also with implementing, on a large scale, policies that ensure that any individuals' personal information subject to processing would be treated properly. Dealing with these issues is urgent, considering the innovative nature in these fields is closely related with the collection of personal data; in fact, a lot of the services/applications that these fields are offering are based on the massive collection of such information. In this essay we will try to outline those issues and discuss possible solutions.

Table of Content

1. Privacy and Data – A Brief Introduction

“The world’s most valuable resource is no longer oil, but data” (Economist, 2017) is the title of an article published in the journal Economist, one of the largest financial newspapers of the world. A rather ground-breaking statement, yet, it only represents the exponential increase of data created, distributed and stored especially in the past decade, playing a crucial role in today's industry. A lot of the largest firms that exist nowadays acquire and manage zettabytes of data, making the collection and centralization of data one of the most significant assets of the modern financial and industrial world. As the centralization of data gets bigger and bigger (mostly managed by business giants) and as the bytes of data created every day are also rising in dizzying numbers - estimated at 463 exabytes every day by 2025 (World Economic Forum, 2019), there is an ongoing greater need for better protection, especially when a vast amount of critical and sensitive information is going “online”, or being managed by “3rd parties”. From that perspective, the introduction of IoT and the vision for a rapidly growing “connected world”, not only broadens the range of ICT solutions that could be applied to more aspects of everyday life, but also can be quite “intrusive” regarding an individual's “privacy”.

In order to analyze the impact that IoT has in terms of privacy, we first have to briefly discuss what “privacy” exactly means and represents, both from a practical and a legal point of view. Also, we have to examine why privacy is a matter of great importance and in which ways technical and institutional regulations and procedures are affected and/or need to be adjusted, with the introduction of these technologies and their wide adaptation.

Personal data, confidentiality of information and privacy are related concepts all dealing with similar issues. But what exactly we mean with the term “privacy”? At first it was described as an individual's “right to be left alone” (R. Kerr, 1850). In general, the right to privacy is determined when a person (and any information about him/her) is protected from interference or intrusion. Various legal acts have taken place since one of the very first definitions we mentioned above, enriching the term “privacy”, defining more precisely its context, and making its boundaries clearer. One of the most significant acts regarding privacy was the publication of the article “The Right to Privacy” where privacy is defined as: “The common law secures to each individual the right of determining, ordinarily, to what extent his thoughts, sentiments,

and emotions shall be communicated to others.. and even if he has chosen to give them expression, he generally retains the power to fix the limits of the publicity which shall be given them... No other has the right to publish his productions in any form, without his consent. This right is wholly independent of the material on which, the thought, sentiment, or emotions is expressed" (Warren-Brandeis, 1890). In fact, one of the statements of this article is that the publication of any information about an individual and its products (as described above) relies solely on its free will (or it is enforced in a court). It is considered a milestone in the law perspective of privacy and although a lot of discussion and research has taken place since then in this field -and many other law cases-, it had a crucial impact on how the scientific society and the institutions understood and got involved with the concept of privacy. Based on that assumption, nowadays, information privacy (and security) - when it comes to the digital world - is the right of a person or a group of people to have control over how their personal information is collected, managed, used and published. Privacy and personal data are also protected by various legislations and regulations, both on state and on EU level. The intention of these regulations is to protect the individual against those organizations that collect his/her data and may misuse them. These regulations enforce the organizations to apply policies for the correct collection, management and distribution of the data, ensuring that privacy is maintained and any "leakage" of private information between the parties involved is preserved at low risk.

How IoT affect Privacy? Entering the new decade in an ever-changing digital world, the introduction of these technologies opens a whole new range of possibilities and benefits in everyday life of citizens and also in public governance and the enterprise world. However, all the services that come along with technological progress serve at a cost; there is a great amount of work to be done to ensure that collective and individual privacy will not be harmed by over-abundance and rampant distribution of data collected and managed, and also for over-watching organizations (intelligent agencies, governments, large organizations) not to gain too much power. There is a soft point of balance that needs to be worked on between providing these services and the benefits that come along them (increasing security, better public governance, smart cities, targeted advertising and a lot more) and the possibilities of data leaks, data-analysis failures and, primarily, closing the distance between one's private life and public life -which distance is a prerequisite for privacy. We will examine IoT

regarding the capabilities and the challenges as far as security and privacy are concerned, and then we will discuss the existing regulatory implications and the measures that need to be taken to ensure -at some extent- that modern and future implementation of these technologies will provide better solutions for the users but also preserve privacy, both individually and collectively. Last but not least, we will examine how these general challenges and “best practices” apply in modern “smart home” architectures, thus providing an essence of how measures of privacy can be contained in a seemingly “intrusive” environment.

2. Internet of Things

2.1 IoT: Greater Steps to “Digital Society”

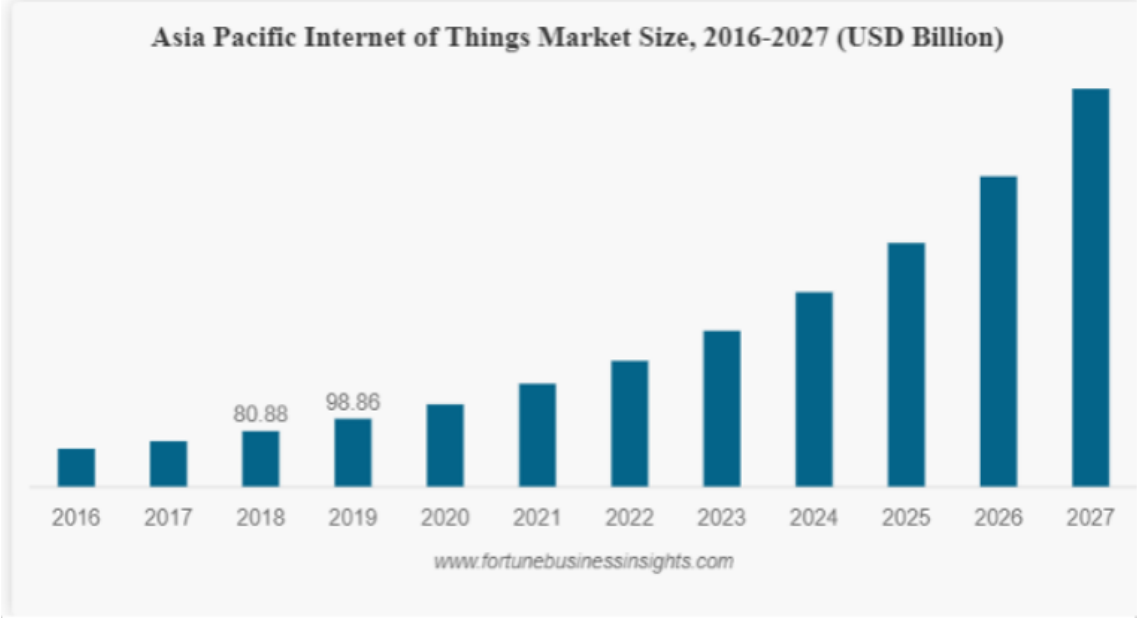
“The Internet of Things (IoT) describes the network of physical objects—a.k.a. "things"—that are embedded with sensors, software, and other technologies for the purpose of connecting and exchanging data with other devices and systems over the Internet” (Wikipedia). There are various definitions that scientists and researchers have given over the years referring to the Internet of Things, emphasizing on various aspects of its use and nature. Yet, Internet of Things cannot be examined as “one” technology, as it contains various different technologies, techniques and devices – varying from “standard” mobile devices to “smart thermostats”, for example. In fact, every single “thing” that can access a network or the Internet can potentially be part of the so-called Internet of Things. Generally speaking, IoT shall be examined as more of a concept; the rapid advancement to totally interconnected systems that can communicate and exchange data broadly and universally, leading to huge improvements for everyday life and well-being. There are various approaches to examine this matter, both from a “pure” technological but also from a social perspective. Taking this into account, a framework examining IoT was proposed, indicating 5 variants that combined can entitle most part of the wide area that consist the IoT: Social Actors, Things, Data, Networks and Processes (Lynn et al., 2020). Analyzing these factors furthermore is out of scope for this essay; however, it gives us a general sense of how complex and broad this field is.

The massive introduction of IoT solutions and the rapid growth of “smart, connected devices” (which “go online”, and give us more and more data), adds some more problems to the equation, from a privacy perspective. The problem of the society and the technological world to be able to adjust to the occasion, from the one hand embracing the new technologies (which as we explained before are quite “intrusive”) and from the other hand, at the same point preserving and ensuring privacy for all the individuals-members of the “connected world”. There is one major drawback to the IoT – along with problems deriving from that; the access they provide to an individual's personal data is enormous, with (if not protected and regulated properly) possible major consequences. The possibilities that IoT offers are also great, from Wireless Sensor Networks that could track and prevent disasters, to automated driving and navigation, and in general the whole deployment of the concept of “smart cities”.

In order to better understand the bigger picture, the growth of connected devices is rapid; “Figures show that the connected products and devices already exceeded the global population and is expected to reach 50 billion by 2020, up from 25 billion in 2015” (M. Unver, 2018) and alongside the reach and the variety of the possible applications. These massive (and in short time) growth is leading to a rather “uncontrollable” development and promotion of IoT solutions, taking over a large piece of the ICT market, either applied to small-scale home solutions (e.g., an online CCTV monitoring system) or other large-scale solutions (e.g., the whole navigation of a fleet with GPSs etc.) Of course, this rapid growth brings along wide security and privacy gaps (lots of attacks on IoT systems have taken part until today) and/or conflicts against more traditional concepts and approaches of privacy; and the legal framework around it.

2.2 Defining the field: Current Status and development

As previously discussed, the development in the field is enormous. Huge vendors and enterprises invest and produce more and more ICT solutions in the context of IoT, making it one of the extremely developing markets among the technological world. One survey and estimation about the future development of the IoT market growth is indicative:



Picture 1: Fortune Business Insights, July 2020

The vast majority of the predictions about the market of IoT agree on the same estimation; a rapid exponential growth on the products and solutions provided.

Also, there is a wide variety in the market of the IoT, affecting lots of “traditional” industries. Indicatively:

- Healthcare
- Agriculture
- Manufacturing
- Energy
- Fitness & well-being (gyms, personalized fitness and nutrition advice etc.)
- Autonomous driving and public transportation
- Smart homes and cities
- Public Sector (Firefighting assistance, military purposes, surveillance)

and many more.

As we can easily assume, the invasion of Internet of Things' solutions in both industrial and every-day life will change instantly the amount of data collected by different vendors, large and small, putting security and privacy into perspective. The amount of data created and transferred every day through mobile, smart and electric devices, make them one of modern world's most valuable assets, both from an industrial but also from a public governance point of view. Furthermore, if not properly collected, stored and regulated, the same data can put privacy and security at great risk.

2.3 Internet of Things and Personal Data

In order to examine and evaluate best practices and guidelines for security and privacy-preserving IoT technologies, we first have to analyze further what kind of data is processed and analyzed by the emerging IoT solutions, and, from that point on, how the personal data of an individual or a group of people could potentially be exposed and the corresponding impact of security and data breaches. It is easy to assume that different types of data are collected by the different IoT solutions, due to the various areas that IoT is affecting, and the wide scale of solutions. We will try to

look generally at the categories of data collected by different IoT devices, in order to better explain the reason why security and data privacy in these environments are of great importance. The data collection itself, divided by each different category/industry and service provided – especially in industries that deal with critical infrastructures or handle huge amounts of personal data - can clarify why IoT should be protected efficiently and effectively. For example:

1. In terms of healthcare use, IoT is developing rapidly offering a number of services and solutions. Remote health and monitoring services, assisted living and elderly care, managing and monitoring chronic diseases, personalized medication, reducing emergency waiting times, monitoring equipment, personnel and environment in healthcare establishments are only a portion of the solutions aiming to improve and automate healthcare. Therefore, a lot of sensitive data is transmitted, stored and analyzed such as sensor Ids, geolocation data, security numbers and patients' conditions, drugs' intake or even information about the status and monitoring of a hospital, all of which could be collected and exploited for malicious purposes.
2. Regarding public services, as we can easily assume from the wide variety of issues IoT could be used (means of public transport, military and national security, surveillance, fire prevention), huge amounts of personal data are created in order to enhance those services, such as: multimedia containing personal data, geographical data, personalized information that could be used for profiling etc.
3. As of smart homes, it is probably the most “intrusive” - if misused – type of IoT services, as it could not only expose an individual's public life or habits, but also collects data inside their most personal space. CCTVs, smart electronic devices, create a rather complicated environment privacy-wise.

