



**University of
Zurich^{UZH}**

*Burkhard Stiller, Muriel Franco, Christian Killer, Sina Rafati,
Bruno Rodrigues, Eder John Scheid, Rafael Ribeiro, Eryk Schiller
(Edts).*

Internet Economics XIV

TECHNICAL REPORT – No. IFI-2021-01

June 2021

University of Zurich
Department of Informatics (IFI)
Binzmühlestrasse 14, CH-8050 Zürich, Switzerland



B. Stiller et al. (Eds.): Internet Economics XIV
Technical Report No. IFI-2021-01, June 2021
Communication Systems Group (CSG)
Department of Informatics (IFI)
University of Zurich
Binzmühlestrasse 14, CH-8050 Zürich, Switzerland
URL: <http://www.csg.uzh.ch/>

Introduction

The Department of Informatics (IFI) of the University of Zurich, Switzerland works on research and teaching in the area of computer networks and communication systems. Communication systems include a wide range of topics and drive many research and development activities. Therefore, during the autumn term HS 2020 a new instance of the Internet Economics seminar has been prepared and students as well as supervisors worked on this topic.

Even today, Internet Economics are run rarely as a teaching unit. This observation seems to be a little in contrast to the fact that research on Internet Economics has been established as an important area in the center of technology and economics on networked environments. After some careful investigations it can be found that during the last ten years, the underlying communication technology applied for the Internet and the way electronic business transactions are performed on top of the network have changed. Although, a variety of support functionality has been developed for the Internet case, the core functionality of delivering data, bits, and bytes remained unchanged. Nevertheless, changes and updates occur with respect to the use, the application area, and the technology itself. Therefore, another review of a selected number of topics has been undertaken.

Content

This new edition of the seminar entitled “Internet Economics XIV” discusses a number of selected topics in the area of Internet Economics.

The first talk provides an overview of virtual private networks (VPN) services, especially discussing topics related to privacy and the influences of VPN services in businesses during pandemics. Talk 2 discusses the market behind Software Defined Radio (SDR) and its practical aspects. Talk 3 explores the economics of vote buying and combines an overview of the real-world appearance of vote buying and its implications with technical countermeasures and the debate on this topic. Talk 4 discusses the wireless sensing market share, providing an in-depth analysis of existing tools and approaches and their economic impact. Talk 5 brings the state of blockchain-based banking services, discussing blockchains implemented by Swiss banks nowadays and what laws are being made to facilitate the use of blockchain technology in the financial sector. Talk 6 provides an analysis and comparison of Blockchain-as-a-Service (BaaS) providers, first giving an overview of the BaaS market and then comparing the BaaS providers based on various criteria. Talk 7 discusses public Proof-of-Stake (PoS) blockchains from an operational perspective. Furthermore, this talk provides an in-depth review of the requirements and the security risks for operating a PoS validator based on the analysis of six different PoS protocols. Talk 8 introduces the economic advances of blockchain-based trading platforms, focusing on

the commercial component of blockchain-based trade networks and how blockchain leads to economic development in multiple business sectors. Talk 9 is a survey of models for cybersecurity economics. The main metrics and state-of-the-art approaches available for cybersecurity economics were investigated and described in this talk. Talk 10 elaborates on the business landscape of IoT (Internet of Things) Security, investigating challenges in the security of IoT devices by examining characteristics of these devices. Talk 11 investigates the randomization of MAC addresses, bringing market opportunities in light of various use cases and a discussion about privacy concerns regarding the data collection. Finally, Talk 12 provides an analysis of company investments in cybersecurity, presenting details on how companies are protecting themselves against cyberattacks. Furthermore, this talk presents the current cybersecurity market and security technologies.

Seminar Operation

Based on well-developed experiences of former seminars, held in different academic environments, all interested students worked on an initially offered set of papers and book chapters. Those relate to the topic titles as presented in the Table of Content below. They prepared a written essay as a clearly focused presentation, an evaluation, and a summary of those topics. Each of these essays is included in this technical report as a separate section and allows for an overview on important areas of concern, sometimes business models in operation, and problems encountered.

In addition, every group of students prepared a slide presentation of approximately 45 minutes to present its findings and summaries to the audience of students attending the seminar and other interested students, research assistants, and professors. Following a general question and answer phase, a student-lead discussion debated open issues and critical statements with the audience.

Local IFI support for preparing talks, reports, and their preparation by students had been granted by Muriel Franco, Christian Killer, Sina Rafati, Bruno Rodrigues, Eder John Scheid, Eryk Schiller and Burkhard Stiller. In particular, many thanks are addressed to Sina Rafati and Rafael Ribeiro organizing the seminar and for his strong commitment on getting this technical report ready and quickly published. A larger number of pre-presentation discussions have provided valuable insights in the emerging and moving field of communication systems, both for all groups of students and supervisors. Many thanks to all people contributing to the success of this event, which has happened in a lively group of highly motivated and technically qualified students and people.

Zurich, June 2021

Contents

1 Virtual Private Networks (VPN) Goals: Privacy or Profit	7
<i>Ivan Allinckx, Simon Ammann</i>	
2 The Market behind Software Defined Radio (SDR)	26
<i>Jan Schnyder, Aline Schaufelberger, Ülkü Karagöz, Aljoscha Schnider</i>	
3 The economics of vote buying	61
<i>Domenic Luca Fuerer, Lennart Lou Jung, Sascha Deboni, Tony Ly</i>	
4 Wireless Sensing Marketshare	106
<i>Imami, F., Kaushik, R., & Zimmermann, T.</i>	
5 State of Blockchain-Based Banking Services	142
<i>Marion Dübendorfer, Ramon Solo de Zaldivar, Raphael Imfeld</i>	
6 An Analysis and Comparison of Blockchain-as-a-Service (BaaS) Providers	167
<i>Felix Hoffmann, Charlotte Eder, Alain Küng</i>	
7 Public Proof-of-Stake Blockchains From an Operational Perspective	205
<i>Bill Bosshard, Simon Bachmann</i>	
8 Economic Advances of Blockchain-based Trading Platform	231
<i>Saiteja Reddy Pottanigari, Ankan Ghosh</i>	
9 Survey and Analysis of Models for Cybersecurity Economics	262
<i>Jan Bauer and Yung-Hsin Chen</i>	
10 The Business Landscape of IoT Security	287
<i>Andy Aidoo, Jara Fuhrer, Jonathan Stahl, Michael Ziörjen</i>	
11 Randomization of MAC Addresses: Market Opportunity or Privacy Violation?	319
<i>Csanad Erdei-Griff, Anna Jancso, Andris Prokofjevs</i>	
12 Is it All About Money? An Analysis of Company Investments in Cybersecurity	352
<i>Adrianna Marszal and Pascal Marty</i>	

Chapter 1

Virtual Private Networks (VPN) Goals: Privacy or Profit

Ivan Allinckx, Simon Ammann

This paper provides an overview of what a VPN is, how it is working and for what a VPN can be used. As VPNs can be used to be more anonymous on the internet security aspect will be discussed as well as the breaches that happened as an example one from NordVPN and how to find a more secure VPN than others. As this paper was written during the pandemic of the coronavirus, there is a section about which influences this had on VPNs in general such as an increase in usage and the possibility to work from home and having access to the resources of a business.

Contents

1.1	Introduction	9
1.1.1	VPN Types	10
1.1.2	Early VPN uses	10
1.2	Why Would One Use a VPN?	11
1.3	The Covid-19 Pandemic Impact on the Usage of VPN	12
1.4	Providers	14
1.5	What Data can be Collected?	15
1.6	Breaches that happened	16
1.6.1	Hola	16
1.6.2	Hotspot-Shield	16
1.6.3	Nord VPN	17
1.6.4	Data handed over to authorities	17
1.7	Market Behind VPNs	18
1.8	Discussion on Best Practices to Use a VPN	19
1.9	Conclusion	20

1.1 Introduction

Virtual Private Networks (VPN) are not a novel concept, being discussed and described since 1998 [7]. A VPN can be viewed as a temporary and secure connection between two points via the Internet, often used to access private or internal content [17]. However, as a more detailed definition, one can define a VPN by decomposing the term in three words and define each one individual to have an overview of what a VPN is, as performed in a paper from P. Ferguson and G. Huston[7].

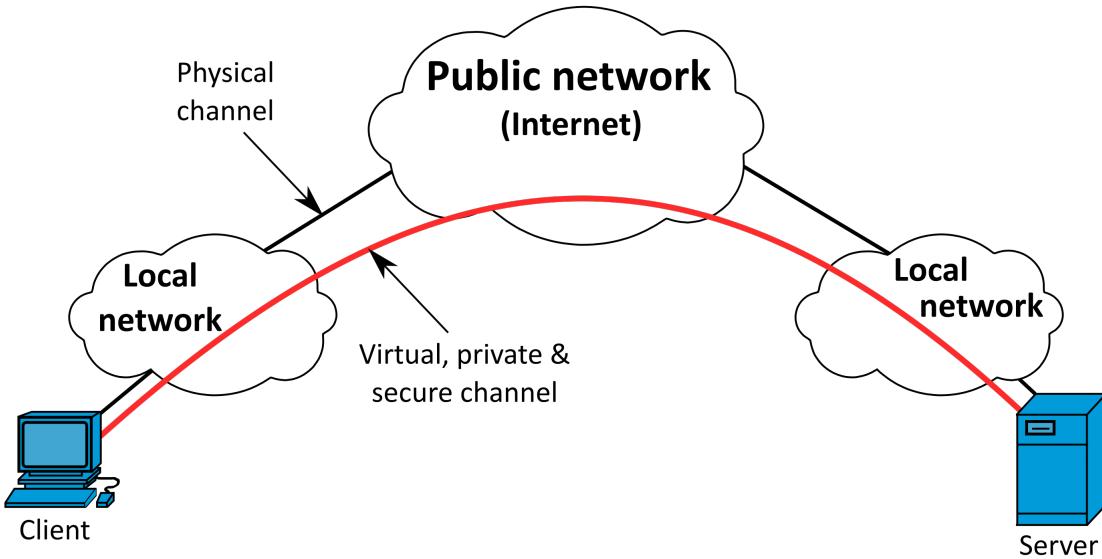


Figure 1.1: Representation of a simple VPN [10].

A computer network is known nowadays as it is used in our day-to-day. However, in technical terms, a network consist of a number of devices that are connected to each other using network devices (*e.g.*, routers and switches) and that can therefore communicate with each other. There are different ways for those networks to communicate with each other (for example different Protocols). Those devices are usually routers, printers and computers. By being connected to the internet, these devices do not need to be physically close to each other to communicate, being able to be placed in different countries or continents [7].

In the computer networks context, private means that communication between two or more devices is encrypted. Which means that other devices in the same network can not listen to a private channel between other devices in the same network. Another way to define this is by looking at the antonym of private, which is public. A public conversation can be listened to by everyone, for example, a speech from someone that is transmitted by television. Privacy can be created by using protocols or encryption which are generating a tunnel between those devices. [7].

The last word is virtual. Virtual in the context of a VPN means to simulate something that is actually not real or visible. For example, a virtual machine is a software that is emulating or simulating hardware that is actually not real, to then run an Operating System (OS) or other software on it that would require this hardware. In a VPN the virtual part is recreating the smaller private LAN (Local Area Network) over a WAN (Wide Area Network). A good example of a LAN is a home with a router that is connected to all devices of this house. The WAN is then the internet that is connecting all public servers to those homes and creating a bigger network. [7].

To summarize, a VPN recreates a smaller private network over a publicly available bigger network. This is achieved by creating a tunnel-like connection between all those dives or

LAN's by encrypting the communication between them. A straightforward visualization is depicted in Figure 1.1.

1.1.1 VPN Types

VPNs can be categorized into different types. They can differ in several aspects, such as how they are implemented as using different protocols or for what purpose they are used. One can categorize VPNs in three main groups: remote access VPNs, intranet-based site-to-site VPNs, and extranet-based site-to-site. The most common type used by the individual user is the remote access VPN [8].

Remote access VPNs connect a single user device to a remote server. From this server, one then can have access to the local network of this server. This is also the central aspect of the commercially used VPNs. The commercial VPN services are allowing one to use their network to connect to the internet. Therefore, when one is surfing the internet everything that is sent and received first passes through the VPN service. Now one can browse the internet as is if one would be the server or services him/herself. This means that ones IP address is now basically the services provider one which leads to more private surfing in comparison to browsing without a VPN. Those VPNs are usually relatively easy to install if one has basic computer and networking knowledge as one has only to install the launcher to those VPN services [8].

A different type of VPNs is site-to-site VPNs which work in a different manner than remote-access VPNs. The goal of this VPNs is to provide multiple users in different locations the access to each other's resources. This is mainly to share resources and information securely without passing through a second party. One can find those VPNs mostly in larger businesses that are nationally and internationally operating. These types of VPNs are more complicated to implement than the first type as they require specialized equipment and more in-depth knowledge about VPNs and how they work [8].

Another aspect in which VPNs can be differentiated is the protocols that are used to create the encrypted communication between those servers and the devices. A comprehensive representation of all the different classification for VPNs is illustrated in Figure 1.2. The most used protocols used for those VPNs are [14]:

- Internet Protocol Security (IPsec)
- Point-to-Point Tunneling Protocol (PPTP)
- Layer Two Tunneling Protocol (L2TP)
- Secure Socket Tunneling Protocol (SSTP)
- Transport Layer Security (TLS)/Secure Sockets Layer(SSL)

1.1.2 Early VPN uses

At the beginning of the 2000s, VPNs were mostly used by companies and not by individuals. Moreover, universities were using VPNs so that their students could join the university intranet and have access to resources that were usually only accessible to devices connected to the network of the campus [17].

This was and still is especially helpful for digital libraries as now students are able to get access to digital resources that were usually only available if one was connected to the campus network as those licenses to get access to those resources are often coupled to the public IP-address of the university [17].

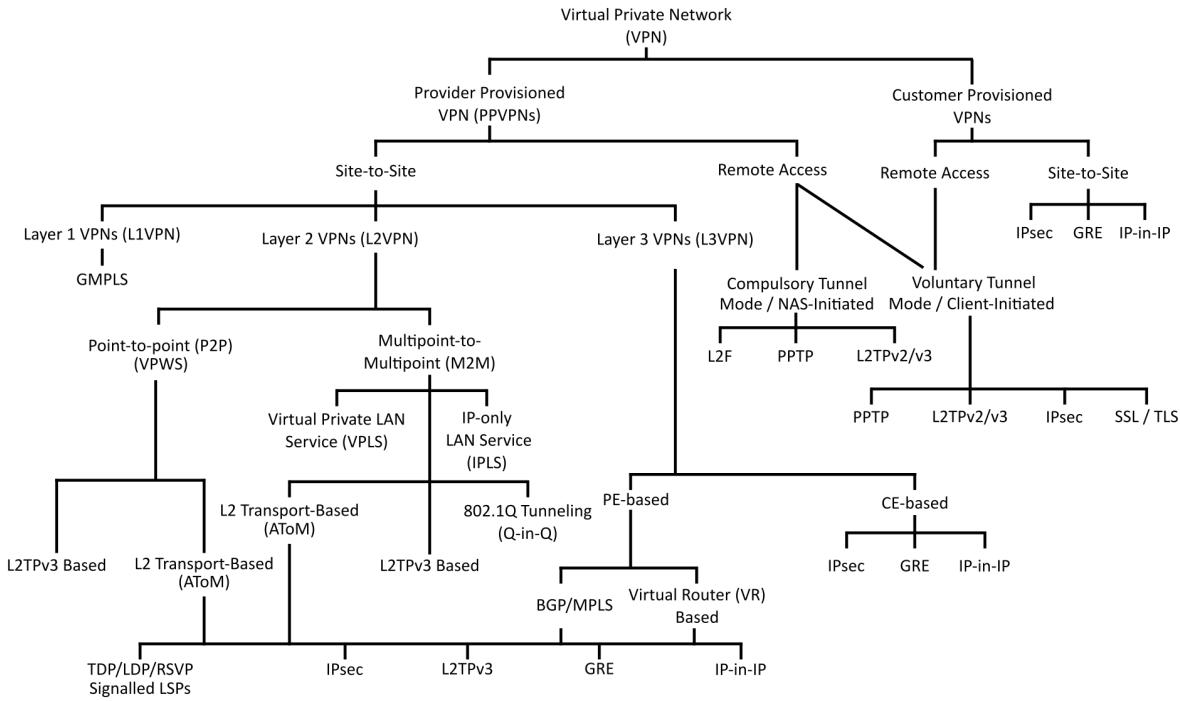


Figure 1.2: VPN classification based on the topology first, then on the technology used [10]



Figure 1.3: Student connecting to IEEE Xplore via VPN of the University of Zürich.

For example, if a student connects to the IEEE xplore website [42] through the VPN of the university, the student is automatically recognized as a member of this university. This can be seen in Figure 1.3 when a student registered at the University of Zurich (UZH) and using a VPN to connect to the IEEE Xplore website.

1.2 Why Would One Use a VPN?

There are different motivations for a user to use a VPN. In this paper, there is a focus on commercial VPNs, for example, NordVPN. However, most reasons are valid for private VPNs and business VPNs as well. The most common reasons to use a VPN are privacy and security, to bypass area restricted content, to bypass restrictive networks, to save money, and to connect to the intranet of a business or school.

A commercial VPN works as follows. The VPN server is processing all request for one and sending back the results to one. This means that, for example, one's Internet Service Provider (ISP) can not know on which sites the user went or which data were sent. The only information that the ISP can see is that there is a connection from the user to the server and that encrypted data is sent to and from the user. This leads to an increase in privacy. Also, because all traffic is encrypted directly on the devices and the server itself a man-in-the-middle attack is not possible, this can be particularly helpful when connected to an open WiFi. Nevertheless, the question here is it really necessary to use a VPN to encrypt all connection data as many pages have switched to Hypertext Transfer

Protocol Secure (HTTPS) instead of Hypertext Transfer Protocol (HTTP) which is by itself already encrypted [22].

Another use case for VPNs is to bypass area restricted content. Media available on the internet might be only available from a certain country due to legal reasons. Providers of such content usually use the IP-address to block the communication of one as the IP-address can be quite easily matched to an area on the globe. By using a VPN one can change his IP-address and appear as being in another area. This only works because the VPN server is passing the request the user made to the content provider. The content provider then assumes that the VPN server made this request by its one and cannot know that it was actually made by another device somewhere else [22].

In countries where the government controls which information is available to his people by the internet, the government can with the help of ISPs restrict the websites one is allowed to visit. For example, Facebook is blocked in China, with a VPN one could bypass this restriction [36]. There is not only governments and countries censoring the internet, but also some universities or schools are restricting access to the internet. One can also bypass this by using a VPN. Because most universities are using only a blacklist and not a whitelist - with blacklist one mean adding websites to a list that are prohibited from accessing, with a whitelist one mean having to add a website to a list to be able to access it all other are blocked by default - one can still connect to a VPN server, and because everything is encrypted from the device to the VPN server the university or school can not block the website that was searched for as the university or school is not even able to see that one is trying to access this website [15].

The fourth reason why one could use a VPN is to have access to different product prices on marketplaces or services. As explained above, one can simulate the area one is by accessing the website by a VPN. Some websites are changing their prices for people coming from a different location as some countries are richer than others and the people living in those regions are usually willing to pay more for a specific service/product than those people coming from a poorer country. This way of saving money was working good a couple of years ago, but it gets more and more difficult to use this method to save some money as websites are changing the prices at the end of the process when one is entering his or hers personal data at checkout [21].

Lastly, one can also use a VPN, as explained in the section 1.1.2 to connect to the intranet of their company. Especially nowadays the people that are allowed to work from home (*i.e.*, home office) is increasing and therefore need to connect to theirs company intranet to access the resources of those companies.

1.3 The Covid-19 Pandemic Impact on the Usage of VPN

Before the coronavirus, there were already many people who used a VPN. According to Proofpoint's User Risk Report 2020, a survey of 3,500 adults working in the United States, Australia, France, Germany, Japan, Spain and the UK, 39% of these people were using a VPN before Corona. However, this report does not differentiate whether these VPNs are business VPNs or commercial VPNs. 29% had the feeling not to use a VPN and 32% did not know what a VPN is. This could be because they do a job that does not require a VPN or because these companies do not use VPNs [20].

However, if the people have already installed a VPN, then 50% of these people use their VPN all the time. 33% said they use their VPN frequently and 12% said they only use their VPN when necessary. Again, this report does not mention what the report means by only when necessary. However, it can be assumed that this is meant when these people use the VPN for business purposes and need to access internal business resources. Only

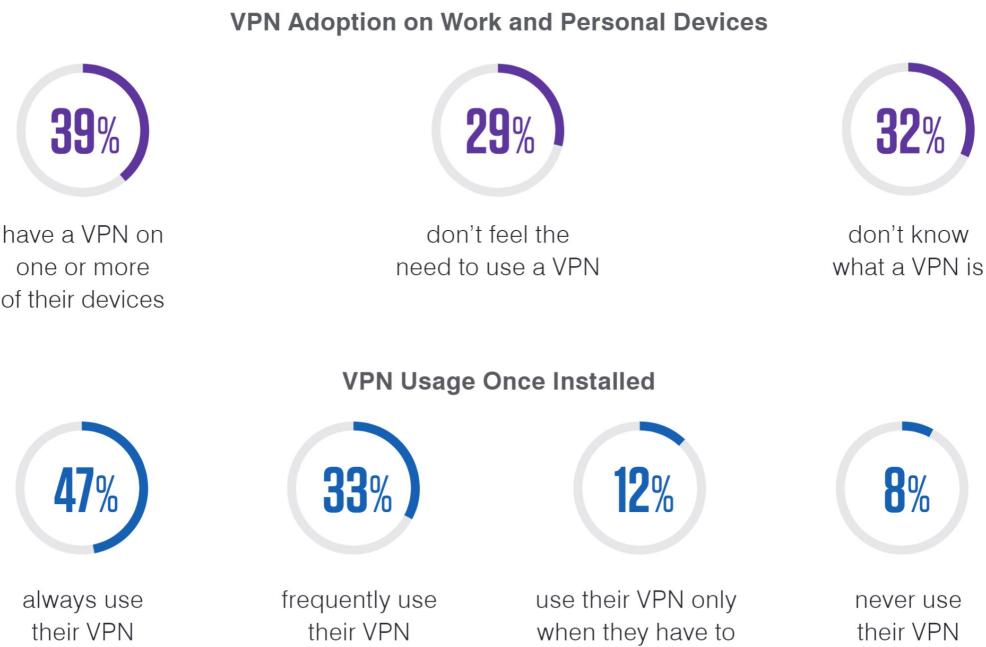


Figure 1.4: VPN Adoption on Work and Personal Devices [20]

8% would not use their VPN. Here it would be interesting to know why these people have downloaded a VPN at all [20].

Due to the Corona pandemic in spring, an increased number of people had to stay at home. This had a major impact on people's behaviour. They spent noticeably more time with digital devices. This is not surprising, as it was one of the only ways they could still communicate with other people who did not live around them. This is also reflected in the numbers. In a survey conducted by Globalwebindex's [19], 47% of respondents said they spent more time with social media. 57% said they were using more shows and streaming services. Those changes in media habits can be seen in Figure 1.5. This could lead to an increase in the use of VPN, as for example, certain series are not available everywhere. And as already discussed in the section 1.2, one can access this content with a VPN.

During the pandemic, significantly more people also had to work from home. This challenged the companies, as the vast majority now had to switch to home offices. The usage of VPNs increased so that employees could access company resources. Since this is a secure way to connect to the company. Small as well as large tech companies such as Facebook, Amazon, Google or Microsoft had to rely on such a solution [4].

This change has also been seen by Atlas VPN as they derived data from their 53'000 weekly users. The usage of their VPN between March 8 and March 22 increased in the United States by 124% and 160% in Italy. At the same time, they had an increase of 6% on customers. The interesting part here is that the usage in most countries increased by more than 6% which mean that not only they gained on the number of customers but the customer that already were using Atlas VPN used it more often and for a longer time of period [2].

As can be seen in Figure 1.6 the orange curve that shows the usage increase in two weeks by countries matches the bars that shows the increase of Covid-19 cases. An exception here is Russia as they had a quite big jump in usage of a VPN but not for the cases. An explanation that Altas VPN is giving for this is that the Russian government is covering up coronavirus cases as ordinary pneumonia. This was also a statement from Russia's Alliance of Doctors union. Atlas VPN stated that they would give from there on

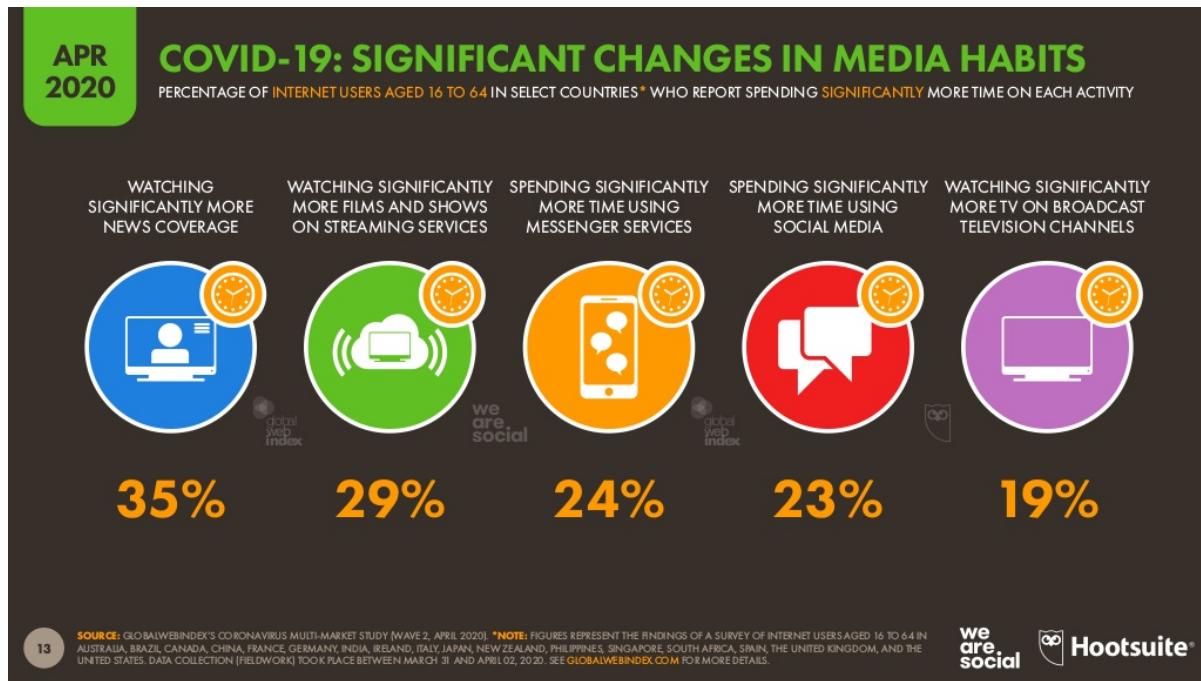


Figure 1.5: Covid-19: Significant changes in media habits [19]

their premium subscription for 3-month free of charge as the user could with their VPN overcome content restrictions and get better information about the coronavirus [2]. The coronavirus had a big impact on how people and companies have a view on remote work. The coronavirus showed to many people that home-office was actually working. In a study conducted by OpenVPN, they have found out that 68% of the employees said that their company expanded the usage of a VPN as a result of the coronavirus and 29% even said that their company used a VPN for the first time. In the study 99% of the person said that they think that their company will continue to offer the possibility for remote work after the Covid-19 pandemic. Thus, for OpenVPN, it is quite clear that the need for VPNs for companies will continue to exist even after the pandemic. This is no surprise as there was already an increase of 159% in remote work from 2005-2017 [18].

1.4 Providers

Table 1.1 providers an overview and comparisons of selected providers in terms of subscription type offered, prices, supported devices, and if a data breached has occurred or not. Such data breaches are discussed in Section 1.6.

Table 1.1: VPN Providers Comparison

Provider	Subscription Type	Price	Supported Devices	Data Breached
IPVanish	Monthly/Yearly	39\$/year	Router/Smartphone/PC	No
TunnelBear	Monthly/Yearly	60\$/year	Smartphone/PC	No
NordVPN	Monthly/Yearly	50EUR/year	TV/Smartphone/PC	Yes
Hola	Monthly/Yearly	80EUR/year	Console/TV/Smartphone/PC	Yes
Hotspot Shield	Monthly/Yearly	95EUR/year	TV/Router/Smartphone/PC	Yes

There are multiple providers in the VPN market. Most of them are similar, such as NordVPN, Tunnelbear, ExpressVPN, Surfshark, Hotspot Shield, IPVanish. Most of them are not free and are based on a monthly or yearly subscription. They are mostly compatible with all sorts of devices, such as smartphones, laptops, computers and televisions.

There are exceptions to these rule, *e.g.*, ProtonVPN, that has a fully functional VPN service for free. They say they created their VPN because of the reason for free speech

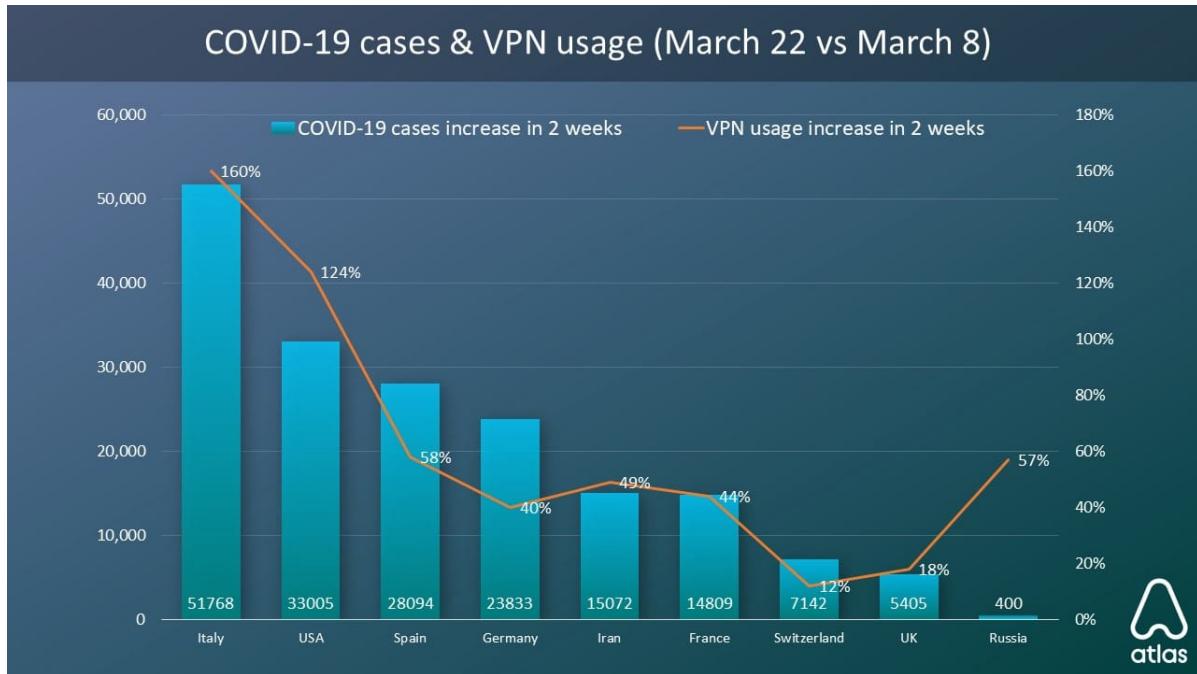


Figure 1.6: Covid-19 cases and VPN usage (March 22 vs March 8) [2]

and for journalist or activists. However they still have different option to opt-in if one wants more services *e.g.*, being able to access blocked content or to connect to a server in other more than 3 countries [9].

1.5 What Data can be Collected?

According to the Unisys Security Index [39] data collection, especially on a personal level, is a growing concern. The issue can be seen from different perspectives. Data regarding an Internet user can be collected by websites or services used, the ISP and the VPN provider. Using a VPN can be a way to circumvent data collection by the ISP. Because as discussed in previous sections, a VPN, in essence, redirects the internet traffic through a secure and encrypted connection, which in most commercial providers connects the device with a remote server by the VPN-company. This means that the internet traffic passes through this server which in most cases is controlled by the commercial provider or some third party. So all data that the ISP otherwise could see is now in the hands of the VPN company. Although most commercial VPN providers advertise to have a so-called "no-log-policy", in theory, they could log various data and information which include [38]:

1. The true client IP address (issued by the ISP)
2. Internal customer ID in the system (usually the user's login credentials for the VPN provider)
3. Internal IP address of the client when connecting to VPN (on most VPNs, the client will be assigned a "non-routable" IP address)
4. Connection statuses, control, and error messages (these are important for technical support and troubleshooting purposes)

Further, it can be differentiated between connection logs and usage logs [40]. Connection logs include connection timestamps of the users VPN session, the amount of data, sometimes the outgoing IP-address. They are mainly used for technical troubleshooting

purposes. Usage logs on the other hand include more personal data including software being used and the browsing history including visited websites and downloaded files. It has also be kept in mind that it is more economically viable and easier to keep logs than to keep minimal or no logs. Because to keep no logs the company has to run the entire network to run in RAM-disk mode [41]. Also, logs in the form of connection statuses, control, and error messages are important for tech support and troubleshooting purposes. If a person uses a VPN to improve your privacy he/she may get privacy by policy but not by design. All that the VPN does, is taking away the data from the internet service provider and hand it to the VPN-provider. Thus, it is crucial to keep in mind that privacy is not fully guaranteed just by using a VPN alone, one must rely on encryption mechanisms.

1.6 Breaches that happened

A data breach, as TechTarget defines it, is a confirmed incident in which sensitive, protected, or confidential data has potentially been viewed, stolen, or used by an individual unauthorized [33]. Data breaches have high costs and often affect millions of users in case a big company is affected. In the Year of 2019, according to the Verizon Data Breach Investigations Report [37], there have been 2013 confirmed data breaches. For example, in 2013, the retail company Target Corporation suffered a data breach where names and credit card information of 110 million costumers got exposed.

Using a VPN can be seen as a tool to mitigate the risk of intrusion attacks and data breaches. But on the other hand, using a VPN can also pose a risk itself in case there is a data breach on the side of the VPN-provider. Security has been one of the main selling points of the commercial VPN industry. For the most part encryption and IP-masking are mechanisms to allow more secure connections while using a VPN. At the same time, as discussed before, the industry mostly promises no log practices. Because of the commercial nature of the VPN ecosystem and the lack of independent and peer-reviewed evaluation of the providers, it is really hard to provide positive evidence for all those claims, contrary it is easier to provide negative evidence of some cases.

Therefore, the following sections will explore selected incidents on the side of VPN providers. Which by no means represents a complete list of all incidents they are just a selection of cases that serve to illustrate different problems in the industry.

1.6.1 Hola

Hola is a popular and free browser extension, which functions through a peer-to-peer network and lets you select the country through which you will access the internet. In 2015, it came to light, that Hola takes the internet of Hola users and makes it available to whoever wants to route queries through residential IPs [31]. This is achieved through a parent company called Luminati. Thus, in essence, they were operating a worldwide botnet in which Holas users exist as exit nodes. This could easily be abused for cyber-criminal activity. This botnet has been used for several Dos attacks. Luminati the parent company has an estimated yearly revenue of 23 million dollars. Thus, free services might rely on the use of selling user data or advertisements for profit.

1.6.2 Hotspot-Shield

In 2017, a group of privacy researches accused the provider to sell users data to advertisers, including IP addresses, unique device identifiers, and other “application information” while in use [25]. Further, they were accused of redirecting user traffic to partner websites.

The company still claims to have a no-log policy and has denied all these claims including that they do not make any money with advertisers. But at their corporate website, they promise to connect advertisers with unique users [25]. Taking a closer look at their terms of agreement it can be seen that Hotspot-Shield does not consider IP-address and unique device identifiers as personal data.

1.6.3 Nord VPN

Moreover, there are potential weak points when it comes to the servers itself, as the case of Nord VPN shows. Nord VPN is a paid service with 14 million users. To run part of their servers they use third parties. In March 2018, hackers gained access to one of the servers in a data centre in Finland through an insecure remote management system [27]. Nord VPN stated that no users data was affected as they keep no logs, but there were an estimated 20 to 70 active sessions in time of the attack with which the hackers could have interfered. Other than that only an expired private key got exposed. More interesting is the timeline of the incident, has nearly passed one year until the company itself has learned about the event and nearly another year passed until they publicly admitted the incident and informed users [27]. The company as stopped working with the affected data centre and claims to have learned their lesson.

1.6.4 Data handed over to authorities

Another good indicator to check if providers keep logs is to look at cases where information have been handed out to the authorities. Further, these cases give insight into where the servers are located because the laws and practices for privacy differ between countries. In this context, it is important to be aware of the 5-9-14 eyes surveillance alliance, which is an international alliance that is collecting and sharing surveillance data with each other. For decades this network was and is spying on people. The ISP of those countries are often working with those state agencies but also big tech companies are involved in collecting data. This process is documented in the PRISM surveillance documents. These main countries participating in these alliance are the ones contained in the Five Eyes which are Australia, Canada New Zealand, United Kingdom and the United States. The nine eyes are the five eyes and Denmark, France Netherlands and Norway. The 14 eyes are the nine eyes countries and Germany, Belgium, Italy, Sweden and Spain[13].

The countries in the five eyes are also those countries with laws that are “abusers” of online privacy. *e.g.*, for the United States have implemented mass surveillance with the help of large ISPs. Those ISPs are since march 2017 even allowed to record user activities and sell those to third parties. Another example is the United Kingdom that passed the Investigatory Powers Act in 2016 which allows ISPs to record browsing history, connection times and text messages, those have then to be stored for two years and UK government agencies have access to this data without warrant [13].

This means that if one is using a VPN that is located in one of these countries that they by law have to give access or to be obligated to collect data from users. This is why one should avoid such a VPN if privacy is a concern [13]

As an example, the UK-based provider HideMyAss even handed over data to the authorities several times [29]. But not only countries located in the five eyes countries give out data to foreign authorities. PureVPN located in Hon Kong for example also gave out data to US-authorities despite claiming a no-log-policy [34]. There are also counterexamples of VPN companies where authorities tried to confiscate information but did not find any logs. In 2017 Turkish police seized a server of the provider Express VPN in the course of investigations but were not able to find any connection logs [35]

Other data breaches of VPNs is the research by Noam Rotem and Ran Locar at vpnMentor found personal data for over 20 million VPN users of VPNs who claim to have a no-log-policy in the form of user logs in open databases [28]

1.7 Market Behind VPNs

Historically, VPNs were mostly used in enterprises to securely share resources and access internal resources while offsite. With the rise of the internet individuals discovered VPNs to circumvent local restrictions and mask their IP addresses [23]. Which was supported by a growing number of VPN providers that made VPN more user friendly and easy to use. Currently, there exists a large range of commercial VPN providers. The global VPN market size is valued at 25 billion dollars and is expected further to expand [30]. Reasons for the expected growth are the rise in cybercrime and data breaches, the proliferation of internet services, increasing adoption across multiple business verticals, regulations requiring organizations to implement additional security, growing adoption of Bring Your Own Device (BYOD) and enterprise mobility and of as mentioned in a separate section 1.3 the recent pandemic and the increase in work from home [30].