As we can easily understand, securing all emerging IoT services is of great importance, but also much more challenging than more “traditional” information security approaches. The nature of these technologies, the rapid growth and competition between industry partners that offer these solutions and the lack of regulatory and legal frameworks, leaves security and privacy as “the last wheel of the wagon” and only recently researchers in the academic and industrial world began to

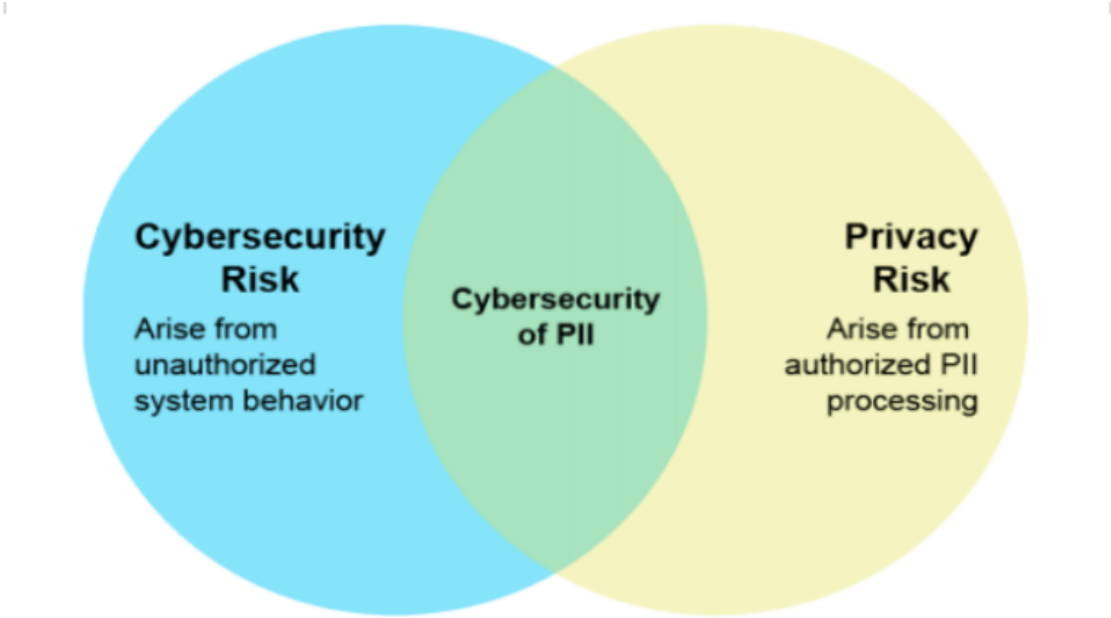
consider those issues and suggest solutions. The purpose of this part of the essay is to further examine privacy and security challenges in the rapidly growing environment Internet of Things, and then especially in smart homes, and also discuss measures and guidelines about how authorities and regulatory bodies, industry partners and end users could prevent security breaches and personal data exposures/leaks.

3. Internet of Things VS. Privacy & Security

3.1 Privacy and Security Risks in IoT

As we already described, IoT is greatly “closing the distance” between one's private and public life, creating a lot of data that if misused (either by the vendor of the product, or leaked) can lead to privacy and security violations with a potential significant impact both for companies and individuals. First of all, we need to clarify

that Privacy and Security Risks are related, although they are not the same and that's why we need to develop strategies to secure both “ends”. According to NIST IR 8228: “(Risk is) a measure of the extent to which an entity is threatened by a potential circumstance or event, and typically is a function of: (i) the adverse impact, or magnitude of harm, that would arise if the circumstance or event occurs; and (ii) the likelihood of occurrence. For cybersecurity, risk is about threats—the exploitation of vulnerabilities by threat actors to compromise device or data confidentiality, integrity, or availability. For privacy, risk is about problematic data actions—operations that process personally identifiable information (PII) through the information lifecycle to meet mission or business needs of an organization or “authorized” PII processing and, as a side effect, cause individuals to experience some type of problem(s)”. Indicatively:



Picture 2: NIST IR 8228, "Considerations for Managing IoT Cybersecurity and Privacy Risks"

There are numerous challenges that comes along IoT, taking into consideration the approach described above and also special aspects that diversify IoT technologies than more traditional IT devices and infrastructures.

We can see above a matrix where we can observe some of the main differences between the functionalities and considerations for IoT security in contrast with traditional IT security:

Traditional IT security	IoT Security
The content of the IT security is created by the humans.	The content of IoT is created by the machine itself.
We have to Add-on the security features in traditional IT.	It has built-in security i.e. security is added during the designing phase.
The content is consumed by the request.	By triggering actions and pushing information the content is consumed.
Using explicitly defined links the content is combined.	Through explicitly defined operators, content is combined.
It uses complex algorithms for the devices.	Lightweight algorithms for resource-constraint devices are used.
Privacy is user control that means the data is collected after getting the user's permission.	IoT collects user information automatically without getting the permission from the user.
Low heterogeneity devices are used or device heterogeneity is less.	Technological heterogeneity is more in this case as compare to traditional IT.
IT devices are located in the closed environment.	IoT devices are located in open environment.
More number of security guards.	Less number of security guards.
Content creation (HTML) and content consumption (search engines) both is achieved.	In IoT mainly content creation is achieved.

Table 1: T. Varshney, "Architectural Model of Security Threats & their Countermeasures in IoT"

3.1.1 Different interaction with the physical world

As explained above, the many different devices that consist IoT (e.g., sensors) collect and create a lot more data than traditional IT devices (PCs, laptops etc.) In the public opinion, it seems "common" – which is still not right – that when someone surfs the web, he gives away personal information, depending the service. With IoT, data is collected "noiselessly"; data flows are created and used without or with minimum human interaction. In addition to that, the great diminishment of private space could lead to massive extraction of personal information, not only regarding names, SSNs, bank accounts etc., but also habits, physical expressions or even sentiments. As we can see, and due to the nature of these devices, traditional technical measures of protection are not enough, as there is so much "live" and "intrusive" information created that make IoT really attractive for hackers, either at the endpoints (devices) or at the relevant enterprises (vendors of the corresponding services). The problem gets even more complicated when various devices are used in a certain environment (e.g., a "smart home") that also could co-ordinate one another; making it impossible to control what data is created, and especially the collection and processing of that data.

3.1.2 Technological constraints and limitations

IoT solutions are greatly susceptible to hacking attacks. Many attacks were launched and reported at various IoT devices, but in this case, every attack takes a much greater toll regarding privacy. Cybersecurity risks that emerge (in that case primarily at the edge points-devices), mostly arise from two reasons:

1. The IoT devices themselves cannot support fully all the security features that traditional IT devices can. Due to the limited resources these devices possess (battery, processing power, memory), vendors and developers are trying to "find the middle ground" in terms of efficiency and security. For example, using traditional encryption algorithms to ensure confidentiality in device-to-device communications can lead to unacceptable delays. Also, lots of traditional cybersecurity appliances could be of little usefulness in IoT environments (e.g., data protection techniques in devices with little capacity and where data does not stay at rest) or even could not handle the level and type of data created (e.g., a simple firewall for home use). Of course, each device does not have the capacity to handle classic end-point security solutions. Finally, things get a lot more challenging when various IoT devices (from different vendors) operate in the same environment-infrastructure, where managing them under a common framework is – in the majority of these cases – impossible. Last but not least, IoT environments contain a lot more interfaces (beyond just PCs, laptops etc.) that could lead to security breaches; an attacker can take advantage of a single vulnerable thermostat to have complete control over the whole home network. These problems in IoT environments also remain an open challenge both technically and organizationally.
2. The growing competition of companies and large IoT service providers, along with the growing need for IoT solutions in the industry, led to the point where the growth and expansion of the offered technologies and their usability, comes before security and privacy. Therefore, a lot of big and small size companies are entering the field of IoT, offering various solutions, trying to reach higher levels of automation and facilitation as their first priority. In this rapidly growing industry, security and privacy are often set aside. In fact, lots of issues arise from that exact case; the mass production and research in this

field without taking account of the problems that emerge due to the nature of these technologies. For example, lack of interoperability between the different devices can lead to data leakage or loss. Also, especially in smaller scale environments (e.g., a small network of sensors), there are a variety of devices trading data that, since they cannot be managed under a central framework, it is difficult to be managed by the same security appliance-infrastructure. Finally, while most IoT devices are built as “plug and play” devices – meaning that they do not require excessive configuration by the end user – they often act as “black box” for the users; they have little or no control about the state of the device, the logs and data created, the security features etc. Furthermore, often these devices come with built-in default configurations (e.g., user 'admin', password 'admin') and even worse, post-market settings are very difficult to change.

Leaving apart the technical and organizational matters that are described above, there are also open challenges regarding privacy and data handling, besides the obvious data leakage that can occur after a direct security breach.

3.1.3 Identification, linkability, profiling and mass surveillance

As we already described, a great amount of the data that are collected, stored and processed in the IoT environment contain PIIs and sensitive information. That leads to the case where a person could easily be identified and being profiled, given the fact that not only the websites or applications that they are using are exposed, but also their habits, personal beliefs, political views, health condition, religion could easily be extracted. In these circumstances, if the data is not protected adequately – at the user end – we can easily understand how one's private data could be exposed; leading to the problems of “identification” and “profiling”. Now, in the vendor-side – if not regulated – manufacturers (or even law enforcement and governments, public services etc.) have access to huge amounts of data and, by using modern Big Data analytics and AI, they could easily extract information about individuals, making profiling and (at scale) mass surveillance a reality. In IoT environments, these situations can occur even “indirectly”, combining data from different sources and devices to create a profile about someone, which could be “useful” for various purposes: personalized advertisements, tracking, insurance purposes etc. Also, traditional privacy-preserving techniques and measures such as “de-identification of data” and “anonymity” seem

inefficient in these environments, exactly because the collection and combination of certain data (even anonymized) could link the actions and activities of a specific person to their real identity. For example, if we examine a smart-home environment, data extracted from sensors that measure humidity, temperature, light, CO2 and additional information of this kind, could make possible the estimation of the exact location even without a tracker/GPS or an IP. Additionally, a wearable fitness device could also give away the habits of the person, or his medical condition. All these data could be exploited for the purposes mentioned above, giving overpower to vendors and certain authorities (imagine data collected from sensors in public spaces) and potentially become major privacy threats for the end users of those services.

3.1.4 Purposes of data processing, consent and transparency

One of the primary concepts of privacy is the consent of the subject for letting the provider of the services offered to collect, store and process the PIIs that are gathered when a person is using a service. The collection of personal information is a great enhancement to the services offered, taking into account that these technologies offer a whole new range of possibilities, such as personalized recommendation, adjustment to the needs of the user (e.g. a wearable device that delivers personalized training programs or a medical device that monitors the patient's state and adjust accordingly) but also, since these services require personal information to reach this level of personalized assistance, the consent of the person is obligatory – and protected by recent legislation and regulatory frameworks. Furthermore, the consent is given at an extent, depending on the service provided and the choices and preferences of a person; the provider should give options to the subject about the exact processing of his/her personal data, meaning that the person could allow the collection and process of his/her personal data but could deny to let the vendor to give them to a third-party, or he/she could deny any collection of personal data, and so on. Also, the subject always should retain the right to revoke his/her consent, or change the extent of it. The first challenge is for the users to understand and define their consent, but taking into account that we are examining an environment which is mostly “plug and play”, with little or no user interfaces, the consent of the users and moreover the control they have over how their personal information is used is a “grey area” most of the times. Additionally, the key concept of consent is that there are some prerequisites in order to effectively preserve privacy; an “I agree” button on an interface, or a sign in a

multi-paper document when a user buys a product is not enough. So, the challenge is to achieve “meaningful consent” that not only ensures that the user fully understands the extent of the use of his/her personal data, but also can monitor exactly what data is collected and how it is processed – transparency –, revoke or change their consent, and how and who exactly is responsible for the processing of their personal data – accountability. This remains an active and problematic challenge for nowadays IoT solutions and technologies, as they cannot meet those basic privacy requirements. The concept of “meaningful consent” could be summarized in 5 principles:

- Capacity; ability to give consent.
- Voluntary; free expression of consensus (if one device or service is inaccessible in case consent is not given, this principle is violated).
- Current; the collection and processing of personal data should not be a one-off procedure, nor the data should be kept for an indefinite amount of time.
- Specific; the consent is given at some extent, and for clearly defined purposes.
- Informed; the subject must be fully aware who is responsible for his/her personal data, why and with whom they are going to be shared.

These principles challenge the very nature of nowadays approaches in developing and launching IoT services. Key privacy concepts and their conflict with IoT remain an “intractable puzzle” for nowadays researchers, developers, enterprises and regulatory bodies.

3.1.5 Regulatory Issues

Last but not least, a great role in nowadays security and privacy risks plays the fact that the according legislation and standards are yet to meet the expectations. First of all, unlike the Internet in general, there is a lack of pre-defined standards that every offered service should align (e.g., transfer protocols, encryption algorithms etc.). The wide variety of vendors, the rampant growth of many different approaches and technical methodologies, the effort of vendors to balance between user-friendly and secure technologies, define a field where a holistic approach about standards for technical measures and according privacy regulation frameworks is quite ambiguous. Note that, at least in EU, the main framework for Privacy (General Data Protection Regulation) came a significant amount of time 'later'; meaning that the need for a

legislation about data privacy, that would combine best practices and guidelines, emerged as soon as massive amounts of sensitive data began to travel through the Internet. Proportionally, there is a great need nowadays for guidelines regarding IoT and approaches that could include IoTs specific features to the current regulations and legislations.

3.2 Attacks and data breaches over the years

To showcase the great security and privacy risks that come along IoT, we will examine some security and data breaches that took place over the years. One of the most notorious recent data exposures happened in Peloton, a fitness company. In this case, an unauthenticated API discovered by a security researcher, led the way for the internal network and a significant amount of personal data of specific users such as names, weight, gender etc. This matter showcases the need for authentication mechanisms and management frameworks in these environments, along with the proper segregation between the sensors and the internal network.

Shodan case was a great example where a lot of people were able to monitor for example the living room of some apartments because there was a hacked camera. Shodan is a search engine for connected devices to the Internet; including IoT devices. Due to the lack of security mechanisms and authentication methods, it was a matter of minutes for one malicious actor to discover connected cameras and gain a direct view on the inside of an apartment.