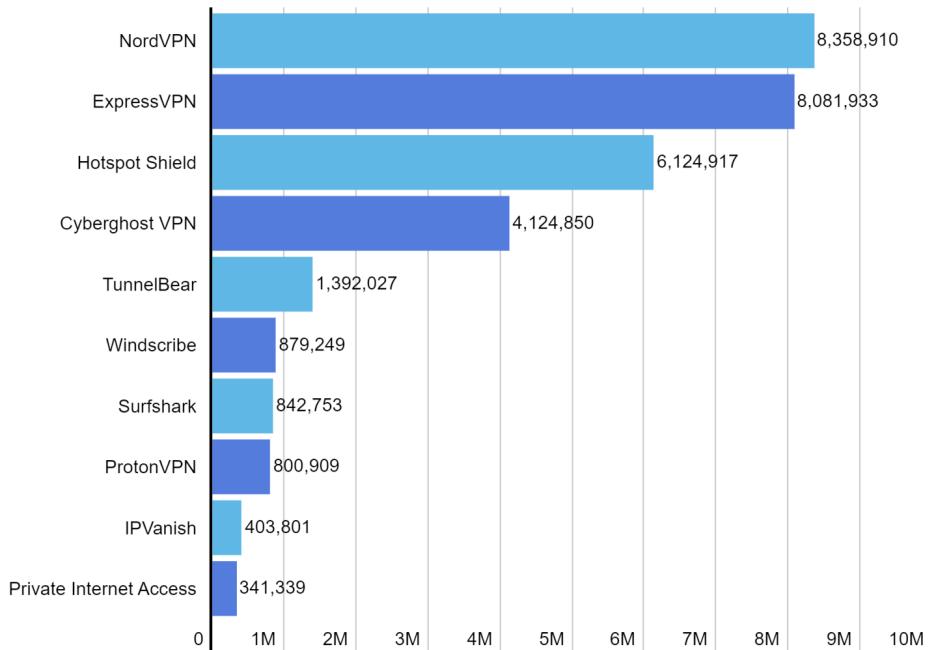


Figure 1.7: Top VPN providers by number of installs, January to May 2020

When it comes to providers, a web search will result in several lists of the "best" VPN providers, many of which affiliated with one or the other provider. About the market share of the providers in terms of users and revenue, a lot less information is to be found. Providers such as Hotspot Shield boast themselves with an incredible number of total 600 million downloads [43], but its highly questionable how much download of a free application tells us about the user base. According to atlasvpn [32] in the period between January and May of 2020, the three biggest players in terms of installs 1.7 and revenue 1.8 were the three providers NordVPN, Hotspot Shield and ExpressVPN, with NordVPN having the most installs 8,358,910 and Hotspot Shield generating the most revenue at 8,816,492 USD. The ranking only considers data from Google Play Store and Apple App Store. Interesting to note here is that the service generating the most revenue, namely Hotspot Shield, is a *freemium* service, while the other two are paid-only.

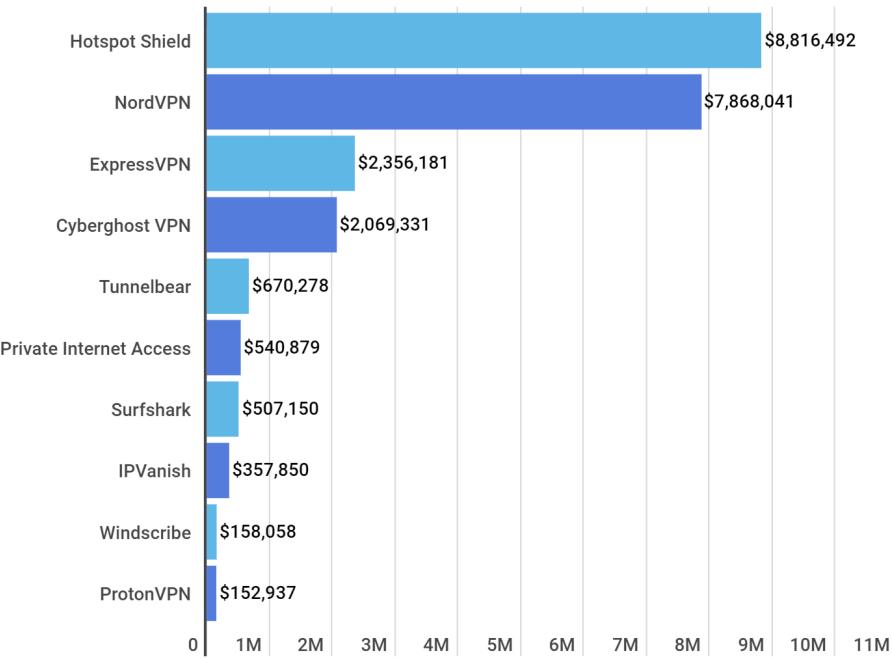


Figure 1.8: Top VPN providers by revenue, January to May 2020

One reason why the market is so heavily advertised, a large portion of it being affiliate marketing, might be that there are so many providers with no one clearly dominating the market, all selling a fairly similar product. Although the players try to market and distinguish themselves with better security protocols, number or locations of servers, special features or cryptocurrency as a payment option. Another strategy of players is to go into partnerships with various regional companies to penetrate foreign markets [30]. Although the technology of VPN was not originally designed for privacy, privacy has become the industry's main focus of advertising and selling point [23]. Privacy is a key motivation for many users, but the sole focus of advertising on privacy might also be due to the fact that other motivations of usage, such as bypassing geo-restrictions or pirating content are not that socially desirable and suitable for advertising.

1.8 Discussion on Best Practices to Use a VPN

There are best practices that one can have to find the right VPN for his/her needs. One way to do it is by searching for comparison online. By doing so one will probably land on a website titled "The Top 10 VPNs you can buy" or "The best VPN of 2020!". These can be websites that were created with the goal to make a lot of money as the affiliate programs of VPNs are usually quite high. As an example, if one buys a month at ExpressVPN through their affiliate link they would get a commission of 13\$. This is quite surprising as a month at ExpressVPN only costs about 11.22 Euro[11].

Therefore, one can never be sure that those websites are actually ranking the multiple VPNs presented on those websites correctly. It could also be that the VPN with the highest affiliate program, meaning having the highest commission rate, is actually ranked first only to increase the money made by this method.

One can spot this by clicking on one of these VPNs links. If then in the URL there are words *e.g.*, affiliate, campaign or coupon, as can be seen in Figure 1.9 and Figure 1.10, then this website is probably using this method to make money and one can not be sure if this website is not biased.



Figure 1.9: Example of an affiliate link

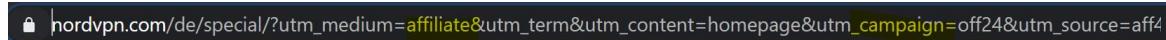


Figure 1.10: Example of an affiliate link

Another way of finding a suitable VPN for one is by creating an own list of VPNs and comparing them by attributes that are important for one. Select attributes will be discussed in the following paragraphs.

One of these attributes would certainly be the cost for a VPN as one usually wants the best or cheapest price for a product.

Another reason to chose a certain VPN is if this VPN has a server located in the area I want. As discussed in Section 1.2, one can use a VPN to access media that is restricted to a certain area or country. If now the VPN provider has no server in this country one can not access the restricted media and therefore this VPN is then useless for one.

Another attribute could be if a breach already happened in the past as seen in the section about breaches that already happened. The question that one could then ask is if a breach already happened can one then trust this provider that it will not happen again as maybe they have learned from there mistakes in the past. Or have VPNs that had no breaches more chance to have a breach happening in the future?

Two further attributes that are also correlated with the attribute before is if the VPN is collecting data about the user and which data exactly is collected and if the VPN is located in a country belonging to the 5-9-14 eyes surveillance alliances. If a VPN is not collecting any data or only data that is not sensitive to a single user then a data breach can not happen or when it happens then the data that is now exposed has no effect on a user.

Further if one is using a VPN in regards to privacy and anonymity then one can ask him/herself if a VPN is actually the best solution that one has considered so far. As an example, the Tor Project or Brower could be more suitable for this concern [12].

Those best practices can of course change or depend on if one can trust his ISP or if the laws for privacy are quite strict in the country one is living and of course for what purpose one need a VPN. In regards to commercial VPN as NordVPN or ExpressVPN, one can never be sure to 100% that no data is collected by them or that they are not forced by laws to hand in some data. Further, one can host a VPN server in his/her premises and then be sure that at least the connection from this private server to the dives one is using is secure. But then to browse the internet one is still not anonymous as the connection from this VPN server to the internet can still be monitored by the ISP the server is connected to.

1.9 Conclusion

To conclude, are VPNs a tool for Privacy or is it all just about profit? VPNs are not a technology invented to improve privacy, still, they can be one tool in the toolbox to improve privacy. They mask a users IP-address and hide its internet traffic from your ISP. It definitely can be argued that there are better alternatives to achieve those same features, but for many people, they seem to be a simple and effective solution. Commercial VPNs have undeniably become a big business with sometimes questionable practices where one does not always get what they promise. So much, that using some providers might pose higher privacy risks than just using no VPN at all. In this context, it can be said that

fee-based providers tend to be better than free providers, but even a fee-based provider alone does not guarantee that your data will be handled securely and trustworthy [24]. In the end, it is and remains a question of trust in the provider.

The market is projected to further grow [30] and the worldwide pandemic will further support this trend. With this massive adoption of the technology, it has to be questioned if all those users need a VPNs for their privacy needs or if this need was just created through fear and advertising. There are as mentioned also other more pragmatics reasons to use a VPN that does not concern privacy, but many users use VPNs mainly for privacy and it is the industries main selling point. So when it comes to commercial VPNs it is about profit and if you want to use a VPN for privacy you should carefully evaluate what your concrete expectations are if there might be another better solution and which provider you want to give your trust.

Bibliography

- [1] C. Davenport, “Flash VPN, UFO VPN, and Five Other Services Leaked 1.2TB of Private Information,” July 2020. <https://www.androidpolice.com/2020/07/19/flash-vpn-ufo-vpn-and-five-other-services-leaked-1-2tb-of-private-information/>, last visit July 23, 2020.
- [2] Atlas VPN, “Lockdowns and Panic Leads to a 124% Surge in VPN Usage in the US,” March 2020. <https://atlasvpn.com/blog/lockdowns-and-panic-leads-to-a-124-surge-in-vpn-usage-in-the-us>, last visit July 23, 2020.
- [3] R. Hodge, “VPN Use Surges During the Coronavirus Lockdown, But So Do Security Risks,” April 2020. <https://www.cnet.com/news/vpn-use-surges-during-the-coronavirus-lockdown-but-so-do-security-risks/>, last visit July 23, 2020.
- [4] OpenVPN Staff, “Why Companies Are Turning To VPNs During The CoronaVirus Outbreak,” 2020. <https://openvpn.net/why-companies-are-turning-to-vpns-during-the-coronavirus-outbreak/>, last visit July 23, 2020.
- [5] J. DeMuro, “Four Ways That a Free VPN Can Profit From Its Users,” March 2019. <https://www.techradar.com/news/four-ways-that-a-free-vpn-can-profit-from-its-users>, last visit July 23, 2020.
- [6] C. Metz, “The Latest in Virtual Private Networks: Part I,” *IEEE Internet Computing*, vol. 7, pp. 87–91, January 2003.
- [7] P. Ferguson and G. Huston, “What is a VPN?,” April 1998. <https://www.potaroo.net/papers/1998-3-vpn/vpn.pdf>, last visit October 04, 2020.
- [8] R. Greenberg, “Different Types of VPNs and When to Use Them (Updated 2020),” August 2020. <https://www.vpnmentor.com/blog/different-types-of-vpns-and-when-to-use-them/>, last visit October 04, 2020.
- [9] ProtonVPN <https://protonvpn.com/>, last visit November 05, 2020.
- [10] Virtual private network From Wikipedia, the free encyclopedia https://en.wikipedia.org/wiki/Virtual_private_network, last visit November 05, 2020
- [11] ExpressVPN Affiliate Program: Frequently asked questions <https://www.expressvpn.com/de/affiliates/faq>, last visit November 10, 2020
- [12] A. Macrina, “The Tor Browser and Intellectual Freedom in the Digital Age,” *Browser and Intellectual Freedom in the Digital Age*, vol. 54, no. 4, pp. 17–20, Summer 2015. <https://www.jstor.org/stable/refuseserq.54.4.17>, last visit November 11, 2020.
- [13] S. Taylor, “Five Eyes, Nine Eyes, 14 Eyes - Explained,” September 3, 2020 <https://restoreprivacy.com/5-eyes-9-eyes-14-eyes/>, last visit November 10, 2020

- [14] W. B. Diab, S. Tohme and C. Bassil, "VPN Analysis and New Perspective for Securing Voice over VPN Networks," *Fourth International Conference on Networking and Services (icns 2008)*, Gosier, 2008, pp. 73-78.
- [15] G. Aceto, A. Botta, A. Pescape, M. F. Awan, T. Ahmad and S. Qaisar, "Analyzing internet censorship in Pakistan," *2016 IEEE 2nd International Forum on Research and Technologies for Society and Industry Leveraging a better tomorrow (RTSI)*, Bologna, 2016, pp. 1-6.
- [16] I. I. Savchenko and O. Yu. Gatsenko, "Analytical review of methods of providing internet anonymity," *Automatic Control and Computer Sciences*, vol. 49
- [17] X. Bai, F. Zhang and D. Wang, "The application of VPN technology in the university's library," *2011 IEEE 3rd International Conference on Communication Software and Networks*, Xi'an, 2011, pp. 563-566.
- [18] OpenVPN Staff, "COVID-19 Fast Tracks Virtualization - OpenVPN Study Reveals Remote Work is the Future", <https://openvpn.net/covid-19-fast-tracks-virtualization-openvpn-study-reveals-remote-work-is-the-future>, last visit November 11, 2020
- [19] S. Kemp "DIGITAL 2020: APRIL GLOBAL STATSHOT", April 2020, <https://datareportal.com/reports/digital-2020-april-global-statshot>, last visit November 11, 2020
- [20] proofpoint., "2020 User Risk Report: Exploring Vulnerability and Behaviour in a People-Centric Threat Landscape, April 2020
- [21] G. Ray, "7 Ways to Save Money With a VPN", February 2019 <https://surfshark.com/blog/6-ways-to-save-money-with-a-vpn>, last visit November 12, 2020
- [22] S. Larsson, M. Svensson, M. d. Kaminski, K. Roenkkoe and J. A. Olsson, "Law, norms, piracy and online anonymity: Practices of de-identification in the global file sharing community", *Journal of Research in Interactive Marketing*, 2012, vol. 6, pp. 260-280
- [23] J. Longworth, "VPN: from an obscure network to a widespread solution", *Computer Fraud & Security*, 2018, vol. 4, pp. 14-15
- [24] K. T. Mohammad et al., "An empirical analysis of the commercial vpn ecosystem", *Proceedings of the Internet Measurement Conference 2018*, 2018, pp. 443-456
- [25] Center for Democracy & Technology, "Complaint, Request for Investigation, Injunction, and Other Relief", *Proceedings of the Internet Measurement Conference 2018*, April 2017 <https://cdt.org/files/2017/08/FTC-CDT-VPN-complaint-8-7-17.pdf>
- [26] Statista, "VPN market size worldwide 2016-2022", <https://www.statista.com/statistics/542817/worldwide-virtual-private-network-market/>, last visit November 08, 2020
- [27] K. Holt, "NordVPN admits to 'isolated' server breach in Finland | Engadget", 2019 <https://www.engadget.com/2019-10-21-nordvpn-server-breach-data-center-finland.html>, last visit November 11, 2020

- [28] G. Fawkes, “Report: No-Log VPNs Reveal Users Personal Data and Logs”, <https://www.vpnmentor.com/blog/report-free-vpns-leak/#Timeline-of-Discovery-and-Owner-Reaction>, last visit November 08, 2020
- [29] J. Leyden, “HideMyAss defends role in LulzSec hack arrest - The Register”, 2011 https://www.theregister.com/2011/09/26/hidemyass_lulzsec_controversy/, last visit November 08, 2020
- [30] Grand View Research, “Virtual Private Network Market Share Report, 2020-2027”, 2020 <https://www.grandviewresearch.com/industry-analysis/virtual-private-network-market> last visit November 09, 2020
- [31] trendmicro, “Shining a Light on the Risks of HolaVPN and Luminati - Security News - Trend Micro HK-EN”, 2018 <https://www.trendmicro.com/vinfo/hk-en/security/news/cybercrime-and-digital-threats/shining-a-light-on-the-risks-of-holavpn-and-luminati> last visit November 11, 2020
- [32] Alex T., “Top 10 VPNs had 31 million installs YTD, generating \$23M in revenue - Atlas VPN”, June 23, 2020 <https://atlasvpn.com/blog/top-10-vpns-had-31-million-installs-ytd-generating-23m-in-revenue> last visit November 11, 2020
- [33] M. Rouse, “What is a Data Breach? Definition from WhatIs.com”, <https://searchsecurity.techtarget.com/definition/data-breach> last visit November 12, 2020
- [34] S. Taylor, “VPNs are Lying About Logs (Alarming Info) ”, 2017 <https://restoreprivacy.com/vpn-logs-lies/> last visit November 11, 2020
- [35] P. Bischoff, “ExpressVPN server seized in Turkey turns up no info in assassination case”, 2017 <https://www.comparitech.com/blog/vpn-privacy/expressvpn-server-seized-in-turkey-verifyies-no-logs-claim/> last visit November 12, 2020
- [36] K. Zucchi, “Why Facebook Is Banned in China & How to Access It”, October 22, 2019 <https://www.investopedia.com/articles/investing/042915/why-facebook-banned-china.asp> last visit December 8, 2020
- [37] Verizon, “2020 Data Breach Investigations Report”, 2020 <https://enterprise.verizon.com/resources/reports/2020-data-breach-investigations-report.pdf> last visit December 10, 2020
- [38] Khan, M. T., DeBlasio, J., Voelker, G. M., Snoeren, A. C., Kanich, C., & Vallina-Rodriguez, N., “An Empirical Analysis of the Commercial VPN Ecosystem”, 2018 <https://dl.acm.org/doi/pdf/10.1145/3278532.3278570> last visit December 10, 2020
- [39] Unisys, “Unisys Security Index”, 2020 <https://assets.unisys.com/Documents/Microsites/USI2020/UnisysSecurityIndexReport2020.pdf?v=2> last visit December 10, 2020
- [40] S. Taylor, “VPN Logs - What You Need to Know”, May 16, 2018 <https://restoreprivacy.com/vpn-logs/> last visit December 10, 2020
- [41] S. Taylor, “No Logs VPN Services That Have Been Verified”, December 10, 2020 <https://restoreprivacy.com/vpn/no-logs/> last visit December 13, 2020

- [42] “ieeexplore” <https://ieeexplore.ieee.org/Xplore/home.jsp> last visit December 17, 2020
- [43] A. Terekhov, “AnchorFree’s Hotspot Shield app surpasses 600 million worldwide downloads”, 2020 <https://www.hotspotshield.com/blog/anchorfree-hotspot-shield-600-million-downloads> last visit December 13, 2020

Chapter 2

The Market behind Software Defined Radio (SDR)

Jan Schnyder, Aline Schaufelberger, Ülkü Karagöz, Aljoscha Schnider

It is explained how software defined radio works and what it can be used for. It also gives a brief insight into the structure of such a system. A small experiment is described as well as a short review of the history of software defined radio. The practical aspects include some examples where this type of system is used. In addition, some of the latest products on the market based on Wi-Fi 6, which in turn use SDR, are presented. The professional as well as amateur use of an SDR system is highlighted and explained. An important point is attacking other technologies using SDR products. Some selected products for beginners are also presented and explained. Finally, some companies that manufacture the products mentioned above are introduced.

Contents

2.1	Introduction	.	.	.	28
2.2	Architecture	.	.	.	28
2.2.1	Antennas	.	.	.	29
2.2.2	Filters	.	.	.	29
2.2.3	Analog-to-digital converter	.	.	.	29
2.2.4	Modulators	.	.	.	30
2.2.5	Mixers	.	.	.	31
2.2.6	Amplifiers	.	.	.	32
2.2.7	Digital Signal Processor (DSP)	.	.	.	32
2.2.8	History and transition of the general architecture	.	.	.	32
2.3	Software - GNU RADIO	.	.	.	33
2.4	Experiment	.	.	.	33
2.5	History behind software defined radio	.	.	.	34
2.5.1	Early developments	.	.	.	34
2.5.2	Reginald Fessenden	.	.	.	35
2.5.3	Lee de Forest	.	.	.	36
2.5.4	Radio pact of 1912	.	.	.	36
2.5.5	Radio pact of 1927	.	.	.	37
2.5.6	Birth of software defined radio	.	.	.	37
2.6	Modern Communication Protocols	.	.	.	38
2.6.1	Mobile wireless generation	.	.	.	38
2.6.2	Wi-Fi 6	.	.	.	39
2.7	Use Cases of Software Defined Radio	.	.	.	41
2.7.1	Spectrum Monitoring	.	.	.	41
2.7.2	Amateur Radio	.	.	.	42
2.7.3	Software Defined Radio in the Military	.	.	.	44
2.7.4	Tracking ships and aircraft	.	.	.	44
2.7.5	Attacks using SDR Technology	.	.	.	45
2.8	Products	.	.	.	50
2.8.1	HackRF One with PortaPack	.	.	.	50
2.8.2	Ubertooth One SDR	.	.	.	50
2.8.3	Lime SDR	.	.	.	51
2.8.4	Seeedstudio KiwiSDR Kit (with BeagleBone Green)	.	.	.	51
2.8.5	DSP Receiver	.	.	.	52
2.8.6	FM Radios	.	.	.	52
2.8.7	EM Smartwatch	.	.	.	54
2.9	Companies that provide SDR-Products	.	.	.	54
2.9.1	Amarisoft	.	.	.	54
2.9.2	Aselsan	.	.	.	55
2.10	Conclusion	.	.	.	56

2.1 Introduction

Software Defined Radio (SDR) is a rather new topic in the radio communication sector, gaining in popularity in several different modern applications. Despite it currently being fairly unknown to most, a lot of underlying implementations in modern systems is, amongst other things, based on SDR. This paper shall give insight to this uprising technology, what Software Defined Radio is, where it gets used the most, how it works and where it came from. Furthermore there will be economical aspects as in market analysis, insight into companies related to the topic and how they monetize this technology.

2.2 Architecture

This section will engage on the base architecture of Software Defined Radios and Radios in general. Furthermore we will have a look at the main parts of the Radio architecture and how they differ from a Software Defined Radio (SDR) system and compare their digital and analog counterparts as seen in Figure 2.1 and 2.2.

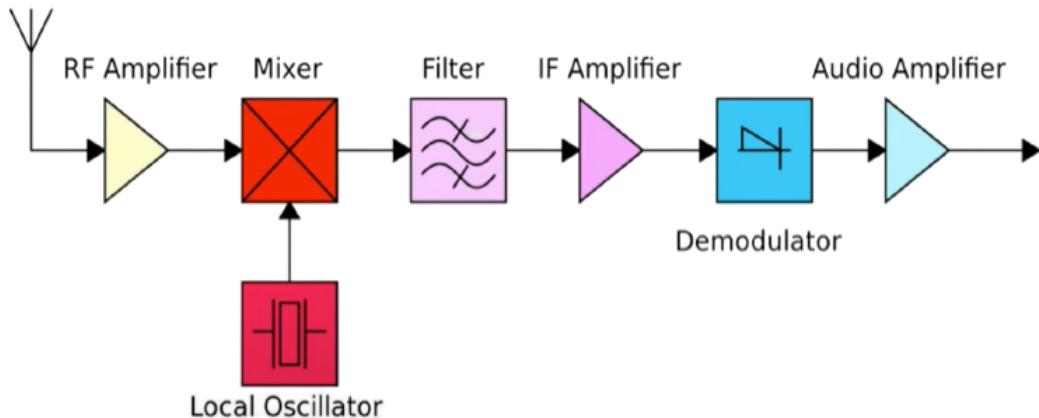


Figure 2.1: Simple radio architecture

The signal Processing has the same steps on mostly every radio; the main difference is that software defined radios use digital signal processing (DSP) and not analog signal processing (ASP) for most sections of the process. To further elaborate on this we have to look at the components in detail and their differences.

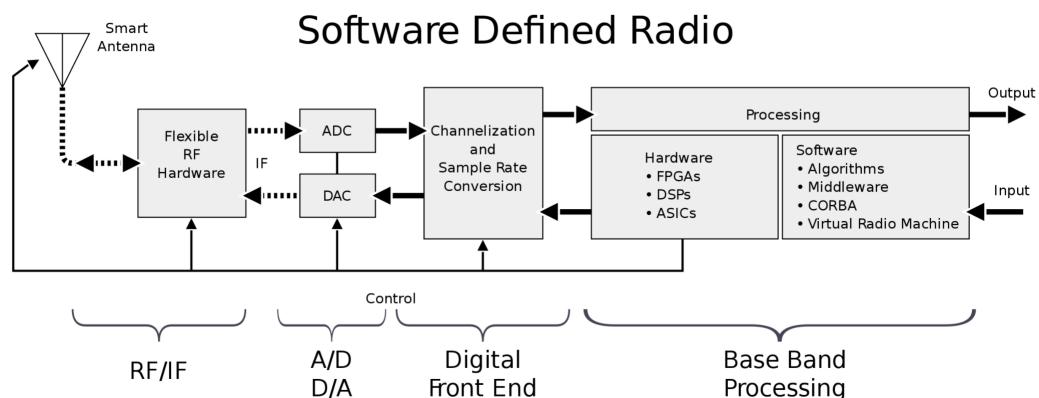


Figure 2.2: Simple software defined radio architecture

2.2.1 Antennas

Most people know that radio signals start with an antenna. An antenna can send and receive electromagnetic waves. You will find it on every SDR or Radio in general. There are even two major distinctions in antennas : analog and digital antennas. However, this distinction makes no difference regarding this subject, since digital antennas just have the additional capabilities to send and receive digital signals in contrast to analog ones, since they are limited by their physical capabilities. Since digital antennas are the new standard now, most Software defined radios will be equipped with one of them. Rather older radios will not have such capabilities. This way the antennas can produce or receive radio waves by regulating electric flow and the respective electromagnetic fields.

2.2.2 Filters

Filters are one of the first and main components of a radio architecture that a signal reaches. Filtering a signal removes specific, mostly unwanted characteristics and qualities of a signal. In usual radios, where the Filters are still part of the hardware, the built in analog filters can mostly only remove anything above or below a certain threshold. Digital filters can be programmed to filter very detailed and specific traits of the signal. Furthermore there is a clear classification among filters: For example, High pass filters remove all signals below a certain frequency and whereas Low pass filters remove all signals above a certain frequency. This way the signal can be perfectly cut into the type we need. One big downside for analog filters is that they decline in precision and quality as they age, and for precise filtering a lot of components are needed, which is expensive. Digital Filters are more precise than simple analog filters, however they also need a digital signal to work with. To achieve this we need to convert the analog signal into a digital one.

2.2.3 Analog-to-digital converter

An Analog-to-digital converter (ADC) is a device or system that converts analog signals into digital ones. This is one of the main parts of a software defined radio, since the software that runs on the SDR needs a digital signal to operate on. The converter samples the signal according to two factors: time and amplitude. The time component is computed from the signal frequency and the amplitude of the signal determines the bit width of the digital values.

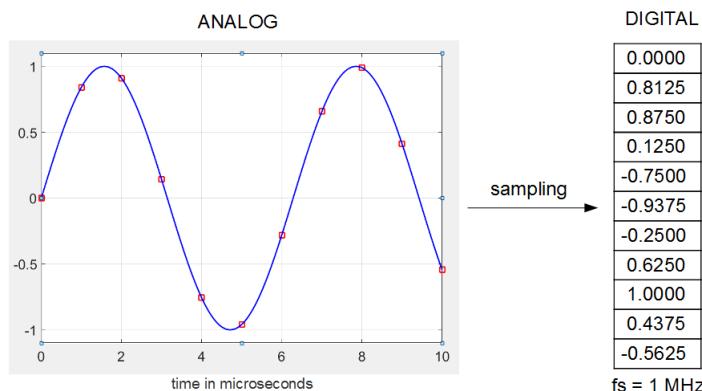


Figure 2.3: Analog to digital conversion with sampling

Sampling a wave is achieved just by looking at specific points at a certain time. The quality of the signal is, therefore, determined by the sampling frequency. If the frequency is too low, information gets lost and the signal decreases in quality, if the frequency is

too high a lot of the data would be redundant or even errors could occur when trying to represent the data due to over-fitting.

2.2.3.1 Nyquist-Shannon sampling theorem

The question of how many samples to perfectly reconstruct the signal without any loss of information is answered by Shannon's theorem [4]. In short: The sampling frequency f_s must be higher than twice the frequency (or highest frequency component) of the analog signal f_{analog} .

This main rule, originated from formulas that would exceed the scope of this paper, gives major guidelines for the sampling of any analog signals. If this rule is fulfilled, the digital version after the conversion will have no information loss.

This theorem also yields one important detail: There is a strong relationship between SDR bandwidth and the data rates (corresponding to the sampling frequency). The larger the bandwidth, the higher data rates generated are in the digital domain. Therefore high bandwidths pose a challenge not only to AD conversion, but also to the succeeding DSP. [5]

2.2.4 Modulators

With the help of modulation one can transform a signal that carries the information to be transmitted (music or similar) into a high frequency signal (f.e. radio waves). It does this with help of carrier signal, which gets modified according to modulation technique. The device that performs the modulation is called modulator. The demodulator performs the demodulation, which is the inverse of modulation. A device that can do both is called a modem. Once again the distinction between analog and digital modulation methods is made.

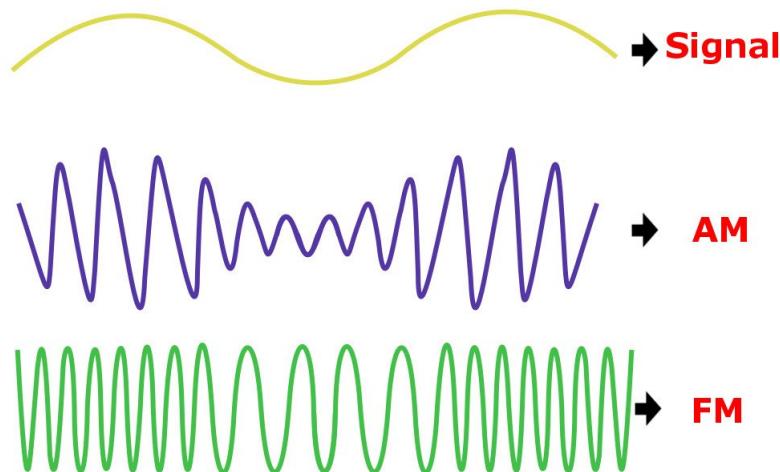


Figure 2.4: Carrier wave with amplitude and frequency modulation

Simply put, analog modulation continuously modulates the carrier signal according to the signal that has the information to be transmitted. There is no digitalization of the signal. The most common analog modulation methods are frequency modulation and amplitude modulation.

Frequency modulation (FM) is the current modulation standard for radio communication. Here the carrier wave gets modified in its frequency domain according to the information (baseband) signal. The amplitude and the phase remain the same. So when demodulation the signal the change and values of the frequency matter.

Amplitude modulation (AM) is a rather similar procedure and it was invented even before frequency modulation. Instead of changing the frequency, the amplitude gets altered according to the information signal. Here the frequency and phase remain the same. With modern technology, FM has clearly dominated AM in use, since FM has higher bandwidth and therefore also better quality and is less prone to interference. AM however remains cheaper and is less impacted by physical barriers. One major point is also the noise, which mostly affects the amplitude and therefore makes AM more vulnerable to it. However there are still various application sectors for AM (f.e. LTE).

2.2.4.1 IQ signals and modulation

There are many different techniques used for modulation. The technique most software defined radios use is the in-phase and quadrature (IQ) modulation. If we look at our radio frequency signals as simple sine waves, we can describe quadrature signals as 2 sine waves that are 90 degrees apart. So a sine and a cosine wave would be in quadrature Q . By convention, the amplitude of the in-phase signal is called I , Hence its called IQ signals.

2.2.5 Mixers

Mixers are mostly used for frequency conversion of electrical signals. Mixers are also, as the name implies, used to create a new output signal combining two or more input signal. The most common way of doing this is taking the sum difference of the input signal frequencies. Therefore, since mixing mostly consists of just arithmetic expressions, a digital mixer perform these operations on the signal without causing any spurs or unwanted modifications, where as analog mixers merely emulate the operations, providing the signal unwanted noise.

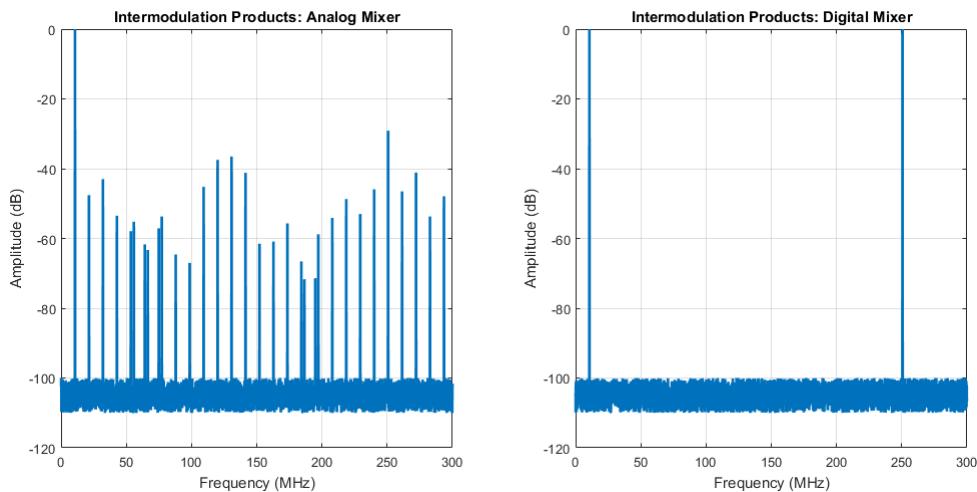


Figure 2.5: analog and digital mixer

Another main usage of mixers is called heterodyning, where the signal gets shifted from one frequency range to another. With newer technology mixers also have been digitalized. With digital mixers you can have programs using multiple mixers on the fly. This allows smaller mixers with greater functionalities, all embedded in software. Analog mixers however are still in use to some extent, due to the much cheaper costs in usages where the signals do not vary too much.

2.2.6 Amplifiers

The other main component altering the signal is called an amplifier. Amplifiers usually increase the magnitude of a signal, with the output signal having more power. There are an abundance of different amplifiers, since they are widely used in electrical devices. The one of interest here is the RF amplifier, which can amplify radio frequencies to a desired extent.

Today most amplifiers use transistors, which are specifically made to amplify electronic signals.

2.2.7 Digital Signal Processor (DSP)

While all the above components are part of the architecture of both radio and software defined radio (in hardware and software respectively), the digital signal processor or DSP is mostly used in modern SDR. It is a microprocessor optimized for digital signals and really the heart of an SDR. Since SDR also receives analog signals, the DSP is usually preceded by an analog-to-digital converter.



Figure 2.6: Simple time steps for conversions

Traditional radio uses analog processing, which means using components(filters, amplifiers, mixers, etc.) implemented in **hardware** to modify and convert the incoming signal from the antenna to the desired output signal (f.e. audio). However, software defined radios rely on digital signal processing methods.

The DSP defines the software in SDR. All components, "The DSP functionality includes the application of software running on a processor or algorithms implemented as digital circuits in microchips (f.e. FPGA/ASIC), that implement the receiver functions, like mixing, filtering, etc. in the digital domain." In all applications, digital circuits have widely replaced their analog counterparts in modern technology.

2.2.8 History and transition of the general architecture

Nowadays most receivers are compact, digitalized circuits with little hardware components left. The analog input signal from the antenna gets small modifications before getting digitalized by a converter which is followed by a digital signal processor. The converter can be placed at any point of the architecture and is therefore deciding which elements get handled by software and which need to be implemented in the hardware.

At the start of the radio age, everything was done analogously with every component being embedded in hardware. There is one pattern or trend that clearly develops: since the invention of an analog-to-digital converter, the position of the ADC is gradually moving further and further to the front, letting the architecture consist of mostly digital processing, its proven to be superior. Therefore, it would make sense to move the converter as close to the antenna as possible, which is referred to as direct sampling.

2.2.8.1 direct sampling

In a direct sampling architecture, the analog-to-digital converter gets connected directly to the antenna, leaving no space for analog processing and therefore making it "all-digital".

However, previous implementation has shown that the direct sampling without any prior analog processing leads to aliasing. Aliasing occurs when the sampling of signals makes them indistinguishable or loss of information while the sampling of signals leads to a low quality reconstruction. Therefore most direct sampling receivers still have some analog processing with a filter or amplifier prior to the converter.

There are still several advantages using direct sampling, since the signal gets digitalized almost right, the only threshold here is the converter itself and the complexity and optimisation of the software running on the DSP.

2.3 Software - GNU RADIO

Since the DSP builds the heart of software defined radios, its also important to look at the actual software running on them. Probably the most popular program working with software defined radios is called GNU radio [10]. GNU Radio is a libre programming tool for SDRs and allows implementations of different of modulation or demodulation techniques, processes to manipulate signals, filters and any kind of other type of software that got written to run on software defined radios.

GNU radio can also be used for digital signal processing without any specific attached hardware, which makes it also a popular tool installed on personal computers.

GNU Radio applications are as *flowgraphs*, which are a series of signal processing blocks connected together, thus describing a data flow.”

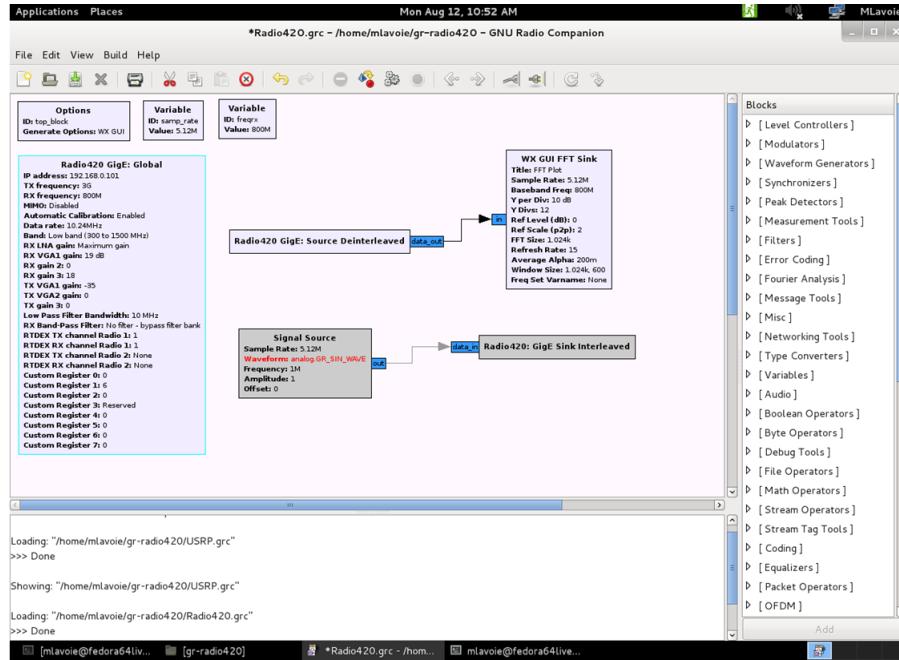


Figure 2.7: flowgraphs in GNU radio

2.4 Experiment

To understand and visualize the whole concept of a software defined radio ourselves, we were lucky to test and play with a simple version of it. Therefore we did some short and simple experiments regarding SDR, LoRa and other basic principles [1] [2]. The SDR had two main parts: the transmitter made out of arduinos [12] and a simple Lime-SDR dongle acting as a receiver.

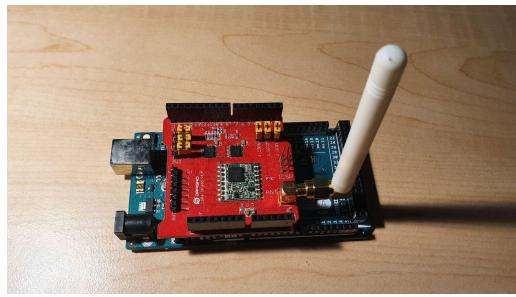


Figure 2.8: arduino with LoRa shield

The Lime-SDR is a software defined radio by itself, since it even is a transceiver, which means it can receive and send different kinds of signal. For the purpose of the experiment we assigned the job of the transmitter to the arduino, which continuously sent out LoRa signals which the Lime-SDR should capture. Using additional software like GNU radio and others, we were then able to modify and look at the received signal. Together with our instructor, we were able to run simple test cases and successfully capture the signals.