Another interesting case is about Ring, where cybercriminals were able to successfully hack smart doorbells and CCTVs, due to default and recycled credentials installed in the devices, and they even managed to verbally harass the users of the service. The same scheme appears also in this case; weak password policies and authentication protocols.

Our last, and probably most terrifying example, is a vulnerability discovered at smart heart plants in St. Jude Medical. Due to this vulnerability, if exploited, a hacker could turn-off the device with, as we can easily assume, devastating consequences.

As a result, we see that, apart from deeper and more foundational issues that Internet of Things face and also could be exploited, there are more vulnerabilities that a hacker could exploit: weak IoT governance, weak credentials, insecure communication, lack of segregation and layers of security.

3.3 Protecting the perimeter: Security concerns

We examined in some scale what are the main security and privacy challenges in the IoT ecosystem, but in order to enhance the protection of those environments we need to further investigate which exactly are the ways, weaknesses and manners that an attacker could use to launch an attack in an IoT environment. In the matrix below, we can see a concentrating attempt examining each security characteristic (e.g., integrity) in device, network and cloud/server level:

Security Characteristic	Device / Hardware	Network	Cloud / Server-Side
Confidentiality	A. Hardware attacks	B. Encryption with low capability devices	C. Privacy concerns
Integrity	D. Lack of attestation, illicit updates	E. Signatures with low capability devices	F. Unchanged
Availability	G. Physical attacks; Radio jamming	H. Unreliable networks	I. Unchanged
Authentication	J. Lack of user input; Hardware retrieval of keys	K. Challenges of using federated identity	L. Lack of widely implemented standards around Device Identity
Access Control	M. Physical access; Lack of local authentication	N. Lightweight protocols for access control	O. Requirement for user managed access controls
Non-Repudiation	P. No secure local storage; Low capability devices	Q. Signatures with low capability devices	R. Unchanged

Table 2: Paul Fremantle, " A security survey of middleware for the Internet of Things"

As we can see from the matrix above, there are various issues that tackle traditional concepts and principles of cybersecurity: confidentiality, integrity, availability, authentication, access control, non-repudiation. In order to further examine these threats, we are going to divide IoT functionality to three general subdivisions that – although there are attacks that can target a combination of these subdivisions – give the general sense how even the existence of IoT itself creates multiple factors that are susceptible to a potential attack, or else, the “attack surface” magnifies for a possible malicious actor:

- Perception/Device Layer; where the primary phase of information collection and handling take place – that means the various types of sensors that consist

IoT (Wireless Sensor Networks, Smart Home networks etc.) and collect and process information.

- Network level; where data collected and processed by the devices in the perception layered is transferred through the network – mostly through the air - to a server/application or another node of the network.
- Application layer; whether the data processing and the services provided happen locally (application, server, laptop as a server etc.) or in the cloud, this layer is where information coming from the network becomes useful for the end user.

Also, we could add a fourth layer that faces security concerns and risks; that would be if the data processing and the delivery of the service comes directly from the vendor of the IoT solution. The risks associated with this layer would be either mostly Privacy issues that could potentially expose personal identifiable information about the end user or people associated with the sensor (a topic that we already discussed earlier in this essay) or classic security threats that an organization can face – or even cloud security issues, that are beyond our scope. We are going to examine the threats of each level separately to better understand the large variety of attacks that a hacker could launch against an IoT environment.

3.3.1 Perception/Device Layer

The very first type of attacks that can be launched in this layer of IoT are physical attacks. There are many attacks that can be launched due to the physical presence of an attacker. One of the easiest methods is to directly steal device's data from physical access (e.g., via a port or a USB device). Moreover, that would be the least. Due to the very nature of IoT (especially in large networks of sensors, or in sensor networks that exist in an open surface), all the nodes are greatly exposed to this kind of attacks. Physical destruction of a node/sensor, node tampering, injecting malicious code to a sensor to take over it and then manipulate the flow of data or even the whole network, physically alter the receiving information of the node, node cloning, and more, all are versions of various physical attacks that can be launched in the perception layer. Additionally, Denial of Service attacks could be launched, either by physical destruction, or by exploiting the low memory and processing capacity of one or more nodes – an attacker could overflow the network with overabundance of information

that could not be handled by the infrastructure. Of course, apart from the forementioned attacks, since the nodes are widely exposed, there is a lot more to it; jamming attacks, man-in-the-middle attacks, packet sniffing that challenge in a direct manner the confidentiality and integrity of a network. Confidentiality could be violated for example in a Wireless Sensor Network that uses sensors with RFID (Radio Frequency Identification) by replay attacks, replaying or spoofing device data. Also, last but not least, routing attacks and impersonation attacks could be launched – as a form of man-in-the-middle attacks – if there is not proper validation of the nodes consisting an IoT network, where an attacker can be part of the network and presents himself as a legitimate node, or by manipulating the valid routes for packet transfer. These are only an overview of the possible attacks that could be launched on the perception/device layer.

3.3.2 Network Layer

Network layer contains all the possible channels of communication that are used by an IoT infrastructure for the transportation of information and data collected - usually the air. As we can understand easily, this layer has a large attack surface, that even expand the known vulnerabilities and possible exploits a bad actor could launch against conventional wireless systems. That derives from multiple reasons; more and diverse communication protocols that are not combined and configured properly (e.g. WiFi, Bluetooth, satellite and others), the necessary evil of lightweight protocols of cryptography, access control, digital signatures, the possible physical access of the attacker inside the network (consider, for example, an IoT environment that monitors remotely a farm) and the unreliability of low capability devices to properly enforce a safe and sound holistic security methodology, because they cannot support it. As a result, multiple different attacks can occur. First of all, DoS attacks are easy to target these environments; either by flooding the network with unnecessary data thus making the processing nodes unable to handle the volume of incoming information, or by targeting specific crucial nodes for the network (for example a central node that communicates with the cloud) and taking them out of the network, making the whole network nonfunctional, and many more ways. Secondly, multiple attacks that take advantage of the lightweight nature of these technologies themselves. Eavesdropping and man-in-the-middle attacks, unauthorized access, manipulation of routing

protocols are only some examples of potential threats that can arise depending the situation.

3.3.3 Application layer

In a way, in this layer, IoT "threat map" resembles closer the situation of a traditional IT system. In essence, the security issues here are primarily the secure storage and processing of data, along with the assurance that each application and process can be executed unhindered. Although the attack is surface is somewhat reduced, previous problems that we noticed in the cyber security approaches of the IoT are still present. For instance, this layer is still greatly susceptible to DoS attacks, since the processing of the overhead of the information is taking place here. Overflooding the network, or initiating communication routes to the server/application by a malicious actor (this method takes advantage also of "weak spots" in previous layers) can lead the whole service to unavailability due to the exhaust of processing resources by the attacker. Furthermore, due to the lack of standardization and common protocols about critical security issues in the IoT (e.g., access control, authentication), attackers could substantially benefit by overriding often weak and vulnerable application security mechanisms. Lack of key management, low capabilities' encryption algorithms, usage of default credentials, misconfiguration of transport and access protocols are creating a vulnerable environment that an attacker could exploit. As we can also notice, an attack or a threat in this level are of higher importance because the information that can be extracted in this layer are much larger than the previous layers, due to the fact that, in most cases, the whole data travelling across the infrastructure ends up in the application in order to be useful for the end user, so, taking into account that risks greatly arise even on this layer, a more "end-to-end" approach should be followed.

3.4 Countermeasures, technical safeguards and open challenges

Until this point, we tried to establish a general sense of how the attack surface and the cyber risks that arise in IoT environments are magnified - comparing IoT to "traditional" ICT solutions and infrastructures – and more diverse. Also, we tried to include in our analysis basic privacy implications and concepts that are challenged from the very nature of IoT technologies. We are going to analyze more the "technical" and "cybersecurity" measures that could be used to enhance privacy and security in these environments and try to answer the "question" of the previous part ("protecting the perimeter"). After giving some general directions for better

enhancing the security of IoT infrastructures, we will discuss about proposed solutions and open challenges regarding 2 of the most insisting issues of IoT security: authentication and identity management, and encryption, as we already discussed about. In the following chapters of our essay, we will try to analyze privacy techniques and requirements, along with all the other issues mentioned (data handling, organizational issues, processes, relevant regulations and standardizations etc.).

- Device software capabilities and updates: The problem is that, especially in large scale environments, there is a lack of proper monitoring of the status of the devices – especially the edge devices – and that there are a lot of cases that, using tiny OSs in the sensor-level, or even in the central architecture of an IoT infrastructure, updates and/or constant monitoring are not possible. Even if the device is deployed "flawless" from the vendor, and the code used is state-of-the-art – meaning that no vulnerabilities are present at the chain of development or production, several vulnerabilities could show up over time. In order to address this issue, OSs with capabilities to encrypt information transmitted and enough capacity to install patches and security updates later on, are a prerequisite. Also, in order to maintain the status of the whole network up-to-date, constant monitoring should be established, managed by a central point and with adequate back-up procedures, or cloud support, in order to ensure that 1. The user has real-time notification about the status of his/her devices, 2. Regular security patches can be installed upon arrival, 3. Back-up adequate to retrieve the status of the network in case of hardware/software failure (this step is introduced to avoid "single point of failure" issues).
- Research, continuous evaluation and assessment; especially in industrial environments, specific methodologies and risk assessment frameworks should be adopted. Also, effort must be put at the research level (R&D labs, industrial research, university labs) to address and overcome issues that drawback technical inefficiencies, as mentioned previously. Moreover, the discussion and research about holistic frameworks and controls that can ensure the security of these environments should intensify, along with the directives from the relevant standardization authorities and institutions (state organizations, NIST, ISACA, ISO etc.). As part of this process, enterprises should develop plans to evaluate cyber risks that exist in their IoT infrastructures, prioritize

them, implement risk remediation plans and review them regularly, ensuring that risk is preserved at a tolerable level. Finally, special vulnerability assessments and penetration testing techniques should be executed at regular intervals, to locate security gaps and weaknesses and remediate them.

- Development and release of security guidelines to the end-users; it is of great importance that users should be notified about potential security or privacy risks that derive from the usage of IoT devices and services. Also, directives should be given to increase the security awareness of the consumers and a baseline of measures that need to be implemented in the first installation of the product (e.g., enforcing a change of the default password, and also some basic actions that can be made in case of a security incident or a malfunction of the service provided (e.g., communication channels with technical support, pre-built quarantine capabilities of the device etc.). In these terms, customer support is very important, and feasible organizational measures must be implemented to ensure that IoT devices are not left completely technically "unattended" - meaning that there should be cost-effective ways of communication with the vendor or support contractors after the products are bought, yet neither violation of users' personal data should be facilitated.
- Physical measures and security hardening; Along with all the proposed measures above, physical access to the equipment should be restrained to unauthorized personnel. The point is, that this element can only be accomplished in private or restricted areas. So, we come to the point where, even if a malicious actor has physical access to a device, it would be extreme difficult to obtain useful information or cause malfunctions to the infrastructure. Adequate encryption techniques, both on the network layer and on the device layer, monitoring, access deprovisioning capabilities if suspicious network behaviour is detected, disabled usb ports etc. are only a few of the measures that ensure security hardening in the device level, even in cases that physical access is possible.

Of course, as we already stated before, these measures are inducted under the prism of the special circumstances that exist in the IoT infrastructures, and are of no use if the capacity, encryption, interoperability, and access control challenges in these environments cannot be addressed efficiently. These are some of the most critical

issues that contemporary researchers have to deal with, and several approaches have already been proposed. We are going to examine a few of them below.