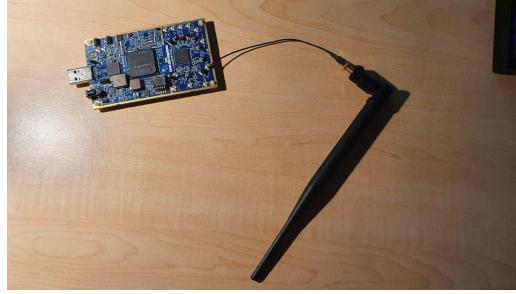


Figure 2.9: Lime-SDR with antenna

The arduino is equipped with a LoRa (long range) shield on top and sends out signals in a specific interval. Additionally, some packets require acknowledgment, which means that the Lime-SDR has to send a confirmation first before receiving the next packet. After sending the request for the acknowledgement request, the arduino transmitter waits some time for a response. If it gets the confirmation, it may continue its job as a transmitter, otherwise it sends another request until it gets confirmation. The signals always start with a certain pattern, to signal where one can start decoding the signal. For example signals please refer to the mentioned papers [1] [2].

2.5 History behind software defined radio

2.5.1 Early developments

Between 1887 and 1888 Heinrich Hertz (seen in Figure 2.10) was the first one who could prove that electromagnetic waves existed and that they behave like light waves in means of travelling through free space, as James Clerk Maxwell claimed in the theory of the electromagnetic field. But at this time, he did not believe that those in that time called Hertzian waves can be used to transmit electric power and telephone signals. As a result, preliminary developments and discoveries were halted by Hertz itself, delaying the use of radio waves longer than necessary. He was much more interested in the fundamental laws of nature than in possible commercial applications, which was of course a hindrance at the time. [13]

Shortly after Hertz's death, Guglielmo Marconi (seen in Figure 2.11) began his own experiments based on Hertz's research. Through countless experiments and investigations, the Italian-born researcher improved Hertz's equipment and was thus able to transmit signals over a much greater distance than before. He did it by using an antenna. In 1896, Marconi applied for a patent for his apparatus which was accepted in June of the same year. The patent is still considered today as the first wireless telegraph. Marconi was then considered by many to be the inventor of wireless transmission, even though he was by no means the only one at the time to have experimented with Hertz's discoveries. Many wanted to buy Marconi's patents, market them, and thereby bring them to the people, but different to Hertz, Marconi wanted to gain money by himself. He and his cousin founded The Marconi's Wireless Telegraph Company of America in 1897, which was the first to sell wireless technologies at that time.

Although at first it did not bring any great financial success and Marconi would have

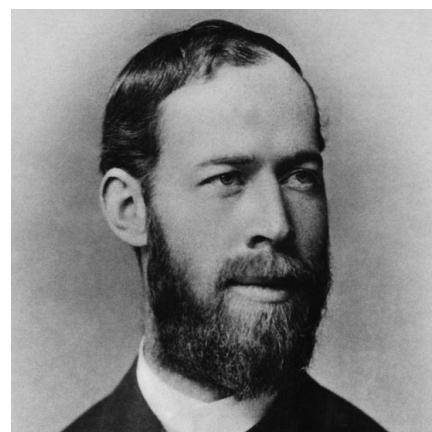


Figure 2.10: Heinrich Hertz



Figure 2.11: Guglielmo Marconi

gone completely bankrupt without the support of financial backers, it was thanks to them possible to advance the experiments further. In 1905, this finally paid off. After Marconi had found the magnetic detector as a replacement for the coherer, his devices finally became economically viable. Especially for ships and, of course, for the army, inventions in the field of wireless information transmission were extremely important, which is why Marconi also specialised in producing wireless information transmission devices. Although it is not clear whether Marconi invented his devices all by himself, he sure did patent them. In stark contrast to Heinrich Hertz, for Marconi commercial success was more important than pure science, perhaps that is why he was more successful. At this point it was only possible to transmit and communicate with simple clicker sounds, which is known today as the Morse code.[14]

2.5.2 Reginald Fessenden

Marconi did not show much interest in the development of voice transmission as he was not sure whether it would be financially viable. That is why he left this specific field to others. One important experimenter in this area was Reginald Fessenden. He quickly realized, that for transmitting a voice, there must be a continuous-wave signal. He then imitated this signal by ramping up the spark rate. After numerous failed attempts, Fessenden was

the first one who's words where heard through thin air. "One, Two, Three, Four, - is it snowing where you are Mr. Thiessen, if it is telegraph back to me! "

The voice was heard quite clearly, the only thing that spoiled the victory was a loud and uncomfortable sound in the background, probably due to the lack of modulation. A bit later, Fessenden was the first man who played live with his violin, so he practically gave the first ever live instrumental performance as well.

Fessenden (seen in Figure 2.12) was able to further contribute to the birth of today's communications by discovering that by adjusting frequency, several messages may be sent simultaneously. This means that he generated different sounds with different frequencies and transmitted them all together at once.

Fessenden continued to improve his machine by using a higher frequency, more power, and the combination of two different frequencies. The method of combining two frequencies is still used today. With this device, he managed to transmit a voice so clearly, that the receiver even recognized the person speaking without any disturbances. On Christmas day 1906, he transmitted the world's first broadcast, which amazed all who heard it. This was the birth of the so called Amplitude Modulation (AM).[15] [16]



Figure 2.12: Reginald Fessenden

2.5.3 Lee de Forest

At that time, radio was recognized as another way for point-to-point connection rather than a method for one person to access multiple people all at once. Lee de Forest, an inventor as well and at some time even a rival to Marconi, saw that potential and pursued it at any cost (seen in Figure 2.13). In 1906, he invented the so-called three-electrode vacuum tube, which is now seen by many as the most important single invention in the history of wireless transmission. It allowed him to amplify and strengthen incoming signals. Using this newfound technique de Forest broadcast music regularly starting in the year 1916, heard by everyone using a radio apparatus at home. He even advertised his own equipment, making him one of the first marketers who used wireless techniques. He gets the credit for being the father of American radio. De Forest had to stop broadcasting during the first world war, which was due to the military banning civilians because all the different waves disturbed their own communication signals. That means only the army was allowed to use any radio systems.

2.5.4 Radio pact of 1912

This was not the first time that the government interfered in the usage of private radio operators, as the so-called radio act was signed in 1912. This act allowed the amateurs to only use a limited range of the radio spectrum, in order not to disturb the professional broadcasters on the one hand, but on the other hand to allow the navy in particular to operate without interference. But what led to this? It was the sinking of the Titanic in the same year. The company behind the Titanic did not include enough lifeboats in their calculations, because they assumed that she ship could simply call for help. This was more difficult than expected because the amateur radio operators on land interfered with the transmissions. Although radio technology allowed more than 700 people to be rescued, many more could have been saved. The US government also recognised this and issued the radio act.[18]

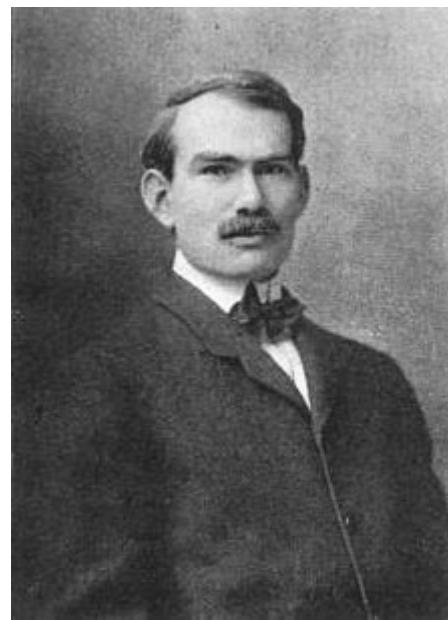


Figure 2.13: Lee de Forest

2.5.5 Radio pact of 1927

However, the radio pact enacted in 1912 was not strict enough to control the radio spectrum for long. Already in 1927, the act of that time was no longer sufficient for the vast number of hobby radio operators, commercial stations and military users, because the government had no right to deny a licence. So, everyone who wanted a license did receive one. This in turn led to a great confusion of messages and signals sent as the number of users grew, partly because not everyone followed the guidelines agreed on in 1912. This, of course, greatly limited the potential for economic growth and a new solution had to be found. The elaboration of a new regulation turned out to be more difficult than expected, because nobody could say how the development of radio would proceed. A lively discussion broke out in the US senate as to who was allowed to broadcast and on which spectrum, but the content was also a major point of discussion. May the freedom of speech be restricted? First of all, the improper speech was banned, especially “obscene, indecent, or profane” language. Those who did use such kind of language could be banned and silenced.

It was still possible to a certain extent to control, as most people have a rough idea of what a curse-word is. It became more difficult when it came to what could and could not be said. Who should control that? The US senate regulated this with a Federal Radio Commission (FRC), which solved all those opinions that did not share the view of the FRC. In other words, all those who rejected the progressive views of the time, such as socialists, non-Christians, immigrants as well as evolutionists were banned. In the FRC’s interpretation this was still valid within the framework of regulated free speech. They justified this with “public interest, convenience, and necessity”. What the US senate basically did is to give all the power over the radio waves to the FRC. This basic version of the radio act is still valid in the USA, but of course people are now free to express their opinion.[19]

2.5.6 Birth of software defined radio

Joseph Mitola is considered the godfather of software radio; he was the first to publish on this subject in 1992. There is a difference between software radio and SDR. SDR is

basically a radio that has its functions defined by software. Software radio is a single wideband analogue to digital converter (ADC) with a number of programmable digital signal processors (DSP). A software radio is always software-defined, but not every SDR is a software radio. As an example, if several narrowband ADCs are used, this by definition is not a software radio, but still an SDR.

Mitola contributed to the US military program named Speakeasy Multiband Multimode Radio (MBMMR). The problem this program had to face was how to reduce the cost of the radio for the military. At that time the military had to regularly change the hardware of radios whenever a new improvement was released. As one can imagine for a whole army that was costly. Another thing Mitola had to worry about was that he has to ensure that the US army was able to communicate with their allies and if they had a new apparatus, he had to be certain that the communication was still possible. Mitola did it by making signal processors programmable through software implementation of waveforms. In this way, it was possible for one radio station to perform rapid changes to waveform modulation which was not possible in analog radios before. But that was only the first phase. In the second phase, Mitola managed to emulate over 15 existing radios of the army. With the software he helped to create, the army was able to communicate with their allies without having to change the hardware. It was the first time in history when this became possible. Unfortunately, the system was still enormous, even filling a whole back of a truck. But over time and due to major new technology, it was possible to scale to whole thing down. It is even used in the mobile telephony sector today. [20] [21]



Figure 2.14: Joseph Mitola

2.6 Modern Communication Protocols

2.6.1 Mobile wireless generation

In this section, an overview of the history of mobile telephony is provided. This section focuses on the mobile wireless generation (G), because the SDR plays a big role there.

2.6.1.1 1G

Starting with 1G, the very first mobile wireless generation, was introduced in 1982 and allowed to make phone calls. The system used was frequency modulated but still analogue. In addition, the device was expensive, heavy and of short lived battery. The restriction allowing using the device in one country was another disadvantage. Especially from today's point of view, one detail is particularly of high importance: there was no security whatsoever. It was easy to eavesdrop on the conversations, something that probably nobody wanted to do at that time, just like today. The transfer rate was about 2.4 kilobits per second.

2.6.1.2 2G

With the introduction of the second generation (2G) in the late 1980s, however, devices have already become much faster. Up to 64 kbps were possible. For the first time, digital signals were used for the transmission of the calls. Besides the obligatory improvements

in call quality, it was also possible to send text messages and even pictures. A great innovation, as this was not yet possible with 1G. Videos could not (yet) be sent. One large improvement was security; it was much more difficult than before to spy on somebody using 2G than 1G. The intermediate generation between 2G and 3G, called 2.5G, made it possible to send and receive email messages. Furthermore, it was already possible to use the web, but although the speed of data transfer increased with up to 144 kbps, it took between 6-9 minutes to download an mp3 song of 3 minutes length.

2.6.1.3 3G

Introducing the third generation, (3G), in the early 2000. The increase in speed was immense, nearly 15 times faster than in 2.5G, starting at 2 Mbps but even 14 Mbps were possible on high end devices. This allowed the user to send and receive videos as well as using TV streaming services or even doing video conferences. The example already mentioned above illustrates this well. 3G needed about 11 to 90 seconds to download the same mp3 song, with 2.5G it took over 6 minutes at least. But this immense increase had its price because the devices were still not very light and quite expensive. That is when they started being called smart phones.

2.6.1.4 4G

The fourth generation (4G) brought some innovations such as increased security and privacy, but most importantly the increase of data transmission was the main focus. There is even talk of up to 1Gbps. But to be honest, that is in fact never the case. The latency on the other hand was reduced dramatically. Latency refers to the time it takes a device to send a data packet to another device. The latency rate in 4g is about 50 milliseconds. The costs was reduced and the quality of the video streaming was improved. Unfortunately, the devices need more power to handle all this than in previous generations. Because of that the energy consumption was more than before, the battery life was lowered.

2.6.1.5 5G

4G is used by the vast majority of devices today, but there are some models of the latest generation of smartphones that can handle 5G. One example is the Samsung galaxy s20+. Countries are also continuing to expand their infrastructure, so it is reasonable to expect that 5G will replace 4G in the future. But what can we expect? Of course, the almost obligatory increase in data transmission, but above all we want to reach everyone everywhere and at all times. The coverage should therefore improve drastically. But the latency should be reduced as well. In short: Faster, better and best of all cheaper. One point is mentioned again and again: Support of the wireless world wide web (WWW). This is considered the most important point of the new 5G technology, as the use of the world wide web has shifted more and more to the smart phone in recent years. In the future, the www should therefore be available wirelessly at all times and in all places.[22]

2.6.2 Wi-Fi 6

The new Wireless-standard Wireless Fidelity 6 is not only faster but has also more power in heavily frequented areas like schools, big offices, companies or airports etc. It is a good basis for building up an intelligent network which provides better networking with a continuous connection [23]. In comparison to the older Wi-Fi 5 (released in 2014) Wi-Fi 6 is up to 40% faster. This is possible through more efficient data, encoding thus, resulting in higher throughput. With SDR more data can be sent with the same radio waves which is exactly how Wi-Fi 6 works. The encoders and decoders are getting more powerful and

therefore can handle the extra amount of data [24]. Wi-Fi 6 brings also other mechanisms like the Target wake time (TWT). TWT is a mechanism to save the power of devices like smartphones by scheduling a wake-up time for devices in power-save mode. This means that the client does not have to be awake all the time because TWT tells it when to wake up. Therefore, the device can save energy by sleeping longer [25]. There are some key technologies that are important for understanding Wi-Fi 6: MU-OFDMA, MU-MIMO and spatial reuse. Multi-user (MU) means that transmissions of data between several devices and an access point are possible [26]. Orthogonal frequency division multiple access (OFDMA) allows multiple clients with different bandwidths to be handled simultaneously. We will not go further in detail how this technology works exactly but it is important to know it leads to better frequency use, reduced latency, and increased efficiency [27]. With Multiple-input multiple-output (MIMO), it is possible to transmit multiple frames to several devices at the same time and on the same channel by using multiple spatial streams. This leads to greater efficiency and spatial diversity but the technology is rarely used and implemented [28]. Figure 2.15 shows the main differences of MU-OFDMA and MU-MIMO.

MU-OFDMA	MU-MIMO
Increased efficiency	Increased capacity
Reduced latency	Higher data rates per user
Best for low-bandwidth applications	Best for high-bandwidth applications
Best with small packets	Best with large packets

Figure 2.15: Comparison

Important: MU-OFDMA and MU-MIMO are different from each other. OFDMA allows MU by subdividing a channel whereas MIMO allows it by using different spatial streams [29].

2.6.2.1 Internet Box 3



Figure 2.16: Internet Box 3

Swisscom is a company in Switzerland that provides IT, communication and entertainment and also offers Fastweb in Italy [57]. In Figure 2.16 the new Internet Box that supports Wi-Fi 6 can be seen which Swisscom presented back in 2019. Clients can now surf the internet with a speed of up to 10 Gbit/s. The Internet Box 3 uses GNU-software from Vestiacom which is based on software defined radio [58].

2.6.2.2 RT-AX88U

The RT-AX88U router (Figure 2.17) from ASUS also supports Wi-Fi 6 and uses OFDMA technology. It has a bandwidth of 160MHz and uses 1024-QAM modulation for faster wireless connections. ASUS is specialized in the Gaming-industry which is also apparent when looking at Figure 2.32 and promises a better gaming experience without latencies and lags with the new router [59].

2.7 Use Cases of Software Defined Radio



Figure 2.17: RT-AX88U

monitoring, in amateur radio, military and mobile radio, for hacking purposes and increasingly in civil applications such as digital radio receivers. A good and illustrative example is the implementation of base stations of cellular networks with SDR. These could easily be upgraded to new standards within a very short time and at low cost. This and further examples will be covered in the remaining part of this paper.

2.7.1 Spectrum Monitoring

With conventional radios it was only possible to listen to individual radio transmissions. Thanks to SDR, there is a possibility to analyze the complete spectrum simultaneously, make recordings and replay them later. Having access to the whole spectrum also gives a much better overview of radio communication than with a conventional receivers.

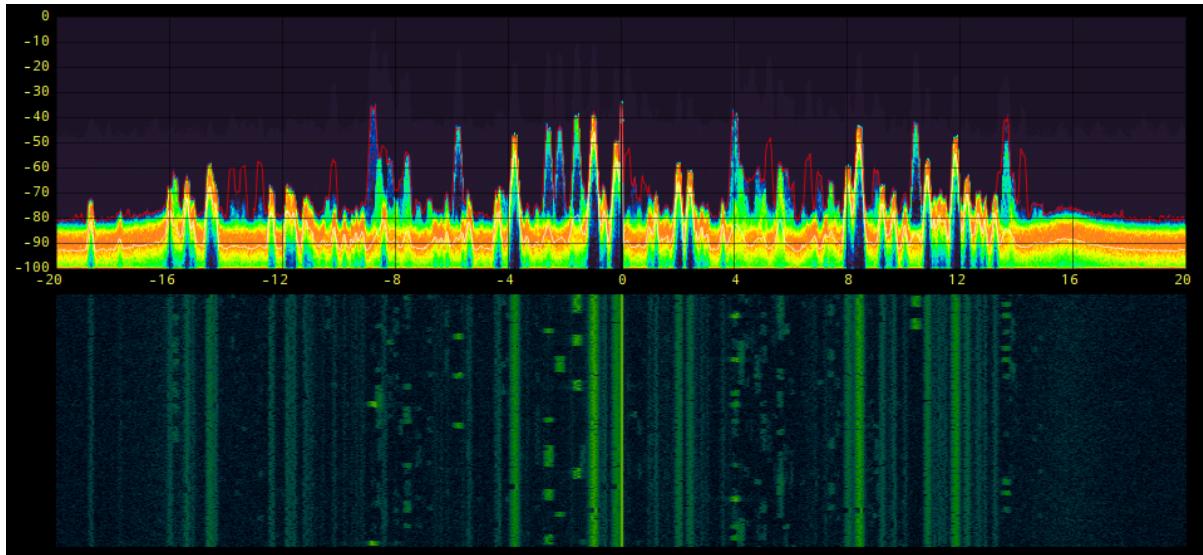


Figure 2.18: Spectrum using Fosphor, a GNU Radio spectrum analyzer

The user interface of this *GNU Radio Spectrum Analyzer* [30] on Figure 2.18 consists of two main parts. The top shows a histogram with the current frequencies with the red line marking the highest recorded signal on a given frequency. The bottom part contains a time axis and shows the frequency over time.

A lot information can be read from the spectrum. By the transmission frequency one can observe in which frequency range the signal of interest is located. One can also

An SDR system performs most of the signal processing within a single device where no dedicated hardware is required for specific signal processing. An essential feature is that the different parameters of the radio system such as modulation, different bandwidths, temporal behaviour and different channel coding methods can be implemented in software. This makes products cheaper and drastically increases the application possibilities.

SDR can be used in a large variety of cases. It can be used for spectrum

discover how much bandwidth the transmitted signal occupies. Digital voice radio and FM need a relatively large bandwidth of about 88 to 108 MHz, the signal will be a few kHz wide. Much less bandwidth is needed for single-sideband modulation, which is an energy-efficient modulation type for speech transmission and is commonly used on analog radio links such as the shortwave range for maritime radio, long-range aeronaautical radio, military applications and amateur radio. The more bandwidth a signal occupies, the higher the transmission speed.

2.7.2 Amateur Radio

Amateur radio is a field of experimentation from which in the past groundbreaking technologies and methods of information transfer have repeatedly emerged. This has not changed since the beginning of wireless communication until today. A list of most inventions and discoveries can be found in the history of amateur radio [31].

2.7.2.1 Setting up new Records on Ultra-Long Wave

As a driving force in radio technology, amateur radio has already set many records. This January, for example, three radio amateurs from Germany succeeded in receiving amateur radio signals from the USA - on ultra-long wave, a very weak frequency below 9 kHz, making it the most distant station to decode the message of a radio amateur from North Carolina. They have covered a total of 7257 km [32].



Figure 2.19: Markus Vester, Tom Kölpin and Bernd Wiesgickl - The radio amateurs, who set a new distance record below 9 kHz.

2.7.2.2 Amateur Radio Satellites in Space

Amateur radio operators have been involved in spaceflight since the 1960s. They construct small or medium-sized satellites that are carried as secondary payloads on commercial or scientific flights.

Amateur radio satellites are built or operated by radio amateurs that use the amateur radio bandwidth for communications. Some of these satellites contain transponders, that are used as relay stations for various modes of operation. Others contain experiments or cameras whose data are transmitted to earth. Amateur radio satellites with an earth orbit are mostly called OSCAR (Orbiting Satellite Carrying Amateur Radio).

The first amateur radio satellite was launched on December 12, 1961, under the name OSCAR 1, just 4 years after the launch of the Soviet first satellite, Sputnik 1. OSCAR 1 was the first satellite to be launched as a second payload along with another satellite, that was carried into an independent orbit. Although the satellite remained in orbit for only 22 days, the project was a great success; over 570 radio amateurs in 28 countries reported their observations to the OSCAR project.



Figure 2.20: Es'hail-2 Geostationary Satellite [35]

most of Europe, Africa and the Middle East. One of the aspects of Amateur Radio is to provide help in the form of radio communications when disasters occur [36].

Over the years, more than 100 amateur radio satellites have been launched. These satellites have often contributed to significant breakthroughs in satellite research, since it has been possible to try out new techniques safely in an amateur radio satellite. They have significantly fewer constraints to provide a replacement satellite at short notice in the event of a failure and have a large number of expert observers working free of charge in the form of radio amateurs. Innovations include the launch of the first satellite with a voice transponder and the development of digital store-and-forward messaging via satellite. In 2018, a Space-X Falcon 9 rocket launched the satellite "Es'hail-2". It carried an amateur radio linear transponder as a secondary payload, giving radio amateurs access to satellite communications over a wide area. Es'hail-2 is at 35,786 km above the equator at 25.8° East and covers

2.7.3 Software Defined Radio in the Military

Communication plays a key role in networked operations and civil-military cooperation. Reliable and efficient systems for mobile communication networks, which enable correct and timely command and control even under difficult conditions, are required. With the constant evolution of communication technology, hardware-based implementation methods show significant drawbacks in terms of cost, production cycles and compatibility. In order to avoid these disadvantages, a new procedure was introduced to the military; the SDR.

Software defined radios can be used to determine the waveform and frequency assignment according to the application and electromagnetic environment. With new antenna technologies such as smart antennas, an increase of power and spectral efficiency as well as reduction of the detectability is possible.

The SDR technology also makes it possible that a hardware platform set up for general purposes offers compatibility to different systems of wireless communication by updating software configurations. This means flexibility for the wireless communication systems and allows new features to be added and the system to be easily upgraded. This can save money that would otherwise be spent on expensive hardware. For the planned improvement of voice and data transmission, the Swiss Armed Forces will spend around 600 million Swiss francs on new radio and directional beaming equipment. However, it is not yet known exactly which products will be acquired [37].

The rapid pace of technical development and the associated short System life cycles of civil telecommunications equipment always result in the use of new concepts and technical systems for military applications. Therefore, investigate new methods like SDR in real environments plays an important role for the military worldwide.

2.7.4 Tracking ships and aircraft

Ships use an Automatic Identification System or short AIS. It's an equipment of great importance to the safety of navigation. With the help of AIS, all ships worldwide identify themselves and send travel-related data for all participants of the shipping industry. A distinction is made here between static and dynamic information. Static information is assigned to the ship, e.g. ship name and call sign. Dynamic information is variable information such as the current course and speed. This means, that it is possible to track all ships and their positions in real time. The only drawback is that AIS devices cost about 800. Using an SDR instead is much cheaper and would make it more accessible for everyone.

It is possible to set up an AIS using an SDR. To do so, the following four things must be present [41]:

1. An SDR dongle working with SDRSharp.



Figure 2.21: Radio in the Swiss Military [38]

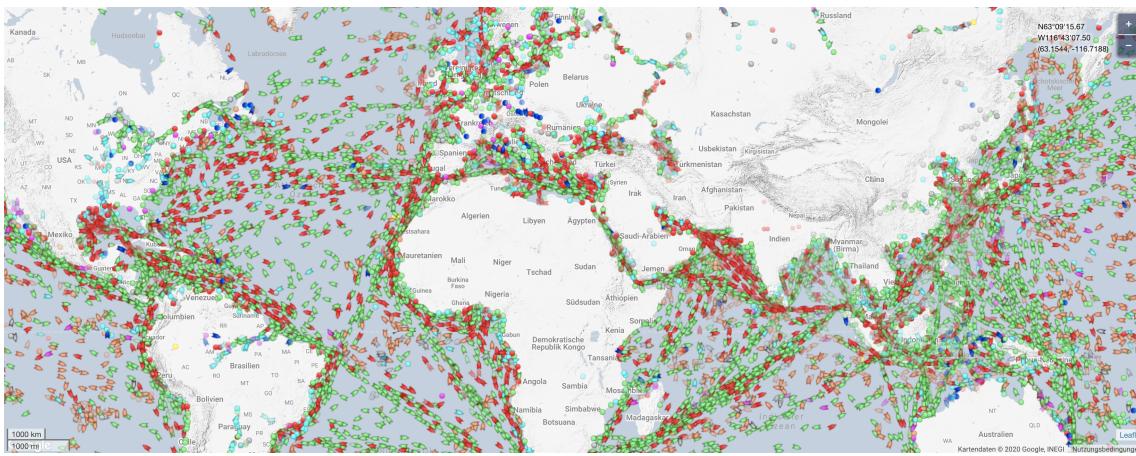


Figure 2.22: Tracking Ships on the Sea with "Schiffs-Radar" [39]

2. An audio piping method to decode the audio.
3. A vertically polarized antenna tuned to 162MHz.
4. A software for decoding the AIS signals.



Figure 2.23: Tracking Airplanes using "Flightradar24" [40]

Airplanes have a similar system for tracking and Identification: The ADS-B (Automatic Dependent Surveillance - Broadcast). For Example: Flightradar24 operates the world's largest network of ADS-B receivers. This network, together with government air traffic control and other data sources, is how the website is able to track aircraft around the globe. The ADS-B signals from the airplanes and the AIS signals from ships, can also be received with an SDR. Sharing this data with Flightradar 24 will be rewarded with a free Business plan subscription. Through advertising banners, app sales and the trade with raw data, the website flightradar24 has made about 10 million Euro revenue in 2018.

2.7.5 Attacks using SDR Technology

Two different types of threats to the internet of things, or short IoT, can be distinguished. On the one hand, known attacks like the replay attack are more frequent because the hardware for such attacks becomes cheaper and is more adaptable. On the other hand, the rapid development of the technology allows new combinations, which can lead to unpredictable attacks. For example, drones can be equipped with SDRs. This allows to overcome physical access restrictions such as fences and walls and finally, with the SDR device, wireless communication on previously inaccessible terrain can be intercepted or disturbed.

2.7.5.1 Hacking the Internet of Things

In our modern world, smart devices connected to the Internet of Things are virtually ubiquitous. Manufacturers equip ordinary objects such as toys, furniture, cars and medical devices with smart features to make them more attractive and competitive. Even manufacturers of water bottles are beginning to network their bottles[42]. As with many new applications, manufacturers are focusing on the benefits and costs rather than the safety aspects when developing the devices. It is therefore not surprising that security problems in IoT applications are becoming more frequent and that attacks can be observed in which weaknesses are exploited in a targeted manner. According to the Swiss government, it is assumed that in future 200 things will be connected per person[43].

The IoT hides many dangers that the average user is often not aware of. This list shows only a few of these risks.

1. Easy or guessable passwords are set as default.
2. It may happen that the network service is not secure.
3. Security lack in the ecosystem interfaces
4. No secure update mechanism
5. Components are outdated

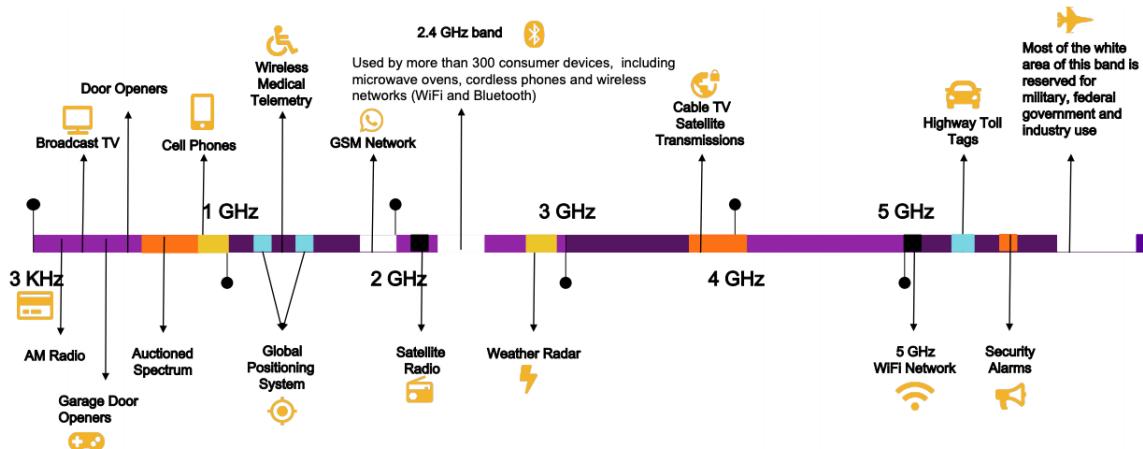


Figure 2.24: Radio Wave Spectrum of the Internet of Things [44]

Figure 2.21 Shows different devices on their frequencies in the radio spectrum reaching from 3KHz to 5GHz. With an SDR device like the HackRF One, which will be described in more detail later, it is possible to hack the complete spectrum without having to spend money on expensive hardware.

2.7.5.2 Replay Attacks

In a replay attack, a signal is recorded and sent again. For example in the case of the radio signal issued by a remote control to open the remotely operated garage door. If the signal was recorded for opening and then the garage was closed by the user, the attacker can later use the recorded signal to open the door again. In this type of attack, the actual signal is not examined or interpreted, but simply retransmitted 1:1. As an example, it can be accomplished on the HackRF One [11], with the following lines of code.

Recording a signal:

```
hackrf\transfer -r 43378000.raw -f 43378000 \\
```

Transmitting the recorded signal:

```
hackrf\transfer -t 43378000.raw -f 43378000 \\
```

Replay attacks are a real threat to network security if they are successful. Unlike many other types of attacks, replay attacks do not rely on decrypting data, making them an effective means of workaround for malicious actors who are increasingly confronted with secure encryption protocols. However, replay attacks also show vulnerabilities:

1. Createing a valid message from scratch is not possible
2. A message can't be modified

There are several ways to protect oneself from replay attacks. Some of them are listed here:

1. Rolling code

A rolling code or hopping code is often used in radio-based vehicle opening systems or entrance control systems such as garage door openers to authenticate a legitimate user wirelessly. Based on a shared key between sender and receiver, called the symmetric key, and a suitable cryptographic algorithm, a constantly changing so-called 'next code' is transmitted from the sender to the receiver for verification.

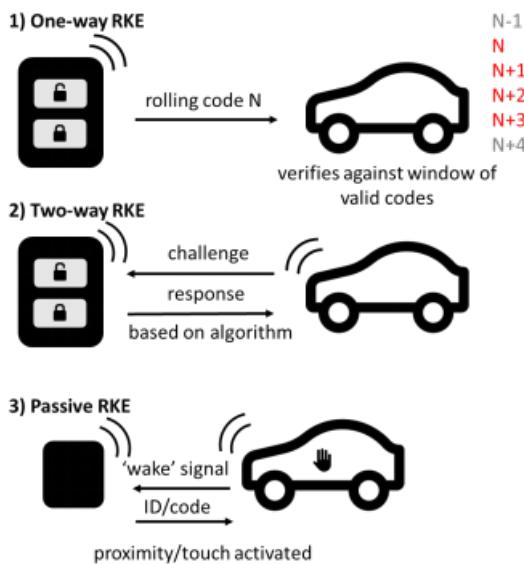
2. One-time passwords

One-time passwords are similar to session tokens in that the password expires after it is used or after a very short time. They can be used during the authentication process to establish trust between two parties communicating with each other.

3. Time stamp

Another way to prevent repeat attacks is using time stamps. The trade-off is that replay attacks might succeed if they are executed fast enough (within a reasonable limit).

2.7.5.3 Hacking Remote Keyless Entry systems uning Jam and Replay Attacks



Remote Keyless Entry systems, short RKEs, can be categorized into three broad types as seen on the illustration 2.25.

1. A one-way RKE requires a manual button press to perform an action. The vehicle receives the signal and confirms that it is a valid code, then performs the requested action like opening the car. With a rolling code system, a number generator installed in the vehicle and the key fob is used to periodically change the required code after a keypress (usually with a small buffer for keypresses, that are out of range).
2. Two-way RKEs require a response from the key fob given a certain challenge from the vehicle.

Figure 2.25: Different kinds of Remote Keyless Entry Systems [45]

3. Passive RKEs automatically unlock within a certain radius or upon the user touching the doorhandle.

One-way RKEs are the simplest and most common form of keyless entry, and its security has some serious drawbacks. The most common attack carried out against one-way RKEs is called "jam and replay attack". The attacker uses an SDR device with full-duplex capabilities, which means that the device may transmit and receive signals simultaneously. This is used to produce a jamming signal, in order to prevent the car from receiving the valid code from the key fob. The device then simultaneously intercepts the rolling code and stores it for later. When the user presses the key fob again, the device captures the second code, and transmits the first code. With this method, the attacker always possesses the next valid rolling code and can easily enter the car as soon as the owner leaves the parking lot.

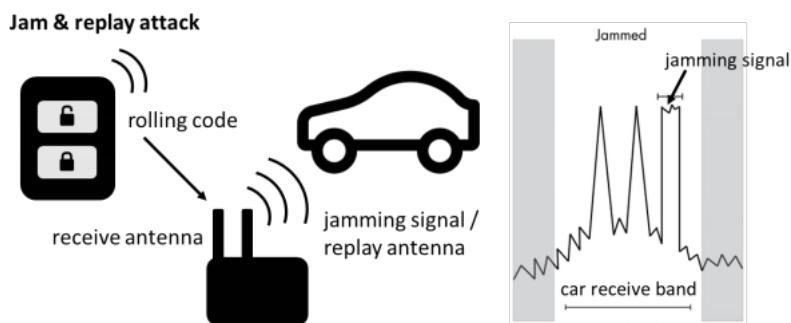


Figure 2.26: A Jam and Replay Attack using a SDR device with full-duplex RF capatabilities to hack a one-way RKE with a rolling code [45]

2.7.5.4 Copy RFID Tags with SDR

Radio Frequency Identification (or short RFID) technology is omnipresent. Passive versions are used in identity cards, access cards, electronic cash cards or car keys. The functional principle is usually the same: on the inside is a coil connected to a microchip. The reader generates a magnetic field which induces a voltage in the coil and thus supplies the chip with energy. The chip uses the coil as an antenna to send information to the reader, such as an identification code. The range is deliberately limited to a few centimeters. But the reader can be replaced by additional hardware like the Proxmark3[46], which can increase the range, such that all kinds of attacks on RFID technology can be carried out.



Figure 2.27: Products using Near Field Communication[47]

Especially the widely used smartcards of the MIFARE Classic type are affected. They contain 1 or 4 KByte of EEPROM memory, which is divided into blocks and can only

be read and written with knowledge of the respective keys. Due to a weakness, however, knowledge of a single key is sufficient to access blocks (this is called a nested attack). Unused blocks often use default keys. These are already known and can be tested quickly. Thus, it is possible to read out a canteen card for example and copy it to a second card within a few minutes. If the credit is stored locally on the card, it would now be possible to reload the card over and over again for profit.

The problem also affects many access cards. In a careless moment, an attacker could snatch a card, clone it and bring it back unnoticed. In theory, a full lane is enough to get close enough to an NFC card. Newer card types, which are also used in identity cards, are considered secure up to now due to improved crypto procedures.

2.8 Products

2.8.1 HackRF One with PortaPack



Figure 2.28: HackRF One

The PortaPack is being connected to the HackRF One which is seen in Figure 2.28 and has a touchscreen-display and navigation elements but also a headphone jack, a real-time clock and a Micro-SD-card slot. The PortaPack can be charged and there is therefore no need for a power connection. There is also no additional computer required because the firmware runs on the fast ARM-processors (<https://archive.org/stream/AcornUser102-Jan91#page/n8/mode/1up>) in the HackRF One unless the user wants to reprogram the firmware. The HackRF One can transmit or receive 1MHz to 6 GHz. It is designed to be able to test and develop modern radio technologies.

Specifications:

- The device has a half-duplex SDR transceiver which means that it cannot transmit and receive at the same time.
- It can record up to 20 million 8-bit quadrature samples (8-bit I and 8-bit Q) per second.
- It is compatible with GNU Radio, SDR and more [11].
- It has a baseband filter and a software-configurable RX and TX gain (Transmit and Receive)
- The antenna can be controlled by the software (power: 50 mA at 3.3V)
- It has a SMA (SubMiniature version A) female antenna connector as well as a SMA female clock input and output for synchronization.
- There are also several buttons for programming, internal pin headers for expansion and Hi-Speed USB 2.0.
- The device is powered by USB and consists of a open source hardware [11].

There were some critics about the HackRF One because it can be used to “hack” wireless devices such as cars [49]. The device costs about 250 CHF or a bundle with antennas, a clock and a USB-Cable about 150 CHF on www.aliexpress.com.

2.8.2 Ubertooth One SDR

The Ubertooth One which can be seen in Figure 2.29 is able to receive and send 2.4 GHz signals through Bluetooth and is an open source solution. Apart from this it is able to operate in monitor mode and can thus monitor Bluetooth traffic in real-time. This type of mode has already been present in WiFi for a long time and has been used in the research, development and security fields but was not compatible with Bluetooth. Not only the software but also the hardware are both open source solutions which makes the radio a fully open-source platform [50]. The Ubertooth One can be purchased for 150\$ on www.amazon.com.



Figure 2.29: Ubertooth One

2.8.3 Lime SDR



The LimeSDR in Figure 2.30 which we used in our experiment can send and receive LTE, LoRa and Bluetooth. The advantage of this board is that it can be used by anyone who know how to use an app store. Through the store you can upload and download apps which others can use. It's a good device to start with learning SDR.

Figure 2.30: LimeSDR

Here are some possible applications listed:

- Radio astronomy
- RADAR
- 2G to 4G cellular basestation
- Media streaming
- IoT gateway
- Wireless keyboard and mice emulation and detection
- Tire pressure monitoring systems
- Drone command and control
- Test and measurement

The board has 256MBytes of memory and a frequency range of 100kHz - 3.8GHz. The bandwidth is 61.44 MHz and it also uses 2x2 MIMO multiplexing [52].

2.8.4 Seeedstudio KiwiSDR Kit (with BeagleBone Green)



Figure 2.31: KiwiSDR

The KiwiSDR shown in Figure 2.31 is an SDR which supports shortwave, longwave and AM broadcast bands. Besides that, it also covers different utility stations and amateur radio transmissions. The radio can transmit from 10kHz up to 30MHz. The KiwiSDR is a circuit board and can be connected to the BeagleBone computer (green part in Figure 2.28) by adding an antenna, an energy supply and a network connection. The software is provided with a micro-SD card. With a browser that supports HTML5, it is possible to listen to a public KiwiSDR everywhere in the world. Up to four people can listen to the radio at the same time and tunes independently of each other. The radio is also able to automatically calibrate frequencies via received GPS timing and further decoders and other utilities can be added as an extension [50]. The KiwiSDR is web-based and shared. This means that anyone who has access can listen via a web browser [51]. This SDR costs about 300\$ on www.seeedstudio.com.