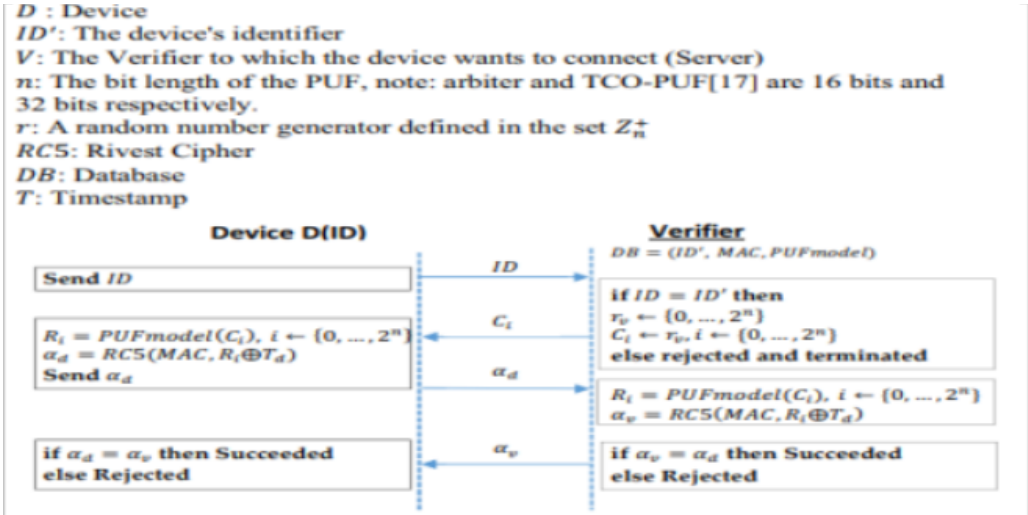
3.4.1. Authentication techniques and encryption methods

First of all, we will examine authentication protocols and models and encryption methods used, as we already examined the authentication issues and the attacks that derive from them in an IoT environment (MITM Attacks, eavesdropping, DDOS, manipulation of routing routes etc.). Taking into account critical drawbacks that exist in IoT in terms of authentication – low device capabilities, lack of standardization protocols, limited interfaces and management capabilities, several methods are currently examined and deployed and could be used. We should note that, as long as each "category" of authentication models and schemes contain various proposals and different affiliated works, concepts are going to be inducted in a high-level manner in order to include and present different approaches in a field that currently has ongoing research:

- One approach that has been proposed is the authentication model that uses PUF (Physical Unclonable Functions), that try to address the challenge of low power and computational resources combined with authentication issues and weak protocols. Although many different models and many different assumptions are introduced with the PUF method, we are going to investigate the main idea of PUF: the need to develop a way of secure authentication and key generation that is cost-effective power-wise, meaning that we have to develop a way to compute keys and establish a network without the same consumption of energy as in conventional IT infrastructures. Briefly explained, a PUF is a physical object inside a physical structure (nowadays, mostly integrated circuits that are part of an IoT appliance) that, when included in a challenge-response authentication scheme, can be used to create a "digital fingerprint" of the corresponding device, that can be unique and used to refer to that exact device. The main concept is, because the challenge-response pair are physical-dependent (and the physical structure assembled at the development phase is random and cannot be duplicated), that PUF-based authentication cannot be tampered, or eavesdropped. The described solution considers a client-server architecture where the pairs are stored on the authentication server:

Picture 3: An Braeken, "PUF Based Authentication Protocol for IoT", August 2018

Also, there are models of this approach that can be used without the storage of the pairs in the authentication server, only the computational method. The general scheme contains two phases: the enrollment/registration phase and the verification phase, as we can see below in the method described:



Picture 4: Yildiran Yilmaz et al., "Lightweight PUF-Based Authentication Protocol for IoT Devices", July 2018

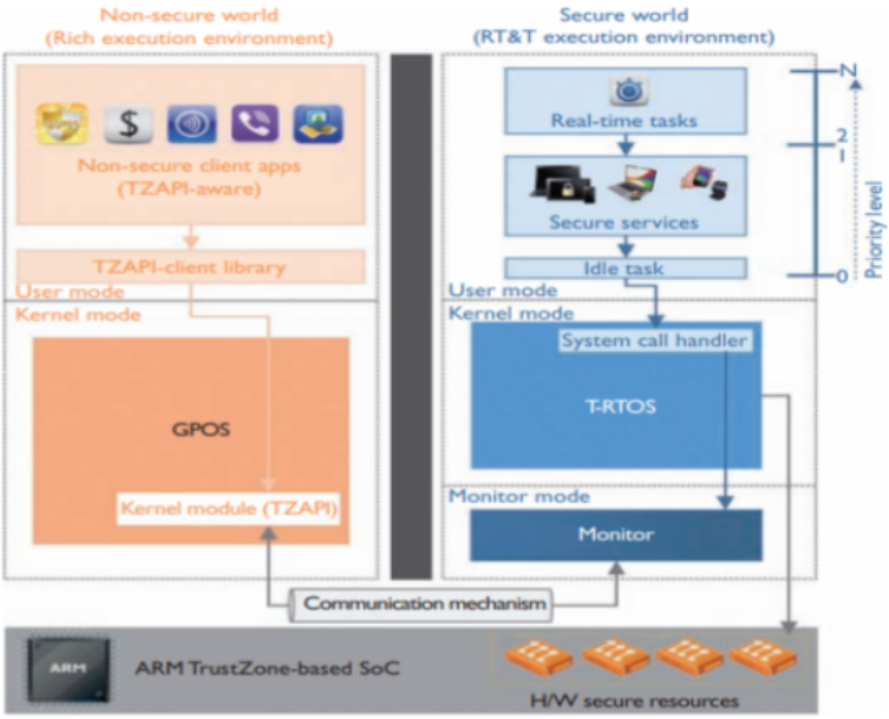
The proposed models that use PUF authentication are much less power consuming. Although there is a high level of security – in general – in these type of deployments, machine learning techniques have been developed that can lead to certain attacks. Therefore, it is suggested that PUF-based should be

properly implemented by the vendors and put through excessive testing against known vulnerabilities, to ensure their proper function and integrity.

- Another approach that has been proposed and used is a Public Key Infrastructure using X.509 digital certificates, hashing, and encryption for better safekeeping of private and public keys. A central management entity (most times in the cloud) is used as a Certification Authority (CA) and Registration Authority (RA), in order to maintain the certificates of the entities connected to the IoT Network. Although it is a challenging task – especially in large-scale environments that are consisted by thousands of devices – regarding the maintenance of the state of each entity and the digital certificates that are only visible from the CA – the approach can be used in a very straightforward way to authenticate the devices and also monitor and understand a possible falsification in the process. Another drawback in this method is the time and power needed for these calculations – increasing the latency in the communication between the different layers of the infrastructure. For that reason, research is made to deploy faster computational models than X.509 certificates, such as Elliptic Curve Qu-Vanstone (ECQV) implicit certificates and a key agreement protocol based on Elliptic Curve Diffie-Hellman (ECDH) key exchange.

- Other methodologies also have been proposed, such as authentication schemes based on hardware, basing security on isolation techniques inside the hardware. These authentication techniques are using either Trusted Execution Environments (TEE) or Trusted Platform Modules (TPM). This technique also has a variety of different implementations and ideas that have been worked upon the “isolation” solution. The main concept is briefly described as follows: the authentication data is kept separately in the device's processor, and encrypted with adequate encryption algorithms. The authentication takes place alone, as the OSs and the other functionalities are being executed parallelly, so the workload power-wise is greatly reduced and the procedure is highly secure – due to the fact that the authentication data is not involved with the other components of the device or the network. TPM works in a corresponding way; each device contains a chip where authentication data is stored and software has no access. When the device enters a network, since the

network identifies the device and match its key, the authentication is complete and secured. An example of such authentication techniques is described, using ARM Trustzone to create an architecture configured in this way; IIoTEED, an architecture intended to provide Trusted Execution functionalities in industrial – scale infrastructures. In this architecture, security operations and real time processing is happening inside the "trusted zone", and all the other processes are executed in the "untrusted zone" - that needs to be protected with other measures, if necessary. An overview of this proposed methodology is shown below:



Picture 5: Sandro Pinto et al., "IIoTEED: An Enhanced, Trusted Execution Environment for Industrial IoT Edge Devices. IEEE Internet Computing", February 2017

Of course, a lot more authentication schemes and methodologies are proposed by recent research and growth, in order to provide the best solution to deploy, depending on the situation. These techniques combine traditional IT authentication and cryptography methods, along with recent progress: ID/Password-based authentication, MAC address-based authentication, One Time Password usage, authentication with gateways (especially when communication with the cloud or the Internet is demanded – e.g. smart homes), Constrained Application Protocol and Datagram Transport

Layered Security certificate-based protocols, or even distributed authentication models such as blockchain, each of them is a great field of further investigation to define the strong and weak spots of their usage, and to determine which solution (and in what form) suits best for different occasions. In 2017, Atwady Y. and Hammudeh M. made a comparison between several of the different models presented or mentioned above, summing up the work of a lot of other researchers:

Proposed Technique	Strengths	Weaknesses
Certificate-based DTLS handshake delegation method	Reduce overhead of certificate based authentication	Need change to be done on DTLS Protocol
A middle gateway devices pass content from the IoT device to Internet and vice versa	Node is isolated from security attacks	New layer of hardware must be implemented
Gateway, controller and central data store authentication architecture	Use IPv6 address as a node identifier	Single point failure in case of CDS failure
ID-based multiple authentication scheme	Strong ability to prevent different attacks	Requires changes to current IoT architecture
Session key issued after RA and HRA negotiation	Remove the overhead of authentication from the node	Single point of failure
Four-way handshake added to CoAP protocol	Distributed-base solution	Vulnerable to Sybil attack
Nodes are grouped into layers managed by layer manager	Provide simple efficient authentication using DH key exchange	Not tested against known attacks
One time password using IBE-ECC	Smaller key size and no need for storing passwords	Requires changes to current IoT architecture
Use PUF and PKG to provide security and authentication	Abandon factory-deployed static keys needs	Requires hardware change
PUF-based protocols over elliptic curves	Low computational and storage requirements and low-cost tamper	Requires hardware change

Table 3: Atwady, Y and Hammoudeh, M, "A survey on authentication techniques for the internet of things", July 2017

Last but not least, lack of interoperability and standardization measures in these environments intensifies the problem of compatible and convenient solutions of authentication, when various devices (from different vendors) are used. Imagine devices that are contained in a smart home and authenticated through a central gateway, to interact with other devices from another vendor that use different communication and encryption protocols. Not only compatibility, but also security is challenged in these situations, as described previously. In order to address these issues, lots of regulation bodies are trying to set “benchmarks” or “baseline” requirements that the vendors should comply, in order to ensure that a minimum set of security goals are met when a product reach the consumers. To get a brief overview of this discussion, below are presented the basic requirements that NIST issued in August 2021 -as a draft- that needs to be met for IoT products. We can notice that several of the following requirements/principles – stated in a general manner that sets the vendor in the position to choose the best implementation of each control – already have been discussed as critical issues previously:

1. Asset Identification: The IoT product can be uniquely identified and can inventory all of the IoT product's components.
2. Product Configuration: The configuration of the IoT product can be changed, and such changes can be performed by only authorized individuals and other IoT product components.
3. Data Protection: The IoT product can protect the data it stores (across all IoT product components) and transmits (both between IoT product components and outside the IoT product) from unauthorized access and modification.
4. Logical Access to Interfaces: The IoT product can restrict logical access to its local and network interfaces, and to the protocols and services used by those interfaces, to only authorized individuals and IoT product components.
5. Software Update: The software of all IoT product components can be updated by authorized individuals and other IoT product components only by using a secure and configurable mechanism, as appropriate for each IoT product component.
6. Cybersecurity State Awareness: The IoT product can detect cybersecurity incidents affecting or effected by its components and the data they store and transmit.
7. Product Security: The IoT product can perform other features and functions across some or all of its components to make IoT products minimally securable for the sector.
8. Documentation: The ability for the manufacturer and/or the manufacturer's supporting entity, to create, gather, and store information relevant to cybersecurity of the IoT product and its product components prior to customer purchase, and throughout the development of a product and its subsequent lifecycle
9. Information and Query Reception: The ability for the manufacturer and/or supporting entity to receive information and queries from the customer and others related to cybersecurity of the IoT product and its product components.

10. Information Dissemination: The ability for the manufacturer and/or supporting entity to broadcast and distribute (e.g., to the customer or others in the IoT product ecosystem) information related to cybersecurity of the IoT product and its product components.

11. Education and Awareness: The ability for the manufacturer and/or supporting entity to create awareness of and educate customers and others in the IoT product ecosystem about cybersecurity related information, considerations, features, etc. of the IoT product and its product components.

Finally, we can easily assume that cyber security research and implementation of protective measure are an ongoing process, where lots of effort needs to be put in order to achieve the desired result in terms of security. The question, beside the technical requirements that need to be met, is to examine how data (and especially personal data) should be treated in all layers of the infrastructure, in order to be controlled properly and regulated, taking into consideration the special circumstances of IoT. From this point of view, and as the privacy issues are already described above, we are going to briefly study data handling measures in IoT, along with the regulative and legislative framework that engages with personal data in these environments.

4. Data Protection in Internet of Things

Still, the general picture remains to that; Internet of Things are yet to expose its whole capabilities, but, as “invasive” as these technologies may be, there is a great amount of work that needs to be put in terms of enhancing the measures required in order to maintain and ensure security and privacy. These privacy issues, regarding data handling and privacy requirements that are of high priority in these systems, and in many occasions the functionality of these systems is in the opposite direction, is the next part of our essay.

4.1 Law and Regulations: The Existing Framework and GDPR

There is a great deal of research in recent years, examining and analyzing the aspects of IoT that are “covered” by the existing laws and regulations (constitutional protection of fundamental rights, common law, GDPR and others) and what other aspects -or, in other words, conflicts among existing and traditional concepts of privacy- should be taken into consideration, that may lead to regulatory failure in these environments. In the previous sections we discussed about various privacy challenges that change the perspective that the organizations and the state face privacy, exceeding the part of security and touching also fundamental topics such as the distance between private and public life, the access and control of an individual's personal data, the combination of the data to depict behaviors and habits, and many other paradigms that impose the same controversy; what is the room of improvement and in which direction in order to maintain consumers' welfare and ensuring privacy at the same time? We will try to examine briefly the existing framework and in what part it matches the challenges that arise in the era of Big Data and IoT.

The main guidelines we will look briefly are the principles described in GDPR, ePrivacy Regulation (which refers to all electronic communications) and the various improvements (along with the local legal frameworks) in these texts which are the main frames that try to cover the existing Privacy challenges.

This framework outlines the basic principles that the parties involved in any data exchange and processing must comply with. These guidelines refer to the ethical and socially aware use of data collected, the need for preventive and risk assessment

policies (that need to be implemented before any data is collected or exchanged, especially personal information), the data collected should be limited to just the information that is enough for the according service or use, the adoption of the Privacy By-Design approaches, the correlation between the automation of the analysis (and the result) and the human responsible for the actions required, open data approaches to tackle data discrimination and proper education about the exact purposes of the data collection etc. These guidelines in essence are also part of the GDPR, attempting to combine classic privacy methodologies with the new emerging challenges.