2.8.5 DSP Receiver

Figure 2.32 shows a Digital Signal Processing (DSP) SDR which has a frequency range of 50KHz to 200MHz. The LCD touchscreen is very practical and big enough to be able to read it without problems. The radio is built on SDR and is compatible with all types of analog modulations like AM, Sideband (SSB), Narrowband Frequency Modulation (NFM) and Wideband FM (WFM). The filter width can be changed and it has an adaptive and threshold noise suppressor as well as a noise blower. The Automatic Gain Control (AGC) and equalizer can also be counted to the important functions. The radio can be connected to the computer via USB to transfer IQ and also audio. Like the HackRF One, this device also provides an SMA antenna socket [53]. It can only receive but transmission of signals is not possible. This product can be purchased for 150 CHF on www.aliexpress.com.



Figure 2.32: DSP SDR

2.8.6 FM Radios

2.8.6.1 NX-3720E

Figure 2.33 shows the NX-3720E which is a multi-protocol digital radio designed to operate with Next Generation Digital Narrowband (NXDN), DMR (Digital Mobile Radio) digital and FM analog. A mixed operation between digital and analog is possible which means it can be used universally.



Figure 2.33: NX-3720E

There is also an advanced version which is the NX-3720GE. Additional to all basic features, it also has built-in features like Bluetooth and GPS for providing position data which makes the radio very flexible and thus more practical. Through open-source software based on SDR, it is possible to customize the device. The cost for the radio in Figure 2.28 is at 630 CHF. The advanced one costs about 680 CHF [54].

2.8.6.2 NX-5700E

The NX-5700E (Figure 2.34) is designed for radio networks with up to 1024 channels or optional in total 4000 radio channels and also for the transfer protocols NXDN, DMR and P25 as well as analog QT (Quiet Talk)/DQT (Digital Quiet Talk)/5 sound. The radio uses 56-Bit-DES-encryption and is therefore proof against eavesdropping. The screen is a thin-film-transistor-display and is even readable by sunlight. Like the radio above the NX-5700E consists of a Bluetooth-hands-free kit which also can be used with the BT-transport-protocol. Through the integrated DSP and two microphones this radio has a noise suppression. It also consists of a GPS-receiver which allows the localization of other devices. Additionally, it has a microSD card slot and can save up to 32GB of data. It costs about 910 CHF [55].



Figure 2.34: NX-5700E

2.8.7 EM Smartwatch

Disney-research has published a paper about a smartwatch seen in Figure 2.35 with EM-Sense in 2015. The paper describes how many objects emit electromagnetic (EM) signals and how a watch can be used to detect the objects by using these signals. When touching an object, the body of the user works like an antenna and sends these signals to the watch. By using a software defined radio these EM signal can be detected in real-time which allows to identify the object on-touch. Figure 2.36 shows the electromagnetic waves when an object is being used [56].



Figure 2.35: EM-Smartwatch

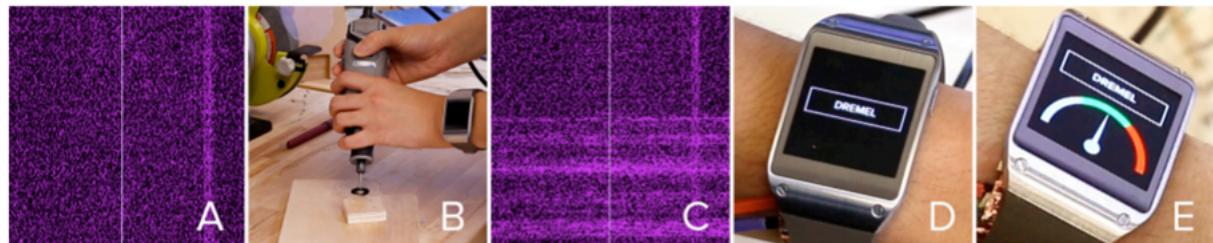


Figure 1. Spectrogram of ambient electromagnetic noise (A). When a user operates an electrical or electromechanical object, such as a Dremel (B), it emits EM noise (C), which we classify (D) and use to enable rich contextual applications (E).

Figure 2.36: Waves

2.9 Companies that provide SDR-Products

2.9.1 Amarisoft

Amarisoft was co-founded by Fabrice Bellard in 2012 and is a company that provides software solutions with SDR especially for the telecoms industry [60]. Amarisoft offers a variety of products to install LTE, Narrowband IoT (NB-IoT) [61] and 5G Non-Stand Alone (NSA) and Stand Alone (SA). Some products:

- Small cell is based on 4G, 5G or NB-IoT and runs a software. It consists only of a Radio Front End which is added to a mother board.
- Macro cell: This cell supports TDD and FDD and reaches 400Mbps per sector by using 256QAM modulation.
- LTE access point: It can act as an LTE or 5G access point [62].

Amarisoft also provides products for testing and measuring 5G, LTE and Cellular IoT like the Amary Callbox. There are three different subproducts which are all a box running a software. With that software logs from the physical to the IP layer can be generated [63].

2.9.2 Aselsan

Aselsan is a company located in Ankara, Turkey and was founded in 1975. Their main goal was to improve the communication in the Turkish Armed Forces. Aselsan provides many products from communication and information technologies to weapon systems, air defense and even medical systems [64].

One of them is the in Figure 2.37 seen GRC-5220 which is based on Orthogonal Frequency Division Multiple Access (OFDMA) that is explained further below in this paper. The GRC-5220 is an Ethernet radio and is used for reliable communication. The product also uses MIMO technology to provide higher throughput [65].



Figure 2.37: GRC-5220



Figure 2.38: Sahara PRC-5333

Another product is the SAHARA PRC-5333 (seen in Figure 2.38), an SDR that provides continuous audio, high speed data and even video communication. Its frequency bandwidth is between 30 and 512 MHz and both AM and FM modulations are supported. By using different frequencies on the same radio, different units are able to communicate with each other on the field. The device can be updated with new software and new frequency rates [66].

2.10 Conclusion

There are many cheap software defined radios for amateur users that can start from 20 CHF but can quickly get up to several hundred swiss francs. Although most people do not know about SDR it is already of big importance and is being used more and more in our everyday lives. Not only became Wi-Fi 6 the standard Wi-Fi in 2020 but also 5G is getting more and more popular (and there is already work for 6G). Companies that offer software defined radio no matter if they provide solutions for communication or for the military will gain more importance in near future. The EM-smartwatch is a good example that soon everybody will be wearing a software defined radio around their wrist. This is how SDR can simplify our lives.

Using SDR not only allows more flexible and cheaper devices, it also enables to use the own computer as a complex analysis and decoding tool. The free software GNU Radio already offers the toolbox for generating and decoding radio waves on the PC and paves the way for a multitude of possible projects, which so far could only be realized with complex and expensive hardware. Whether this is for the better or for the worse is up to the user.

Bibliography

- [1] B. Stiller, E. Schiller, C. Schmitt: *Design and Evaluation of an SDR-based LoRa Cloud Radio Access Network*, September 2020. <https://www.zora.uzh.ch/id/eprint/189754/>.
- [2] S. Ziegler, M. James: *An Overview of Netwrok Communication Technologies for IoT*, 2020, Springer.
- [3] J. Machado-Fernández: Software Defined Radio: Basic Principles and Applications; http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0121-11292015000100007, October 2020
- [4] Nyquist-Shannon; <https://www.allaboutcircuits.com/technical-articles/nyquist-shannon-theorem-understanding-sampled-systems/>, October 2020
- [5] panoradio; <http://www.panoradio-sdr.de/>, October 2020
- [6] Software defined radio: Operation, challenges and possible solutions; <https://experts.syr.edu/en/publications/software-defined-radio-and-software-radio-technology-concepts-and>, October 2020
- [7] Software Defined Radio and Software Radio Technology: Concepts and Applications; <https://experts.syr.edu/en/publications/software-defined-radio-and-software-radio-technology-concepts-and>, October 2020
- [8] Fourth Generation Systems and New Wireless Technologies; <https://www.sciencedirect.com/topics/engineering/software-defined-radio>, October 2020
- [9] Learn the Fundamentals of Software-Defined Radio ; <https://www.digikey.ch/de/articles/learn-the-fundamentals-of-software-defined-radio>, October 2020
- [10] GNURadio; <https://www.gnuradio.org/>, October 2020
- [11] Great Scott Gadgets; <https://greatscottgadgets.com/hackrf/one/>, October 2020
- [12] Arduino; <https://www.arduino.cc/>, October 2020
- [13] D. J. Cichon: *The Heinrich Hertz wireless experiments at Karlsruhe in the view of modern communication*, September 1995, 7.
- [14] Probir K. Bondyopadhyay: *Gugilelmo Marconi: The father of long distance radio communication - An engineer's tribute*, September 1995, 7.
- [15] Mervyn C. Fry: *Radio's First Voice...Canadian!*, March 1973. https://www.ieee.ca/millennium/radio/radio_birth.html.

- [16] John S. Belrose: *Reginald Aubrey Fessenden and the Birth of Wireless Telephony*, April 2002.
- [17] James A. Hijiya: *Lee De Forest and the Fatherhood of Radio*, 1992.
- [18] S. Khrystyne Keane: *100 Years of Amateur Radio Licensing*, 2012.
- [19] Mark Goodman: *The Radio Act of 1927: Progressive Ideology, Epistemology, and Praxis*, 2000.
- [20] A. S. Margulies: *Software Defined Radios: A Technical Challenge and a Migration Strategy - Spread Spectrum Techniques and Applications*, 1998.
- [21] Raymond J. Lackey: *Speakeasy: The Military Software Radio*, 1995.
- [22] Lopa J. Vora: *EVOLUTION OF MOBILE GENERATION TECHNOLOGY: 1G TO 5G AND REVIEW OF UPCOMING WIRELESS TECHNOLOGY 5G*, 2015.
- [23] Netzwoche; <https://www.netzwoche.ch/news/2020-09-07/wi-fi-6-fragen-und-antworten-fuer-cios-und-netzwerkmanager>, October 2020
- [24] How To Geek; <https://www.howtogeek.com/368332/wi-fi-6-what%E2%80%99s-different-and-why-it-matters/>, October 2020
- [25] David Coleman: 802.11ax For Dummies; Book, Hoboken New Jersey, 2018, 23-24, <https://www.lantech.nl/wp-content/uploads/2019/02/E-book-802.11ax-For-Dummies%2ae-Aerohive-Special-Edition.pdf>
- [26] David Coleman: 802.11ax For Dummies; Book, Hoboken New Jersey, 2018, 11, <https://www.lantech.nl/wp-content/uploads/2019/02/E-book-802.11ax-For-Dummies%2ae-Aerohive-Special-Edition.pdf>
- [27] David Coleman: 802.11ax For Dummies; Book, Hoboken New Jersey, 2018, 12, <https://www.lantech.nl/wp-content/uploads/2019/02/E-book-802.11ax-For-Dummies%2ae-Aerohive-Special-Edition.pdf>
- [28] David Coleman: 802.11ax For Dummies; Book, Hoboken New Jersey, 2018, 16, <https://www.lantech.nl/wp-content/uploads/2019/02/E-book-802.11ax-For-Dummies%2ae-Aerohive-Special-Edition.pdf>
- [29] David Coleman: 802.11ax For Dummies; Book, Hoboken New Jersey, 2018, 18, <https://www.lantech.nl/wp-content/uploads/2019/02/E-book-802.11ax-For-Dummies%2ae-Aerohive-Special-Edition.pdf>
- [30] Arik Yavilevich: *Why use Fosphor, a GNU Radio real-time spectrum analyzer?*, August 2016. <https://blog.yavilevich.com/2016/08/why-use-fosphor-a-gnu-radio-real-time-spectrum-analyzer/>.
- [31] LUXORION: *The History of Amateur Radio*, July 2015. <http://www.astrosurf.com/luxorion/qsl-ham-history5.htm>.
- [32] BR Fernsehen: *Heimat der Rekorde*, Weltrekord im Empfangen einer Amateurfunk-Nachricht unter 9kHz, Februar 2020. <https://www.br.de/mediathek/video/heimat-der-rekorde-weltrekord-im-empfangen-einer-amateurfunk-nachricht-unter-9khz-av:5e17b21df59d22001af2cab7>.
- [33] Norsk Romsenter: *NORWAY'S SATELLITES*, June 2019. <https://www.romsenter.no/eng/Norway-in-Space/Norway-s-Satellites>.

- [34] EE Publishers: *Amateur radio geostationary satellite to support disaster relief communication*, August 2019. <https://www.ee.co.za/article/amateur-radio-geostationary-satellite-to-support-disaster-relief-communication.html>.
- [35] AMSAT-UK: *Radio Amateur Satellites*, Es'hail-2 Geostationary Satellite, November 2018. <https://amsat-uk.org/satellites/geo/eshail-2/>.
- [36] International Radio Amateur Union: *Emergency Communications*, January 2020. <https://www.iaru.org/on-the-air/emergency-communications/>.
- [37] Aargauerzeitung: *Ständerat gibt grünes Licht für höhere Militärausgaben*, June 2020. <https://www.aargauerzeitung.ch/schweiz/211-milliarden-franken-parlament-stellt-sich-hinter-hoehere-militaerausgaben-139252109>
- [38] Schweizer-Soldat: *Radio in the Swiss Military*, April 2019. <https://www.schweizer-soldat.ch/2019/04/hptm-sarah-brunner-f%C3%BCrcht-als-erste-frau-eine-inf-kp.html>.
- [39] Schiffsradar: *Schiffe weltweit in Echtzeit verfolgen*, December 2020. <https://www.schiffs-radar.de/schiffsradar.php>.
- [40] Flightradar: *Live Air Traffic*, December 2020. <https://www.schiffs-radar.de/schiffsradar.php>.
- [41] RTL-SDR: *RTL-SDR TUTORIAL: CHEAP AIS SHIP TRACKING*, April 2013. <https://www rtl-sdr.com/rtl-sdr-tutorial-cheap-ais-ship-tracking/>.
- [42] HidrateSpark *Smart Water Bottle*, June 2019. <https://hidratespark.com/>.
- [43] Bericht des Bundesrates: *Sicherheitsstandards für Internet-of-Things- Geräte (IoT)*, April 2020. <https://www.melani.admin.ch/dam/melani/de/dokumente/2020/29042020-Bericht-IoT-d.pdf.download.pdf/29042020-Bericht-IoT-d.pdf>
- [44] M. Himanshu, A. Harshit: *RSA Conference 2019*, RF Exploitation: IoT/OT Hacking with SDR, March 2019.
- [45] Awesome Open Source: *Rf Jam Replay*, Jam and Replay Attack on Vehicular Keyless Entry Systems, June 2019. <https://awesomeopensource.com/project/trishmapow/rf-jam-replay>
- [46] Hacker Warehouse: *Proxmark3 RDV4 Kit*, July 2019. <https://hackerwarehouse.com/product/proxmark3-rdv4-kit/>
- [47] DoTheBest: *RFID Products using NFC Applications*, June 2020. <https://www.dtbrfid.com/nfc-applications.html>
- [48] Luca Roselli: *Reverse Engineering and Security Evaluation of Commercial Tags for RFID-Based IoT Applications*, January 2017. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5298601/#sec2dot5-sensors-17-00028>
- [49] RTL-SDR; <https://www.rtl-sdr.com/defcon-23-some-software-defined-radio-related-talks/>, Otober 2020
- [50] Bliley Technologies; <https://blog.bliley.com/10-popular-software-defined-radios-sdr>, October 2020
- [51] RTL-SDR; <https://www.rtl-sdr.com/tag/kiwisdr/#:~:text=The%20KiwiSDR%20is%20a%20US,web%20browser%20over%20the%20internet.>, December 2020

- [52] Lime Microsystems; <https://limemicro.com/products/boards/limesdr/>, November 2020
- [53] Swling; <https://swling.com/blog/2019/11/the-new-malahit-dsp-a-portable-all-in-one-wideband-sdr-receiver/>, October 2020
- [54] Kenwood; <https://www.kenwood.de/comm/nexedge-dpmr/nexmobilfunk/NX-3720E/>, October 2020
- [55] Kenwood; <https://www.kenwood.de/comm/nexedge-dpmr/nexmobilfunk/NX-5700E/>, October 2020
- [56] Gierad Laput, Chouchang Yang, Robert Xiao, Alanson Sample, Chris Harrison: EM-Sense: Touch and Recognition of Uninstrumented Electrical and Electromechanical Objects; Paper, Pittsburgh, November 2018, 1, <https://la.disneyresearch.com/publication/emsense/>
- [57] Swisscom; <https://www.swisscom.ch/de/about.html>, October 2020
- [58] Swisscom; <https://www.swisscom.ch/de/res/hilfe/geraet/open-source.html?campID=shortcut-opensource>, October 2020
- [59] Asus; <https://www.asus.com/ch-de/Networking/RT-AX88U/>, October 2020
- [60] Amarisoft; <https://www.amarisoft.com/about-us/>, October 2020
- [61] GSMA Internet of Things <https://www.gsma.com/iot/narrow-band-internet-of-things-nb-iot/>, October 2020
- [62] Amarisoft; <https://www.amarisoft.com/products/network-deployment/>, October 2020
- [63] Amarisoft; <https://www.amarisoft.com/products/test-measurements/amari-lte-callbox/>, October 2020
- [64] Aselsan; <https://www.aselsan.com.tr/en/about-us/who-we-are>, October 2020
- [65] Aselsan; <https://www.aselsan.com.tr/en/capabilities/military-communication-systems/military-broadband-multimode-radiolinks/grc5220-tactical-broadband-ethernet-radio>, October 2020
- [66] Aselsan; <https://www.aselsan.com.tr/en/capabilities/military-communication-systems/vuhf-military-radios/prc-vuhf-sdr-handheld-radios>, October 2020

Chapter 3

The economics of vote buying

Domenic Luca Fuerer, Lennart Lou Jung, Sascha Deboni, Tony Ly

Abstract This paper focuses on vote buying and combines an overview of the real-world appearance of vote buying and its implications with technical countermeasures and the philosophical debate on vote buying. The first part focuses on the basic concepts of vote buying. Firstly, it describes the concept of ballot secrecy and provides a model explanation why people vote. Secondly, it presents various ways of how vote buying is done in practise. Furthermore, the section presents some incidents of vote buying. Additionally, a macroeconomic perspective examines the phenomenon with the extended liquidity during the voting month. The insight into the law regulation of Switzerland and the US will provide insight into the basic legal aspects. A small excursion about the e-voting discussion in Switzerland will conclude the first part about politics. The second introduces the technical countermeasures against vote buying in voting systems. With the goals of coercion-resistance and receipt-freeness at hand, it shows theoretical ideas, cryptographic and blockchain solutions to achieve those. In the last part, the paper examines the discussion about whether vote buying is ethically acceptable. Some philosophers who state that vote buying per se isn't necessarily problematic or should even be organized in a voting market are presented and encountered by their opponent scholars. Furthermore, the paper will show how the philosophers writing in favour of vote buying can be located in the epistemic field of libertarian thinking. Referring to critical theory, the paper will finally conclude that the premise of vote buying can't hold in realistic circumstances with differences in prosperity.

Contents

3.1	Introduction	.	.	.	63
3.2	Global overview of Vote Buying	.	.	.	63
3.2.1	Ballot Secrecy and Vote Buying	.	.	.	63
3.2.2	Vote Buying in the past	.	.	.	68
3.2.3	Circulation of money	.	.	.	76
3.2.4	Legal aspects of vote buying	.	.	.	79
3.2.5	Excursus: e-voting in Switzerland	.	.	.	81
3.3	Countermeasures of vote buying	.	.	.	82
3.3.1	Privacy and Verifiability	.	.	.	82
3.3.2	Receipt-freeness and Coercion resistance	.	.	.	83
3.3.3	The ThreeBallot voting protocol	.	.	.	83
3.3.4	The VAV protocol	.	.	.	84
3.3.5	Receipt-Free Voting Based on Homomorphic Encryption	.	.	.	84
3.3.6	The JCJ scheme	.	.	.	86
3.3.7	The Estonian scheme	.	.	.	87
3.3.8	Voting systems based on blockchain	.	.	.	87
3.4	Ethics of vote buying	.	.	.	88
3.4.1	General Legitimisation of Vote Buying	.	.	.	89
3.4.2	Theoretical Counterarguments	.	.	.	91
3.4.3	Tradition of thought: When Philosophy becomes political	.	.	.	94
3.4.4	Is there a real democracy?	.	.	.	96
3.5	Conclusion	.	.	.	96

3.1 Introduction

Vote buying is an existing phenomenon. In the public debate, it is considered as a danger to democratic systems since monetary transfers, coercion and corruption seems to be involved. The growing demand on democratic participation and the development of new technologies pushes the debate as well as scientific research in the fields of social and political science as well as computer science where cryptographic tools are developed. New techniques should enable modern societies to participate with their electronic devices without making any compromise or even with improving security and control. This paper aims to give an overview of vote buying by firstly examining some theory from political scientist research, applying them to incidents on vote buying and reflect the frame modern democracies provide with their laws. Also, technical developments shall be addressed. Because of the demand in electronic voting systems, cryptographic theories and voting protocols have been researched a lot in the last three decade. Particularly systems which are robust and verifiable and not vulnerable to attacks. Forming the basis of voting systems, further work have been done to mitigate vote buying by achieving receipt-free and coercion resistant means. Finally, the premise of the debate considering vote buying as bad per se should be questioned. An insight to the philosophical debate presents the most important philosophers promoting vote buying and encounters them with their opponents. What is vote buying? Where can it be observed? And how can it be encountered? Should it even be encountered or are there reasons to defend vote buying? These questions shall be addressed on the following pages.

3.2 Global overview of Vote Buying

Whenever people have the possibility to vote, the phenomenon of vote buying seems to occur as well. This was observed in the past and is still present in modern times. The goal of this section is to give the reader a broad overview of the topic. Firstly, the definitions of ballot secrecy and vote buying are presented. This information builds the fundament for the reading of the later sections. Afterwards, some examples of the past and the present are named in order to affirm the real presence of this phenomenon which is not just of theoretical nature. The focus on the economic measurements of increased liquidity in the market during voting months and manipulations of the monetary and fiscal policies shall then be discussed as evidence for real-world vote buying. As there seems to be a probable connection to vote buying. Subsequent, the law concerning the topic of vote buying in the United States and Switzerland are discussed and compared with each other. To complete this part, there will be an excursus to the intensively discussed topic of E-Voting in Switzerland.

3.2.1 Ballot Secrecy and Vote Buying

This first section presents the fundamental concepts concerning vote buying. The goal is to give an overview about the most relevant topics by providing some key definitions and aspects which could influence a voter's choice. This is achieved by first introducing the concept of the secret ballot. What is followed by a theoretical concept which tries to explain why people even vote. With these two first concepts, it is possible to get a closer look at the topic of vote buying itself in which two different perspectives are shown and the strategies and problems of vote buying are highlighted.

3.2.1.1 Secret Ballot

The concept of the secret ballot is essential for vote buying. A secret ballot means that the voter's choice is anonymous. A first form was already introduced by the Romans in the second century and had the goal of decreasing the power of the upper class and improve the voter's freedom of choice [103]. Another hope people had by introducing ballot secrecy was that bribery, intimidation and coercion should be reduced and with it, potential vote buying should be prevented. Nowadays secret ballot is one key instrument in fair elections of modern democracies and even the United Nations mention that secret ballot is a crucial component in elections. It is important to remark that ballot secrecy is affecting a voter's choice as one can be sure that what has been voted will have no consequences.[39]

To deeply understand the concept of ballot secrecy, one can distinguish two different forms of it. The first form describes that the ballot is psychologically secret if the voter actually believes that the election is held in such a way that her ballot choices are secret. If people do not believe that their choice is held secret, for example if they are not well enough informed about the election process or if they simply do not trust the election institutions, the potential benefits of secret ballot can be lost. Secondly, social factors are important for the social secrecy of the ballot. People may feel a need to discuss their votes with each other as elections can often be a widely discussed topic over a long period of time. These discussions may influence a voter's opinion as one may fear to reveal his true opinion in such a discussion. People could, just to avoid conflict, change their mind and thus vote for something they did not initially want and with this, lose the concept of secret ballot again.[39]

These two forms are both important for the secret ballot in order to exist and work as intended. The secret ballot is a first early countermeasure against possible vote buying. Furthermore, it still comes into action in nowadays democracies as later the discussion about the current legal situation will show.

3.2.1.2 Why do people vote

To fully understand the concept of vote buying, it is important to not only know the concept of secret ballot, but also to understand why people vote. To discuss the nature of voting in a scientific way, the calculus of voting model is used [49]. This model follows an economic approach and suggests that people will only vote when the expected benefits outweigh the costs [32] [90]. To explain a social process such as voting by using economical terms has implications on the way one thinks about this phenomenon. This will be discussed in the section about philosophical considerations. This The motivation to vote can be described with the following formula.

$$pB > C$$

Where p is the probability of being a pivotal voter, B is the benefit the voter gains if his favourite candidate wins, and C describes the cost to undertake the effort to vote. Considering this formula, the act of voting seems to be unworthy as the chances of being a pivotal voter are almost zero. The larger the size of an election gets the more dramatical this effect becomes. This means that the presumed benefit of voting is almost nonexistent and rational individuals would neglect voting. Thus, an extension of the basic calculus model tries to explain why people vote nevertheless. Various studies have extended the model with an additional parameter D. For the American economist Downs, this parameter signifies the additional value, a continuing democracy can bring to any voter [32]. Another point of view is made by the political scientists Riker and Ordeshook, who see in D the duty of each citizen to vote [73]. Yet another study argues that social pressure

on a voter can be a significant part of D. This hypothesis is supported by experimental evidence, made in the United States, that social pressure increases the probability that a voter actually will vote. [38]

But there are also some other studies which had a look at empirical data to make some statement about turnout. Thus, they take the view of a more general case and not of a single individual. In this case, the turnout describes the percentage of voters who took part in the election. One study conducted by Baek in 2009 showed that the turnout tends to be higher, if the costs of gathering information are lower [5]. Another study of the American political scientist Nicter, investigating on voter behaviour in Argentina, shows that the voter turnout tends to be higher when voters expect material rewards [64].

To sum up, it is crucial trying to understand why people vote, as it affects the topic of vote buying. As shown, D can stand for many different objects and is not limited to only the described ones above. Thus, D can possibly be generated through other things, as in Argentina the material rewards, which also a vote buyer could provide. If vote buyers are aware of these theories, they know how they can access people and can develop new strategies to buy votes more effectively. While governments with the equivalent knowledge can better try to prevent it. Finally, it is about setting the best incentives to get motivated voters. [49]

3.2.1.3 Vote Buying

After some premises of voting were explained, this section should deepen the concept of vote buying. This is done by concentrating on some key categories. First of all, vote buying is explained out of different views, followed by the strategic advantage vote buying can bring. Afterwards, the problems with vote buying will be discussed.

As mentioned before, the concept of vote buying can be investigated with the use of different views, more precisely it can be divided in two groups. The first group are the candidates, their campaign team, politicians who want to pass some proposal or even very motivated persons who strongly believe in the superiority of one side. All these can also be named as givers or buyers as they buy votes. These are the people who want others to vote of their own preference and offer some goods for it. The second group is the group of the voters who are also called recipients or sellers. This group reflects the people who possibly sell their vote and would receive a specific good. It is important to make this distinction to fully understand the processes around vote buying from both sides. [78]

Givers

A giver is someone who wants to buy a vote. This is often a candidate himself or herself or an agent working for some candidate. In the view of these people, vote buying is some kind of a contract. Thus, the purchasing is only a purely economic exchange of a vote for a specific good. For a giver, there are three different ways to try to change the voting behaviour of a voter. The first way is described as instrumental compliance. This means, that the recipient changes or does not change his electoral behaviour due to a tangible reward. In this context, not changing the behaviour does mean that the voter already wanted to vote in the givers favour and the reward even tightened this opinion. Secondly, a giver can generate normative compliance. In this case, the voter does change or does not change the behaviour because the offer actually convinces him of the worthiness or goodness of a specific candidate. Lastly, a giver can try to generate coercive compliance. Coercive compliance means that the giver bullies recipients into changing or not changing their behaviour. [78]

Even though there is the concept of secret ballot in theory, there are various strategies available to generate the named forms of compliances above. One can either try to generate only one form of compliances, but it is also possible to combine them. It is important

for givers to check whether a voter actually voted as promised or whether he changed his mind again and only accepted the offer. A first strategy, and probably the easiest one, is to monitor each individual and its vote. This can be done with the help of the election officials, who then have to count how the voters filled out their ballots [40]. If no direct observation is possible, the giver has also other possible strategies to conduct vote buying. A giver approaching a recipient can deliver the voter with a carbon paper. The voter fills in the ballot on top of this carbon paper with the effect that the vote is recorded and still visible. A giver can later check the carbon paper to verify it [78]. The following strategy is more elaborate. It requires that the giver possesses a stolen or fake ballot. This ballot is then prefilled by the giver and handed out to the first voter. The voter enters the polling station and casts the prefilled ballot. The real ballot which is received is not filled in but is taken outside to another voter waiting to enter the polling station. Afterwards this voter fills in the form to the givers satisfaction and enters the polling station, casts the prefilled ballot, and takes with him or her the blank one while leaving. The whole process can be repeated several times. This strategy is also called the Tasmanian dodge [77].

To create instrumental or coercive compliance it is also possible to monitor the outcome of the vote of a whole city. This strategy can be helpful if the givers offer goods, especially in material form, to whole cities. This needs less effort and is done easily for the givers, what means that with this strategy they can survey large parts of a country [19]. A strategy which generates instrumental compliance but follows a completely different way is to pay voters to stay away from the polling stations at all. This is done in order to prohibit any votes for the opponent, which eventually leads to the same result as collecting votes for the preferred candidate. The strategy is also called negative vote [78]. Other strategies to create instrumental compliance, make use of the condition that the proper candidate wins. One of these strategies is to split the tangible reward. This means that one half of the reward is given to the voter before the election, but the other half is only distributed to the voters if the favoured candidate has actually won. In this scenario, the rewards are selected in such a way that half the product is not useful. An example would be half a banknote [78].

These strategies exist not only for instrumental compliance, but also for normative compliance. As mentioned above generating normative compliance means to convince the voter of the worthiness or goodness of the candidate. Hence, one strategy to create normative compliance is giving some small election gifts to voters or paying recipients wages for doing some ordinary election campaign as hanging posters or delivering messages to others even though they are not in the official campaign team [78]. A further possible strategy is trying to instil a personal obligation to vote for a particular candidate. One way to create this obligation is to recruit supporters. These newly recruited people should match some specific characteristics like being popular in a community or have a large influence on many others. Therefore, many other voters will listen to these new givers and will feel obliged to respect the vote promise they made. It is not fully clear whether this strategy is considered as vote buying or just as normal conviction. But the reason of completeness, the strategy has been shown. [102]

It needs to be mentioned that the possible strategies above do not define all existing possibilities. There are other strategies and some of them may not even be known at the moment. Another point to mention is that these strategies only might be effective. In some cases, it can also lead to the opposite, meaning that a person who originally wanted to vote for the candidate gets turned away by an immoral offer [78]. Moreover, new technologies could bring new strategies. Especially e-voting from home via the internet could create new risks.

Voters

The perspective of voters can differ heavily from the perspective of givers involved. A

voter can interpret an offer of a giver completely different than a giver. This means, that the presented good, which the giver understood as a binding gift to vote for a specified candidate, can be understood as a non-binding gift by the recipient. [78]

To better understand the different meanings an offer can have, one first needs to have a look at the different underlying goods and what meaning these can carry. First of all, the giver can pay the recipient as it would be a usual relation. This type of offer is impersonal and can lead to a recipient accepting the payment, but later on not voting as agreed in the relation. Another possibility is a gift or a favour. In contrast to a payment, this kind of offer is more personal. Thus, it could be that this offer produces feelings of obligation or gratitude. Additionally, there is a wage. This is earned for services which the voter has done in order to advertise a candidate. A wage then can produce feelings of gratitude or obligation to the candidate just as with a gift or favour. [78]

Apart from the more obvious meaning, there can be additional subtle meanings along to the others. For example, there could be a threat alongside the good. The voter has no choice between accepting or declining the offer, as declining could lead to retaliation. Another signal can be evidence of winnability. This means that the offer shows the voter that the candidate has powerful forces and is able to win. A more negative meaning can be affront, what means that accepting the offer could damage the self-respect of the voter and would make more voters to decline it. For recipients, accepting or rejecting offers, respectively changing or sticking to their opinion, can have different causes. Some reasons can come with the offer itself as described above as fear or gratitude. But others may be duty, indignity, righteousness, or self-interest. The meaning carried by an offer has an impact on how effectively a giver can affect the electoral behaviour of a recipient. What means, that a voter who sees the offer in a positive way is more likely to accept it and will also vote for the candidate. On the other hand, voters who feel offended or threatened by the offer will get pushed away and probably not accept it. [78]

As discussed in the previous section, why people vote, it has been shown that there is an additional D which makes it rational for people to vote. It can be observed that people who are contacted by a giver are 15 percentage points more likely to vote than the others. What means that somehow their benefit has increased. An analysis of this states that through approaching, normative compliance is generated. This means that the approached voters are more likely to believe in the credibility of politician campaign promises. What could mean that vote buying actually transfers credibility to potential voters and shows them that this candidate is actually changing something. One final remark needs to be underlined on this point. Real world evidence shows that out of the approached people, especially the uneducated, are more likely to be influenced than the more educated groups.[49]

3.2.1.4 Strategies of vote buying

There are different points of interest depending on the strategies which are applied in order to try to buy votes. A key question such strategies should answer is, who should be targeted when one tries to conduct vote buying. One model describes that vote buyers should target the core supporters. These are the people, the buyers are already familiar with and of whom information already exist. This information is a strategic advantage what means that attempting these people is less effort than foreign people, but could lead to less success as many of the supporters would have voted for the favoured candidate anyway [26]. Another model, presented by the political scientist Stokes, says that vote buying will only occur in the context of machine politics. In this context, machine politics describes a political organization in which a small group or only one person with authority who is popular enough to have control over political administration. He reasons that people could accept a bribe or gift but then still not vote as promised. For this reason,

vote buying would only happen if the political machine had enough capabilities to monitor all the individuals from whom the votes were bought and ensure their compliance [85]. A last strategy which should be discussed is presented by Nichter [64]. Here, vote buying is targeting party supporters which were not voting previously. What would mean that parties try to buy turnout from these people by mobilizing them with the offer.

Of course, there still exist some other theories about who should be targeted and a final conclusion has not been found yet. But these three build a first explanation trying to explain why specific groups of citizens are more likely to be targeted. [49]

3.2.1.5 Problems of vote buying

There are many potential problems of vote buying. Often, there are three different main reasons presented why vote buying is problematic and thus forbidden. These three are equality, efficiency, and inalienability which will be discussed in the following section in more detail. Some of these arguments are also connected to the philosophical debate. In this part, the counterarguments shall be embedded in real-world economic, social and political reasoning.

The first argument often mentioned is the argument of the equality. The poor are considered to be more likely to sell their vote in contrast to the rich people. This is explained by the fact that a dollar is of bigger value to a poor than to a rich person. This would mean that the political equality is not given anymore. In elections in which the wealthy buy the votes of the poor, the views of the wealthy are more supported than they would be without vote buying. What would be bad as in democracies, the poor and the wealthy should have equal influence on political outcomes. [41]

The second argument is made on the basis of an efficiency analysis and asks whether the prohibition of vote buying increases or decreases social wealth. To fully understand this argument, one has to shortly focus on the exchange of the vote. A person A buys the vote of selling person B. This would mean that both of them economically profit from it. While this looks reasonable at first glance, it turns out that this exchange does not necessarily produce efficient results. There is always the risk that negative externalities could be created which could affect third parties later on. This would mean that in the view of efficiency, vote buying might lead to a decreased social wealth and is therefore not beneficial. [41]

The last main argument against vote buying is the inalienability argument. This argument is based on the moral judgement that it should not be possible to sell or buy votes. This argument is based on the moral ideas in democracies but cannot be proven in a social scientific way. It will be further discussed when some philosophical objections are examined in the last section. [41]

To sum up, these problems are the most commonly named arguments why vote buying should be prohibited. They are valid in many of the modern democracies. As announced, more reasoning about why vote buying is intrinsically good or bad is made in the last part of this report, where the philosophical arguments and problems are presented and discussed.

3.2.2 Vote Buying in the past

Vote Buying is a part of electoral fraud and its used quiet often. Not just on small elections, but also on big, important ones like the presidential election in the United States of America (USA) in 2000. This paper shows a collection of different examples of vote buying. The examples will show different ways of vote buying, which are considered closely. The first two examples are examples which negotiated in front of the law and has classic behavior. Then the paper takes a closer look on the history of vote buying

with examples from the distant past. The presidential election in Kenya in 2002 is the next point made. This example is not reported from the law side, it is reported from the scientific side and therefore, is more analytical. After Kenya, this paper focusses on the USA. First on the presidential campaign of 2000, and second on the presidential campaign in 2016

3.2.2.1 The sheriff who betrayed democracy

A perfect example for a classic act of vote buying is former Dodge County Sheriff Michael Lawton Douglas Junior. He and his former colleague Norman Gibson bought votes in the election for the Dodge County sheriff in 2004 in the southern district of Georgia in the USA. Gibson helped Douglas to buy votes in the community. He has bribed hundreds of voters with the money of Douglas. They should vote for Douglas in exchange for money. After he won the election, he successfully served for his country for four years until 2008 he was defeated in the re-election process. In 2010 Douglas was sentenced to 18 months in prison to be followed by a three-year term of supervised release. In addition he had to pay a fine and had to work for the community service. Also, his colleague and helper Gibson were sentenced to four months in prison, a three-year term of supervised release and a fine as well as to work for the community. [95] [42]

3.2.2.2 Cigarettes for votes

In the USA, a total of 68 cases of vote buying have been successfully negotiated in court in the last few years since 1968 [42]. The newest case is about Richard Howard and his three friends (Louis Wise, Christopher Williams and Nickey Huntley). They payed hundreds of homeless people \$1 and cigarettes to sign ballot petitions and voter registration forms. The prosecutors called this action as an assault on the democracy of the USA. Therefore, Howard has received a three-year suspended sentence. [81]

3.2.2.3 History of vote buying

In this chapter should be shown that not only nowadays vote buying is a phenomenon. It also was long time ago. To show the development of vote buying, the paper takes a closer look at three examples from different time episodes.

3.2.2.4 History of vote buying

In this chapter should be shown that not only nowadays vote buying is a phenomenon. It also was long time ago. To show the development of vote buying, the paper takes a closer look at three examples from different time episodes.

Alexander VI - 1492

This example is a classical vote buying one. It is the first case, which is well documented of vote buying. Alexander VI's real name was Roderic Llançol I de Borja, he lived from 1431 to 1503 and was Pope from 11. August 1492 until his death in 1503. He was elected after the death of Pope Innocent VIII. There were three major candidates to win the election, the sixty-one-year-old Roderic Llançol I de Borja (seen as the independent candidate), Ascanio Sforza (for the Milanese) and Giuliana della Rovere (seen as a pro-French candidate). During the election, there were rumors that Roderic Llançol I de Borja succeeded in buying votes. Sforza was bribed with four mule-loads of silver. After the election (Roderic Llançol I de Borja won and was then called Alexander VI. Alexander VI started to distribute the benefits after the election. Since this act, it was

known that Alexander VI won the election by buying votes. He could stay as pope since it was normal at this time and his competitors also bribed cardinals to get their votes [27].

So, it is shown that vote buying has a long history of success and unsuccess. Alexander VI was the first well documented case, but he could stay in his function as a pope.