In addition, GDPR and relevant national regulations affect widely IoT and in reverse. The whole change that these technologies bring in the current perspective of privacy leads to two assumptions: 1. There is a need to regulate the rampant collection and usage of data (that seems to outpace the according legislations), 2. The regulatory and law acts need to take into account the core principles that make these technologies unique; and try to balance these two needs with the purpose of ensuring privacy on the one hand, and on the other hand not to limit too much the capabilities of these technological environments.

In many ways, IoT and the relevant data analysis is affected by GDPR. The limitation on storage time (therefore the need for increasing the real-time analysis - anonymized), social media engagement, greater control over the personal data by the users and the choice to exchange the data to another vendor or permanently delete them are only some of the adjustments that need to be done. Additionally, the proposed guidelines and principles that GDPR enforces the vendors to comply with, have various conflicts with the existing IoT solutions. So, the IoT vendors and the existing technologies have to adapt in a variety of ways: choices about consent of the user (which data is collected, for what purposes, for how long, user's ability to withdraw the consent), data minimization and purpose limitation, enforcing the “right to be forgotten” and transparent processing (how the data is handled, the ability of the user to ask for the erasure of data), accountability of the vendor for possible data leaks or breaches. There are major privacy principles (along with the others we stated before such as the distinction between private and public information, the “right for distance” etc.) that need to be ensured by every IoT proposed solution and service. Also, these principles and their implementation in these emerging technologies is also

a matter of a more precise set of regulation and also a matter of sufficient supervision by independent and supervisory authorities.

Furthermore, as these technologies are yet to be implemented on large-scale there is still low levels of trust, both from the consumers and from the enterprises themselves, as the exact definition of when a technology solution is "compatible" with GDPR and many other aspects of Privacy are "blurring" in the context of IoT and the enormous amount of data created every day. Of course, from a privacy perspective there are fields that need to be looked upon first, when it comes to protecting sensitive information.

4.2 Personal data protection principles and IoT

In order to be more specific, we will try to examine basic privacy implications and principles that are outlined in the relevant regulations, and what challenges arise in their implementation in IoT:

- **"Distance between private and public life" -** IoT specifically is coming to assist not only in industrial and grid settings, but it is intended for wide personal use (e.g., smart homes, which we will discuss in the next sections). Its functionality is aiming to make services that provide personalization, based on our needs and habits. In order to do so, a lot of private/sensitive data is collected and analyzed.

- **Consent and transparency –** it's a fine line whether the consent of the user is given in a meaningful way (meaning that the user understands the range of the data collected, how it is going to be used and for what purpose, who will have access to that information etc.) and not in a typical manner, just as a GDPR requirement – supposing that even the consent is asked from the IoT vendors, in compliance with the regulation. Furthermore, it is not always clear who has access in the data collected, who is the owner of the information etc.

- **Data retention and "right to be forgotten" -** along with the previous matters of consent and transparency, it is at question whether the data collected are retained for a pre-set amount of time, and it is deleted without leaving traces after that period of time. In essence, the requirement here is that the data should be stored only for the amount of time that is necessary, but that does

come to a conflict with IoT analysis methods; in order for decisions to be optimal, former behavioral patterns must be used.

- **Data minimization and choice –** In most cases, it is not technically feasible for the user to have a choice over his/her "exposure", to what extent and what data actually will be recorded and stored. Typically, he/she cannot choose, for example, that a sensor will capture his fitness status or the kilometers he walked, but not the exact location. Also, it is a challenging issue of how "minimized" is the collection of the data; meaning that the data collected are the absolute minimum required.
- **Linkability and anonymization:** Anonymization techniques can be used in IoT environments, although this is challenged by the fact that, due to the massive collection of data that refer to one individual, it is possible to extract who that individual is. It is quite easy to suppose that, if a person has access to someone's location, activities, habits, health and fitness status, beliefs and cultural preferences, the exact matching with the actual person is not a very difficult task.

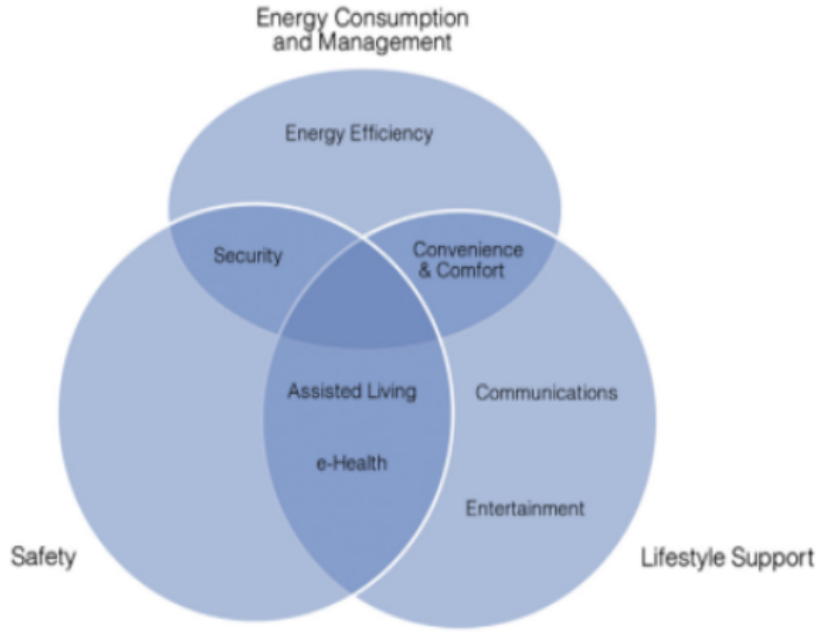
5. “Smart Homes”

One special and also major part of the Internet of Things “ecosystem” is that of the smart homes. Although many of the technology, protocol, security and privacy issues that we examined before are still present in this specific category of services and infrastructures, we should also examine the special circumstances that occur in this part of IoT, that make things – from a security and privacy perspective – a little “trickier”. There are lots of reasons why this is happening; the goal of the next chapters of this essay is to investigate further on the nature of smart homes' architecture and growth, particularly in how the general security and privacy challenges and concepts are imported in the “narrow space” of a smart home, and furthermore define and propose a methodology to engage with these concerns and structure a holistic approach to implement security and privacy controls in this environment: a security and privacy risk framework for a smart home. First of all, we need to examine the scope of this task and the characteristics of this family of services and technological solutions.

5.1 Defining the context

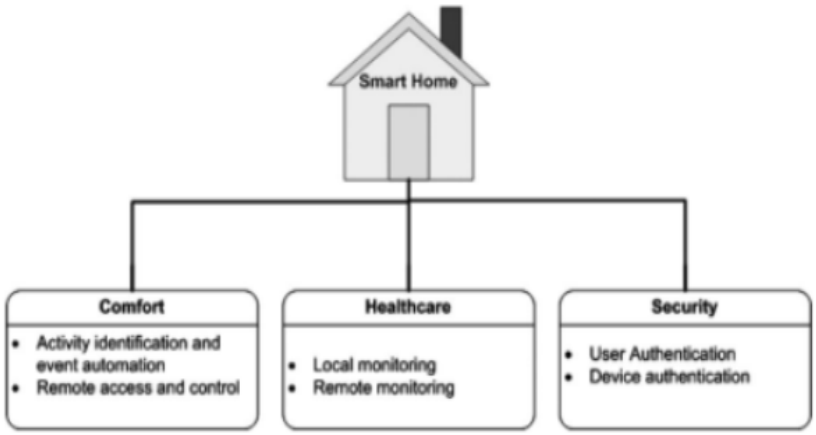
The first issue we have to clarify is, what is a smart home? Smart homes are a “new age” technology, which utilizes computing and information technology techniques to provide automation and control through smart appliances in traditional home functions and functionalities. Or else, smart homes is our effort, us a society, to improve life and well-being at home, using technology as our aid, engaging “smart” traditional electronic devices, robotics, automation and AI to add comfort and assistance in our everyday life – electricity, comfort, healthcare, safety and security. In this new approach, every device is converted to a “smart” device, with distinct features: 1. its ability to make computations and operate with simple commands through interfaces and 2. its ability to interconnect with other devices. Furthermore, contemporary smart home solutions engage Artificial Intelligence and Machine Learning to provide a third feature: 3. The ability of the device to learn from their environment and to adapt to the home's resident needs, habits and will. Remote control is present in most cases, meaning that it is very common for these services to be deployed and controlled remotely (via a mobile device in most cases), that usually

has complete control over the whole network – the whole house. Of course, some activities are more essential than others, along with the fact that smart homes try to cover critical aspects and needs of people's everyday life. Below we can see 2 figures where the basic aspects of smart homes' functionalities are presented:



Picture 6: Nazmiye Balta-Ozkan, "The development of smart homes market in the UK", 2013

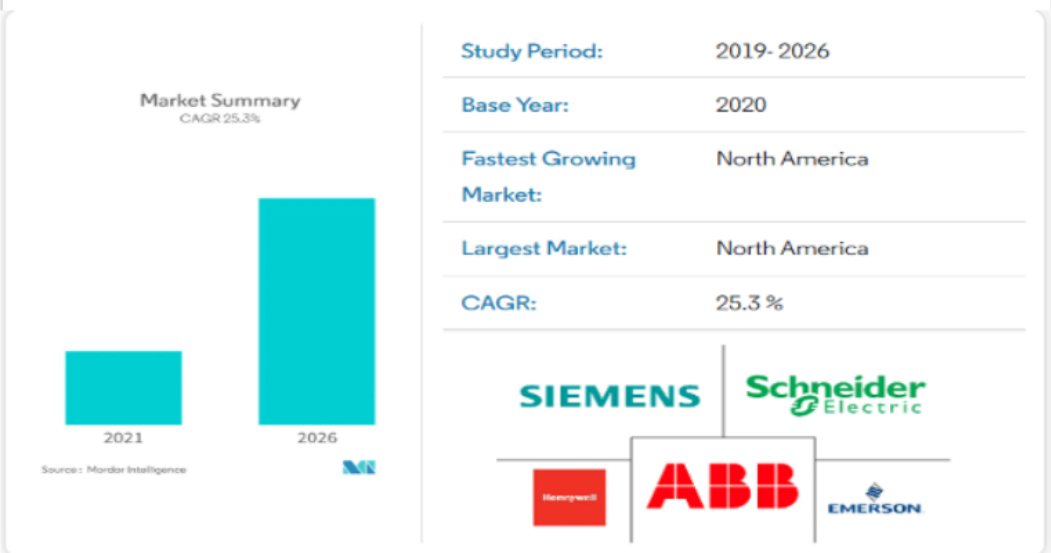
As we can see, three major categories of services are contained in this approach: Energy Consumption and Management, Safety and Lifestyle Support, that interoperate one another. Likewise:



Picture 7: Muhammad Raisul Alam, "A Review of Smart Homes—Past, Present, and Future", 2012

We can see a similar approach containing Comfort, Healthcare and Security. As we can easily understand, the main concept of smart homes is to provide assistance for a more sustainable and comfortable living, along with the engagement with fundamental problems of our days: safety and monitoring of the home, along with its residents, along with their health status and living assistance where needed. Many different smart home services derive from these directions, and the interconnection and management of these services inside a common infrastructure – in a house, a residence – makes the smart home.

First of all, similarly with the general Internet of Things industry, the situation in smart homes looks quite alike. An exponential growth of the market and the products provided to customers, and, due to this rapid growth and the rampant race between the different competitive enterprises (but also in the research field between the different solutions and techniques provided), a great gap and room for development in terms of cyber security and privacy. Indicatively, according to a research by Mordor Intelligence, was valued in 79.13 billion USD in 2020, and it is estimated to reach 313.95 billion USD by 2026 and register a CAGR of 25.3% over the forecast period (2021 – 2026). A vast growth, comparable and accompanied by the general growth of IoT, Big Data analytics, Artificial Intelligence and Machine Learning:



Picture 8: Mordor Intelligence, "Global Smart Homes Market - Growth, Trends, Covid-19 Impact, and Forecasts (2021 – 2026)", 2020

Along with the above data, according to Times, the consumers will spend 123 billion USD by 2021 and 7 billion devices will be connected.

There are some main characteristics of a smart home:

- An established network through which the different components of the network communicate with each other.
- Intelligent and management controls.
- Sensors.
- Smart features (that do not necessarily use sensors – e.g., a smart boiler) that respond accordingly to commands coming either from the sensor or the user.

In order to further investigate the broad and diverse nature of these systems – in order to better understand 1. what kind of data they collect and 2. how they should be protected – we will examine the different lines of services that are utilized in a smart home infrastructure, along with the communication protocols and network architectures that are used in these environments.

5.1.1. Smart Home Services and Technologies

As we already discussed, we will adopt the approach that distinguishes smart home services into 3 categories; comfort, healthcare and security, even though at certain occasions they are linked with each other. The reason is that, regarding the service that we are taking into account, different volumes and types of data are collected and of course, cyber risks that occur can be prioritized differently, along with the relevant risk acceptance. Roughly speaking, an insecure smart heart pacemaker can do more damage in an individual than a vulnerable lightbulb, for example.

1. **Comfort:** Applications and services that target the well-being and more comfortable living of the residents of the smart home. This is mostly achieved by 2 different types of applications; applications or sensors that are triggered by certain events or signals (e.g., a smart coffee machine that turns on when it is 7am.) or applications that enable remote control management of some activities that happen inside the home (e.g., turn up the water heater via a mobile device while outside of the house). Of course, the capabilities are even broader, while the focus today is that models are created that can learn user behavior, track and identify the user and automate their behavior depending on

the habits of the user. To use our previous example, the smart coffee machine has already developed an understanding to turn up from Monday to Friday at 7am., but 10 am., at weekends, or, to extent our paradigm further, when it is interconnected with another device – a wearable device – just a little before the user is awake. We already can see the great opportunities, but also the very sensitive information that these devices receive. Of course, using decision-making models, this can be extended to almost every electric device: lightbulbs, heat, windows, TV, radio etc. that can learn from the user's behavior and automate all of their tasks at the correct time and way, that suits the resident.

2. **Security**: lots of services at the security and safety of the home can be automated in a smart home. Alarms and CCTV cameras, remote control of the door of the house, remote access and monitoring of kids' or pets' activities, sensors that detect the moisture, gas or heat level of the home (to prevent a natural disaster), instant communication with the authorities or security agencies if an incident occurs, are all services that add to the value of the security and safety of the smart home, making it difficult for someone to break-in, or for any other disruption to happen. Of course, security risks arise already at the installation point of this services; a camera that monitors the entire house is a great target for a hacker or an intruder.

3. **Healthcare:** Perhaps one of the most important aspects of smart home capabilities, and at the same time the more sensitive one. There are many applications and services that can be run locally or from distance. An assistance tool for elderly people, or a fitness tracking app that is connected with gym facilities inside the home and gives personalized information and feedback. On the other hand, monitor and response services offered by a health provider. Specific tracking and monitoring services are developed that provide assistance with certain chronic diseases, or for elderly people and people at risk. The sensitive data collected; enormous. Healthcare services are probably the most “invasive” ones, that can be utilized for the greatest good, but also can be very harmful if not treated and protected properly.

5.1.2 Communication Protocols

As we briefly examined the services and devices that consist the home network, we have to also take a closer look at how the information collected from the different devices is transmitted through the network and of course how it is sent to the component where it can be useful; another device, a management framework/application handled by the user, a gateway or the cloud.

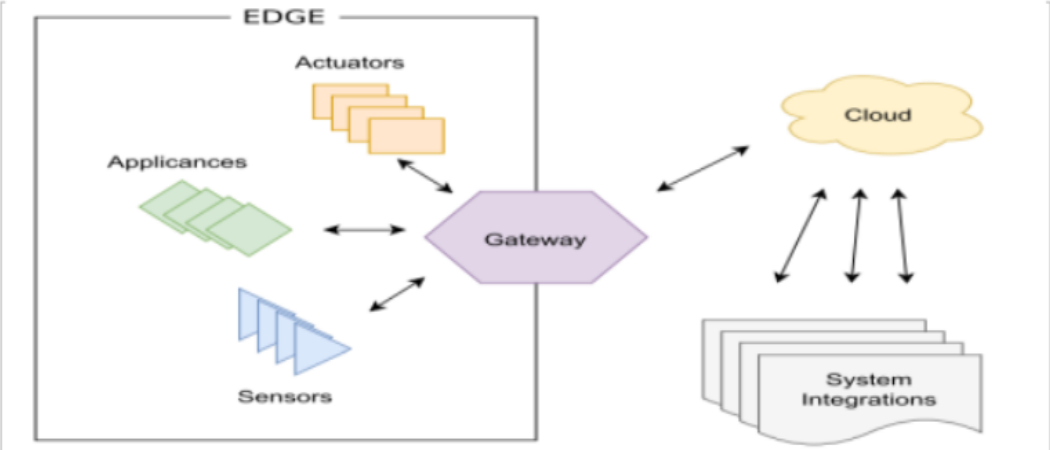
- **Wi-Fi:** Probably the most utilized communication protocol that is used for WLAN home networks, supporting various network devices (PCs, laptops, tablets, mobile phones etc.) is also the most used for the deployment of smart home appliances, providing connectivity with the same interfaces and way as all the already known connected devices.
- **Bluetooth:** Another widely utilized protocol, Bluetooth greatly serves the home networks as the limited area that it can cover is not a prohibitive factor because the connected devices are close to one another. It is also a very cheap and convenient choice, with simple connectivity options, functioning at the 2.4 GH, creating a very flexible and easy to use Personal Area Network (PAN).
- **ZigBee:** It also creates a PAN, using the IEEE 802.5.14 protocol for communication. Very limited cover area (from 10 to 100 meters), but also reliable, low energy and cheap communication solution. It supports multi-hop routing and demands one ZigBee coordinator.
- **Low Power Wide Area Networks (LPWANs):** This class of network protocols has many technologies that evolved in the recent years. It covers the need of lower bit-rate communications, where only signals and small transmissions of data are taking place, and very low power consumption is needed. Narrow band IoT, Sigfox, Weightless, LoRaWan are different protocols that belong to LPWANs.
- **RFID and NFC:** Radio Frequency Identification and Near Field Communication are of great use, especially in Wireless Sensor Networks (which can be a component of a smart home) that are deployed in different variants inside a home network.
- **Z-Wave:** Z-Wave is another wireless communication protocol, which is almost exclusively used in smart home infrastructures. It utilizes low

frequency magnetic waves to enable communication – direct or indirect – between the different nodes of the smart home, ensuring great interoperability between the various network components and appliances.

These are only the main communication protocols that are used in smart home infrastructures. There are lot more hybrid models that combine the above networks, depending the use case. After we overviewed the communication techniques and the different categories of applications that consist the IoT, we are going to analyze the different data collected in these environments, and how this data is used, transmitted, stored, and processed in smart home appliances.

5.2 Data collection and management: current "state-of-the-art"

First of all, we have to note that the fact that houses already possess automated features, that does not necessarily mean that they can be considered as "smart". For example, a camera that is recording the entrance of our home is an automated procedure, yet that is not a characteristic that distinguish the smart home from a conventional one. It is that exact device interaction (through the network that we already described) and the events, triggers and signals transmitted that enable the capabilities of the network to act as "smart". Moreover, even though the interconnections between the devices and sensors exist, it is not adequate yet. The part where the user does not necessarily need to intervene, and data is collectively stored and analyzed, to create behavioral patterns and automate decisions on behalf of the user, that is the point where the capabilities of the network are magnified. There are a lot of components in a home network: sensors, devices, servers, gateways, web applications utilized as a user interface, and the cloud, or even the web in general. Nowadays, the vast majority of applications are cloud-based, meaning that the pattern analysis and concentration of data happens mostly there. That means that data is, regardless of where the concentration and computational part is happening, transferred inside the network, to the gateway or server and to the cloud creating a large attack surface, that we are going to examine later on in this essay. In the picture below we can see a generic architecture of a smart home network. Of course, some components may vary depending on the occasion, but in general terms a smart home is consisted of these components:



Picture 9: Dragos Mocrii et al., "IoT-based smart homes: A review of system architecture, software, communications, privacy and security", 2018

As we can see, the data flow is pretty clear, large and contains different components that need to be protected in every level. In this process, a special part plays the gateway of the smart home, where multiple communication protocols are supported and also an extra layer of security is added. Furthermore, if the internal part of the network (devices, sensors, appliances) is not accessible from the internet, the gateway is a critical asset for network's security. Having said that, it is critical to see where the main part of the computing is happening and examine the according criticality of each component in the infrastructure. There are mainly 3 different architectures;

5.2.1 Cloud Computing

The main architecture/design of the smart home is similar with the generic architecture that we showcased before. In cloud-based architectures, which were the most typical example of a smart home architecture in last years, the whole aggregation of data coming from the whole network happens in the cloud, providing massive storage and processing options and capabilities. The cloud computing solution is highly reliable and scalable, though it faces the current cyber security risks that concern cloud-based aggregations, and another one problem; nowadays smart homes have the ability to create enormous amount of data. The cloud technologies have developed enough to provide adequate storage and computational resources, though it has not been the same for central network services – the world wide web speed and bandwidth. So, the problem is the cloud-based designs are very much dependent on the Internet speed, which brings us to the second architecture.

5.2.2 Fog computing

Fog computing is willing to address this exact issue; bridging the distance between the generation of data and the processing and concentration of data. This is possible via the gateway of the network. In fog computing methods, the gateway of the network acts as an “intermediate” in the computational part, between the devices and the cloud. That means that not the whole bunch of data are sent to the cloud, but data incoming from the sources are pre-processed and compressed in the gateway, so information sent to the cloud is smaller, not “raw” and ready to be processed. The benefits are that much less bandwidth is needed, computations in the cloud level are faster, and also are sent back faster, reducing the overall latency. Also, the reduction in the send-response time creates a pretty much “live” environment, where the sensors collect data and get responses almost instantly, so the decision making and adaptation procedures are less time-consuming. Last but not least, it is a better approach to address “single point of failure” issues, because, even in case of a disconnection from the Internet, the data is still gathered in the gateway and could be sent when the line is up again.

5.2.3 Edge computing

The approach of edge computing is similar to the fog computing method, but with a different implementation. It is a more decentralized approach, because in this model, each device (depending on the situation - devices are trained to do so with ML and AI methods) is able to make a decision whether data should be trashed, kept locally, processed or sent to the cloud. It is a technically more complex methodology, though it opens up a lot more options and possibilities for the home network, and also reduces “single point of failure” risks.

5.2.4 Data Collection and PIIs

As we can see, smart homes are basically defined by the ability of the network to act “alone”, without the intervention of the user. In order to do that, smart homes are based on a handful of incoming data, tracing back a lot of the residents' activities and habits, actually extracting an exact and specific view of almost every part of their everyday lives. It is, as already said, the most “invasive” set of technologies that exists until now, and if misused, it can be a great threat to one's personal life and activities; actually, it can lead to surveillance, either directly – through a hacked camera, a case which we already discussed – or indirectly, through behavioral

analysis based on the type of aggregated data. Which data types are collected, and – more importantly, from a privacy and security perspective – what actual information about one person or family can be extracted if a privacy violation occurs in such an infrastructure:

- Through the "comfort" line, there are various sources of information: light bulbs, heating, TV, radio etc. which not solely reveal the habits and activities of the residents, but also can be quite dangerous. A malicious actor can extract audio and video recordings, specify the hours that the residents are off-site, estimate house's exact location, analyze residents' cultural preferences, political and religious beliefs etc. The data collected and processed in the cloud must be carefully handled in order to ensure that the information coming from these sources cannot be exploited in any way – neither from a hacking attack, nor from the vendor itself or some third party.
- Through the "security" line, things are becoming a lot more obvious. It is easy to assume that – a hacked camera or a hacked perimeter surveillance system can give very important information to any interested party. This sort of information can result in a range of cyber-physical issues; a remote "shut down" of the system can lead to a physical entrance in the home, 24X7 surveillance of the residency can be achieved, a clear "mapping" of the house and the everyday life and social life of the residents, all these are very sensitive information that can be exposed with vulnerable smart security systems.
- In the "healthcare" section of these infrastructures, we can locate even more "intrusive" data coming to the processing facilities (gateways and cloud). The very nature of these devices, due to the fact that they intend to provide personalized healthcare monitoring and care, captures and processes a lot more information regarding the health status of a resident/patient. In fact, the whole clinical status of the resident can be accessed and can be used for a handful of reasons: profiling, blackmailing or even more "dystopian" causes – social security reasons, pharmaceutical big data analytics etc.

We can see that each and every part of a smart home creates valuable assets for malicious actors, in a more direct and targeted manner than most of the other typical

IoT environments. This situation creates a wide attack surface for the attacker, and specific potential security and privacy risks that need to be addressed, in order to establish that these technologies can provide the full capacity of their capabilities and not becoming a threat themselves in the same time, for the privacy of their users.

5.3 Privacy & Security Risks in Smart Home Infrastructures

In order to examine the emerging threats in smart home infrastructures, we are going to follow the same approach as we used in the general IoT architectures. We will divide the three layers that we identified in our previous analysis, the perception/device layer, network layer and cloud/gateway level, and furthermore examine the threats that occur in each level accordingly. Note that, we will include the cyber-physical threats that occur due to the presence of third-party malicious actors or actors that are involved in the data processing procedures or potentially could have access to this data (vendors, authorities, automation third-parties, governments etc.) but not internal intentional security issues imposed by the residents or visitors. In our threat model we will include also physical threats, and data misuse issues that occur in these environments, in order to try to compile a holistic view over the subject, along with the proposed countermeasures and guidelines that need to be established to address the privacy and security risks discussed.

5.3.1 Threats in the perception/device layer

As smart home infrastructures are set in a similar way as the other IoT networks, we can notice a great similarity in the threat factors that apply to these cases. That means that the specific features that characterize these networks, in comparison to more conventional IT infrastructures, are in force also in smart homes largely. Resource constraints play a major role, especially in cryptography and computational capacity, due to small-throughput of RAM memory and small batteries. Additionally, we also have a constrained number of user interfaces in these devices, which not only troubles the user in terms of operability, but questions significant issues of privacy, such as alerts when a potential breach occurs, or in consent-related matters. Of course, physical tampering of the devices is quite possible, given the fact that the devices is out in the open, inside a house or in the backyard. Maybe a more secure setting than IoT environments out in the open, but still susceptible to physical attacks. Last but not least, privacy-related matters also arise in this layer, due to the exact process of data creation, which in most occasions is rather “unclear”, speaking from a data

minimization and transparency point of view. The users are not quite sure what is the extent of their consent, meaning that more data could be created than the original (and theoretically agreed with the user) purpose. This includes also technical limitations, how much personalized a service can be that takes into account its functionality but also users' consent – which may vary depending the occasion. The general picture is that vendors override this in favor of the “usability” of the service – same as other IoT services and technologies – and devices create more data than the purpose (e.g., tracking information in cameras).