George Washington 1756

When the father of the USA [62] - George Washington was at age 24, he bought votes in exchange for alcohol. Daniel Okrent wrote in his Book Last Call [66]: The Rise and Fall of Prohibition: "When twenty-four-year-old George Washington first ran for a seat in the Virginia House of Burgesses, he attributed his defeat to his failure to provide enough alcohol for the voters. When he tried again two years later, Washington floated into office partly on the 144 gallons of rum, punch, hard cider and beer his election agent handed out - roughly half a gallon for every vote he received. The practice was at this time widespread and accepted but technically it was illegal [31]. Even the father of the USA bended the law and bought votes.

George Washington Plunkitt - around 1900

"What tells in holdin' your grip on your district is to go right down among the poor families and help them in the different ways they need help. I've got a regular system for this. If there is a fire in Ninth, Tenth, or Eleventh Avenue, for example, any hour of the day or night, I'm usually there with some of my election district captains as soon as the fire engines. If a family is burned out I don't ask whether they are Republicans or Democrats [...] I just get quarters for them, buy clothes for them if their clothes were burned up, and fix them up till they get things runnin' again. It's philanthropy, but it's politics, too - mighty good politics. Who can tell how many votes one of these fires brings me? The poor are the most grateful people in the world, and, let me tell you, they have more friends in their neighborhoods than the rich have in theirs." [74] This is a quote from George Washington Plunkitt. He was state senator and party leader of the democrats. It was first written down by William L. Riordan in 1995 in his famous book: "Plunkitt of Tammany Hall: A Series of Very Plain Talks on Very Practical Politics". The quote was recorded around 1900 and it shows that already then, the law was bent and back then happened controversial forms of vote buying. Mr. George Washington Plunkitt used a clever way to buy himself votes. Also, in this case could be argued that he just gave his social help, but as he says by himself, he knows that the poor are so grateful that they will vote the next time for him. He got this gratefulness by buying stuff for them and it is just a part of his will to help the people. The bigger and more important part for him is the egocentric one. He wants the votes of the poor in exchange for his financial help. If he just wanted philanthropy, then he could give the money to charity as well. [87]

3.2.2.5 The presidential election in Kenya 2002

Kenya's 2002 presidential and parliamentary election was the third election since the transition to a multiparty politics. It was the first peacefully turnover of executive power since the described transition. Eric Cameron a Ph. D. Candidate in the Department of Political Science at the University of California in Los Angeles took in his paper in 2009 a closer look on this election. He investigated in his paper about vote-buying and its effect on the turnout. [49]

In 2005 a data collection started to have a closer look on this subject. The data collected has a sample size of 1'120 people. The effective turnout was 57% for this election, while

the turnout calculated with the sample size was 63%. Voter turnout is therefore a little bit higher in the data collection than in the real one. The statistic, which is very important for this paper, is showed by the following graphic:

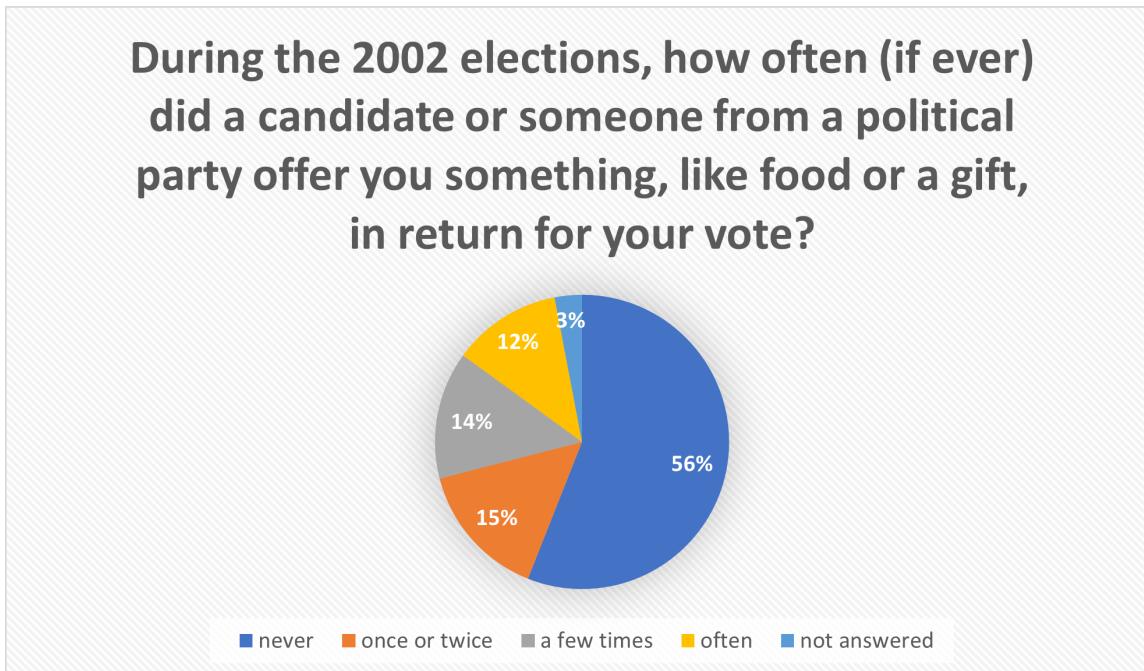


Figure 3.1: How often a voter was asked to sell their votes in the 2002 presidential election in Kenya

It is remarkable that around 40 percent of the interviewed persons was asked by a political party to sell their votes. Also remarkable is that around 12 percent were offered multiple times (often) such bribes. The election was never contested, and the winner could serve for Kenya. Some of the people, whom were offered a bribe, talked in interviews about their experiences[49]: "A NARC [national Rainbow Coalition-Kenya, a political democratic party] agent stopped me at a bus stop and asked me who I was voting for. When I said KANU [Kenya African National Union, a conservative national party], he offered me 500 shillings [about \$6] for my vote ". Or another example [49]: "A man approached me in Naivasha at a bar and asked me what party I'm from. He said he is an agent for KANU and would buy my vote for 700 shillings [about \$8] ". It is proved that candidates in the two weeks leading to the election spend about 60'000 to 80'000 shillings (about \$720 to \$920) per day on vote buying. [101]

The parties had a strategy in distributing the money. Their strategy was to search for poor men that are willing to vote for a minor political party. The idea behind this strategy is: the poor are willing to take their money in exchange for their vote. If their party (a minor party) has no chance to win, then it does not matter if they get a vote or not. They will lose in both scenarios. And the last point in the strategy is that in Kenya the men are often the determining power in the family - so the rest of the family will vote the same as its family leader.

In retrospect it was shown that the strategy was good. The biggest influence if a person let vote-buying impact its decision to go to vote is education. Without formal schooling and after an offer for vote-buying the persons are willing to go voting with a higher probability of up to 25%. When only visiting the Primary Education of Kenya, then the probability to vote is up to 20% higher in comparison to when the person is not offered a bribe. The better the person is educated, then it is more likely that the person is not going to be influenced if he is going to take part at the election day. Cameron thinks this is because the more educated people learn in school how important an election without monetary influence is. Less education and poorness are often highly correlated to each other, the

strategy of the buyers was therefore good. The calculation of the statistics shows that a vote-buying attempt increases the probability of voting by about 14%.

It also was shown that not only less educated persons can be target of such bribes. Also, people who do not think that democracy is the best form of government are strongly more willing to vote if they get money or gifts in exchange. Also shown was that a lot of persons that would vote for a minor party are not going to vote. Because their vote does not matter. Therefore, they are such a good target for buyers. On this perspective the best strategy to buy votes would be poor educated, against the current system and elector of minor parties. [101] [53] [49]

3.2.2.6 Voteauction

Another approach of Vote Buying is that the initial contact is not coming from buyer, in this case the seller of the vote actively searches for a value of his vote.

A user of eBay tried to sell his vote on the platform. Five more people followed this user instantly. Because to sell a voice on eBay is problematic – eBay deleted this offer after a day. [30]. Before deleting, the auction reached \$10'100 and found 200 takers in a single day. A student (James Baumgartner) saw the need of the users there and developed a platform called Voteauction. This platform used real-time technologies and gave its users the possibility to sell their votes in the Presidential Election 2000, A. A. (Al) Gore vs. G. W. Bush. All users who took their vote up to the platform were separated by the constituency. The buyers of the votes then could bid on these pools of votes. So, in the end of the auction there was one buyer per state and the voters just knew for whom they must vote – but they did not know who bought their vote, because the buyers got an anonymous number. Compare for this the table below. [68]

J. Baumgartner got immediate ration to his Voteauction platform. The developer of the website once said: "I got 80'000 hits on Thursday and Friday alone. I think that that along with what was going on with eBay and Yahoo auctions shows that this is something people are really concerned about: if the politicians are selling their votes - and they clearly are - then the people should be allowed to as well ". [68]

After the first attention Baumgartner handed the webpage over to the art group Ubermorgen.com. Before the website was taken down, the price ranges of the votes were between \$2.20 and \$157.03 (compare table below). Eight states (Missouri, Wisconsin, Chicago, Arizona, Nevada, California, Massachusetts, New York) took law enforcements against the website. [92] As an example, the election manager in Chicago took a legal attempt to take down the website. The judges in Illinois gave him right - first Ubermogen.com just deleted the data from the users of Illinois however they did not take down the website. [83] Later, they had to take down the website. After taking down the website, Ubermorgen.com just bought a new domain in Switzerland. This domain only lasted for a few days. Then the Swiss domain registrar cancelled the domain. They cancelled it without a court injunction ordering it. They took it down after some telephone and E-Mail discussion between them and the Chicago Board of Election and Commissioners. They could do this because it was written down in their terms and conditions.[6] So, the website was highly controversial and discussed in the media. Over 2'500 global and national news features in online media, print, television, and radio have been reported. [92] It was clear that Vote Buying in this form is forbidden. But the art portal said that they did not have done something illegally. Their argumentation was that this only was a form of satire. This kind of art is protected under the first amendment ("The First Amendment provides that Congress make no law respecting an establishment of religion or prohibiting its free exercise. It protects freedom of speech, the press, assembly, and the right to petition the Government for a redress of grievances. "[100]). The problem with this argumentation was that the website really worked and users could register - another problem were

statements from Ubermorgen.com like this one [92]: "the only platform in the world that provide the final consumer an effective role in the American election industry. A true interchange system that finally brings capitalism and democracy closer together "or their headline on the platform [92]: "bringing democracy and capitalism closer together." The portal contradicts itself with statements like this. The portal was hoping that they could win the argument in front of the law, because there is a precedent case from Maryland, where a voter offered his vote for sale. He also commented his act as a political prank. He tried with his action to prank the fact that the right of every citizen is to use their vote in the manner of their choosing. His argumentation was followed by the law and he was not prosecuted.[63] As already mentioned, the example discussed here in the paper, the website vote-auction was shut down at 7. November 2000 after an injunction ordering from the government. [6]

Ubermorgen.com collected the statistical data about the website and published the data after their shutdown. So, the data was collected from the beginning of the 2000 election campaign until the 7th of November 2000. The data is shown in the table below. The data is sorted in descending order, beginning with the highest price per vote. Next to the data of Ubermorgen.com are the median incomes per state displayed. This numbers are also from the year 2000 as well as the data from Ubermorgen.com. The data about the median income is from the US Department of Commerce. This data is also ranked. It is ranked in the last column beginning with the wealthiest. [92] [29] The last column has also some additional information. Then the last column is compared with the first column. So, the rank of the price per votes and the rank of the income are compared. Five groups (1-10, 11-20, 21-30, 31-40, 41-51) were formed under all lines. When the number of the income rank is in the same group as the price per vote rank, then the cell appears in green. This is to investigate a possible correlation between income and price of a vote.

Table 3.1: Price per vote compared to average media income in 2000

Rank	Price per Vote	State	Total votes to buy	Highest bid	average income (1999–2000)	Median Income Rank
1	\$ 157.03	New Jersey	519	\$ 81'500	\$ 51'320	4
2	\$ 114.68	West Virginia	109	\$ 12'500	\$ 29'737	51
3	\$ 78.16	Washington	435	\$ 34'000	\$ 44'598	18
4	\$ 62.50	Ohio	386	\$ 23'000	\$ 42'421	22
5	\$ 52.54	Arkansas	295	\$ 15'500	\$ 30'527	50
6	\$ 46.95	Indiana	426	\$ 20'000	\$ 41'010	27
7	\$ 45.23	Illinois	995	\$ 45'000	\$ 47'193	10
8	\$ 37.56	Oklahoma	772	\$ 29'000	\$ 33'235	46
9	\$ 35.60	Utah	371	\$ 13'500	\$ 46'436	13
10	\$ 36.39	Arizona	293	\$ 10'500	\$ 39'911	31
11	\$ 35.84	North Carolina	309	\$ 11'000	\$ 38'712	35
12	\$ 35.84	Tennessee	397	\$ 13'500	\$ 35'824	42
13	\$ 33.42	Oregon	389	\$ 13'000	\$ 42'260	23
14	\$ 29.58	Vermont	710	\$ 21'000	\$ 40'589	29
15	\$ 28.97	Virginia	1208	\$ 35'000	\$ 48'678	8
16	\$ 26.05	New Mexico	902	\$ 23'500	\$ 34'410	45
17	\$ 24.83	Connecticut	302	\$ 7'500	\$ 51'432	3
18	\$ 24.23	California	5077	\$ 123'000	\$ 46'008	16
19	\$ 18.96	New York	4299	\$ 81'500	\$ 41'504	26
20	\$ 18.93	Wisconsin	819	\$ 15'500	\$ 46'357	14

Table 3.1: Price per vote compared to average media income in 2000

Rank	Price per Vote	State	Total votes to buy	Highest bid	average income (1999–2000)	Median Income Rank
21	\$ 17.41	Louisiana	402	\$ 7'000	\$ 32'006	49
22	\$ 17.27	Alabama	550	\$ 9'500	\$ 35'267	43
23	\$ 16.50	Hawaii	394	\$ 6'500	\$ 46'945	12
24	\$ 16.41	Georgia	457	\$ 7'500	\$ 41'822	24
25	\$ 16.00	Kansas	500	\$ 8'000	\$ 38'220	36
26	\$ 15.31	Missouri	784	\$ 12'000	\$ 45'160	17
27	\$ 15.12	Texas	2977	\$ 45'000	\$ 40'065	30
28	\$ 13.48	District of Columbia	408	\$ 5'500	\$ 39'369	32
29	\$ 13.43	Massachusetts	1117	\$ 15'000	\$ 46'312	15
30	\$ 12.84	Florida	6231	\$ 80'000	\$ 37'540	37
31	\$ 11.14	Iowa	853	\$ 9'500	\$ 42'808	21
32	\$ 9.98	Alaska	601	\$ 6'000	\$ 51'993	2
33	\$ 9.95	Delaware	553	\$ 5'500	\$ 49'283	6
34	\$ 9.80	Idaho	765	\$ 7'500	\$ 37'287	39
35	\$ 9.80	Kentucky	612	\$ 6'000	\$ 36'113	41
36	\$ 9.43	Pennsylvania	2015	\$ 19'000	\$ 41'507	25
37	\$ 9.43	Rhode Island	583	\$ 5'500	\$ 43'676	20
38	\$ 8.76	Colorado	970	\$ 8'500	\$ 49'238	7
39	\$ 7.44	Minnesota	1209	\$ 9'000	\$ 49'846	5
40	\$ 6.64	Nebraska	753	\$ 5'000	\$ 39'332	33
41	\$ 6.59	Michigan	4933	\$ 32'500	\$ 46'986	11
42	\$ 6.27	South Dakota	319	\$ 2'000	\$ 36'681	40
43	\$ 6.04	Maryland	2566	\$ 15'500	\$ 52'881	1
44	\$ 5.62	South Carolina	1067	\$ 6'000	\$ 37'455	38
45	\$ 5.57	Mississippi	988	\$ 5'500	\$ 32'581	47
46	\$ 5.02	North Dakota	697	\$ 3'500	\$ 34'665	44
47	\$ 4.77	Wyoming	524	\$ 3'500	\$ 38'839	34
48	\$ 4.58	New Hampshire	982	\$ 4'500	\$ 48'323	9
49	\$ 3.59	Maine	835	\$ 3'000	\$ 40'918	28
50	\$ 3.24	Montana	618	\$ 2'000	\$ 32'169	48
51	\$ 2.20	Nevada	1589	\$ 3'500	\$ 43'918	19
\$ 24.37			56865	\$ 988'000	\$ 41'595	—

In the first four groups are only two matches per group. In the last group there are three matches. In neither of the groups are more than 30% a matches. Therefore, there is no real correlation between price per vote and income. Quite the opposite it is - the wealthiest state was in 2000 Maryland but their price per vote is only the 43rd. The same happens in the opposite site. In the first group, the group of the most expensive votes are represented with two of the poorest states (regarding to the Median Income of 2000). West Virginia which has the second highest price per vote is regarding to the income the poorest state in the USA and Arkansas which after all has the 5th most expensive price per votes is the second poorest state. The paper cannot find a correlation between Median income and price per vote. It is to say that there also cannot be found a correlation when the groups are not compared equally but when the first group of prices per vote is compared

to the last group of median income. Also, there is the correlation under 30%. It seems more than when they both are compared the income is distributed equally.

3.2.2.7 Discussion about Cambridge Analytica and how they proceeded

A part of the past is the case of 2016 presidential elections of Donald Trump. He won the election in 2016 although he had less votes than Hillary Clinton. He won with a percentage of 45.9%, while Clinton won 48.0% of the votes. [21] He won among other things because of Cambridge Analytica. Cambridge Analytica was a mix of a data science and marketing company. They were specialized to win elections. They understood how to analyze user data, create personality profiles from it and so influence election campaigns. They influenced with these skills a lot of election campaigns. In this chapter it should be discussed if it was a danger to democracy and if it is a kind of vote buying. [58]

The procedure in the elections which Cambridge Analytica was involved was always similar. In the case of the Trump election Cambridge Analytica collected an enormous amount of Facebook profiles. They got a data on diverse ways. The biggest amount of data, around 50 million of profiles, have been received from a Cambridge Professor. He did a survey on the social media platform and paid 320'000 users to take part on this survey. Only in the small print it was mentioned that these people sold not only their Facebook data, but also all those of their friends. Their friends did not know that their data reached this Cambridge Professor. Also, they were not informed about this circumstance. They not only collected Facebook data, but they also collected datasets from the physical world (e.g. magazine subscriptions). When Cambridge Analytica had enough data, then they matched the data to the physical address of the people, as well as to their e-mail address. Then they match it all together into personality profile. With this personality profile they knew who was voting for whom. [58] Then, the next stage of their plan began. They knew which state is going to be a swing state, they even knew which constituency was highly competitive. Because they had such a lot of data about the persons, which lived in one of this constituencies. They targeted them with personalized ads. They targeted persuadable voters as their main target. E. g. Cambridge Analytica could target people high in neuroticism and send them ads on Facebook with images of immigrants swamping the country. They only could do this because they knew exactly by what the persuadable voter is triggered. The clue about the targets was that they not only targeted swing voters. They also targeted democratic voters and influenced them to stay at home at the election day. Tamsin Shaw is an associate professor of philosophy at New York University. She has researched the US military's funding and use of psychological research for use in torture - says [36]: "the capacity for this science to be used to manipulate emotions is very well established. This is military-funded technology that has been harnessed by a global plutocracy and is being used to sway elections in ways that people can't even see, don't even realize is happening to them. [...] It is about exploiting existing phenomenon like nationalism and then using it to manipulate people at the margins. To have so much data in the hands of a bunch of international plutocrats to do with it what they will is absolutely chilling.[...] We are in an information war and billionaires are buying up these companies, which are then employed to go to work in the heart of government. That is a very worrying situation. "As she says, this situation is worrying, but is it that worrying as vote buying? It is reasonable to think about this. There are tons of opinions which argue about this topic. [36] [18] [56] [58]

3.2.2.8 Controversy on 2016 US presidential election

The presidential election 2016 in the USA, Trump vs Clinton is considered as controversial[33]. It is considered as controversial for multiple reasons: first of all, the mainstream media

had a much less important role, then in other years. The role of social media platforms, like Twitter or Facebook has increased magnificently. [33] Second there were a lot of comments from Clinton but especially from Donald Trump which were very controversial (e. g. This is a quote, which Donald Trump said on his presidential bid on 16. June 2015 [98]: "When do we beat Mexico at the border? They're laughing at us, at our stupidity. And now they are beating us economically. They are not our friend, believe me. But they're killing us economically. [...] When Mexico sends its people, they're not sending their best. They're not sending you. They're not sending you. They're sending people that have lots of problems, and they're bringing those problems with us. [...] They're bringing drugs. They're bringing crime. They're rapists. And some, I assume, are good people."). Third the Russian hacker attacks on the democrats. [67] This paper concentrates about the first controversy.

Not only but also thanks to social media a lot of swing voters, who normally stayed at home, could be convinced to go to the polls. [33] In the 2016 elections were far more undecided voters than in the recent past elections. On election day still 12.5% of the voters were undecided. As a comparison in 2012, when the voters had the choice between Barack Obama and Mitt Romney, on election day only 4.3% were undecided for whom they are going to vote. The swing voters took so a very important role in the election campaign of 2016. Trump won four (Wisconsin, Florida, Pennsylvania, Michigan) swing states in the last week before election day thanks to them. Then he won of the undecided voters in Wisconsin 59% of the votes while Clinton only won 30% of the votes. In Florida he won 55% in comparison of 38% for Clinton. In Pennsylvania he won with 54% to 38% and in Michigan 50% to 39%. The rest of the percentages were still undecided or voted for a third party. [22] [25] [23] [24] This are 75 combined electoral votes. [82]

Since one of the reasons (the other big reasons were some bad news about Hillary Clinton short time before the election [82]) that swing voters decided to go to the ballots was social media; social media was important in this election. [33] While Clinton only ran 66'000 ads on Facebook, Trump ran 5'900'000 ads on social media platforms. [10]. Trump not only ran more ads, but he also ran them with a clear social media strategy. He and his election team used for this Cambridge Analytica. They posted on the social media platform with a strategic plan to overturn swing voters. For different mindsets of vote buyers, they had different social media strategies – they even used propaganda strategies from the U.S. military to infiltrate the people's minds. [58] [36] These methods were not considered illegal. Therefore, it cannot be considered as vote buying, but it is morally questionable. Mitch Vidler, head of marketing technology and digital analysis at Jaywing underlined this statement once in an article with the words: "Cambridge Analytica circumvented Facebook's terms of use, wrongfully collated information on consumers, and ultimately sold questionable data off the back of it.[...]It's a matter of morals. Just because we have the capabilities to harvest consumer data and create assumptions off the back of it, it doesn't mean we should.[97]

3.2.3 Circulation of money

This chapter focuses on the topic of monetary political cycles. There are two different possible views on the topic which will be distinguished. The first view is rather long term based and focuses on possible manipulation of monetary and budget cycles in the months or years before an election takes place. The other view is a short term one and shows that the monetary aggregate M1 is increased during the voting month.

3.2.3.1 Long term view

In contrast to the following short-term view, the long-term is not considered to be classical vote buying. But it shows by practical examples and theories how candidates try to circumvent vote buying by simply manipulating the fiscal and monetary policies prior to the election in order to get better reputation. Thus the topic is still strongly connected to vote buying itself. In this view, the term political business cycle is often used. This term describes that politicians stimulate the economy just before the election. To understand this, one needs shortly to look at the theory. There exists the theory of opportunistic political business cycles of the American economist Nordhaus [65]. In this theory, some of the assumptions made are that the economy can be described as a Phillips curve, that expectations are adaptive, that politicians are in control of a policy instrument which can affect the demand, that politicians are opportunistic, that voters are naive and that the timing of elections is exogenously given. Based on these assumptions it was possible to determine some implications. The first one is, that each incumbent government will expand the economy before elections. Because of the expansion described before, the inflation will be increased. Then, after the elections, the inflation is reduced again by a contraction of aggregate demand. And the last implication is, that the economy has an inflation bias, what means that the inflation is higher than it would be in the socially optimum. [3]

There also exist basic theories about the rational monetary and budget cycles. One paper of the economists Rogoff and Sibert focuses on the competence of incumbent. It states that these people or governments reduce taxes, raise expenses, or monetize deficits before elections in order to appear more competent. This is done with the idea that voters prefer competent governments and would then vote for the current incumbent and not for the others. The effects of these methods on the inflation and on the tax-loss will be perceived by the voters with a delay, what means only after the election [76]. Another paper of Rogoff states that an incumbent person or government is able to temporarily increase the output of a country and the employment. This can be done by raising the money supply growth during the year prior to the election. The voters will react positively to these two effects. But as before, after the election the inflation would increase what would bring output and employment back to the natural level. [75]

To verify whether political business cycles have an influence on the GDP and unemployment, a study has analysed different datasets of OECD democracies and statistically tested them. It turned out that there cannot be found any statistically significant evidence of such an influence. Similar tests were made depending on the inflation. As mentioned earlier, the inflation should increase after the election. Here, it can be shown that the effect on the inflation is short-lived what means it lasts more or less a year that immediately follows the election. But in contrast to growth and unemployment, the effect seems to be more sensible. However, it has to be mentioned that the effect is not very strong. [3]

Additionally, if there was an electoral business cycle, there should be an observable manipulation of the macroeconomic policy instruments. Namely, these are the monetary and the fiscal policy. So, a political business cycle would affect the monetary policy. It can be shown that money growth, what is defined as the difference of M1 to the previous year, is higher in the year, or up to one year and a half, before the election takes place. [3]

Subsequently, the effects of elections on fiscal policies can be investigated. The theories mentioned above would see a fiscal deficit before the election as said because of tax reductions or increased spending. It has to be mentioned, that tax reductions or increased spending of the government follow economic cycles in general. Because, if growth decreases the government has to pay more for the unemployment compensation or social welfare programs. This fact has to be considered when analysing the fiscal policies. It

turns out that the effect of elections on budget deficits is observable in the year prior to the election and is statistically significant. But the question remains whether the budget deficits come from the increased spending or reduced taxes. This question could not be answered conclusively. [3]

There exist not only the OECD democracies, but other democracies which are of interest. Another study provides evidence for political business cycles in low-income democracies of the developing world. It is based on data from Sub-Saharan African countries and shows the before named effects. A first observation can be made depending on the monetary policy interventions in these countries. In contrast to many OECD democracies, the central banks are not considered to be independent what leads to a larger potential influence. It can be shown that some African governments used different tools of the monetary policy. The reduction of the interest rates of about 1-1.5 percent can be observed. But also, the inflation gets influenced. While the effect on pre-election inflation is not especially high, it tangents the post-election phase heavily. An increase of about 6 up to 8 percentage points can be determined. [9]

Apart from the monetary policy interventions, also an intervention depending the fiscal policy can be measured. The fiscal deficit can be observed to be even more deficit in the pre-election year and less negative in the post-election year to complete the cycle again. Furthermore, proof shows that the public expenditure increases significantly during the election year by roughly 2 percentage points of the GDP only to be corrected in the post-election year by about 3.3 percentage points less expenditure. To finance these expenditures, the governments withdraw its deposits about 2 percentage points more than in non-election years. But in contrast to the other fiscal manipulations, this increase is not compensated in the post-election year. [9]

To sum up, it can be observed that the influences of political business cycles are existent but not in all suspected areas. There are no signs of cycles on GDP and unemployment, but signs of cycles on monetary and fiscal policies and on inflation are observable. It has to be mentioned that these effects are not observed to be very strong in OECD democracies. But in non-OECD democracies as presented in the Sub-Saharan countries, the effects tend to be very influential.

3.2.3.2 Short term view

On the other hand, it is possible to look at the short-term effects. To do so, one can focus on the growth rate of M1. In this scenario, M1 is defined as the cash and overnight bank deposits. A study of Aidt, Asatryan, Badalyan and Heinemann investigated this topic and shows evidence, that growth rate of M1 can be increased by 0.6-0.7 percentage points during the election month. It is important to mention, that this increase can only be observed in low- and middle-income democracies and not in OECD democracies. Out of the democracies where the effect can be observed, it is strongest in low income countries, where not all people are well educated and many live below the poverty line as for example the Sub-Saharan Countries or some in East-Asia [2]

The main explanation for this effect is, that it is a case of vote buying. In this case, vote buying is conducted with payments or gifts for the recipient. This form of vote buying, needs a lot of liquidity and hence influences M1. In this scenario, M1 can be affected through two different channels. The first one is, that illiquid assets are sold and thus made to cash. This transformation affects M1. The second channel can be, that the required funds might come from the shadow economy. This would mean that the money has not been officially in the system before but is brought into it again. In both different ways, M1 increases. Additionally, vote buying requires an environment of weak democratic institutions, elections which are not supervised enough and many voters which do want to sell their votes. The thesis that the increased M1 comes from vote buying is

thus even more probable as the properties named before are more likely to occur in the low-income democracies instead of OECD democracies. This matches the fact that the increase is only observed in these countries. [2]

But there exist also four other possibilities why M1 could be increased which have to be considered. Firstly, the central banks could be accountable for the increase because they could have changed their liquidity policy just before the election happens. This point, however, tends to be unlikely as there is no increase in M1 in the months before but only in the election months. If it was for the central banks, then one had to find the increase also in the other months, as it is not in the interest of central banks to intervene in such short terms. A second explanation might be that the increased M1 comes from the election campaigns. This could actually influence M1, but as before it is unlikely that it would happen only in the election months. Normally, these campaigns take months of preparation. Hence, in the months before, M1 should have increased as well, but it actually does not. Another remarkable point depending this possibility is, that no such effect can be found in OECD democracies. This makes this possibility yet more unlikely, as in OECD democracies large and very costly campaigns are run in general. Thirdly, the increase could come from general economic activity. It would be likely, that such an event, especially the national ones, increase the economic activity. But it is not possible to find such an increase in the growth rate around other events as national celebrations or bank holidays. Thus, it seems unlikely, that the election should increase the activity. The last possibility is that the government might pay salaries and repay credits before the elections. This would not directly influence M1 but could affect the money multiplier. But considering the deposit-cash ratio, which normally is higher for governmental institutions than for private credit grantors, the effect on M1 would be negative and not positive. This again makes this possibility very unlikely. [2]

To conclude, it is not possible to fully eliminate the probability of other possibilities. But they are all very unlikely leaving the explanation of vote buying for the effect as the most probable.

3.2.3.3 Comparison

As written in the section before, there exist both long- and short-term view effects. In contrast to the short-term view, the effects on the long-term view are also observable in OECD democracies, however they tend to be not very strong. But both views observe effects in the low-income countries as the Sub-Saharan. For the long-term effects, the manipulations are also considered to be statistically significant in these democracies, as various examples show. While the effects on the short term are most probable because of systemic vote buying, as all other possibilities can be eliminated with a high chance.

3.2.4 Legal aspects of vote buying

This section tries to introduce the legal aspects of vote buying. For this reason, the legal situation of vote buying in the United States and in Switzerland is investigated more. The goal is to provide the legal basis for all the examples shown before. Additionally, the section talks about the consequences selling, respectively buying votes can have in these countries.

3.2.4.1 USA legal aspects

In the United States of America, vote buying is prohibited by Chapter 29 Elections and political activities, Part I. Crimes of Title 18 in Crimes and Criminal Procedure of the U.S. Code. which in numbers are paragraphs 591 up to paragraphs 612 [96]. Out of special interest are the following paragraphs

- §594. Intimidation of voters
- §597. Expenditures to influence voting
- §598. Coercion by means of relief appropriations
- §599. Promise of appointment by candidate

The first paragraph of these handles with the intimidation of voters. Apart from intimidation it also includes threatening, coercion, or the attempts of doing these. So, any person using the named methods before to interfere with the right of the voter to vote as he wants can be punished. The punishment can be a fine or imprisonment for at most one year or both.

Paragraph 597 deals with the expenditures to influence voting. This essentially means that anyone who offers an expenditure to any voter to vote or not vote for a candidate, but also any person who accepts such an offer, can be punished. In this case, the person can be fined or imprisoned for maximally two years or both.

Paragraph 598 talks about a special coercion by means of relief appropriations. This means that anyone who uses any appropriation for work relief, relief or for increasing employment in order to coerce any person to vote as he wants can be punished. The punishment in this case can be a fine or imprisonment not exceeding one year or both. The last paragraph of interest is paragraph 599. It deals with candidates who use their influence to promise an appointment to a position or employment to any person in order to get this person's vote, can be punished. The person can be fined or imprisoned for at most one year or both. If the infringement was done wilfully, the duration of the imprisonment can be extended to maximally two years. [96]

3.2.4.2 Switzerland legal aspects

In Switzerland as in many other countries, there is a regulation to prohibit vote buying. In the case of Switzerland, this topic is addressed in Article 279 and the successive of the Swiss Criminal Code. In detail, there are articles 279-284 covered in title fourteen which concern the topics of vote buying.

- Art. 279 Disruption and obstruction of elections and votes
- Art. 280 Attacks on the right to vote
- Art. 281 Corrupt electoral practices
- Art. 282 Electoral fraud
- Art. 282¹ Vote catching
- Art. 283 Breach of voting secrecy

The first of these, article 279 talks about the disruption or obstruction of an election. This means that any person using violence or threat to disrupt an election or vote can be punished. Article 280 concerns an attack on the right to vote. This means that any person, who uses violence or threats to prevent a voter from voting at all or uses violence to make the voter vote in a specific way as the intimidator wants, can also be punished. In terms of vote buying article 281 is probably another main point. This article specifies that any person who offers a promise, a gift, or any other advantage in return for voting in a particular way or to not participate in a vote can be punished.

In article 282 the topic of electoral fraud is discussed which says that any person who forges, falsifies, removes, or destroys an electoral register can be punished. This also

means that counting the ballots wrongly or omitting some can be punished.

The second last article is about vote catching which means that it is illegal to systematically collect or complete ballot papers for others or distribute prefilled ballot paper can be punished.

The last article is concerning a breach of voting secrecy. What means that any person who obtains knowledge in an unlawful way can be punished.

All articles apart from article 282¹ share the same punishment. This means that for all the others a custodial sentence not exceeding three years, or a monetary penalty is possible. For article 282¹, only a monetary punishment can be imposed. [88]

3.2.4.3 Comparison

It can be observed that the regulations depending vote buying are similar between Switzerland and the United States. Although, the states do not name it exactly the same, the idea behind the law is equal. Both countries prohibit vote buying. As seen, both laws talk about buying votes through coercion. And both countries know some laws against vote buying with expenditures, job promises or any advantage in another way. But it has to be mentioned, that the punishment is not always the same. The Swiss criminal code knows imprisonment for at most three years or a monetary penalty for most of the violations apart from vote catching for which only a fine is possible. The united states criminal code also has the possibilities of imprisonment or a fine, but it can also be both. In contrast to the Swiss criminal code, the length of the imprisonment in the United States is maximally one year for most of the mentioned offences and two years for wilful promises of appointments. A last observation to mention is, that regarding the length of the imprisonment the Swiss criminal code is stricter. To conclude, one can observe that the law resulting in the political reality is very clear. There is no discussion about the legitimacy of vote buying. It is strictly forbidden.

3.2.5 Excursus: e-voting in Switzerland

Because of the national background of the authors, this paper discusses how far Switzerland is processed in switching to e-voting.

E-Voting in Switzerland was tested with two systems. One of the canton Geneva and the other one from the Swiss Post, which let the system create in Spain. [51] Geneva announced on the 28th of November in 2018 that they will stop their e-voting system because the upcoming updates are too expensive. They researched and developed the program for a total of 15 years. [45] e-voting in Switzerland has a good standing as polls in 2019 showed. Nearly 70% has in 2019 desired that the e-voting system is offered to the public and not only to some test users. 47% of the voters would take part in more elections than they did before and only 8% want that E-Voting is forbidden in the future. [17]

10 of the 26 cantons in Switzerland used one of the both mentioned E-Voting systems in a test phase. Geneva stopped its system in 2018, the post had to stop their system in June 2019, but not because of monetary reasons. The Canadian IT-specialist Sarah Jamie Lewis recognized horrendous mistakes in the code. The vulnerability of the code would allow system administrators to manipulate votes and compromise elections without detection.[72]

Since then, the government decided to stop the testing phase. The Swiss Post is still working and developing on the code. They are confident to have a system which is ready for the users to test in 2021. The code should be made public, as it was already before. The biggest change should be that the system is not only individually verifiable (it was only 50% universal verifiable) but it the goal is to have a system which is near to com-

plete universal verifiable [70]. The paper takes a closer look at universal verifiability in the upcoming sections.

3.3 Countermeasures of vote buying

The following section shows countermeasures of vote buying. Especially how to achieve a voting system, where vote buying cannot happen. Nowadays, there are voting systems in use like in Estonia [50]. Even in Geneva, Swiss people had the chance to cast votes electronically between 2003 and 2004 [4]. However, while electronic voting may increase the participation and also simplify the tallying of elections, it also carries the threat of potential abuse at a large scale [4]. In the past 20 years, there has been a lot of research in the field of voting systems, focusing mainly on the robustness, anonymity and verifiability in electronic elections [4].

The following section, as already mentioned, gives an overview of countermeasures against the threat of vote buying in voting systems. Firstly, there will be a short introduction of properties which a voting system should satisfy according to [52]. Subsequently, the notion of receipt-freeness and coercion resistance will be defined. There is a lot of literature which are proposing voting systems that are receipt-free, respectively coercion resistant. Receipt-freeness was firstly introduced by Benaloh and Tuinstra in 1994 [44]. The property will be shown in the ThreeBallot voting protocol and the VAV protocol [52]. Also, Martin Hirt from ETH Zürich and his partner Kazue Sako presented a paper, where receipt-freeness is achieved by homomorphic encryption. Coercion resistance has in the literature not an unambiguous meaning [50]. Many voting systems are using a combination of cryptographic techniques like blind signature, homomorphic encryption, mix networks and further means to achieve their view of coercion resistance under some assumptions and requirements. The definition of strong coercion resistance have been formulated by [4], which will be explained in section 3.3.2. Besides this, [4] also proposed a voting system which is coercion resistant, also known as the JCJ scheme [50] [84]. Based on this scheme Clarkson, Chong and Myers extended it and created the Civitas protocol in 2008 [20] [50], claiming that it is the first electronic voting system that is coercion resistant, verifiable and appropriate for remote voting [20]. However, because the JCJ scheme will be explained in detail, the Civitas protocol will not be further explained. The Estonian voting system will be also presented. Based on a very trivial solution it achieves coercion resistance [50]. Although, as it will be shown, it does not provide the highest level of coercion resistance as defined by [4]. Lastly, there will be a rough overview of voting systems based on blockchain.

3.3.1 Privacy and Verifiability

In the 19th century, the Australian secret ballot has been introduced. As it is shown in section 3.2.1.1, the property mitigates coercive behavior by intimidating the voter, because the coercer should not be able to link a vote to a person in any way. This property is a general central requirement in remote voting systems, which has been researched thoroughly [52]. Another central security property according to [52] is verifiability, which can be divided into individual verifiability and universal verifiability. In an election, a voter wants to know if his/her vote is tallied correctly. Intuitively, this can be done through a receipt which contains a code linked to the vote on the bulletin board. After the election, one can check his/her vote accordingly. However, especially in electronic remote voting this is not done. It follows that the property of individual verifiability should be a security mechanism provided by the voting system [52]. Universal verifiability however, states that

anyone can approve that the outcome of the election corresponds to the bulletin board and that all the votes in the election are legal [52].

3.3.2 Receipt-freeness and Coercion resistance

Voting as intended is one of the fundamental rights in democratic countries, i.e. without being coerced or being able to sell ones vote in favor for the buyer [50]. With the trends of remote voting in the current time, mitigating or even preventing vote buying should be achieved in voting systems. Mostly because the danger of coercion and vote buying are potentially much more troublesome in electronic voting schemes than in normal, physical voting schemes [4].

Benaloh and Tuinstra firstly introduced Receipt-freeness [44] in 1994, stating that receipt-freeness is not to have the ability to prove to a third party how the voter voted in particular e.g. with a receipt from the voting system [4]. This, mitigates vote buying. However, just not being able to construct a receipt, does not prevent from stronger attacks. Imagine a channel where the vote has been sent through can be observed by the buyer. Consequently, the buyer knows how his seller voted and can therefore trust him. However, receipt-freeness does not consider this, while coercion resistance does. The problem that a buyer can observe the channel is a privacy issue. The authors of [28] states that the property coercion resistance guarantees that the buyer does not get any information at all, even when the voter cooperates. From their point of view, coercion-resistance is stronger than privacy and receipt-freeness is in between [28]. [4] goes even further and states that assuming the fullest range of behavior from the coercer, voting systems should also defend from randomization, forced-abstention and simulation attacks [4], referring this property as the highest level of coercion resistance [50].

3.3.3 The ThreeBallot voting protocol



Figure 3.2: Two possibilities to fill out the ThreeBallot in favor of candidate B. x = marked; o = not marked.

ThreeBallot is a multi ballot voting protocol designed by Ron Rivest in 2006 [52]. This protocol was intended to vote from a voting booth where there is a scanner to scan the multi ballot. However, because the concept is very trivial, it can also be implemented as a whole remote voting system [28]. Anyhow, the protocol is used for electing only two candidates. The voter is given three simple ballots, where each candidate is listed in a fixed order. The voter is supposed to mark the candidate in favor with two x and the opposing candidate with one x. It is also possible to mark both candidates in a single ballot as shown in figure 3.2 b). Afterwards, the voter feeds his multi ballot in a machine and the machine prints random numbers on every single ballot. Subsequently, the voter chooses one single ballot as a receipt. The machine keeps now all single ballots separated in a ballot box. When tallying all the votes the machine posts all single ballots on an electronic bulletin board e.g. a web page. Defining the variables first:

- $\alpha \equiv$ number of votes for candidate i
- $\Theta \equiv$ number of single ballots marked for candidate i
- $\tau \equiv$ total number of votes

The number of votes of candidate i can be computed as follows [52]:

$$\alpha = \Theta - \frac{\tau}{3}$$

Because the voter chooses one single ballot as a receipt, he cannot prove to the buyer how he really voted, even if he actually wanted to, because it is the whole multi ballot which states how he votes. Therefore, the scheme mitigates vote buying and also provides some sort of individual verifiability with the single ballot as a receipt containing the code [52]. The authors of [52] also state that it is also coercion resistant to some extent. However, according to the definition of [4], it does not prevent from coercive behavior, which means that a coercer can intimidate a voter to abstain from the election for example. Therefore, it is not seen as coercion resistant.

3.3.4 The VAV protocol

(a)	V 1: x 2: o 3: o	A 1: x 2: o 3: o	V 1: x 2: o 3: o
(b)	V 1: x 2: o 3: o	A 1: o 2: x 3: o	V 1: o 2: x 3: o
(c)	V 1: x 2: o 3: o	A 1: o 2: o 3: x	V 1: o 2: o 3: x

Figure 3.3: VAV Multi ballot. Three possibilities in favor for candidate 1

The VAV voting protocol is also based on a multi ballot scheme. As shown in Figure 3.3 on top of two single ballots is the letter V and on the last one the letter A. The difference between VAV and ThreeBallot is that in VAV a voter can choose between multiple candidates. The voter is supposed to mark the desired candidate on a V-ballot, whereas the remaining two ballots can be marked randomly. However, the voter must mark the same candidate on the remaining A-Ballot and V-Ballot. Figure 3.3 shows three ways for how to vote for candidate 1. Afterwards, the voter feeds the multi ballot in a machine, e.g. a scanner, which also prints some numbers on it. The voter can randomly choose a single ballot as a receipt with the printed number. This number, as in the ThreeBallot, is used for checking and verifying that the voters multi ballot is counted correctly. In the tallying phase, the authorities cast all single ballots on an electronic bulletin board e.g. a web page. The result can easily be computed by eliminating all A-ballots marking candidate i with all V-ballots also marking candidate i . Because the voter chooses one single ballot as receipt, a buyer does not know that the receipt corresponds to the whole multi ballot, mitigating vote buying. However, as shown in the section of ThreeBallot, the scheme is not coercion resistant [52].

3.3.5 Receipt-Free Voting Based on Homomorphic Encryption

Hirt and Sako presented a paper, where receipt-freeness is achieved by homomorphic encryption. They are stating that classic voting schemes are failing to achieve receipt-

freeness because a voter can prove to a buyer his encrypted vote by revealing the randomness he used for encrypting e.g. by constructing a receipt [44]. This section gives a rough idea behind the scheme of Hirt and Sako by firstly introducing the model with their assumptions, following by the functionality of the scheme and eventually the argumentation why it is receipt-free.

3.3.5.1 Models and assumptions

Entities: The scheme contains N authorities and M voters. A fixed number of t honest authorities are also needed.

Communication: On a publicly readable bulletin board every voter can write onto it. It is not possible to delete something from it. Also, from every authority it is assumed that there exists a one-way channel to a voter, which can not be wiretapped. The voter receives data from the authorities and it is assumed that even if a buyer is physically present at the voter's place, the buyer cannot see what kind of data is received.

Key Infrastructure: Each voter has a secret key and a public key and each voter must know his own secret key according to the public key, which is used for identification. Even when a voter gives his secret key to a buyer, the property of receipt-freeness is still achieved.

Generality: The scheme is a 1-out-of- L voting scheme, which means that a voter may submit one valid vote from the set ν . In an election with two candidates $|\nu| = L = 2$. Allowing the voter to submit an arbitrary string may disturb the property of receipt-freeness, because the voter can mark his vote and then prove it to a buyer.

Security: For the whole protocol to work, it must be ensured that at least t authorities are remaining honest. s authorities, where $s < t$, cannot decrypt any encrypted votes, which provides privacy and therefore receipt-freeness.

3.3.5.2 The Hirt and Sako model

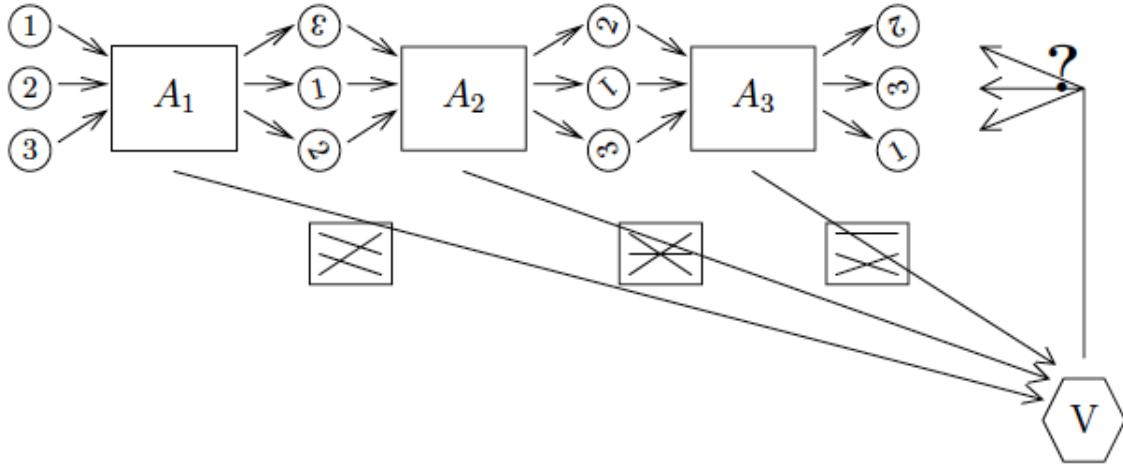


Figure 3.4: Constructing a vote in $\nu = \{1, 2, 3\}$ with 3 authorities

In Figure 3.4 the basic idea behind the scheme is shown. The voters are casting votes from the set ν , which is then encrypted (from the left side). A_1 takes the list of encrypted votes, randomizes it and changes the order of the list by homomorphic encryption, which changes the cipher text but not the content of it. After that, it sends the new list to the next authority, who does the same. Simultaneously, each authority must send in an untappable channel the change function to the voter and reveal publicly that the shuffling is done correctly. [44] argues that this kind of scheme is receipt-free, because of the

verifiability of the shuffling is private over an untappable channel, the voter has no chance to reveal how he voted to a buyer.

3.3.6 The JCJ scheme

The JCJ protocol was published by Juels, Catalano and Jakobsson in 2005 [4]. Since then, the protocol for electronic e-voting is seen as the only known model, which offers individual verifiability, receipt-freeness and strong coercion resistance. In addition their assumptions about the model are in practice feasible to implement. However, in the tallying phase, the cryptographic prove of each vote takes quadratic time, which implies that it can not be used in a large election [4] [84]. Although, the protocol is highly considered in the cryptographic and scientific scene and is taken as a basis for further work and improvement, for example the Civitas protocol [20] [50] or the work of [84]. The main reason why this protocol is coercion resistant, and therefore private and receipt-free, is, because they introduced fake credentials which a voter can use while being under coercion [50]. Through cryptographic mechanism the coercer does not know if the credential is fake or not, providing privacy.

In this subsection, it will be shown how an election takes place and under which kind of assumptions and requirements the protocol works and why it is coercion resistant.

3.3.6.1 Description of the protocol

The following part describes the registration phase, the voting phase and the tallying phase of the JCJ protocol. The cryptographic means are done with ElGamal encryptions, based on multiplicative cyclic groups [84], which will not be further explained. Firstly, the entities will be introduced. After that, it will be shown how an election is executed with the JCJ protocol.

Entities:

1. *Registrars:* The registrars are denoted by $\Pi = \{\Pi_1, \Pi_2, \dots, \Pi_{n_p}\}$, where n_p is the number of the registrars. They are responsible for issuing the credentials to voters.
2. *Authorities:* $\tau = \{\tau_1, \tau_2, \dots, \tau_{n_t}\}$ are n_t authorities, who are responsible for tallying the votes.
3. *Voters:* $\Omega = \{\Omega_1, \Omega_2, \dots, \Omega_{n_v}\}$ are n_v voters, who are entering an election by registering to Π . i is a public identifier for Ω_i .
4. *Bulletin Board:* B is a publicly visible bulletin board where voters and registrars can write on it. However it is not possible to erase something from it.
5. *Candidate slate:* C is an ordered set of n_C distinct identifiers, which means from this set, a voter can choose a candidate $C_i \in C$ and $i \leq n_C$ [4].

With the entities defined composing the voting system, the protocol can be run as followed:

Registration: Π_i creates a random credential σ and gives it through an anonymous channel to a voter Ω_j , $j \leq n_V$. By way of cryptographic means (ElGamal), the registrars encrypts σ to Ω 's entry of the bulletin board B [84].

Vote Casting: To cast a vote the voter Ω_j must submit two encryptions to B . Through cryptographic means, both encryptions will be checked if $\sigma \in B$ and $C_k \in C$. If a voter is coerced, she can fake her real credential σ' . One of the encryption can be used to prove to the coercer that she complies with his demands or let the coercer vote (Simulation-attack) [84].

Tallying: After the voting phase, B contains N votes. First, the authorities are proving that both encryptions of a voter Ω_j are correct. If a prove does not hold for a vote, for

example because it was cast multiple times or was cast using a fake credential, it will be cancelled without any traces. The prove for all votes takes quadratic time [84].

3.3.6.2 Assumptions and Requirements

The JCJ protocol requires an anonymous channel where the credentials are given to the voter by the registrars, whereas other protocols are assuming unrealistic untappable channels. By using mixnets, an anonymous channel can be realized [4]. The authors of [50] are assuming that also visiting the authorities to get the credentials is possible. However, this would not be seen as remote e-voting. Nevertheless, the protocol allows the credentials to be reused in other voting events, providing the citizen to visit the office of the authorities just one time. In the registration phase the credentials will be passed to a voter. It must be assumed that a majority of the registrars Π_i is honest granting that only Ω_j knows σ [84]. Also, it is assumed that a coercer is not able to exchange messages with a registrar, so the voters can give fake credentials to the coercer, which achieves receipt-freeness. To defend against randomization attacks, the protocol filters out multiple votes and does not count $C_i \notin C$. Such that the voter is not attacked by forced abstention, it requires that the voter is given a moment of privacy. Those assumptions and requirements are needed such that the protocol is coercion resistant [84] [4].

3.3.7 The Estonian scheme

Estonian citizens were able to cast binding votes through remote e-voting in October 2005 [61]. It was the first time in the world that a whole country had the possibility to cast the vote through the internet. Estonia is a pioneer in e-democracy. Half of the households are possessing a computer at home and every school is connected to the internet. In addition, it is the only country in the world with compulsory ID cards, used for remote identification and binding digital signature functions. E-voting was therefore the next step to implement [61].

In this section, it will be shown, according to the Estonian scheme, how to mitigate coercion by a very trivial solution. In fact, they are using the meta technique of re-voting [50]. Their voting system are providing a voter with the possibility to change the vote while being under coercion. However, this is the only anti-coercion functionality the scheme offers, which may cause problems. The biggest issue is that a coercer stays with the voter until the end of the election process, making sure that the voter does not re-vote again. This is also called "over-the-shoulder coercion" [50]. However, to lower this threat, remote submissions are ended two hours before the polling stations are closed. The voter can, if coerced, cast a paper vote, which cancels the e-vote, assuming the voter can make it in time to the polling station. This mechanism also allows a simulation-attack to be overwritten [50].

3.3.8 Voting systems based on blockchain

Privacy, verifiability and coercion resistance in voting systems are usually achieved by a combination of cryptographic techniques like homomorphic encryption, blind signatures and mix networks [104] [89]. However, the security is not always given, for example in the 2008 US elections, 197 votes have been erased from the database [104]. Another example is that 66'000 electronic votes have been tampered in the 2015 state election in Australia [104]. Those incidents raised scepticism against electronic voting systems based on bulletin boards which is mostly ran by central third parties. The rather new technology - Blockchain - with growing popularity because of Bitcoin, provides a new way of achieving secure and verifiable electronic voting systems [104]. Blockchain is a decentralised

database, where the append-only data structure is shared among all participants in the network [89]. This new technology supplies three properties [43]:

1. Immutability: A data entry, also called a block, references through a hash the previous block. By induction, an immutable chain is created, which defends from tampering of the integrity.
2. Verifiability: The database is distributed over multiple locations. Each of the locations contains a copy of the database. This implies high availability and verifiability, because all nodes must have the same version of the database.
3. Distributed Consensus: Before a new block is appended to the chain, the majority of the network nodes must have the same version of the blockchain.

Those properties are partly achieved by other cryptographic voting systems. Blockchain is therefore highly considered by researchers and has in their meaning a massive potential for deploying a new voting process [43]. Subsequently, a lot of blockchain-based voting systems have been developed, which can be grouped into three categories [104] [89].

1. Cryptocurrency-based: A vote, in a cryptocurrency based voting system, can be performed by sending a coin to the desired candidate. There are however three major problems which may occur. Firstly, the voter can just keep the coin. Secondly, it lacks receipt-freeness, which is not further explained in [89]. And lastly, it requires a centralised third party to conceal ones vote from the authorities [89] and to coordinate the payment between the voter and the candidate [104]. In addition, the latter problem nullifies the benefit of blockchain by not requiring the need of a third party [89].
2. Smart contract based: There are multiple proposals of voting systems which are based on smart contracts. However, each of them has there flaws. Either it is that voters can just choose from two candidates or there is limited amount of people who can participate [104]. Another smart contract based voting system is using in addition the cryptographic technique of homomorphic encryption to tally the vote from another voter which is then send again to the next voter. Nevertheless, the system does not promise that a peer may send it to a next peer [89]. The authors of [89] provided a very promising voting protocol based on smart contract, which achieves receipt-freeness and coercion resistance. Under practical assumptions the voter does not have to perform a counter strategy e.g. lie or fake credentials to achieve coercion resistance. They just have to vote. Their solution is a randomizer token which is a smart card device that can not be manipulated. The randomness for creating a ballot is not known to the voter and with other cryptographic means provided by the blockchain the system achieves the property of coercion resistance [89].
3. Blockchain as ballot box: Some commercial system has also been developed using blockchain as a ballot box as a basis. However, even if they claim that the system is verifiable and accessible, they have their flaws. Either a central trusted party is needed to hide the link between a voter and its ballot (achieving receipt-freeness) or it does not provide a mechanism to protect voters from coercion [89].

3.4 Ethics of vote buying

The following sections give an overview to the discussions in philosophy on vote buying. Whereas studies in the fields of social sciences and humanities focus on the direct

implication that vote buying has in real democracies, the ongoing discussion in political philosophy is concerned with the more theoretic question of whether or not vote buying is intrinsically wrong. In the public discourse, vote buying is treated as corrupt, problematic, and as something that endangers democracy. [57], [60] Also our contribution is concerned with historically problematic evidence on the one hand, and with the technical prevention of vote buying on the other. The frame in which we treated vote buying remains clearly that vote buying should be prevented. Taking the ordinary discourse as well as the strong moral intuitions serious, this part aims to examine two recent philosophers work who stand for an optimistic position on vote commodification. Working with their arguments, the premise should be questioned. Jason Brennan most prominently wrote a book where he discusses ethical aspects around voting. [11] Based on the conviction that there is a common good, which can be promoted or endangered via voting, he also focuses on vote buying and makes the argument that in some specific cases vote buying isn't problematic per se. Christopher Freiman takes Brennan's reflections as foundation for his demand to legalize voting markets. He argues that vote buying and the complete commodification of this democratic tool would enhance the quality of democratic decision making. [35] To contrast these positions, this section will then present their philosophical opponents and focus on three main objections. These are chosen as most important ones, each of them representing a different approach to formulate the critique. The first argument focuses on the claimed efficiency of democracies with voting markets. It argues that the degree of intensity, that voters want something, could not be better reflected having voting markets. Furthermore, it is accepting the opinion of everyone the will to abstain or the weak conviction of something shouldn't be underestimated. As second argument, the commodification of voting market is criticized philosophically: The economic discourse of voting changes the way we think about core democratic institutions. It adulterates voters from their valuable role in democracies. And as third philosophical argument, the respect of voter's autonomy disallows mechanisms of vote buying, since the act of using money over rational conviction passes over the core process of opinion building. To end this section, a critical glance at the premise of democracy is finally brought up. It will be argued that vote buying can't be of any legitimacy in our existing democracies because of the materialistic circumstances in any existing society. When a philosophical debate is announced, one might think of a highly abstract and theoretical debate going back to thinkers such as Rousseau and other enlightenment philosophers which were working on core democratic theories. That's not, what this sections is about: Rather, a discussion shall be launched which cares about the simple question of whether or not our intuition that vote buying shouldn't be allowed in any case can live up to moral reflections, and to strong logical reasoning.

3.4.1 General Legitimisation of Vote Buying

3.4.1.1 Brennan's reflections

Jason Brennan is a professor in Ethics and Economy at the Georgetown University. In his book "The Ethics of Voting", which was published 2011, he addresses vote buying in one complete chapter and embeds this discussion in a broad field of ethical reflections on voting principles. [37] One core question is concerned about an obligation to vote. Brennan discusses that there's what he calls Folk Theory of Voting which suggests that every citizen has the obligation to vote following his or her good faith. [11] p. 3. As democratic duty, citizens need to vote and therefore participate in the democratic process. Brennan contradicts and the foundation of this contradiction is deeply connected to question on vote buying. Brennan argues, that in each vote there's something as a common good which should be promoted with the right choice in the election. Voters should therefore

"on the basis of sound evidence for what is likely to promote the common good" make their decision. [11] p. 135. This principle is further referred as principle P1. If voters can't make this well-informed decision, it is not immoral to abstain - it is even ethically better for society. For example, if they are corrupted by egoistic motivations or if they don't know about the key facts around some proposals to decide which is the best choice to make, then they should abstain. Brennan therefore concludes that there's no obligation to vote. At this point, one needs to realize, that the conviction that there is always a clear decision for or against the promotion of the common good, is hardly debatable, thinking of different ideologies for example. Of course, there are political debates where one can determine egoistic motivations to vote for or against a proposal. But since there are different views on the functioning of economic as well as social processes and furthermore diverse evaluations of the importance of some core values such as economic liberty versus social control, something as one final view on what promotes the common good rests illusionary. However, to understand Brennan's argument for vote buying, one must accept this premise of a clear decision for or against a common good. Since this discussion is concerned about the general question whether vote buying is a problem at all, there's no need for considerations about cases where the outcome of a vote on the common good is clearly debatable. Accepting Brennan's premise, it follows that if vote buying can be done in accordance with this Principle P1, there's no problem with that. He exemplifies this using some examples. Brennan presents three cases. In two of them people who wanted to abstain are financially motivated to go voting. [11] p. 140ff. First, an unmotivated hero is introduced as abstaining voter: A man, who served his country in military service, worked a lot for society and is well-informed about politics and all possible proposals and candidates. For some reason, he is just not motivated to go voting on the voting day. He instead wants to stay at home and watch a movie. Brennan now argues that this man could be convinced to go voting by a friend who really wants him to take part of the decision making because he knows for sure that his friend would make the right decision. Therefore, he pays him hundred dollars to go voting instead of staying at home. Brennan argues that there's no violation of P1 and therefore no problem with vote buying per se. He anticipates the counterargument that there could be a problem if many people did this and if there was a scaling of this effect. But even in this case, Brennan states clearly the common good would only be promoted more clearly since the voting is done in accordance with P1. As second example, not a hero but an average person is paid to go voting instead of abstaining. The average person is known to also vote for the common good. Therefore, again, there's no harm done. Brennan only states that the money might be spent better in a donation since it's only an average person who is moved to go voting. [11] p. 142. Stating this, Brennan differentiates between the quality of votes depending on personal achievements of citizens. This needs to be stated here for the counterargument in the upcoming part of this philosophical considerations. In the very last example, again a well-informed and morally good person A wants a second person B to vote for candidate X who A considers to be the best choice. Brennan argues that B wouldn't vote but abstain on the voting day, since he or she isn't very interested. Of course, there's the possibility that A could somehow convince person B to vote for candidate X. But only the argumentative conviction isn't enough of an incentive for B. B is just too lazy. Here, the vote buying comes in: A gives B some money to manage to go voting. Also here, Brennan argues, the process isn't problematic since Principle P1 isn't violated and there's no coercion.

3.4.1.2 Freiman's vote market

Based on Brennan's arguments which legitimize a general money transfer on votes, Christopher Freiman presented 2017 a paper in which he goes beyond the theoretical discussion

which is presented in *The Ethics of Voting*. Freiman is professor at the Colleague of William and Mary in the United States. [54]. His work "Vote Markets" is not only interested in the moral legitimization of vote buying, he even promotes a real-world vote buying market: "I argue for the legal permissibility of vote buying and selling." [35] p. 760. He insists on the economic principle of the appeal to exchange. "Under normal conditions voluntary economic exchange is ex ante mutually beneficial. A trade is not consummated unless both parties expect to benefit." [35] p. 761. Freiman argues that the appeal to exchange, meaning that both sides benefit from an economic deal, should also hold in democratic transfers. "I'll sell my vote for n dollars only if I value n dollars more than my vote. The buyer will buy my vote for n dollars only if she values my vote more than n dollars." [35] p. 761. Following these basic economic intuitions which - ignoring any materialistic constraints - understand voting as a normal good, Freiman derives all advantages of a market from the commodification. He further argues that for a democratic system, the strength of political conviction will be better expressed if market mechanisms were introduced. The private benefits resulting from a trade exchange can also be applied to political groups and ideological convictions. In democracies where vote buying is forbidden one can only use means of argumentative and propagandistic work to gain votes. But in the end, each vote is counted equally which means that the one voting for a proposal or candidate out of some poor conviction is counted equally as someone who is deeply convinced. Establishing a voting market would provide the ones who are strongly convinced of the right vote monetary instruments to express this conviction in buying further votes. So, it follows that the intensity of preference will finally be adequately reflected. Freiman illustrates this in the following example: "Suppose 50.01 percent of the electorate just barely supports Candidate A. If they judged Candidate A to be only marginally worse, all 50.01 percent would flip to Candidate B. The remaining 49.99 percent, by contrast, are in such enthusiastic support of Candidate B relative to Candidate A that they will all emigrate if Candidate A wins. ... If vote markets are permitted, those who care more about the election can buy more votes, leading to the election of Candidate B. So, if the people preferring B would do some effort (meaning also some monetary effort to buy votes) they would easily be able to get some votes which would have been otherwise given to A or votes which would have abstained." [35] p. 764. Freiman is convinced that market mechanisms also make democratic processes more efficient. Since there is the possibility of a fine-grained expression of the intensity of one's by adjusting the amount of money one spends on buying votes, the democratic process of voting complemented by market mechanisms optimally reflect the will of the people. In the second part of his work, Freiman encounters counterarguments: There are some heavy weighted counterarguments brought up by a wide range of political philosophers of which three shall be deepened in the following sections.

3.4.2 Theoretical Counterarguments

The philosopher Paul Sheehy argues that indifferent voters should not vote [80]. He mainly argues that it is fair that people who are indifferent to the outcome shouldn't determine the outcome. Each person has the right to participate equally in a vote: If there are some who aren't interested, then their absence of political engagement should be expressed as well. [80] p. 49f. Sheehy clearly states that even if there are voters with strong opinions on something, it isn't legitimate to change abstaining people's position. The abstaining people's absence marks their interest, respectively their lack of interest. Since each vote is counted equally, it would be unfair to change their position other than by conviction. To gain more votes from people who are originally not interested in the outcome would be wrong since it would adulterate the interests of the people living in a community. Brennan argues differently. In most elections, he sees a decision between an ethically better and

worse candidate or proposal. The outcome of most elections, other than deciding the colour of a flag e.g. [11] p. 146, have moral implications: There's always a right and wrong choice to make. Therefore, if one can mobilize the ignorant ones to vote for the better choice, there's no corruption of the democratic expression of the people's will. Based on the belief that there's always such a thing as the right choice, every mean to promote its election should be used. For Brennan, it is therefore legitimate that the votes of abstaining people are bought. This means that a vote can finally express the will of another person than the one originally possessing the vote. However, it should be done in accordance to the right choice. The attitude that there's a morally right choice to make in most decisions is strongly connected to Brennan's general premise P1 that there's a common good, which can be neutrally evaluated. There's another critique concerning this promise. The German Journalist Lenz Jacobsen, who has a focus on political studies and democracy, presents one of Brennan's more recent work 'Against Democracy' which he introduced on a panel with other political scientists in Germany in 2017. [12] It is not necessary to deepen the book's content. However, some background is needed to situate the book in our discussion on vote buying. To provide some necessary insight, Brennan questions the concept of a democracy in which everybody can express its will. Since he considers most people to be not adequately informed or biased by a prefabricated party opinion, the demand of total participation won't necessarily promote the common good. His ideal of a voter is a person who is very well informed and always considers every aspect: She or he would have the intellectual as well as social capacity to make a considerate decision. Brennan's position was encountered by the political scientist Ulrike Guerot. She questions his fundamental conviction of the common good which should be neutrally evaluated. Her critique concerns the premise Brennan also prominently brought up in the context of vote buying. Guerot argues that what Brennan wants to keep out of political decision making, the bias based on ideology or political conviction is the core of politics itself. She emphasises that there's no such thing as an objective truth but a plurality of political judgments which need to be negotiated. Therefore, there is no such thing as objectivity which could be the result of well-informed persons. The core of politics is the battle of approaches to achieve the common good. There is not an objective truth or approach to reach the common good, but different judgments and ideologies which need to be debated in the political process. Politics consists of conflicts between goals and values. [46] Coming back to the question on vote buying, Guerot's critique undermines the fundament of Brennan's argumentation. Since she doesn't accept the premise that the outcome of each decision can be evaluated neutrally and that one can always decide between a choice which is better or worse for society, the decision an individual voter is doing can't be evaluated as being in accordance to P1 or not. Considering indifferent voters being paid to go voting, the simple justification based on the improvement of the common good can no more be given. Ideology and the political conviction which needs to be developed in accordance with one's knowledge and values determine the result of the personal voting. A further critique on Brennan's general thoughts which intrinsically allow vote buying focuses on exactly this aspect of the development of personal conviction. Lachlan Montgomery Umbers is a moral and political philosopher at the University of Western Australia. In his research, he focuses on democratic theory and covers topics such as egalitarianism, applied ethics and climate justice. [94] Umbers strongly argues against Brennan in his paper "What's Wrong With Vote Buying?".[93] He examines different objections being brought up against the allowance of vote buying in the first part: He differentiates some instrumental objections to the legitimacy of vote buying: None of these he considers as wholly sufficient. [93] p. 555. One objection against vote buying is concerned about the exploitation of the worse-off: If vote buying was allowed, Umbers presents the argument, wealthy people could exploit the political force of poor voters by easily buying their votes. The incentive to sell the own vote is much higher under poor financial circumstances. But Umbers

argues that the mechanism of exploitation only applies if you understand poor people as a political force *qua* group - a force that would have a consciousness about its own political power as a group. Since especially the worse-off aren't organized as political force, this doesn't apply, and they can't be corrupted as a group and therefore can't lose their power. They can't be understood as one coherent political group. Individuals can sell their votes to different parties, but this doesn't mean that they won't lose as a social group directly their power. Umbers considers the argument of exploitation as reasonable but not as strong enough to completely reject the whole idea of vote buying intrinsically. Another objection is that people would lose their effective participation in decisions to which they will be subjects. Also this, Umbers argues, can't be applied for the same reason: "Vote buying would almost certainly diminish the political impact of the worse-off, considered collectively. But, to explain the wrongness of vote buying in terms of a loss of effective participation and political impact, we must explain why it is that individuals *qua* individuals would lose out in this respect." [93] p. 557. Umbers stresses the importance of an individual approach to delegitimize vote buying. Since the loss a class consciousness can't accurately describes the problem which results with vote buying on the individual level of the transaction. Umbers considers the argumentation with social groups to be incomplete because the core of the objection lies in the political force *qua* groups and not on the level of a single transaction. His critique deals with the level of single transactions. As fundament, examines first everybody's autonomy participating in a democratic system. Mature individuals as citizens encounter each other based on mutual respect. Respect is defined as the recognition of certain abilities of each other. [93] p. 561. Being respectful means to consider the person opposite as bearer of morally significant properties and to act accordingly. This means that one always acts in acceptance of these properties of his or her opposite. Such properties are e.g. sentience or rationality. In recognizing one's sentience, one is never allowed to cause pain. A further morally significant property is personal autonomy. Every person is bearer of such an autonomy. Umbers understands personal autonomy as self-governance. [93] p. 561. Autonomous agents act based on reasonable decisions. For a political system respectively for the state, a citizen's personal autonomy needs to be encountered in an adequate political system which takes account of everyone's ability to reason and contribute to the state. When it comes to vote buying, personal autonomy means that each individual voter comes to a conviction (which could also be the absence on a vote) by well-reflected reasoning. The acceptance of such an autonomy therefore means that one does not interfere in this process of opinion building in a way that corrupts the opposite's ability to reason. Setting up a monetary contract, which vote buying is - whether direct or indirect -, Umbers argues, corrupts this respect on two levels. First, the person who is paying money to get a vote fails to recognize the political opponent's ability to reason. He or she sets a monetary and therefore not intellectual respectively argumentative incentive to convince a voter. The purchase of a vote means that the voter is paid to ignore the opponent's reasoning at all and vote as directed in the contract: This means that with the buying, the respect towards every actor participating in the debate around a voting process is lost. It is wrong to buy votes and therefore ignoring the conviction work of political opponents. Secondly, the person selling his vote is no more taken serious as an autonomous citizen. He or she isn't voting based on conviction but only based on a direct economic incentive. Therefore, not only the ability of moral reasoning of the political opponent but also the voter's ability is over gone. Umbers concludes that vote buying is therefore not legitimate and underlines importance of the banning of vote buying which was brought up by real world democracies to encounter the phenomena. For sure, Umbers argues, there are cases where the contempt of voter's respect can be morally legitimate because the consequences of a bad outcome of a political decision can be predicted [93] p. 570 - what Umbers considers as an exception is somehow similar to what

Brennan presents as the premise in hardly any political decision: The (non-) existence of predictability of outcomes on political decisions. The final objection which shall be presented in this paper comes back to the efficiency argument most clearly brought up by Freiman. The emerita law professor Margaret Jane Radin formulated a theory on the limits of markets. In her main book “Contested Commodities” [71], which was published in 1996, Radin focuses on the question which spheres of life can't be transferred in a system of purely economic incentives. Radin observes that the dominance of market logic is omnipresent and leads the way people think and speak also of extra economic processes. She understands the evaluation of anything with economic terms and monetary categories as commodification. Commodification describes beyond the real-world trade of goods also the definition of what can be considered as a tradable good and the way of talking about society as a whole in economic terms. Furthermore, Commodification includes that every aspect of society is quantified. Only with a quantified evaluation, decisions can be made in accordance with an economic rationale. This finally changes the way people act in fields where moral convictions and categoric imperatives would determine the decision making. “Economic rhetoric applied to such practices as free speech, tort compensation, and democratic politics is therefore not innocent.” In the first part of this paper where people's motivation to go voting is explained as a formula, also this paper perpetuates what Radin criticizes. The calculus model suggests that there is a trade-off behind the decision of going to vote. There's something like an incentive or benefit who makes people vote. Bringing the process of voting and democratic participation down to an economical formula, is strongly connected to this idea of rationalizing and quantifying social aspects of life. Furthermore, it stands in contrast to the position that the democratic decision making is based on principles such as a civic duty to participate in the society one is living in. Radin argues that the quantified economical view erodes the honest motivation of each citizen to think autonomously about political decisions. This erosion of a principal value of the democratic participation endangers the stability of such institutions. Communities which are based on the political engagement of its citizens who together constitute the community. The stability and legitimacy of the state holding together the community and being defined as a consequence of their decisions is endangered. Radin sets the focus back to principal values and concludes that these are being eroded as soon as a monetary commodification of votes comes in.

3.4.3 Tradition of thought: When Philosophy becomes political

For now, the conclusion can be made that there are strong philosophical counterarguments which are brought up against vote buying. Whereas Brennan tried to form a basis for a legitimate debate on vote buying, Freiman extended the general legitimacy and promoted the idea of a voting market. As explained with three different examples the opponents on vote buying argue on a basis guided by principal convictions on democratic systems. Most strongly, Radin brought up the general critique on commodification. Although Brennan and Freiman aren't directly referred in Radin's critique since she formulated her work in the late 80s whereas the discussion on legalizing vote buying was quite recently launched, their way of thinking is deeply connected to the school of thoughts Radin addressed. For this section, a little insight into the philosophy behind commodification and the networks which promoted those ideas historically shall be given. Although philosophy thinks of politics in a conceptual way and develops theories on a meta level, it isn't disconnected to political ideologies. Brennan's and especially Freiman's theories stand in a line with political philosophers and the founding fathers of neoliberal thought. To get an insight into this world of ideas, some of their platforms shall be presented in the following section. It will be discussed how these platforms provide an epistemic network for the development of ideas which are politically coined in a liberal way. Although Brennan does explicitly not

argue for the legalization of vote buying, he is personally very attached to the libertarian worldview. [11] p. 137. As presented in his theory on vote buying, he underestimates the decisive role of an individual in a regulated community. In states rules defined by the people can lead to a better life of the individual because of its own participation. The perspective on social systems which negates the importance of the individual's importance towards the community is typical for a liberal position. [86] Not only Brennan's ideas can be affiliated to liberal positions, he also personally engages for a liberalism, to be precise for libertarianism. Other than neoclassical liberalism or neoliberalism, libertarians follow consequently the premise of a self-regulated systems and therefore demand for a society which is completely dependent on markets. They negate all state interventions, especially when it comes to taxation. They consider taxation as a theft. The Bleeding Heart Libertarians strongly believe in the compatibility of radical market freedom and social justice being achieved in such a society. [8] It can be shown that that Brennan's postulate of a well-defined common good which we have seen in the explanations on his theory on vote buying resides in a broader and politically coined epistemic field of liberal respectively libertarian thought. Also in Switzerland, Brennan's ideas are referenced in the traditional channels of liberal thinking: The conservative magazine (*Schweizer Monat*) which, having a dark and partly fascist background, promoted liberal ideas after the Second World War as a platform. They brought together national and international expertise in economics. [15] [105] Stars of neoliberal economic thought such as Ludwig von Mises, Wilhelm Röpke and Friedrich August von Hayek were often referenced or even wrote articles by themselves. The magazine prominently references Brennan's person and political theories in order to criticize democratic systems of equally participating citizens. [13] [79] In the same field of thinking, Christopher Freiman needs to be situated. Also, he is an active contributor to the association of libertarian thinkers. [8] Furthermore, he engages, as Brennan does as well, on the knowledge transfer platform Learn Liberty. [55] In his articles, he promotes theories that explain the weakness of democratic voting systems. He explains how politicians need to make promises which aren't in favour of the common good because they must satisfy certain interest groups to get their votes. [34] To do so, he comes back to the well-known economic public choice theory which was prominently coined by thinkers such as James Buchanan and Gordon Tullock who wrote *The Calculus of Consent*. [14] Also in his work on vote buying, Freiman references Tullock and Buchanan [35] p. 764. The public choice theory, reflections on the quantifiable decision making and Freiman's effort for vote buying market stand in one direct line of argumentation. Also, Buchanan and Tullock were prominently engaging in different liberal think tanks. They contributed among economists such as the before named Hayek but also Milton Friedman etc. to the rise of neoliberalism to the most dominant economic way of thinking since the late 70s until the 90s. [69] This excursion serves as historical background to pin down the philosophical ideas which seem to be developed in the void of theoretical thinking on democracy to the reality of politics, interest and networks. All of these are strongly intertwined by strong ideological convictions which can be traced back in some of the premises which were being presented and criticized. As explained before, there are thinkers such as Radin who speak out against these schools of thought, holding up principles of democratic communities against the logic of markets. [99] Of course, also the critique of vote buying and in a broader sense the critique on the rationalisation of social and political processes is grounded in ideological convictions which cannot claim an absolute truth. However, what could be shown in this outline of different views on vote buying, is that the premise of a well-defined common good which probably all claim to be willing to promote can hardly live up to different perspectives. Shall vote buying which results in the dispose of one's vote and the citizen's responsibility be allowed?

3.4.4 Is there a real democracy?

To conclude this section, the discussion on vote buying shall be placed in a lot broader discussion about democracy. One strong argument which is always brought up against vote buying is the one worrying about inequality, as the part to general and real-world reflections already announced. It considers the poorer people to be a lot more prone to be corrupted by vote buyers. This critique suggests that people who are poorly financed are more likely to accept a deal and sell their votes. This important and quite concrete critique needs to be resided in a more extensive critique on existing democracies. Democratic institutions as of today issues concerning unequal chances of democratic participation. Starting by untransparent party finances, ending by the fact that some part of society is in precarious financial need, works hard and doesn't have the time to get properly informed. Neither they understand themselves as a political force in modern society. [47] Not to mention how desperately dependent they could be on the money they get from selling their votes. Adorno and his Frankfurter Schule wrote in the postwar Germany about democracy and consumer capitalism. In a lecture he gave on movements of the radical right, Adorno states: "Dem Inhalt nach, dem gesellschaftlichen-ökonomischen Inhalt nach, [hat sich] die Demokratie eben bis heute nirgends wirklich und ganz konkretisiert, sondern [ist] formal geblieben." The statement expresses how the core of democracies, namely the equality of the people, has not been consequently established in modern states. Adorno isn't thinking of a communist state but rather of a society in which economic differences have no influence on the participation and power distribution within the political system. In his lecture, he concludes that materialistic differences lead to the formation of a group of people being neglected by the economic and social development which have to potential to follow dangerous promises of radical nationalist ideologies. For the question on vote buying respectively on the promise of democracy which implies the same participation rights for all, Adornos disillusioning remark is resounding. Without having market mechanisms such vote buying at hand, existing democracies have a deficit although they allow equal participation in voting. On the other hand, Adorno's critique allows an interesting thought: Might the discussion on vote buying be a different one if we'd live in a democracy as Adorno is indicating it? Given a society of very weak economic as well as social differences, vote buying where everyone could participate as buyer or seller equally might lead to a more fine-grained democratic decision process. Since we are away from having robust democratic systems - not to mention intransparent party financing even in democracies such as Switzerland - these question remains hypothetical.