Taking into account these challenges, we can sum up the main threats in this layer:

- **Identity theft:** This type of attack covers a lot of cases and scenarios, where the primary goal of an attacker is to target the integrity of the network and establish himself (in the device level) as a legitimate user of the network, taking control of a smart device. This can happen due to open and unprotected ports in a device, hardware attacks, key theft or generation etc.
- **Software and hardware code execution:** This type of attacks can happen mainly due to vulnerable software and hardware in the devices, that an attacker can exploit to achieve the execution of malicious components, with lots of possible consequences for the system (denial of service, eavesdropping, manipulation of the data processing, tracking etc.)
- **Data leaks:** Attackers are usually able to obtain lots of information, taking advantage of the low-level of encryption and the use of wireless communication to intercept traffic coming from the devices, for further analysis. In this manner, an attacker can extract a lot of sensitive data about the home's residents, as we showcased before. Network reconnaissance and information gathering, replay of messages, man in the middle and session hijacking are only some examples of attacks that can be launched and result in data leakage.
- **Cyber-physical attacks:** This type of threat mainly takes advantage of the physical specific features of a smart home, to achieve certain alterations in the network. This can be caused due to improper design during the development phase of the product, or with the existence of unprotected hardware

components that an attacker can use to access the OS and data of a device (USB ports, SD slots).

- **Information manipulation/falsification:** This type of attacks can be used to manipulate or alter the information created and processed in the network. It mainly challenges the lack of non-repudiation techniques in a home network, and can lead to errors in the logging of the system, bad computations and decisioning by the system, repudiation faults, overflood of the network etc.
- **Jamming attacks:** Through the radio transmission of a high-power radio source (jammer) the attacker can cause a lot of malfunctions in the network's integrity and availability. It is one of the primary threat factors that can be exploited in the physical layer.
- **Tampering:** Straight-forward type of attack, which takes advantage of the low-level physical security of a smart home network, tampering the device, making it unavailable or completely unserviceable in a lot of occasions.
- **Failures/malfunctions and natural disasters:** This last type is not a specific attack that can be launched, but establishes a threat as long as a device is susceptible to service failures due to technical construction issues, end of life issues or even due to natural disasters (fire, flood etc.).

5.3.2 Threats in the network layer

In the second section of our analysis, we will include the attacks and threats that target the communication part of a home network, mostly between the devices and the gateway and/or the cloud. We still can see that the limitations of the IoT technologies are relevant in this section too; interoperability issues, as there are multiple communication protocols between the different components of the network, as we already described. This matter becomes a a lot more complicated when equipment from different vendors is used under the same infrastructure, and communication between the different devices is needed. Also, there is a “mobility” issue, due to the fact that there can be devices that function both in and out of the home network (e.g., wearable devices), so the encryption and authentication challenges in these situation – taking into consideration the whole “inhomogeneity” of the network, and the lack of adequate standardization protocols – are even more prominent in this layer:

- **Malicious code execution via network links:** This type of threats refers to the general ability of an attacker to execute remote code to one or more parts of the network. It can vary from DDoS attacks from a subnet of the network organized as a botnet, to single command executions in vulnerable parts of the network. Critical for the success of these types of attacks are flaws in the network configuration or weak authentication credentials or schemes.
- **Denial of service:** This threat is established due to the inefficiency of many home networks to allocate adequate computing and power resources. It can be relatively easy to launch DoS attacks in home networks, overflooding the network with consecutive requests, making it difficult for the devices to reject or process the incoming data, causing unavailability of the service provided.
- **Unauthorized access:** These threats comprise all the vulnerabilities that an attacker can possibly exploit to gain access to the network. The threat factors in this type of attacks can be multiple: insecure network services and configuration, software/firmware/hardware vulnerabilities, lack of transport encryption, security flaws in web interfaces etc.
- **Malfunctions affecting the Internet connection:** Not a type of attack, but a threat factor for the availability of the services, due to the fact that lots of smart home applications are widely dependent on the stability and bandwidth of the Internet connection, as already examined.

5.3.3 Threats in the service layer

In this layer we have possibly the more “blurred” lines in terms of where exactly to look for a security gap or threat, due to the wide variety of different architectures and schemes, and the lack of general security by-design development techniques. In this point, we will try to examine the threat factors that may arise in the application layer and the data transport to the cloud and the handling later on. Although some issues may touch topics of cloud security; we will attempt to have a thorough look to the data management and processing issues that can potentially expose users' privacy and security. Of course, the main challenge here is primarily the monitoring, update and support from the vendors' side, along with a lot of challenges considering privacy implications and concepts, that we already tried to describe regarding the general situation of the IoT technologies. In general, the interaction between the user and the

vendor regarding security, technical support, consent, transparency, user control over data handling, third-party data management, secondary use of data, are primary issues in this point:

- **Denial of service:** The same type of attack can also be launched in this layer, taking advantage of weak software components of the initial product bought by the user, overloading the responding server or gateway with packets etc.
- **Sensitive information disclosure:** These types of attacks are generally targeted towards the cloud side of the infrastructure. The main goal is to extract confidential information and PIIs exploiting vulnerabilities in the cloud security measures. Also, we should note that the sensitive information could be misused from an internal bad actor (e.g., an employee that has access to that type of information – either subject to access control or not) or from a security breach that could happen directly to the service provider that infects the cloud services.
- **Unauthorized access – privilege escalation:** One of the most prominent issues in all the layers. A lot of weak spots in the application/service layer can lead to that kind of exploits, leading to great damage due to the capabilities provided to the attacker by this kind of attacks – in many times total control over the service, since he/she gains administrative rights. Insecure web interfaces (this can happen for varying reasons e.g., weak credentials, poor encryption and authentication protocols, web vulnerabilities etc.), lack of authentication or authorization, insufficient protection measures of personal data, social engineering, misconfiguration of interconnections between the components, insecure mobile interfaces, security gaps in the cloud interface and/or infrastructure are a number of vulnerabilities that can be exploited for this kind of attacks.

It is obvious, as we already examined, that all of these potential security breaches and threat factors lead to a direct exposure of personal data. After examining the situation and the possible threats, we are going to move on to describe general countermeasures against these potential threats, and also try to provide a framework of fundamental security and privacy principles that need to be met in smart home infrastructures in order to provide an adequate level of privacy. Last but not least, we will examine

future directions of the ongoing discussion and research concerning cyber security and privacy in the field.

6. Ensuring Privacy in a “Smart Home” environment

6.1 Focus areas and scope

As already discussed, there are many drawbacks and open problems that remain to be addressed in the open discussion about smart homes and their security and privacy. Especially, regarding privacy, many data handling techniques and principles imposed by relevant legislation and GDPR in the EU, are tackled by the rapid deployment of such solutions, the “usability and mass production”-first approach and the limitations of those legislative frameworks. We will try to showcase some countermeasures and techniques to “close” possible security and privacy gaps presented at the previous chapter, and then propose a general methodology of evaluating the security and privacy status of the infrastructure. We should also notify that, many of the issues that we will try to address are still under current research, so we will try to showcase general measures and principles in order to mitigate privacy and security issues.

The problems that arise in this context can be divided in two levels: the data creation, transport and process when data is in control of the home network (that means, just until data reaches the gateway) and the rest of the process that happens in the cloud, the storage, access and processing of data. The scope is to propose considerations that can implement security and privacy end-to-end, during the whole lifecycle of the system. The model that we will try to analyze is as presented above; for that reason, we will suppose that wearable and mobile devices that are components of the network are managed under the same infrastructure, meaning that they interact with the devices and the cloud either to exchange data as the other devices, or for user and admin activities (e.g., view history, set settings for the smart home etc.). Along with that, note that the technical and organizational limitations that were described before (resource constraints, lack of standardization, multiple communication protocols, data handling and access issues, mass production not “privacy-by-design”-based, privacy constraints) are still present in the proposed principles.

In order our handbook to become as much “holistic” as it could be, we will try to include users' perception about vulnerabilities and threats in their home network. According to a relevant survey, the tables below present the main threats and vulnerabilities that are considered by smart home users:

Vulnerabilities	Concerned	Mentioned but not concerned
Data at risk in the cloud	1/15	5/15
Weak passwords	5/15	0/15
Lack of transport level security	4/15	0/15
Insecure devices	4/15	0/15
Malicious devices	3/15	0/15
Unsecured Wi-Fi network	2/15	0/15
Devices can be unpaired	1/15	0/15
No identified vulnerabilities	3/15	

Table 4: Eric Zeng, "End User Security and Privacy Concerns with Smart Homes", July 2017

Threats	Concerned	Mentioned but not concerned
Continuous audio/video recording	3/15	5/15
Data collection and mining	1/15	5/15
Adversarial remote control	4/15	1/15
Network attack on local devices	3/15	1/15
Spying by other user in home	3/15	0/15
Account/password hacking	2/15	0/15
Network mapping by mal. devices	1/15	0/15
Re-pair device with attacker's hub	1/15	0/15
No identified threats	1/15	

Table 5: Eric Zeng, "End User Security and Privacy Concerns with Smart Homes", July 2017

We can notice that the threats and vulnerabilities mentioned are pretty similar with the cases we examined in previous chapters. In order to address the issues and challenges above, we will use the more "conventional" scheme about cybersecurity, in order to cover all the subjects: Confidentiality and Data Protection – Integrity, Authenticity and non-Repudiation – Availability and specific Privacy controls. The last part will try to specifically address privacy issues that derive in all the data flow of a smart home infrastructure and are not covered in the previous "categories".

6.2 Proposed principles and methodology

We will now develop our model-handbook for an end-to-end security and privacy approach, as we tried to describe it in the previous section:

6.2.1 Confidentiality and data protection

1. **Identify local storage and processing:** Determine whether data are stored or processed at the device level. If yes, measures about local data storage (encryption) and measures against tampering and physical threats should be

taken (safekeeping and restricted access to the devices, locked USB ports while changes are not necessary etc.)

2. **Identify communication protocols and narrow interoperability issues:** Secure and encrypted protocols should be used between the devices, protected wireless access with strong passwords, compatible components with security features should be used. Additionally, configuration issues should be dealt at the start of the usage of the network and trusted links between the heterogeneous devices should be established, if applicable.
3. **Establish a secure encryption scheme:** Possibly one of the main challenges. Identify the protocol and methodology used (static/dynamic key creation, private or public key infrastructures, Elliptic Curve algorithms for managing resource constraint), consider infrastructure limitations, establish measures for session keys if applicable, consider apply additional privacy measures (e.g., homomorphic encryption and others).
4. **Establish a secure link between the network and the cloud/aggregator:** Ensure that the link is not over the public network, establish VPN links and other encrypted methods of communication.
5. **Create a secure authentication mechanism for the users:** Measure password's strength for authentication, consider applying Multi-factor Authentication for privileged use, restrict the possible log in/log out areas (e.g., SSO methods via a central interface – a mobile app), establish identity management for different users if applicable.
6. **Restrict Internet access for the meters and devices:** Ensure that the devices' communication is happening only via the gateway and secure links with the internet, properly filter incoming traffic to the devices, apply properly tuned firewalls.
7. **Establish monitoring methods for the network / security logs:** Identify a manner to keep the user up-to-date about the status of the network, consider secure logging for actions happening inside the network, establish intrusion prevention techniques in the gateway to locate any suspicious activity, enable alerts about possible threats, establish technical support and maintenance procedures for updates and patches.

8. **Apply additional privacy-preserving techniques:** Anonymization techniques before data is send to the cloud, “trusted aggregators” method (where the computation is made by a third party, and the vendor only receives the processed data to send it back to the home network), perturbation models and data obfuscation techniques, zero-knowledge and verifiable computational models.

6.2.2 Integrity, Authenticity and non-Repudiation

1. **Establish an appropriate cryptographic hashing technique:** Identify secure hashing techniques (e.g., SHA-3) and implement integrity verification across the network.
2. **Maintain a clear network topology:** Establish an inventory of the devices in use, consider identification issues, establish authorization for the connected devices, ensure only authorized devices are part of the network. Establish digital signatures or other techniques for maintaining the status of the network (e.g., digital watermarking, PUFs etc.).
3. **Consider IPS/IDS monitoring:** Establish effective mechanisms to monitor and trace anomalies in network traffic, use anomaly-based detection schemes (this also applies to Availability issues).
4. **Use verification schemes and techniques:** Apply package “homomorphy” to the network, verify the integrity of the message on each device and in the gateway by using timestamps, sequencing, checksums, session keys etc.

6.2.3 Availability

1. **Establish effective Internet connection to support the infrastructure:** apply high availability to the connection between the network and the cloud.
2. **Establish physical security measures:** restrict access to the devices used and the components of the network, attempt to narrow the usage of the devices to external territories. If necessary, apply additional security measures to those devices.
3. **Apply additional availability techniques:** Use power management techniques to reduce the chance of an internal failure, use multiple alternate frequency channels to change frequencies when anomalies are detected (e.g.,

possible jamming attacks), or alternative schemes (e.g., Frequency Quorum Rendez-Vous etc.).