3.5 Conclusion

This last section concludes all of the topics discussed before. It could be revealed that the secret ballot was introduced to circumvent vote buying but is not always perfectly successful in doing so. By introducing the calculus of voting model, it was possible to show why people vote and that people may see various different reasons in the additional parameter D. But the thing they have all in common is that the benefits outweigh the costs of voting. Hence, they vote which implies vote buying is possible. With that knowledge it can be stated that the different participants in buying a vote have different conceptions of such an exchange. While the givers try to enforce different forms of compliance with various strategies, the voters are more interested in the different meanings the offer can carry. With vote buying being practiced, also different problems arise. There are three different major problems which are the inequality, the question whether vote buying increases social wealth or not and the inalienability argument discussing the correctness of the whole process. To sum up, the inequality argument talks about the fact that a selling a vote brings more benefit for the poor people than for the wealthy and thus the election could

be biased. The social wealth argument addresses the question of increasing or decreasing social wealth of vote buying. It turns out, that there is the risk of negative externalities and thus, it is not eligible.

In the following subsection the paper took a closer look at vote buying in the past. For this, it takes several examples. Starting with a sheriff in the USA who won an election by classical vote buying. Subsequently, the paper shows some historical examples of tangible rewards as exchange for votes. Afterwards, the focus changes to the history of vote buying itself. It is shown that vote buying exists since 1492 and that even George Washington, the father of the USA bought votes in 1756. His namesake George Washington Plunkitt shows that not every good action is philanthropy. For the help in the community G. W. Plunkitt demanded a return service in the form of votes. The historical subsection is followed by two large examples of vote buying in the more modern past. First a paper which took a closer look at vote buying in Kenya in 2002 is analyzed. It could be shown that 40% of the probands were asked to sell their votes. Additionally, it is shown which kind of strategies the political parties in Kenya in 2002 applied to buy votes. The paper analyzed the strategies and proofed that they were good, but not perfect. In the last two parts of the past subsection, the paper focuses on the US presidential elections from 2000 and 2016. In 2000, an art project showed that there was a large market to buy and sell votes. They proofed this by a seller platform, where voters could sell their votes. All the votes were taken together and sold to the highest bidder. It has become clear that in the 2000 presidential election there was no correlation between wealthiness of a state and the value of a vote. Afterwards, the paper analyzes the 2016 presidential election. It focuses on the case of Cambridge Analytica and Donald Trump. It has become clear that in this election there was no vote buying, but there happened a lot of ethically controversial aspects. It has been shown that social media and the campaigns using it, had a massive influence on the undecided voters in swing states. Donald Trump won the most of them shortly before the election, thanks to social media.

Another topic which addresses vote buying in the present and the past is the circulation of money. Depending this topic, two different views can be distinguished which are the long- and the short-term view. In the long-term it has been shown that signs of intervention with the fiscal and monetary cycles have been recorded. However, they tend not to be very strong in OECD democracies but are observed to be influential in countries with not that developed democracies. In the short-term, it looks similar. No influence can be seen on M1 in OECD democracies, while in the low-developed countries an increase of 0.6-0.7 percentage points of M1 is observable. This increase is best explained with it being considered as systematic vote buying.

In the last subsection of the first part, an excursus concerning e-voting in Switzerland has been made. It is shown that the e-voting system tested in 10 cantons had a broad acceptance in the public. Nevertheless, the Swiss Post had the government to end the testing phase as the used code had enormous security vulnerabilities. Nowadays, the Post is looking forward to 2021, when they are going to test the new advanced program, which should have closed these vulnerabilities.

Countermeasures against vote buying in voting systems are heavily researched in scientific and cryptographic scenes. It has been shown that using a process like the ThreeBallot or the VAV protocol are mitigating vote buying by achieving receipt-freeness. By showing the voting protocol of Hirt and Sako and the JCJ scheme, it gets clear that cryptographic solutions achieving receipt-freeness, respectively coercion resistance, some assumptions and requirements has to be done beforehand. Also, it is shown that implementing a secure, verifiable and coercion resistant voting system is very difficult to achieve. Sometimes trade-offs has to be done between security and anonymity. The research in blockchain voting systems is situated in a very early stage, already promising secure, private and coercion resistance voting processes.

The philosophical reflections on vote buying are disparate and diverse. There is a strong school of philosophers, which promotes firstly the principal legitimacy of vote buying by trying to prove that there's not an intrinsically valid ban on selling and buying one's vote. The philosopher who most prominently worked on this, Jason Brennan, argued on the premise of being able to define a common good which needs to be promoted by any means. If vote buying takes place in accordance with the promotion of the common good, it can be done. On this principal allowance, the idea of installing real-world voting markets was defended. Based on theories of mutual profit and the advantages of commodification vote buying was extended to a concept of improved democratic decision making with market mechanisms at hand. These concepts are encountered by diverse criticism. Firstly, vote buying ignores the statement which is done with the voting behaviour of uninfluenced people. Whether one votes or abstains, he or she is participating through this behaviour. Secondly, the idea that there is a predefined common good, which to promote by buying votes, is completely legitimate needs to be questioned, since democracy is about different ideologies and reasoning, which in a permanent discussion result in politics. Thirdly, the respect for people's autonomy is brought up as a strong argument against vote buying. Since mature citizens owe each other respect in the social coexistence and political formation of the state, which means that one is aware of and recognizes the ability to reason autonomously, vote buying isn't legitimate. As a purely financial incentive it ignores the political opponent's ability to convince as well as the seller's ability to build up an autonomous decision. Fourthly, economic terms erode on the democratic principles of participation. Market logic as it would be introduced to voting if a vote buying market was installed corrupts the way people live together in democracies.

The discussion which promotes vote buying could then be situated in a long tradition of economic thought. It is strongly intertwined to theories of commodification and the public choice theory. Furthermore, it could be shown that the actors appearing in the debate work together on different platforms to bring forward libertarian ideas. Finally, the considerations on inequality were again brought up to conclude the discussion on vote buying. As referenced in the practical part as well, inequality remains one of the core worries for the effectiveness of democracies. Referring to Frankfurter Schule's most prominent scholar, the premise of existing democracies was questioned. Reflecting it on the surface of the discussion of vote buying, it could be concluded that real-world democracies are far away from a system where vote buying could be considered realistically as a mean to improve democratic decision making.

Bibliography

- [1] T. W. Adorno: *Aspekte des neuen Rechtsradikalismus. Ein Vortrag*, Berlin, 2019.
- [2] T. S. Aidt, Z. Asatryan, L. Badalyan and Friedrich Heinemann: *Vote Buying or (Political) Business (Cycles) as Usual?*; Social Science Research Network, 2015
- [3] A. Alesina, G. D. Cohen and N. Roubini: *Macroeconomic Policy and Elections in Oecd Democracies*; Economics & Politics, 1992, pp. 1-30
- [4] A. Juels, D. Catalano , M. Jakobsson: *Coercion-Resistant Electronic Elections*, 2011 IEEE Symposium on Security and Privacy. RSA Laboratories Bedford, MA, USA
- [5] M. Baek: *A Comparative Analysis of Political Communication Systems and Voter Turnout*; American Journal of Political Science, 2009, pp. 376-379
- [6] S. Bechtold: *Governance in Namespaces*; paper, <http://dx.doi.org/10.2139/ssrn.413681>;
- [7] Bleeding Heart Libertarians: Author: Christopher Freiman; <http://bleedingheartlibertarians.com/author/christopher-freiman/>, November 2020.
- [8] Bleeding Heart Libertarians: *About Us*; <http://bleedingheartlibertarians.com/about-us/>, November 2020.
- [9] B. A. Block: *Political Business Cycles, Democratization, and Economic Reform: The Case of Africa*; Journal of Development Economics, 67(1), February 2002, pp. 205-228
- [10] I. Bogost and A. C. Madrigal: *How Facebook Works for Trump*, <https://www.theatlantic.com/technology/archive/2020/04/how-facebooks-ad-technology-helps-trump-win/606403/>, 18. April 2018
- [11] J. Brennan: *The Ethics of Voting*; Princeton, 2011.
- [12] J. Brennan: *Against Democracy*; Princeton, 2017.
- [13] J. Brennan: *Wie Demokratie (wirklich) funktioniert*; Schweizer Monat, 2018, <https://schweizermonat.ch/wie-demokratie-wirklich-funktioniert/>.
- [14] J. M. Buchanan; G. Tullock: *The Calculus of Consent. Logical Foundations of Constitutional Democracy*, Michigan, 1962.
- [15] W. Burckhardt: *Unsere Einstellung zu Deutschland*; Schweizer Monatshefte, 1936, p. 187.
- [16] F. A. Burkle-Young: *The Cardinals of the Holy Roman Church*; <http://cardinals.fiu.edu/>, 20. October 2020

- [17] M. Buess, A. Ramsden and O. Bieri: *Nationale E-Government-Studie 2019*; population survey, Demo SCOPE AG
- [18] C. Cadwalladr: *The great British Brexit robbery: how our democracy was hijacked*; <https://www.theguardian.com/technology/2017/may/07/the-great-british-brexit-robery-hijacked-democracy>, 7. May, 2017
- [19] W. Callahan and D. McCargo: *Vote-Buying in Thailand's Northeast: The July 1995 General Election*; Asian Survey, 36(4), 1995, pp. 376-392
- [20] M. Clarkson et. al.: *Civitas: Toward a Secure Voting System*. 2008 IEEE Symposium on Security and Privacy (SP 2008). pp. 354-368.
- [21] CNN: *exit polls 2016*; <https://edition.cnn.com/election/2016/results/exit-polls/national/president>, 23. November 2016
- [22] CNN: *exit polls 2016*; <https://edition.cnn.com/election/2016/results/exit-polls/florida/president>, 23. November 2016
- [23] CNN: *exit polls 2016*; <https://edition.cnn.com/election/2016/results/exit-polls/wisconsin/president>, 23. November 2016
- [24] CNN: *exit polls 2016*; <https://edition.cnn.com/election/2016/results/exit-polls/pennsylvania/president>, 23. November 2016
- [25] CNN: *exit polls 2016*; <https://edition.cnn.com/election/2016/results/exit-polls/michigan/president>, 23. November 2016
- [26] G. W. Cox and M D. McCubbins: *Electoral Politics as a Redistributive Game*; The Journal of Politics, 48(2), May 1986, pp. 370-389
- [27] P. de Rosa: *Vicars of Christ: The Dark Side of the Papacy*; Book, Corgi, 1989
- [28] Delaune Stéphanie, Kremer Steve, Ryan Mark: *Coercion-Resistance and Receipt-Freeness in Electronic Voting*
- [29] C. DeNavas-Walt, W. Cleveland, M. Roemer: *Money Income in the United States: 2000*; report, U.S. Department of Commerce, p.17
- [30] J. Derfner: *Buy This Vote!*; <https://slate.com/news-and-politics/2000/08/buy-this-vote.html>
- [31] R. J. Dinkin: *Campaigning in America*; Book, Greenwood Press, 1989
- [32] A. Downs: *An Economic Theory of Political Action in a Democracy*; Journal of Political Economy, 65(2), April 1957, pp. 135-150
- [33] M. P. Fiorina: *The 2016 Presidential Election - An Abundance of Controversies*; essay, hoover intititution, series No. 10
- [34] C. Freiman: *Why you vote for corn syrup even though it might be killing you*; Learn Liberty, 2017, <https://www.learnliberty.org/blog/why-we-cant-get-rid-of-farm-subsidies/>.
- [35] C. Freiman: *Vote Markets*; Australasian Journal of Philosophy, 2014, pp. 759-774, <http://dx.doi.org/10.1080/00048402.2014.892147>

- [36] R. J. Gonzales: *Hacking the citizenry?*; magazine article, ANTHROPOLOGY TODAY VOL 33 NO 3, JUNE 2017
- [37] George Town University: *Jason Brennan*; <https://gufaculty360.georgetown.edu/s/contact/00336000014RXIUAA4/jason-brennan>, November 2020.
- [38] A. S. Gerber, D. P. Green and C. W. Larimer: *Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment*; American Political Science Review, 102(1), February 2008, pp. 33-48
- [39] A. S. Gerber, G. A. Huber, D. Doherty, und C. M. Downling: *Is There a Secret Ballot? Ballot Secrecy Perceptions and Their Implications for Voting Behaviour*; British Journal of Political Science, 43(1), January 2013, pp. 77-102
- [40] G. B. Gist: *Progressive Reform in a Rural Community: The Adams County Vote-Fraud Case*; The Mississippi Valley Historical Review, 48(1), 1961, pp. 60-78
- [41] R. L. Hasen: *Vote Buying*; California Law Review, 88(5), 2000, pp. 1323-1372
- [42] The heritage Foundation: *A sampling of election fraud cases from across the country*; state publication, 07. October 2020
- [43] F. Hjalmarsson, G. K. Hreijarsson, M. Hamdaqa and G. Hjalmtysson: *Blockchain-Based E-Voting System*; 2018 IEEE 11th International Conference on Cloud Computing (CLOUD), San Francisco, CA, 2018, pp. 983-986.
- [44] M. Hirt and K. Sako *Effcient Receipt-Free Voting Based on Homomorphic Encryption* EUROCRYPT 2000, pp. 539-556.
- [45] F. Imbach: *Schwarzer Tag für das E-Voting in der Schweiz*; <https://www.srf.ch/news/schweiz/genfer-kehrtwende-schwarzer-tag-fuer-das-e-voting-in-der-schweiz>, 28. November 2018
- [46] L. Jacobsen: *Alle Macht für Mister Spock*; Zeit Online, April 2017, <https://www.zeit.de/politik/deutschland/2017-04/jason-brennan-wahlsystem-demokratie-ungebildete-wahl-verbieten/komplettansicht>.
- [47] U. Kadritzke: *Klassenkampf oder Tanz in den 1. Mai?*; Deutschlandfunk Kultur, April 2019, https://www.deutschlandfunkkultur.de/soziologe-ulf-kadritzke-zum-klassenbewusstsein-klassenkampf.1013.de.html?dram:article_id=447609.
- [48] J. Koetsier: *This Big Data Marketing Firm Claims To Have A Perfect Track Record In Winning Elections*, <https://www.forbes.com/sites/johnkoetsier/2017/11/09/trumps-election-data-firm-has-100-track-record-in-winning-elections/#5574f56c5c91>, 9. November, 2017
- [49] E. Kramon: *VOTE-BUYING AND POLITICAL BEHAVIOR: ESTIMATING AND EXPLAINING VOTE-BUYING EFFECT ON TURNOUT IN KENYA*; Working Paper, Afro Barometer, Nr. 114, October 2009
- [50] K. Krips and J. Willemson: *On practical aspects of coercion-resistant remote voting systems*, Estonia, University of Tartu
- [51] R. Kürnzi: *E-Voting-Moratorium vom Tisch*; <https://www.swissinfo.ch/ger/schweiz-demokratie-abstimmung-e-voting-moratorium-volksinitiative-zurueckgezogen/45855362>, 23. June 2020 ”

- [52] Kuesters Ralf, Truderung Tomasz, Vogt Andreas: *Verifiability, Privacy, and Coercion-Resistance: New Insights from a Case Study*, 2011 IEEE Symposium on Security and Privacy.”
- [53] M. Lacey: *Panel Tries Hard to Keep Kenya Vote Aboveboard*; <https://www.nytimes.com/2002/12/23/world/panel-tries-hard-to-keep-kenya-vote-aboveboard.html>, 23. December 2002
- [54] Learn Liberty: *Christopher Freiman*; <https://www.learnliberty.org/speakers/chris-freiman/>, November 2020.
- [55] Learn Liberty: *Chris Freiman*; <https://www.learnliberty.org/speakers/chris-freiman/>, November 2020.
- [56] D. leip and D. Wasserman: *Presidential Election Results: Donald J. Trump Wins*; <https://www.nytimes.com/elections/2016/results/president>, 9. August 2017
- [57] S. Levine: *Florida's attorney general requests inquiry into Mike Bloomberg's voting effort*, The Guardian, London, September 2020, <https://www.theguardian.com/us-news/2020/sep/24/florida-inquiry-michael-bloomberg-voting-effort>.
- [58] P. Lewis and P. Hilder: *Leaked: Cambridge Analytica's blueprint for Trump victory*; <https://www.theguardian.com/uk-news/2018/mar/23/leaked-cambridge-analyticas-blueprint-for-trump-victory>, 23. March, 2018
- [59] K. Lippert-Rasmussen: *Vote Buying and Election Promises: Should Democrats Care About the Difference?*, Oxford, The Journal of Political Philosophy 19, 2, 2011, pp. 125-144.
- [60] R. Maclean, E. Egbejule: *Nigeria election marred by vote buying, tech failures and violence*; The Guardian, London, February 2019, <https://www.theguardian.com/world/2019/feb/23/nigeria-election-goes-ahead-amid-violence-tech-failures>.
- [61] Ü. Madise and T. Martens: *E-voting in Estonia 2005. The first practice of country-wide binding Internet voting in the world*; Krimmer, R. (Hrsg.), Electronic Voting 2006 - 2nd International Workshop, pp. 15-26
- [62] J. H. Morrison *The political philosophy of George Washington*; Book, The Johns Hopkins University Press, 2009
- [63] R. Murray *Protected Political Speech or Treason?*; Journal, Journal of High Technology Law, 5, 357-380
- [64] S. Nicther: *Vote Buying or Turnout Buying? Machine Politics and the Secret Ballot*; The American Political Science Review, 102(1), 2008, pp. 19-31
- [65] W. D. Nordhaus: *The Political Business Cycle*; The Review of Economic Studies, 42(2), 1975, pp. 169-190
- [66] D. Okrent: *The Rise and Fall of Prohibition*
- [67] P. H. O'Neill: *The Russian hackers who interfered in 2016 were spotted targeting the 2020 US election*; <https://www.technologyreview.com/2020/09/10/1008297/the-russian-hackers-who-interfered-in-2016-were-spotted-targeting-the-2020-us-election/>, 10. September 2020

- [68] J. Pilcher: *The Art of Voting*; Paper, January 2015
- [69] D. Plehwe, B. Walpen: *Wissenschaftliche und wissenschaftspolitische Produktionssweisen im Neoliberalismus. Beiträge der Mont Pèlerin Society und marktradikaler Think Tanks zur Hegemoniegewinnung und -erhaltung*; PROKLA, Zeitschrift für kritische Sozialwissenschaft, pp. 203-235.
- [70] Post: *Publikationen und Quellcode*; <https://www.post.ch/de/geschaeftsloesungen/e-voting/publikationen-und-quellcode#offenlegungsquellcode>, Website, 05. November 2020
- [71] M. J. Radin: *Contested Commodities*, Cambridge 1996.
- [72] B. Rigender: *Dieser Code entspricht schlicht nicht dem Standard*; <https://www.swissinfo.ch/ger/direktedemokratie/schweizer-e-voting-dieser-code-entspricht-schlicht-nicht-dem-standard--/44820334>, 13. March 2019
- [73] W. H. Riker and P. C. Ordeshook: *A Theory of the Calculus of Voting*; American Political Science Review, 62(1), March 1968, pp. 25-42
- [74] W. L. Riordan: *Plunkitt of Tammany Hall: A Series of Very Plain Talks on Very Practical Politics*; Book, Penguin, 01. November 1995
- [75] K. Rogoff: *Equilibrium Political Budget Cycles*; American Economic Review, 80, 1990, pp. 21-36
- [76] K. Rogoff and A. Sibert: *Elections and Macroeconomics Policy Cycles*; The Review of Economic Studies, 55(1), January 1988, pp. 1-16
- [77] F. C. Schaffer: *Might Cleaning Up Elections Keep People Away from the Polls? Historical and Comparative Perspectives*; International Political Science Review, 23(1), January 2002, pp. 69-84
- [78] F. C. Schaffer: *What is Vote Buying?*; Center for International Studies MIT, 2007
- [79] Schweizer Monat: *Jason Brennan*; Schweizer Monat, <https://schweizermonat.ch/author/jason-brennan/>.
- [80] P. Sheehy: *A Duty Not to Vote*; Ratio (15, Nr. 1), March 2020, pp. 46-57.
- [81] M. Sheets: *Four men admit to paying homeless people in LA's Skid Row cash and cigarettes in exchange for forged signatures on ballots and voter registration forms*; <https://www.dailymail.co.uk/news/article-8578247/Four-men-admit-paying-homeless-people-LAs-Skid-Row-join-voter-fraud-scheme.html>, 30. July 2020
- [82] N. Silver: *The invisible undecided voter*; <https://fivethirtyeight.com/features/the-invisible-undecided-voter/>, 23. January 2017
- [83] Der Spiegel: *Richter verbietet virtuellen Stimmenhandel*; <https://www.spiegel.de/netzwelt/web/us-praesidentschaftswahlkampf-richter-verbietet-virtuellen-stimmenhandel-a-98857.html>, 19. October 2000
- [84] O. Spycer et. al.: *A New Approach Towards Coercion-Resistant Remote E-Voting in Linear Time* LNCS 7035, pp. 182-189.

- [85] S. C. Stokes: *Perverse Accountability: A Formal Model of Machine Politics with Evidence from Argentina*; The American Political Science Review, 99(3), 2005, pp. 315-325
- [86] J. Strasser: *Keine Macht den Dummen. Warum Jason Brennan die Demokratie abschaffen will*; Neue Gesellschaft Frankfurter Hefte, 2017, <https://www.frankfurter-hefte.de/artikel/keine-macht-den-dummen-2417/>.
- [87] C. Strünck: *Wie viel wert ist eine Stimme? Wettbewerb und Wettbewerbsverzerrungen im Wahlsystem der USA*; anthology, VS Verlag für Sozialwissenschaften, 1. edition, August 2006, 146-169
- [88] Swiss Government: *CC 311.0 Swiss Criminal Code*; December 1937
- [89] Tassos Dimtiriou: *Efficient, Coercion-free and Universally Verifiable Blockchain-based Voting*, Cryptology ePrint Archive, Report 2019/1406 <https://eprint.iacr.org/2019/1406>
- [90] G. Tullock: *Toward a mathematics of politics*; Universtiy of Michigan Press, 1967
- [91] W. Turnherr: *Vote électronique*; <https://www.bk.admin.ch/bk/de/home/politische-rechte/e-voting.html>, 06. October, 2020
- [92] Übermorgen.com: *Vote-Auction*; <https://www.vote-auction.net/>, 24. October 2020
- [93] L. M. Umbers: *What's Wrong With Vote Buying*, Philosophical Studies, 177 (2), 2020, pp. 551-571.
- [94] L. M. Umbers: *Lachlan Umbers*; <https://lachlanumbers.com/>, November 2020.
- [95] U.S. Attorneys Office: *Former Dodge County Sheriff and Deputy Sentenced for 2004 Election Fraud*; <https://archives.fbi.gov/archives-atlanta/press-releases/2010/at062810b.htm>
- [96] US Government: *18 U.S. Code Chapter 29 - ELECTIONS AND POLITICAL ACTIVITIES*; January 1980
- [97] M. Vidler: *Cambridge Analytica: A question of ethics not legality*; <https://www.decisionmarketing.co.uk/views/cambridge-analytica-a-question-of-ethics-not-legality>, 26. March 2018
- [98] Washington Post Staff: *Full text: Donald Trump announces a presidential bid*; transcript, <https://www.washingtonpost.com/news/post-politics/wp/2015/06/16/full-text-donald-trump-announces-a-presidential-bid/>
- [99] D. Wasserman: *Radin, Margaret Jane. Contested Commodities*; Ethics, 1999.
- [100] the white house: *The Constitution*; <https://www.whitehouse.gov/about-the-white-house/the-constitution/#:~:text=The%20First%20Amendment%20provides%20that,the%20right%20to%20bear%20arms.>, 30 October 2020
- [101] J. A. Widner: *The Rise of a Party-State in Kenya*; Book, Berkeley: University of California, 1992
- [102] D. Wurfel: *Southeast Asia. Pollwatching, elections and civil society in Southeast Asia*; The Journal of Southeast Asian Studies, 34(1), February 2003, pp. 159-193

- [103] A. Yakobson: *Secret Ballot and Its Effects in the Late Roman Republic*; Hermes, 123(4), 1995, pp. 426-442.
- [104] Yu, Bin et al.: *Platform-independent Secure Blockchain-Based Voting System*, 2011 IEEE Symposium on Security and Privacy. IACR Cryptol. ePrint Arch.
- [105] A. Zimmermann: *Der "Schweizer Monat". Reaktionär seit 1921*; Wochenzeitung, 2011.

Chapter 4

Wireless Sensing Marketshare

Imami, F., Kaushik, R., & Zimmermann, T.

As the global business interest in Wireless localisation techniques continues to increase, this technology is reaching new levels of accuracy and ease of deployment. This localisation technology enables the near-real-time positioning of a device or physical object by analysing the radio signals in the environment. As the market share of IoT devices and mobile devices continues to increase, companies which provide such wireless localisation technologies promise to disrupt the market by enabling more in-depth data about person or physical object locations and movement throughout their network of sensors. To deliver a better understanding of the current technology's state, this report aims to review the current literature regarding wireless localisation techniques and the current market of such technologies. The presented literature review allows the report to then conclude with a review of a selection of companies and their products in order to observe how their solutions position themselves within the market.

Contents

4.1	Introduction	108
4.2	WLAN based localisation techniques and approaches	109
4.2.1	General attributes and techniques of WLAN localisation	109
4.3	Fingerprinting	112
4.3.1	RSS based fingerprinting	112
4.3.2	Difficulties with WLAN based fingerprinting	114
4.3.3	Accuracy of WLAN based fingerprinting	115
4.3.4	Magnetic field fingerprinting	115
4.3.5	Map-Aided Indoor Positioning	116
4.3.6	Spatial and Temporal Signal Patterns	116
4.3.7	Collaborative localisation Among Mobiles	116
4.3.8	Motion-Assisted localisation	116
4.3.9	Commercial WLAN fingerprinting Systems	117
4.4	Alternatives to WiFi localisation	117
4.4.1	Bluetooth localisation	117
4.4.2	Radio Frequency Identification Service	118
4.4.3	ZigBee based wireless sensor networks	118
4.4.4	Acoustic Signal based localisation	118
4.4.5	High Sensitive GNSS	119
4.5	Marketshare of Wireless sensing applications	119
4.5.1	Current Market & Trends	119
4.5.2	Common Use Cases	121
4.6	Case Study: Companies using Wireless Sensing	125
4.6.1	Locatee AG	125
4.6.2	Onway AG	126
4.6.3	Livealytics AG	126
4.6.4	Skyhook, Inc.	127
4.6.5	CLOUD4WI, Inc.	128
4.6.6	Combain	128
4.6.7	Comparison	129
4.7	Privacy aspect of RSSI fingerprinting	129
4.8	Future Trends in Radio based Positioning	130
4.8.1	Machine learning improvements	130
4.8.2	IoT development	132
4.8.3	UWB	132
4.8.4	Combination of the three	133
4.9	Survey analysis	133
4.10	Final Considerations	133
4.11	Appendix I: Survey results	135

4.1 Introduction

Cisco annual Internet report of 2020 forecasted that the number of WiFi hotspots will grow four-fold from 2018 to 2023, totalling over 628 million public WiFi hotspots and that by 2023 there will be 8.7 billion handheld or personal mobile-ready devices [1]. The demand and usage of WiFi and Bluetooth enabled fitness tracking devices, such as Apple Watch, FitBit or Samsung Gear, is also growing and in 2018 reached a 10% adoption rate in the US. Globally the number of connected wearable devices worldwide has more than doubled in the space of three years, increasing from 325 million in 2016 to 722 million in 2019 [2]. The number of devices is projected to reach more than one billion by 2022. If Cisco 2023 forecast holds, it will mean that each person worldwide will have more than one WiFi capable portable device.

As WiFi devices are getting more widespread, so is indoor location tracking. Indoor positioning technologies are based on both WiFi or Bluetooth, or a combination of the two, the main focus of such technologies is to pinpoint the position of a WiFi-enabled device accurately. Over the years, both software and hardware matured, and the accuracy has quickly increased and is now capable of identifying the position with a 1-meter range of error [3]. This has attracted researchers and commercial developers, and now multiple companies offer commercially viable products for indoor localisation, such as RUCKUS, Extreme Networks, Skyhook. However, commercial solutions provided by these companies have an accuracy that ranges from 5-meters (Skyhook)[4] to 20-meters (RUCKUS)[5]. Indoor localisation products can help companies gather vital data about the customer's movement, position within their spaces, information about the time of arrival and departure or information about the time spent around a certain point. Afterwards, this data can be used to tailor the on-location experiences of the customers in order to drive up sales and improve the customer's satisfaction.

Another use of the data generated by indoor localisation products is to monitor open crowded areas, which is key for policymakers to set up the proper measures for people's security and safety [6]. This type of use is already widely adopted in airports, stations, event venues and other sensitive areas [4].

This report aims to review the current literature about wireless localisation techniques and the current market of these technologies. In order to do so, at first, this report will present the reader the different technological approaches which are used for wireless fingerprinting and localisation. The literature review is then followed by a presentation of selected real-world use case examples and an examination of the current products available on the market. The report then concludes with a brief overview of the future implications and technological developments in order to place the final discussion within a broader and future context.

Starting in section 4.2, this report presents an overview of the existing approaches that use Wireless Local Area Network (WLAN) for localisation and the general attributes that are used in the localisation algorithm. The following section introduces the fingerprinting concept, which aims to estimate location based on signals received or emitted by devices. This is accompanied by a overview of different types of fingerprinting and the annexed challenges which are encountered during the fingerprinting process. The presentation of the current approaches is then completed in subsection 4.3.3, which exhibits the factors that can affect accuracy of fingerprinting.

After focusing on WiFi-based sensing and localisation, in section 4.4, the present report presents a multitude of alternatives such as Bluetooth and RFID. It is also important to understand the topic from current market value perspective.

The literature review is then followed by an analysis and elicitation of developments and applications from the current market of Wireless indoor positioning (section 4.5). Followed by a discussion on the different common use cases of wireless indoor positioning, which

range from elderly care to action classification. This part of the report aims to provide the reader key insight about the current state of the indoor localisation market.

In section 4.6, the present report discuss the approaches taken by companies for indoor localisation and the strengths or weaknesses of their approach. This part aims to further elicit, how companies leverage the data gathered from localisation techniques and their individual approach. The discussion is then pointed towards the current privacy concerns and the future solutions that research is developing (Section 4.7,Section 4.8.).

The present report then concludes with an overview on a survey, conducted at the University of Zurich, and uses its results to add context to the final considerations.

4.2 WLAN based localisation techniques and approaches

In the following subsections, an overview of the existing approaches in WLAN based localisation is provided by analysing and particularising the approaches mentioned in the papers of Pavel & Piché [7], Zafari et al. [8] and He et al. [9]. This section is structured as follows: The specific approaches of WLAN based localisation are explained, several variations are mentioned as well as the potential advantages and disadvantages of the approach. During the explanation and analysis of the various methods of WLAN positioning, this report mainly focuses on indoor positioning. Specific techniques, which are used by these WLAN positioning methods, will briefly be characterised, so that in the end the differences between each of them are intelligible.

This section represents an introduction into the world of WLAN based positioning by focusing on providing the reader an insight in the technical possibilities and approaches. WLAN based fingerprinting will be shortly introduced in this section, but a more detailed analysis of them will be done subsequently in another section. After this section, the reader should have an understanding of the possibilities of today's technology in terms of WLAN based localisation.

4.2.1 General attributes and techniques of WLAN localisation

WLAN is well known for being able to connect several devices with each other and provide an internet access via a gateway. However, the signals, or so-called frames, coming from a WLAN router, which in the following will be referred to as Access Point (AP), can also be used for WLAN based localisation. In order to illustrate the communication between an AP and modern smartphones, the term "device" is used in this report as a synonym for "mobile node". Several possible techniques are using these frames in order to find the position of someone's wireless compatible device [8]. In general, one of the various signal metrics is analysed and converted into distance. This is done by using algorithms. In order to understand these techniques, it is important to have some basic knowledge about the frames, and how the AP interacts with a user device. At first, an AP sends out beacon frames periodically. Every wireless compatible device in range of the signal can monitor this beacon frame. The beacon frame contains the Media-Access-Control (MAC) address of the AP and general network information, which is used for sensing nearby devices and to synchronise with them [10][7].

Even with the high availability of modern smartphones, there is no such thing as the perfect indoor positioning method. Depending on the circumstances given and the main goal, a suitable method needs to be used. Each technology has its advantages and disadvantages in terms of positioning accuracy, availability and costs [11]. In any case, a high density of WLAN routers or beacons leads to increased accuracy in positioning [7]. However, there are as mentioned before, multiple signal metrics on which WLAN localisation can be based. In the following subsections, this report gives a short characterisation

of them, so that a basic knowledge of signals used as potential localisation indicators is provided. This report mainly focuses on RSS based approaches.

4.2.1.1 Angle of Arrival

Wireless localisation based on the Angle of Arrival (AoA) uses the angle, at which the signal from the device arrives the receiver or AP. Therefore, the goal of using this signal metric is to estimate the direction, where the signal transmitting device is located at. By using at least two APs, which have to be able to capture the signal, the location of the transmitting device can potentially be calculated in a 2D area. Consequently, three APs can possibly locate the user device in a 3D area. However, the implementation of AoA requires relatively complex hardware in order to be able to pick up the angle precisely. The larger the distance of the transmitting device, the less accurate this technique might be. This is caused by the increasing error in estimating the position, which already happens with low angle deviation (Fig. 4.1). AoA approach is susceptible to multi-path effects, which may lead to errors in the localisation, especially indoors where the Line of Sight (LoS) is often not available. LoS means, that no walls or other obstacles are located between the AP and the device. Nevertheless, AoA may be used as an alternative to fingerprinting, if optimal circumstances are guaranteed [8].

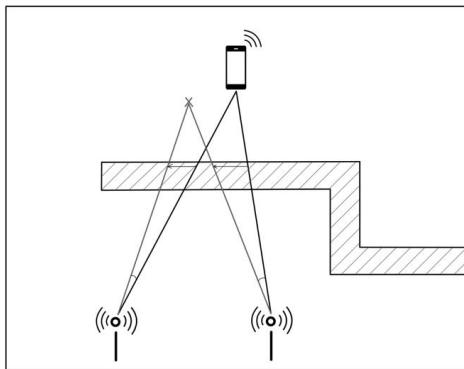


Figure 4.1: Visualisation of minor errors in AoA

4.2.1.2 Time of Flight

In the Time of Flight (ToF) technique, the speed of light is considered. The signal propagation time gets ascertained from the AP, which then is calculated to the distance between the device and AP. Packets are being sent around between the device and the APs, which contain time stamps. The higher the frequency of the sent packages is, the more accurate the time and, therefore, the distance can be estimated. For this technique, at least three APs are necessary, as well as synchronisation among the APs and the device. For achieving a high accuracy with ToF, LoS is required.

Time Difference of Arrival (TDoA) is a variation of ToF, which uses relative time difference of the signal. In TDoA, at least three APs estimate the location of the device by considering the relative time differences of each device's signal. This requires a large bandwidth, but only synchronisation among the APs. Another variation is the Return Time of Flight (RToF), which works quite similar. Here, the time the signal needs to take a round between the device and the AP is measured and used for distance estima-

tion. None of the mentioned variations, including ToF requires fingerprinting to localise a device [8].

4.2.1.3 Received Signal Strength

One of the most commonly used techniques, which also is used in basic fingerprinting methods, is the position estimation by using the Received Signal Strength (RSS). This requires at least three APs. The RSS, that an AP gets from a device, is converted to distance. High RSS defines a low distance and vice versa. The distance estimated defines a geometric circle around the AP, on which the device potentially is located. In the example of three APs using the RSS for estimating the device's distance, the three circles intersect on a point (Fig. 4.2). Localisation by using the RSS defines that this point is the device's location. Inspired by the figure that Zafari et al. used, Figure 2 shows the usage of RSS visually. Zafari et al. [8] mention an indicator based on RSS in their paper, called the Received Signal Strength Indicator (RSSI). That indicator is labelled by them as a "relative measurement", in comparison to using the monitored RSS. The RSSI may vary depending on which WiFi-chipset is used. In indoor environments, where often no LoS is available, and a high amount of noise and multi-path effects occur, the mere use of RSS or RSSI may lead to problems in positioning. Walls and even human bodies can lead to disruptions. Although localisation based on plain RSS values has its advantages, such as the ease of implementation, flexibility regarding the combination with other technologies and cost efficiency, it has room for improvement when it comes to accuracy and consistency. A solution would be the approach of fingerprinting in WLAN based indoor positioning, which will be thoroughly explained later in this report [8][7]. Zafari et al. [8] separate WLAN based localisation into two categories. They use the terms Monitor Based localisation (MBL) and Device Based localisation (DBL). MBL is explained as the approach of localisation, in which APs are passively observing their range, in order to detect devices and localise them. This can work through beacon frames. In DBL, the application of MBL is turned around. The device is the node, which is reaching out for other nodes in its area.

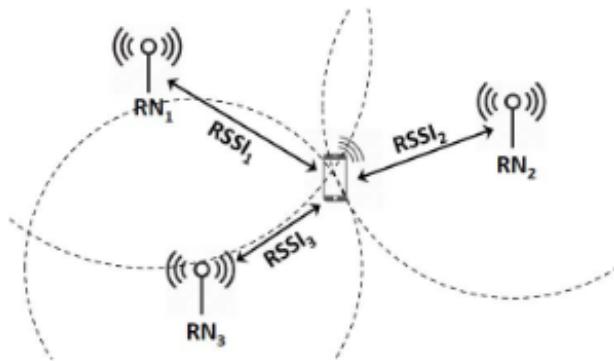


Figure 4.2: RSS based positioning. (Reprinted from [8])

4.2.1.4 Channel State Information

Channel State Information (CSI) includes more transmission data than RSS. It consists of detailed information about the communication channel, used by the receiver and the transmitter. This data includes information about the environmental effects that the signal faces during its way, which in indoor environments has a positive effect in terms of position accuracy. Therefore, CSI is more consistent than RSS, in terms of WLAN

based positioning. This seems to be an ideal solution to the problems, which occur in the RSS based localisation. However, as modern-day smartphones mostly are not build for supporting CSI data collection, this approach fails with WLAN based localisation of smartphones [8].

4.3 Fingerprinting

Most of the previously discussed localisation techniques require LoS. Although the use of trilateration and triangulation is a precise way to localise devices (GPS, for example, works through trilateration), the fingerprinting technique is more suitable for indoor localisation. This section is structured as follows: The RSS based fingerprinting and its basic approaches are explained, as well as some possible additions to the basic approaches, difficulties with WLAN based fingerprinting are mentioned and discussed, possible improvements for fingerprinting accuracy are analysed and in the last subsections, more advanced fingerprinting methods are illustrated.

4.3.1 RSS based fingerprinting

RSS based fingerprinting is a localisation method which uses the RSS of devices. All fingerprint approaches need at least three APs. The RSS, that an AP gets from a device, can be calculated to distance between the AP and the mobile node. Instead of only calculating RSS to get the distance and find the more or less exact location with trilateration or N-lateration (plain RSS method), the location can be found out by using fingerprinting. In RSS based fingerprinting, a so-called "radio map" is created during an offline phase. During this phase, the APs collect "fingerprints", which are the RSS values from devices that are located in the range of the APs. This data gets stored in a database (radio map). After the offline phase, the mentioned radio map contains several Reference Points (RPs). These stored RPs are compared to the RSS of devices passing by. The phase, where the comparison happens, is called the online phase (Fig. 4.3). In simple words, the stored RSS value, that is the closest to a device, determines the position of the device passing by. So it can be assumed that the more data gets collected during the offline phase, the more accurate the localisation during the online phase will be. This rather represents a trade-off between costs and accuracy. The more the preparation of the radio map costs, in order to collect a numerous amount of Reference Points, the higher the accuracy during the online phase will be. However, many administrators of localisation systems might be interested in an easy, fast and affordable solution. On the other side, it is important to pay attention to the changing conditions of the observed area, like changes in the building structure or the infrastructure of the APs. These may affect the localisation accuracy if no regular adjustments are made. For decreasing the impact of these effects, the offline phase needs to run again to achieve adjustments in the radio map. As fingerprinting is a widely and commonly used approach, there are two types of basic fingerprinting methods. Probabilistic and deterministic algorithms lay the foundation for all advanced fingerprint systems [7].