6.2.4 Privacy Controls

In addition to the techniques proposed mainly in the "confidentiality and data protection" section, there are a handful of other privacy issues that need to be addressed, as also described before, that need to be included in a general security and privacy evaluation of a smart home network:

1. **Adopt Privacy by-design approach:** Consider using techniques that do not expose personal data to neither the provider of the service nor third parties. Apply techniques to make computations and provide technical support without the prior knowledge of actual PIIs.
2. **Meaningful and time-specific consent:** Establish ways to explain the extent of usage, storage and processing of personal data, along with the specific amount of time that the data will be stored and used. Make this procedure a formal requirement for the purchase of the service.
3. **Establish off-line modes and utilities:** Implement an approach that the user can still use the network – even with minimum functionalities – without being connected to the internet ("right to disconnect").
4. **Identify channels of communication between the user and the vendor:** Create procedures to allow the user to revoke his/her consent, delete his/her personal data, or change the consent permissions.
5. **Clarify the involvement of third-parties in data processing:** Notify the users if their data are going to be used – and how – by a third entity.
6. **Establish data minimization and transparency:** collect and analyze only the absolute necessary information, give the users a clear view of the whole process, minimize the collection of PIIs.
7. **Increase the levels of user – awareness:** Guidelines with best practices about security and privacy should be introduced built-in to the products, notify the users about the risks, update the instructions and guide-lines at regular intervals.

6.3 Open Challenges

In the last section of this essay, we will try to sum up the challenges that derive from all the previous analysis, and that still remain a subject of research, while the technologies discussed are still expanding.

Although several methodologies already proposed and discussed, a lot of them actually functioning in real environments, encryption and authentication are still long before we could actually end up with a single solution – or some general standards at least. The resource constraint in these environments leads to various different solutions as we already discussed about, which all could suit for a particular model or network architecture. Especially, lack of universal identity management in network and application-level cases, adds complexity to this topic. Along with that, the need for lightweight cryptography and the relevant research for even more secure but also "lighter" from a computational point of view solutions, comes to the discussion if we want to provide some solutions that cover the whole subject. Different architectures are also entering the discussion, distributed environments, blockchain models, in order to identify the best possible solutions in usability and security combined at every occasion.

Privacy issues as discussed in the previous chapter, are also a challenge. The lack – and even contradiction – of these technologies with the existent frameworks, principles and methodologies is under discussion, along with the very little progress done by the manufacturers to introduce adequate privacy measures to the deployment of IoT and smart homes. Also, this comes across with many technological limitations (or not cost-effective, such as the remote effective support of a smart home network for a user) and on contrary with current technological design principles.

Lack of technical standardization widens the interoperability problem. In essence, the vendors select the methods and schemes that already are working for them, and the incompatibility between various components under the same infrastructure remains, importing various security and privacy risks as we already saw.

Lack of monitoring and additional methods of patching and updating the infrastructure are also a great issue. Additional security mechanisms (firewalls, IDS/IPS etc.) add a lot to the technical complexity of each infrastructure, an infrastructure that cannot be monitored and updated regularly by the vendors already.

Users' ignorance of security risks magnifies this issue. Typically, the same way the devices are pre-installed and configured by the vendor, the same way they remain until they are changed (or breached).

Last but not least, the lack of risk and vulnerability assessments for these environments widens the problem of security and privacy in these infrastructures. The lack of frameworks for testing, or holistic methodologies that can include every single possible security gap leads to the fact that – in a great extent – a lot of home networks and other IoT networks cannot be tested and audited appropriately.

Conclusion

Truth is, that the challenges we mentioned in the whole essay are not tackled directly by the existent technical, organizational and regulating approaches and principles proposed until now, so additional work needs to be done in that direction. Implement

end-to-end secure solutions, establish effective encryption and authentication schemes, defining the exact information that is allowed be collected, setting limits on the purposes of the data processing, reducing and regulating the processing that large organizations apply to hundreds of thousands of different people of data, educate the technology users and create "trustful" solutions, minimize the leakage or private information created by IoT, dealing with the problem of profiling and tracking, produce more "unlinkable" ways of anonymizing data, defining a more precise legal framework; these -and all the other situations that are mentioned above- are emerging open risks when it comes to ensuring and enhancing security privacy in these environments. The rapid growth of IoT, the existence of a great number of enterprises, research labs and manufacturers that develop these technologies may seem overwhelming and may initially change some of the accumulated patterns that were used for security and privacy, but obviously the broad adoption and adjustment in these technologies is the only way to go (regarding the capabilities that they offer). Taking that into consideration, we attempted to examine and narrow down the major risks and aspects that are threatening security and privacy, concerning these emerging technological solutions, and propose a basic guidance handbook including the risks and the essential controls that are required to ensure a security and privacy baseline in a smart home environment. In conclusion, we can easily understand that there is a great deal of work that needs to be done to ensure security and privacy, both from a regulatory and technical point of view, at state, multinational, academic, enterprise and social level, in this ever-changing digital environment.

References

[1] L. Mitrou, "*Personal Data and GDPR",* [presentation], March 2020

[2] S. Warren, L. Brandeis, "The Right To Privacy", December 1890

[3] S. Gritzalis, "*Privacy Framework ISO 29100:2011/2017*", [presentation], 2020

[4] Mehmet Bilal Unver, “Turning the crossroad for a connected world: reshaping the European prospect for the Internet of Things”, International Journal of Law and Information Technology, March 2018

[5] Lynn T., Endo P.T., Ribeiro A.M.N.C., Barbosa G.B.N., Rosati P, “The Internet of Things: Definitions, Key Concepts, and Reference Architectures”. In: Lynn T., Mooney J., Lee B., Endo P. (eds) The Cloud-to-Thing Continuum. Palgrave Studies in Digital Business & Enabling Technologies. Palgrave Macmillan, Cham., 2018

[6] R. Mahmoud, T. Yousuf, F. Aloul and I. Zualkernan, "Internet of things (IoT) security: Current status, challenges and prospective measures," 10th International Conference for Internet Technology and Secured Transactions (ICITST), 2015, pp. 336-341

[7] Fortune Business Insights, “Internet of Things (IoT) Market Size, Share & Covid-19 Impact Analysis”, July 2020

[8] So-Eun Lee, Mideum Choi, Seongcheol Kim, “How and what to study about IoT: Research trends and future directions from the perspective of social science”, Telecommunications Policy, Elsevier, 2017

[9] Kanwalpreet Kour et al., “IoT: Systematic Review, Architecture, Applications and Dual Impact on Industries”, IOP Conf. Ser.: Mater. Sci. Eng. 1022 012053, 2021

[10] Akkaş, M. A., SOKULLU, R., & Ertürk Çetin, H., “Healthcare and Patient Monitoring Using IoT. Internet of Things”, 2020

[11] Gilad Rosner, Erin Kenneally, “Privacy and the Internet of Things: Emerging Frameworks for Policy and Design”, CLTC Occasional White Paper Series

[12] Katie Boeckl et al., “Considerations for Managing Internet of Things (IoT) Cybersecurity and Privacy Risks”, NISTIR8228, June 2019

[13] N Alhalafi and Prakash Veeraraghavan, “Privacy and Security Challenges and Solutions in IOT: A review”, 2019, IOP Conf. Ser.: Earth Environ. Sci. 322 012013

[14] Varshney, T., Sharma, N., Kaushik, I., & Bhushan, B.,“Architectural Model of Security Threats & their Countermeasures in IoT”, International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), 2019

[15] Business Insider, “The security and privacy issues that come with the Internet of Things”, 2020

[16] Tiago M. Fernández-Caramés, Paula Fraga-Lamas “Teaching and Learning IoT Cybersecurity and Vulnerability Assessment with Shodan through Practical Use Cases”, June 2020

[17] Emil Sayegh, “Peloton Breach Reveals A Coming IoT Data Winter”, Forbes, July 2021

[18] Jo Vanwell, “IoT Security Breaches: 4 Real-World Examples”, Conosco, January 2021

[19] Paul Fremantle, “PRE-PRINT: A security survey of middleware for the Internet of Things”, July 2015,

[20] Rwan Mahmoud, Tasneem Yousuf, Fadi Aloul, Imran Zualkernan, “Internet of Things (IoT) Security: Current Status, Challenges and Prospective Measures”, The 10th International Conference for Internet Technology and Secured Transactions (ICITST-2015)

[21] ENISA, “Guidelines for Securing the Internet of Things”, November 2020

[22] Lo’ai Tawalbeh, Fadi Muheidat, Mais Tawalbeh, Muhannad Quwaider, “IoT Privacy and Security: Challenges and Solutions”, Applied Sciences MDPI, June 2020

[23] An Braeken, “PUF Based Authentication Protocol for IoT”, Applied Sciences MDPI, August 2018

[24] Yilmaz, Y., Gunn, S. R., & Halak, B, “Lightweight PUF-Based Authentication Protocol for IoT Devices”, 2018, IEEE 3rd International Verification and Security Workshop (IVSW)

[25] Karthikeyan, S., Patan, R., & Balamurugan, B., “Enhancement of Security in the Internet of Things (IoT) by Using X.509 Authentication Mechanism” Recent Trends in Communication, Computing, and Electronics, 217–225, December 2018

[26] Pinto, S., Gomes, T., Pereira, J., Cabral, J., & Tavares, A., “IIoTEED: An Enhanced, Trusted Execution Environment for Industrial IoT Edge Devices”. IEEE Internet Computing, 21(1), 40–47, January 2017

[27] Atwady, Y and Hammoudeh, M (2017) A survey on authentication techniques for the internet of things", International Conference on Future Networks and Distributed Systems (ICFNDS 2017), Association for Computing Machinery (ACM), July 2017

[28] NIST, "DRAFT Baseline Security Criteria for Consumer IoT Devices", August 2021

[29] K. Solins, "IoT Big Data Security and Privacy vs. Innovation", 2018

[30] I-Scoop, "IoT regulation: IoT, GDPR, ePrivacy Regulation and more regulations", 2018

[31] Consultative Committee of the convention for the protection of individuals with regard to automatic processing of personal data, "Guidelines on the protection of individuals with regard to the processing of personal data in a world of Big Data", January 2017

[32] Jenna Lindqvist, "New challenges to personal data processing agreements: is the GDPR fit to deal with contract, accountability and liability in a world of the Internet of Things?", International Journal of Law and Information Technology, 2017

[33] Daniel Bastos et al., "GDPR Privacy Implications for the Internet of Things", December 2018, 4th Annual IoT Security Foundation Conference

[34] Patric Lucas Austin, "What Will Smart Homes Look Like 10 Years From Now?", Times, July 2019

[35] Mordor Intelligence, "Global Smart Homes Market - Growth, Trends, Covid-19 Impact, and Forecasts (2021 – 2026)", 2020

[36] A. A. Zaidan et al., "A survey on communication components for IoT-based technologies in smart homes", Springer, March 2018

[37] Muhammad Raisul Alam, Mamun Bin Ibne Reaz, Mohd Alauddin Mohd Ali, "A Review of Smart Homes—Past, Present, and Future", IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews, vol. 42, no. 6, November 2012

[38] Sharda R. Katre, Dinesh V. Rojatkar, "Home Automation: Past, Present and Future", International Research Journal of Engineering and Technology (IRJET), October 2017

[39] Nazmiye Balta-Ozkan, Rosemary Davidson, Martha Bicket, Lorraine Whitmarsh, "The development of smart homes market in the UK", Elsevier, September 2013

[40] Dragos Mocrii, Yuxiang Chenb, Petr Musileka, "IoT-based smart homes: A review of system architecture, software, communications, privacy and security", Elsevier, 2018

[41] Jordi Mongay Batalla, Athanasios Vasilakos, and Mariusz Gajewski ,"Secure Smart Homes: Opportunities and Challenges. ACM Comput. Surv. 50, 5, Article 75, September 2017

[42] Eric Zeng, Shrirang Mare, and Franziska Roesner, "End User Security and Privacy Concerns with Smart Homes", Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017), July 2017

[43] Yaxing Yao, Justin Reed Basdeo, Smirity Kaushik, and Yang Wang, "Defending My Castle: A Co-Design Study of Privacy Mechanisms for Smart Homes" In Proceedings of CHI Conference on Human Factors in Computing Systems Proceedings, Glasgow, Scotland UK, May 2019

[44] Joseph Bugeja, Andreas Jacobsson, Paul Davidsson, "On Privacy and Security Challenges in Smart Connected Homes", European Intelligence and Security Informatics Conference, 2016

[45] Yan Meng, Wei Zhang, Haojin Zhu, and Xuemin (Sherman) Shen, "Securing Consumer IoT in the Smart Home: Architecture, Challenges, and Countermeasures", IEEE Wireless Communications, December 2018

[46] Antorweep Chakravorty, Tomasz Wlodarczyk, Chunming Rong, "Privacy Preserving Data Analytics for Smart Homes", IEEE Security and Privacy Workshops, 2013

[47] Fraser Hall, Leandros Maglaras, Theodoros Aivaliotis, Loukas Xagoraris and Ioanna Kantzavelou, “Smart Homes: Security Challenges and Privacy Concerns”, October 2020

[48] Cihan Emre Kement, Bulent Tavli, Hakan Gultekin, Halim Yanikomeroglu, “Holistic Privacy for Electricity, Water, and Natural Gas Metering in Next Generation Smart Homes”, IEEE Communications Magazine, March 2021

[49] N. Komninos, E. Philippou and A. Pitsillides, “Survey in Smart Grid and Smart Home Security: Issues, Challenges and Countermeasures”, IEEE Communications Surveys & Tutorials, 2013

[50] Julie M. Haney, Susanne M. Furman, and Yasemin Acar, “Smart Home Security and Privacy Mitigations: Consumer Perceptions, Practices, and Challenges”, NIST, July 2020