4.3.1.1 Deterministic fingerprinting

Deterministic fingerprinting uses Nearest-Neighbour (NN), K-Nearest-Neighbour (kNN) and Weighted- K-Nearest-Neighbour (WkNN) algorithms. With the AP monitoring the RSS, the Euclidean distance gets calculated. The NN or kNN fingerprint is matched to devices observed during the online phase. The NN algorithm uses only the nearest saved RP to determine the position of the device. The kNN algorithm considers several k fingerprints and WkNN assigns different weights to the RPs. For example, fingerprints

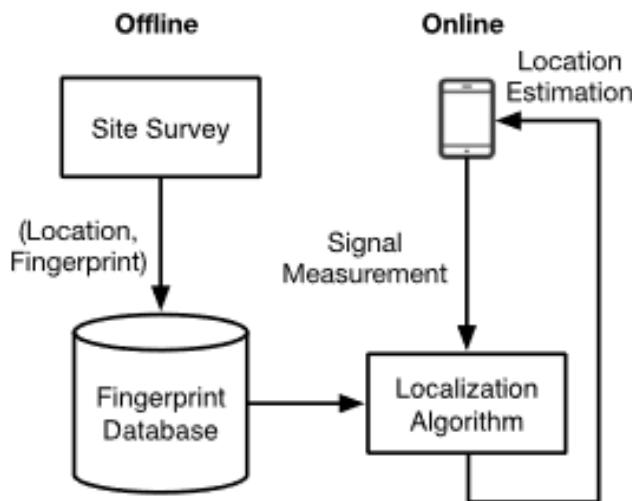


Figure 4.3: Fingerprinting process. (Reprinted from [9])

that are nearer to the device than others, are weighted more in the calculation of the position. The deterministic approach of fingerprinting is simple to use, robust and has reasonably good accuracy [7].

4.3.1.2 Probabilistic fingerprinting

Probabilistic fingerprinting calculates the device's position by considering the Bayes rule and the stored fingerprints. The stored RSS value of a device, which is registered during the online phase, is compared to a probabilistic distribution. That distribution is created based on the stored fingerprints. Using probability algorithms and mathematical concepts like the Gaussian distribution, the location gets estimated. There are two common approaches, which are based on probabilistic fingerprinting: Coverage Area Models (CAM) and Path-Loss Models [7].

Considering a coverage area, where the fingerprints were collected during the offline phase, the position of a device gets estimated here without using the specific RSS value. This makes it less susceptible to changes in the environment and therefore disruptions. But the calculation of probability results in less accuracy than the location estimation based on specific RSS values [7].

In Path-Loss models, the reduction of signal strength is considered to happen logarithmic, depending on how far the device is away from the AP. This approach uses several parameters in its calculation, such as variables, that should define the environmental conditions. It can be assumed, that the Path-Loss model works very similar to the basic probabilistic fingerprinting approach. It shares the same advantages and disadvantages as every fingerprinting method that uses specific RSS values. The location estimation using Path-Loss models is potentially accurate, but its accuracy is dependent on the environmental circumstances and changes in the infrastructure of the APs, as well as to RSS disruption [7].

4.3.1.3 Support Vector Machines and Compressive Sensing

Davidson et al. [7] also mention other fingerprinting methods, which potentially solve the problems that may occur with basic approaches. One of them is called Support Vector Machines (SVM). The SVM technique was created to extend and improve the earlier mentioned fingerprinting approaches. It extends the RSS based fingerprinting with machine learning. Using SVM, the collected data gets structured for finding the best "hyper-plane", which generally is the nearest RP to the device, but considers more

parameters than basic fingerprinting. For example, if D is the distance estimated between the RP and the device, then that RP has the largest "hyper-plane". There is no other RP around the device, which has a smaller distance to it, than D . Nevertheless, here also the process of an offline phase is needed, which leads to losses in accuracy over time when the environment changes.

Also, the usage of multiple weighted decision trees has been discussed. This approach considers that smartphones often have different orientations. The heterogeneous builds in smartphones may lead to problems in localisation accuracy. The approach of multiple weighted decision trees is efficient in terms of ease-of-implementation and execution time [7].

Davidson et al. [7] discuss the approach of Compressive Sensing (CS) in their paper, which only uses a smaller subset of stored fingerprints. A minimisation problem is solved considering a low amount of noise. The RSS of the device gets compared to different clusters, to restrict the area where the device may be located at. Afterwards, the location in that cluster gets defined more accurately by using the theory of compressive sensing.

4.3.2 Difficulties with WLAN based fingerprinting

Localisation using WLAN based fingerprinting can be affected by environmental constraints, human bodies and the different orientation capabilities of smartphones. Every change of these circumstances during fingerprinting would lead to an adaption of the localisation system, if no loss of accuracy should happen. WLAN based fingerprinting systems need regular adjusting and require a lot of time and database management [7].

4.3.2.1 Localisation of heterogeneous devices

The components important for the positioning of smartphone devices are first and foremost chipsets, antennas and the operating system. These components may significantly vary from device to device. This has an effect on the RSS based localisation. Heterogeneous devices may accordingly not be able to be located with the same accuracy. Changing the fingerprint type from RSS to Signal Strength Difference (SSD), as explained in the paper of Davidson et al. [7], is an approach to reduce the problems that occur by localising heterogeneous devices. Instead of using RSS values, the SSD is not affected by the device specifications and therefore this seems to be a possible solution. SSD works with the relative signal strength differences and calculates the position by comparing these differences to other devices. RSS values are not used during the localisation but are nevertheless considered when it comes to comparing the signal strength differences. In the end, the usage of SSD also shares some of the same problems as the RSS-based positioning [7]. Davidson et al. [7] also mentioned, that it was discovered, that the differences in RSS measurements between heterogeneous devices correlate with each other. Therefore, with specific training during the offline phase, the correlations can be used to improve accuracy.

4.3.2.2 Disruptions in WLAN signals caused by people

The human body interferes with the RSS measurements because it absorbs the radio-signals and therefore hinders the signal's propagation. In terms of RSS measurements, also the hand of the user covering the antenna of the smartphone will lead to a lower accuracy [7]. As described in the research conducted by Rosa et al. [12], if a hand is fully covering the antenna of the smartphone, it can lead to a range error of about 9 meters, even if the smartphone is only 3 meters away from the AP.

4.3.2.3 Improving accuracy with AP selection

A research conducted by Youssef et al. [13] showed that the selection of the APs with the strongest signals might lead to a reduction in power consumption; this is caused by a reduction of the computational cost algorithm. Besides the mentioned improvement, Youssef et al. also demonstrated that the selection of signals based on their intensity could lead to increased accuracy.

4.3.3 Accuracy of WLAN based fingerprinting

Several factors can improve the accuracy of WLAN based fingerprinting. A high density of APs and a regular radio map maintenance contribute to a better position accuracy. Also, the choice of the algorithm affects the accuracy [7]. Several existing algorithms, such as WkNN, Path-Loss model (PL), Coverage Area Model (CAM) and the generalised Gaussian mixture (GGM), which considers an approximation of the Path-Loss model, have different performances. Müller et al. [14] attempted an experiment, where the mentioned methods were used for localising a user's device on a university campus. The accuracy of PL and GGM was slightly better than the rest. These algorithms had 50% of the time an error below 8 meters. The CAM algorithm error was slightly higher with 10 meters, but has a lower computational cost. The computational cost of CAM is approximately 30 to 50 times less than the foremost algorithms. "WkNN also showed good performance with position error below 6 meters 50% of time." [14]. In the experiment, they took away some APs, which as expected lead to a higher error with all algorithms. "However, the performance of WkNN algorithms deteriorated significantly exhibiting the largest position errors among all of these algorithms." [14]. Also, when an older radio map was used in this experiment, the accuracy decreased. In the mentioned research paper, an older radio map means a radio map, that was not maintained and therefore does not consider environmental changes over time. The positioning accuracy of the WkNN algorithm suffered the most from having an out-of-date radio map, and when compared to the rest, it was the one which accuracy decreased the most.

4.3.4 Magnetic field fingerprinting

Both Magnetic field fingerprinting and RSS based fingerprinting have an offline and online phase. In magnetic field fingerprinting, not RSS values are stored during the offline phase, but instead magnetic field characteristics in the covered area. Environments may have the same magnetic field characteristics, which is the case if the metal structures of buildings are very similar. This limits this approach to local use only. Compared to RSS, the magnetic fields in indoor environments change much less over time and additional investment in infrastructure, as needed in RSS based positioning, is not necessary. Disadvantages of this fingerprinting method are the reliability of a relatively small number of parameters, which may lead to less accurate positioning. Also, heterogeneous devices are a problem here, as well as magnetic fields, which can be disrupted too. The altitude also matters, because it has an impact on the positioning in magnetic field fingerprinting and therefore a creation of a 3D magnetic field map improves the outcome significantly. The availability of magnetic field based positioning for smartphones is also similarly high, like the RSS based positioning, because the necessary magnetometers are already built-in in any modern smartphone device. Similarities to the RSS based positioning are also the susceptibility to interference, in any way. Parameters used in magnetic field fingerprinting are for example the direction of gravity and the magnitude of the magnetic field intensity. One of these parameters may fail being computed, which affects the accuracy and the performance

of this method. Therefore, magnetic field fingerprinting is potentially a very accurate method, but it is also susceptible to disruptions [7].

4.3.5 Map-Aided Indoor Positioning

The addition of a building floor plan in WLAN indoor positioning systems can decrease the number of errors. Indoor environments are often not only a large hall, but they contain walls and doors and also elevators and/or stairways. There are several approaches for implementing indoor maps into the fingerprinting process: Probabilistic and topological map matching, as well as map matching based on the link-node representation of a building plan [7]. Pavel et al. [7] explained these methods thoroughly in their work. They concluded that the performance of WLAN based indoor positioning is more accurate with the addition of map-aided positioning methods. In some approaches, the already available floor plans can be used. Sometimes, when the floor plan is not readily available, more preparation is necessary. The usage of probabilistic map matching is in general not too complex to implement, but depends on more given parameters, to achieve better accuracy. The map matching based on the link-node representation of a building plan is far more complex, but on the other hand, also works relatively accurate with fewer parameters available [7].

4.3.6 Spatial and Temporal Signal Patterns

Considering spatial or temporal observations in fingerprinting methods can lead to an improvement and higher accuracy. The dependency on RSS signal vectors of the mentioned fingerprinting methods creates a vulnerability to multi-path effects and disruptions. Measurement noise can result in the case of the target being mapped to a distant position of similar signal vectors [9].

Temporal patterns are the WiFi signal sequence patterns that occur during walking in indoor environments. Hence, the route a device takes in the area covered from the APs can be used, to improve the accuracy of localisation. Comparing this "walking" pattern with a single RSSI vector carries temporal information and that information can be used to state the WiFi fingerprint-based localisation more precisely [9].

Temporal patterns require the knowledge of user motion inside the area being covered from the APs, which may not be available often in reality. Therefore, geographical signal patterns (Spatial patterns) can be used to improve the basic fingerprinting instead. "Spatial patterns include RSSI order, signal landmarks (AP locations) and signal coverage" (He et al. [9]).

4.3.7 Collaborative localisation Among Mobiles

As mentioned in the paper of He et al. [9], most of the current works on WiFi fingerprint localisation are based on independent estimation. This means each target device gets located by the fingerprinting system (group of APs for example) independently. The relative location of devices to each other is not considered. Collaborative localisation arises from existing trends in mobile computing as a way to minimise independent estimation errors.

4.3.8 Motion-Assisted localisation

An example of a classical hybrid technique for indoor localisation is Motion-Assisted localisation [9]. Due to the increasing implementation of motion sensors in portable WiFi

compatible devices, this fingerprinting technique is gaining more and more attention. Using motion sensors in smartphones or other similar devices can improve the accuracy.

4.3.9 Commercial WLAN fingerprinting Systems

Companies specialised themselves on creating their own fingerprinting systems. In the following subsections, two real-life examples are shortly illustrated.

4.3.9.1 XPS WLAN by Skyhook

The XPS WLAN fingerprinting system, which is offered by Skyhook, is specialized on localizing mobile positions in dense urban areas. Skyhook has build up and is still maintaining a global database of WLAN APs, which constantly collect RSSI data. When used, fingerprinting does not need to be considered any further, because Skyhook already maintained to build a net of RNs. It is therefore ready to use. It has a accuracy of 10 to 20 meters outdoors, and 30 to 70 meters indoors [5] Later in this report, the company Skyhook will be analysed more detailed.

4.3.9.2 Real-time location system by Ekahau

Ekahau's real-time location system is generally based on RSS based fingerprinting. It uses track history in addition to the stored RSSI vectors [5]. Ekahau's system is, however, quite different from the mentioned XPS system by Skyhook. The RSSI vector database does in this case not already exist, but it still needs to be created during the fingerprinting. [5]

4.4 Alternatives to WiFi localisation

In this section, the report will discuss and describe the functionality of alternative approaches to localisation. They will be compared to the WiFi-based one and existing advantages or disadvantages shown. [5]

4.4.1 Bluetooth localisation

Bluetooth is a wireless standard for Wireless Personal Area Networks which is a managed and owned by the Bluetooth Special Interest Group. It provides better security, cost less and uses less power and space. A maximum power output of 1mW is required at the highest power level of the Bluetooth standard, which enables communication within 5m to 10m depending on the propagation conditions. [5]

Previous work on localisation using Bluetooth have been unsuccessful. The time of flight based positioning method does not work with the standards and original protocol of Bluetooth. By default, Bluetooth also does not allow host to read the receiver signal strength and using variations analysis in signal strength is also not an option. [5]

Because Bluetooth hardware does not provide a trustable receiver signal state information, [15] proposed to used fingerprint based localisation which uses the Response Rate of Bluetooth. ZONITH [16] came up with a product that allows indoor positioning and is a wearable device. The module consists of Bluetooth devices which are similar to those used in smartphones.

4.4.2 Radio Frequency Identification Service

RFID or Radio Frequency Identification system is made up of antenna that probes signals coming from nearby transceivers or tags. It uses radio waves to transfer data from RFID tags to the device with RFID scanner.

The transferred data has an unique id of the tag that was used to transfer the data and this id can reveal the location of transmitter. Cell of Origin or CoO is the commonly used technology to fetch the location or distance of RFID tags. It can be used to detect nearby person carrying RFID device or tags. The accuracy of RFID systems depends on how many tags are used in an area and also on the reading range of receiver. The strength of signal received by the receiver can be used to determine the position based on the time which signal arrives or is received. Making such estimates have proven to be difficult to achieve. To be able to estimate the distance between device and receiver at a better resolution, higher bandwidth waves needs to be used and also estimates over several times needs to be averaged to attain a more accurate estimate. Using localisation using RFID is based on the time of arrival metric and it uses only single RFID tag to make the position estimate. Researchers have also proposed different approach for RFID localisation other than ToA. One of them is Phase of Arrival which is described in the paper [17].

4.4.3 ZigBee based wireless sensor networks

ZigBee is a wireless technology that finds its use in cases where power consumption needs to be low and transfer of data doesn't take place very rapidly. Hence, it is also known as a type of low rate WPAN.

In external environment, a single ZigBee node can have a range of about 100m if there is no obstruction, although this range decreases in indoor environment and is only up to 20-30m. The receiver signal strength between node and transmitters is used to estimate the distance between any two nodes. ZigBee can suffer from obstruction in signal due to interference.

8 ZigBee nodes were deployed in an office space by Larranaga et al [18] to estimate location of moving people. The location of deployed nodes was noted and from this the characteristics of propagated signals were measured with an accuracy of 3m.

4.4.4 Acoustic Signal based localisation

Sound is a type of wave that needs a medium to travel. The propagation is possible because of pressure difference created by vibrations of particles in the medium. Sound based localisation system can use air or walls in building as a medium.

Sound waves can be used for positioning. To perform localisation using sound waves, there is a need to place stationary nodes on the walls or surface of a room and the moving devices carried by human or robots can be located using the multilateration measurement. Generally Ultrasound is the preferred type of acoustic signal used for localisation but some work also makes use of audible sound signals.

4.4.4.1 Ultrasound

The time at which receivers receive the Ultra Sound(US) signals can be used to calculate the relative distance between any two nodes. In comparison to radio waves, the US ToA is 10m or less and doubling the distance causes attenuation in the signal's sound pressure. Active device system is an US signal based localisation system that uses three or more than three receivers that are deployed at fixed location. Such a setup can be used to improve the accuracy of localisation using multilateration. A different approach is to use

more than one stationary emitters rather than receivers which are placed at fixed location. Such a system architecture is known as passive device systems.

Ultrasound based positioning systems can be used for localisation of people and mobile devices in indoor environments but are not so often used in external environment.

As sound waves can propagate through multiple paths within the same medium (Mautz and Ochieng 2007 [19]), it poses a challenge to localisation using ultrasound system. Ultrasonic system also consumes high power and is known to drain battery very rapidly. Hence, it is a challenge to build a system that minimises the usage of battery.

4.4.4.2 Audible Sound

Only a small minority of the localisation system makes use of audible sound waves. The motivation behind using audible sound waves to build a system that can use the sound cards present in common devices.

Filonenko et al. [23] generated audible ultrasonic signals by making use of the built-in speakers in smartphones. It was used for deploying Time of Flight localisation from the common hardwares found on a mobile phone.

4.4.5 High Sensitive GNSS

Global Navigation Satellite Systems (GNSS) can be employed to fulfil the vision of ubiquitous positioning. GNSS is the only technology that satisfies both these conditions:

- operates independently of the local infrastructure
- global coverage

GNSS receivers face a challenge of tracking satellite in an environment where signal fading is severe and hence they hardly cover indoor spaces. To resolve this issue a lot of research has been done in the recent times and they focus on building more sensitive and better GNSS receivers but these advancements still haven't been able to resolve a lot of issues.

4.4.5.1 Assisted GNSS

Assisted GNSS or AGPS is a localisation method that can be used in external environments. It can be used to estimate the location and position of internet enable devices such as smartphones, etc. An additional data link can provide more information from satellite ephemeris which until now was normally obtained from the GNSS satellites directly.

4.5 Marketshare of Wireless sensing applications

In this section, market for wireless sensing is discussed in detail and in particular indoor position market. Some interesting insights into the current market share and also future growth possibility is described. In the later section, some of common use cases of indoor positioning is explained.

4.5.1 Current Market & Trends

The market share for WiFi sensing is growing at a great pace. This has been driven by reducing hardware costs and development of wireless technologies or devices that use wireless technologies, e.g. Internet of Things, WiFi-enabled smartphones and other sensors. The current market for Indoor Positioning is valued at USD 6.1 billion, and during the period of 2020-2025, it is expected to grow at a rate of 22% [20]. Currently, North

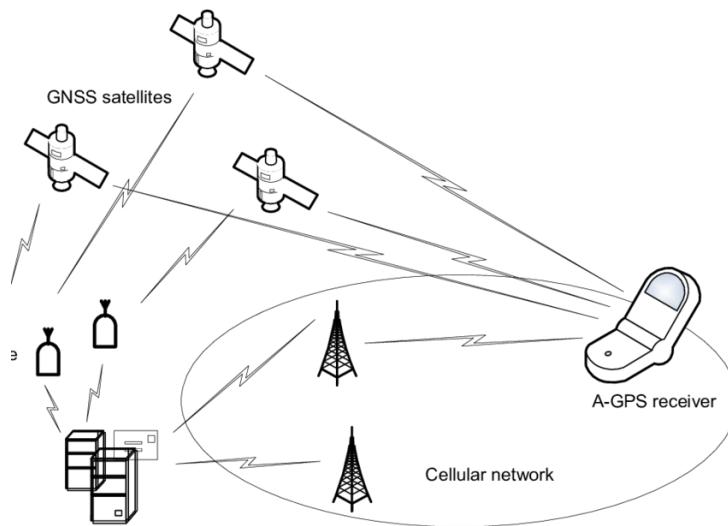


Figure 4.4: Assisted GNSS

America holds the highest share in the market of indoor positioning, but the Asia Pacific is expected to overtake North America to become the largest market by the end of the forecast period[20].

Below some recent developments in indoor positioning markets is described [20]

- Mist System was launched in March 2020, by Juniper Networks which a trusted and popular firm in AI and security. It aims to provide it's customers with insights that allows them to introduced digital solutions in their project and increase the demand for their product.
- Microsoft opened a new data centre in Spain in February 2020 and also made pact with Telefonica to extend it's reach in Spain. The new expansion aims to introduce more digital technologies in both public and private firms that will allow to increase them the operation size and increase innovation.
- To estimate location of a target, the ASSA ABLOY group launched the FiRa consortium in August 2019. The new product is better than all the previous products in the same field and can measure position with best overall accuracy.
- In July 2019, Inpixon was acquired by Jibstream a commpany that works to provide indoor localisation and mapping.Inpixon will work towards creating more smart indoor space by using Jibstream platform.

Top Industries making use of indoor positioning solutions are:

- Hospitals and Health care, where WiFi-based indoor positioning can be used for locating medical equipment's or locating patients and visitors.
- In Defence & Military, indoor positioning can be used for surveillance and security, by locating people inside the base or also for robot navigation which can be used to guard an area,
- In Logistics & Warehousing indoor positioning finds its use for optimising routes and collision avoidance when transporting items within the warehouse.
- Manufacturing can make use of indoor positioning in managing items in inventory and for quality assurance.
- Sports & Entertainment for location-aware AR/VR applications



Figure 4.5: Indoor positioning use cases

4.5.2 Common Use Cases

In the previous section, some insights into the market share of indoor positioning were discussed. In the below section, some common use cases and challenges to indoor positioning is described.

4.5.2.1 Elderly People Monitoring

The older population is increasing, and the ratio of population aged 20-64 is expected to approach 35% in 2030 [21]. The worldwide population over 65 is expected to grow to 1 billion by the year 2030. The majority of the elderly stay at their home for most of the day [22]. Every year 33% of the elderly people over the age of 65 witnesses a fall. Such accidents can often be fatal or cause a reduction in quality of life. In a large of such fall cases, elderly people cannot get up by themselves and about 50% of those people who were a long time on the floor without getting any help to die within next six months [22]. Fall is also one of the main reason for the death of elderly people, and the cost of institutionalised health care is also high. Fall Detection hence becomes an important problem which can save many lives by helping in providing an immediate alert in case of fall.

In the paper [24] a WiFi-based fall detection system, WiFall was proposed. WiFall makes use of the channel state information measurements. Human activities in an environment can affect the channel state information or the receiver signal strength. WiFall makes use of this to classify a change as fall or not. Fall causes a rapid change in the environment

which also means CSI varies a lot for the duration of Fall. Monitoring the change in CSI can then be used to predict if there was a fall or not.

WiFall fall detection system is a two stage algorithm. The first stage in the algorithm is to identify the CSI series and compare it to normal or expected CSI series. If the observed CSI series is different from expected series then second stage is to be performed. In the second stage, the observed CSI series is input to an activity classification based on Support Vector Machine (SVM). SVM is a machine learning-based algorithm which can be used for the classification task. In WiFall it acts as a classifier of CSI series, and it predicts fall or no fall. WiFall achieved an accuracy very close to fall detection system that uses devices carried by the user. It had a 87% detection rate and 18% false alarm rate.

4.5.2.2 Activity Classification

Activity Classification is an area of a use case where the ability to classify human behaviour or activities is necessary. Activity Classification mainly finds its use in surveillance where law enforcement authorities would want to localise areas with suspicious activities. Activity classification approaches use mathematical model or machine learning model to differentiate between normal and abnormal behaviour. The observed behaviour is compared to the modelled or already known expected behaviour. A classification model such as neural networks or SVM can be used for classification of normal and abnormal behaviour.

In [25], Reschke et. al proposed an approach that tries to classify human activities happening in front of smartphones. It makes use of wireless signals and analyses the variation in receiver signal strength. The idea is that similar pattern of activities will have same time series variation. The user does not necessarily needs to carry the mobile phone with him/her, but the approach also works if mobile phone is near to the user. However, the achieved accuracy was lesser than conventional sensors such as an accelerometer.

For the activity classification task, several features have been studied for classifications. Some of them are average magnitude squared, signal to signal noise ratio and signal amplitude. In the paper [26] the proposed algorithm could recognise and count up to 10 moving or stationary users. They further improve the accuracy of the algorithm by making use of additional frequency-domain features.

4.5.2.3 People Counting

People counting as the name suggests is the problem of counting number of people in an environment at any given time. It can be useful in several scenarios including use of guided tours or crowd control [27]. Many applications can benefit from people counting. It can be used in surveillance, disaster management and also in marketing. In surveillance, it can be used to track people movements and activities inside a building whereas in disaster management it can be used to optimise evacuation process where the knowledge of the number of people and their location can be very critical. In marketing, it can be used to identify the number of people actively participating in an event. Other uses can be for regulating room temperature based on the number of people or other smart home applications. While people counting is an interesting problem with many applications, it has its own challenges. The report discusses some of the challenges in the later section. Mostofi et al. [28] proposed a system that helps keeps track of number of people in an indoor environment. The proposed system makes use of WiFi signal and only needs a pair of transmitter and receiver. People can affect the receiver signal strength of the WiFi signal. It can be because the signals can be either blocked off or reflected from the human body. The framework makes use of these two important effect and proposes a

mathematical model that describes the received signal strength as a function of blocking and reflecting effect. The approach was tested in both indoor and outdoor environments and it was able to accurately count up to nine people.



Figure 4.6: People counting experiment

4.5.2.4 Through the Wall Sensing

A novel and rapidly evolving research area, called *through the wall sensing* tries to solve object localisation when its line-of-sight is obstructed by a wall. Other sensing approaches make use of the fact that signal strength is affected due to peoples movement or obstruction. But such approach only work within a single room. If one want locate objects across multiple rooms or within entire building, through the wall sensing approach needs to be used. Through the Wall Sensing find it's use in many emergency situations. Once such is identification of victims under stones or rubble in case of an earthquake. It can also be used in evacuation when there is fire in a building. Hence through the wall sensing remains a useful tool for emergency situations and can be used by police or firefighters and disaster management teams.

Through the wall imaging is also getting considerable attention for security applications, and it uses radio frequency sensors to be able to map areas across walls and identify interesting objects across rooms or walls. One major challenge in through the wall sensing is signal attenuation which can be cause by obstruction such as walls. The signal attenuation increases with increase in signal frequency used for localisation.

In [29] & [30] authors conduct multiple experiments to understand whether Wireless signals can be used for through the wall localisation. In cite, the authors propose an indoor events detection system based on a time-reversal technique to detect changes in indoor multipath environments.

In another paper Wilson et al.[31] proposes a new method for localisation and motion tracking through the walls. People movement can affect the signal received by receivers. The new approach makes use of this to study the variation in signal. They also show that signal strength is highly dependent on moving objects in the environment. A mathematical model was proposed which can relate the location and movement to the RSS variance.

4.5.2.5 Some other Applications

Emotion recognition is among other applications of WiFi-based indoor sensing that is an active area of research and has drawn growing interest recently. Emotion recognition deals with the problem of developing devices which can sense human emotions. A possible application could be the change in music or television based on peoples emotions. The detected emotion can also be interesting to movie makers and can help evaluate people's experience. Other applications could for diagnosing symptoms of anxiety, depression or bipolar disorder. More broadly, emotion could become a form of input to computers.

Body gesture can be used to measure emotion. In the paper [32] Bull et. al. indicate that motion information can be extracted from human body configurations. The paper also shows how motions and positions are related to human states such as boredom or interest. WiFi sensing could be used for body pose and gesture as was seen in other use cases.

Researchers in paper [33] developed a system that enables them to analyse emotions of a user using reflected radio signals from human body. The wireless signal emitted by the system is reflected off the user body, which is then used recognise emotions such as happiness, sadness, etc. The system is based on the key algorithm that allows it to use wireless signal to read heartbeats at an extremely good accuracy.

Nowadays every place is equipped with WiFi, especially home and offices. The paper [34], uses WiFi signals reflected from a hand and keyboard to guess keys pressed by the users. The idea behind this is similar to previous works, that is to detect movement of hands and fingers and which can then be converted into a unique time series CSI information and later classified into different keystrokes. The accuracy of the proposed system was 97.5%; however, the system only worked in a controlled environment.

4.5.2.6 Challenges to WiFi Sensing

In the previous sections, the report illustrated some common use cases and application of WiFi-based sensing and indoor positioning. While WiFi-based sensing has lots of varied application, there are also areas of concern or challenges that limit the efficiency and accuracy of such a system. In this section, some of these challenges will be discussed.

- Occlusion of WiFi signal, occlusion means a hindrance to the receiver and emitter signals. As WiFi signals can be reflected or absorbed by items in the environment and also by the human body [5]; this proves a challenge to WiFi-based sensing and positioning systems.
- Presence of a large number of people is also directly related to the aforementioned challenge, which was occlusion [35]. The human body can reflect or absorb WiFi signals; the presence of a large number of people greatly decreases the accuracy of WiFi-based systems. It was noticed in various research that people counting work efficiently only for up to 10-15 people, and with more people, the accuracy drops greatly.
- Challenges based on the environment, a large number of research studies are done in a controlled environment which does not take into the account the real-world environment where signals can be obstructed and receiving strength varies greatly even without the presence of a large number of people. Hence it is often quite challenging for the proposed systems to work efficiently in a real-world environment.
- Limited range resolution, Although WiFi has the best location accuracy when compared to UWB and Bluetooth, it still falls short by a large amount when compared to other signals such as radio waves [5]. The receiver signal strength might vary a lot when receiver and emitter are located at a large distance.

4.6 Case Study: Companies using Wireless Sensing

In this section, the report gives an oversight of three Swiss companies and three international companies and their approaches, goals and market position. Subsequently, a comparison will be made between the mentioned companies to highlight the crucial differences between them.

4.6.1 Locatee AG

Locatee is a Switzerland-based startup which provides workplace analytics solutions by leveraging WiFi and LAN infrastructure. It was founded in 2013 in order to provide a simple and intuitive solution for managing and optimising corporate real estate [36]. It has over \$3.9 Mio.[37] in founding amounts and provides its analytical solutions to various high profile companies such as *Swiss Re*, *Zurich Insurances* and *Johnson & Johnson* [36].

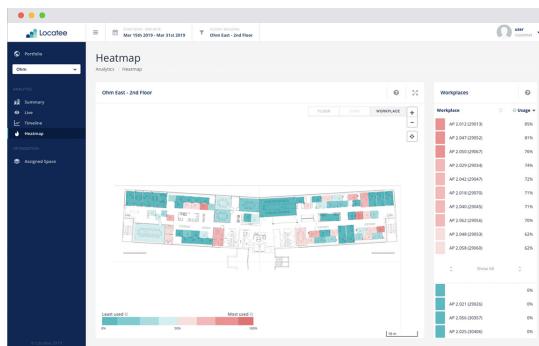


Figure 4.7: Example of *Locatee*'s dashboard [36]

4.6.1.1 Approach

Locatee's solution is to use the company's existing WiFi and LAN infrastructure and collate the data together with floorplans and other location attributes. Together with their proprietary algorithm, the tool can detect a building occupancy rate, in near real-time. Besides using WiFi and LAN, it is also possible to integrate sensors in their net of data-gathering devices. *Locatee*'s approach is a three-step procedure. First, a device connects to the company's WiFi or Lan and the device's MAC address and timestamp of the measurement is collected. Afterwards, in the second step, the data is encrypted and sent to a data centre where it is processed and anonymised in order to create real-time interactive reports on offices' occupancy and utilisation. The third and last step concludes the whole process by aggregating the data into comprehensible visualisations visible on an online dashboard [36].

The software-only solution can be deployed on existing infrastructure and is highly scalable. This approach is also supported by a subscription-based selling strategy, which lowers the initial cost and therefore, also the barrier of implementation of this type of solution. The deployment of such a system is fast but requires a setup process which needs floorplans and the location of the used access-points.

4.6.1.2 Goal

The product developed by *Locatee* aims to provide an in-depth analysis tool to companies. The core product is an interactive dashboard, which provides critical insights into the utilisation of office spaces and other real estate objects owned by the company. An example for these type of dashboard is the heatmap, visible on figure 4.7, The heatmap

dashboard can be used both by the admin's to see office utilisation and by the employees in order to find free meeting rooms and desk spaces.

By using this data, companies can realise savings by aligning the supply and demand for office spaces and reducing the time spent by employees searching for office spaces. *Locatee*'s solution was applied in the Munich offices of *Swiss Re* and allowed them to decrease their office space offering by 10% and thus realising savings for over \$240'000 a year [36].

4.6.2 Onway AG

Onway AG, formerly known as CloudGuard Software AG, is a Swiss company which was founded in 2004. It is specialised in developing and maintaining multifunctional and integrated communication platforms for public transport companies. Their services are used by high profile regional companies such as *PostAuto*, *SBB CFF FFS* and *Zurich Airport* [38].

4.6.2.1 Approach

Onway provides its customers with a variety of hardware and/or software solutions. On the hardware level, they use their in-house build technology to provide vehicle-to-ground communication by using mobile communication networks. This communication channel is then often used to provide WiFi to the transportation users, streamline the sending and receiving of vehicle data, messages, updating eventual information screens and other services. *Onway* therefore usually deploys its own infrastructure in a combination of both software and hardware, based on their customers need [38]. Position analysis is done by using RSSI fingerprinting using GPS, cellular and WiFi networks, for analysing a vehicle position, and within the vehicle fingerprinting is used in order to evaluate the number of passengers are present.

The system is scalable but requires *Onway*'s proprietary hardware and infrastructure; therefore, it is aimed at larger companies and requires a larger initial investment, concerning other analysed companies. *Onway*'s system also requires a general overhaul of the interested customer already implemented infrastructure.

4.6.2.2 Goal

The main goal of *Onway* is not the localisation and monitoring of persons or objects, but a mere by-product they can offer bundled with their systems. By combining their hardware with both external sensors (such as GPS/GPRS antennas), they offer real-time localisation of vehicles. By analysing the internet usage and active connections, they also offer other analytics to the companies, such as real-time internet consumption, customer analysis or passenger information, all of which can be monitored from the cloud [38].

4.6.3 Livealytics AG

Livealytics AG is a Swiss company founded in 2018. The company provides analytics tools and IoT for measuring and benchmarking in-store customers. It counts on multiple market-leading customers such as *Samsung*, *Credit Suisse* and *hp*[39].

4.6.3.1 Approach

Livealytics approaches the tracking and monitoring industry by offering both hardware as software-based solutions. They developed Bluetooth and WiFi-enabled beacons which use passive WiFi fingerprinting to capture devices approximate position together with

a timestamp. This data is then anonymised and analysed in order to preset aggregate data to the company using one of *livealytics* tools. All data gathered is stored on a cloud platform in order to provide ease of access to the shop owner or marketing team.[39]. The system is highly scalable with a low barrier of deployment, as the basic subscription costs 99.- CHF a month. Additionally, the system can be deployed quickly and does not need outside data (such as floorplans) in order to function[39].



Figure 4.8: Example of Livealytics dashboard

4.6.3.2 Goal

Livealytics products are primarily marketing supporting tools, which entails that the data they gather is used to analyse the customers of a given location. As visible in figure 4.8[40], the tool provides in-depth analysis about customers and their interaction within the range of the beacon, that can also be outside the shop perimeter. By setting distance thresholds, the system can define the potential customer opportunities, how many visits there were in a considered shop, and whether customers returned multiple times to the store. [39]

4.6.4 Skyhook, Inc.

Skyhook is an American company, founded in 2003. The company provides location positioning data, context and data intelligence. It is a worldwide leader in this segment with a total funding amount of \$16.8 Mio.[42] Its customer's base is comprised of various high profile companies, such as *Alibaba.com*, *Sharp* and *Sony* [4].

4.6.4.1 Approach

Skyhook started with a purely software-based approach to location tracking; it leverages RSSI fingerprinting [41] and relies on their own database of access points and their corresponding signal strengths. *Skyhook* maintains this database up-to-date by surveying the area with specially equipped cars. As of 2020 *Skyhook* has over 5 billion WiFi access-points and 200 Million cell towers worldwide in their database[4]. The *Skyhook WPS* can be used both passively from the access-points retrieved RSSI fingerprint or actively by the device itself. The latter is done by sending the signal strength of the access points around the device to *Skyhook*'s servers. The software is highly scalable and does not require additional sensors, even though these can be integrated into the company's sensing network. Gallagher et al.[41] found a positioning accuracy of around 10-20 meters of outdoor and 30-70 indoor. As the provider maintains the system, there is no setup required to start using *Skyhook*'s positioning system.

The system is highly scalable but requires *Skyhook* to survey the location in order to create a signal strength heatmap. Once the location is surveyed, and in the database, it is ready to use. Additionally, by using their SDK, it is possible to retrieve the sensing device location within a few seconds [41].

4.6.4.2 Goal

Skyhook provides two core products; one is an SDK which helps devices accurately measure their own position in an environment even without access to the GPS. This solution relies on the client to send signal strength data to *Skyhook*'s data centre which then calculates the position and send it back. Additionally, *Skyhook* sells the aggregate data it collects from both the devices and the access points, as marketing research insights, object tracking and security [4].

4.6.5 CLOUD4WI, Inc.

CLOUD4WI is an American company, founded in 2014, which provides customer insights solutions. It is a worldwide leader in Guest-Wifi deployment and analysis with a total funding amount of \$18.1Mio [47] Its customer's base is comprised of various high profile companies, such as *Armani*, *BMW* and *Swatch* [46].

4.6.5.1 Approach

CLOUD4WI produces networking software which analyses customers behaviour both while connected to the WiFi and when offline. *CLOUD4WI*'s approach is based on passive RSSI fingerprinting. This approach is based on off-the-shelf enterprise WiFi access points. This allows them to track the behaviour of customers' within the access point range and aggregate this data together with guest-WiFi utilisation. Their primary source of localisation data is by customers using the guest-WiFi within stores or transit. Once a customer log-in *CLOUD4WI* can attach the account's data, such as gender, name, language and age to the device's IP and MAC address. This data is then aggregated and provided to the store as a marketing analytics tool. [46]

The system which *CLOUD4Wi* provides is highly scalable and can be deployed on existing enterprise infrastructure. No information about pricing and time to deployment is publicly available.

4.6.5.2 Goal

The main goal of *CLOUD4WI*'s platform is to gather data about the store's customers behaviour and demographics to then aggregate and generate in-depth marketing reports for the store-owners. By offering guest-WiFi, the customers accept the terms and conditions that allow *CLOUD4WI* to also aggregate their navigation and personal information together with the device's location within the access point [48] [46].

4.6.6 Combain

Combain is a Swedish company, founded in 2009, which provides indoor positioning and is a renowned geolocation service provider. It is a worldwide leader in indoor and outdoor WiFi-based positioning with a total funding amount of \$84.5K[44],[43]. Its customer's base is comprised of various high profile companies, such as *Nokia*, *Ericsson* and *Huawei* [43].

4.6.6.1 Approach

Combain developed and currently maintains a database containing billions of positions from crowd-sourced locations, with global coverage with over 120 million cell ids. Similar to *Skyhook*, the device positioning is done on-device as it sends the found AP's signal strength and sends them to *Combain*'s API. *Combain*'s data centres then calculate the

approximate location and send it back to the device, one such API request and reply sequence are visible on the sample request in Figure 4.9[43].

Combain solution is highly scalable and requires no new equipment to be placed on location. The basic access to their API is free, so the barrier of entry is shallow compared to other companies. As the only thing needed for the system to work, is for a device to send all the found networks around it and their respective signal intensity to their API.

Sample API request:

```
{
  "radioType": "gsm",
  "cellTowers": [
    {
      "mobileCountryCode": 240,
      "mobileNetworkCode": 1,
      "locationAreaCode": 3012,
      "cellId": 11950
    }
  ]
}
```

Sample API response:

```
{
  "accuracy": 500,
  "location": {
    "lat": 59.331706,
    "lng": 18.079073
  },
  "logId": 4088671
}
```

Figure 4.9: Sample request and reply to/from a Combain API [45]

4.6.6.2 Goal

The main goal of *combain*'s solutions is to provide precise indoor and outdoor positioning for WiFi, Bluetooth or cellular-based devices. *Combain* can also supplement existing infrastructure with tracking beacons which can be used to track assets and objects along the supply chain.

4.6.7 Comparison

Table 4.1 summaries the wide scope of applications that positioning solutions to cover. Some companies, such as *Locatee* are more of a one focus product company, while others such as *Skyhook* and *Combain* provide geopositioning solutions to a broader audience. These two companies also have a drastically different approach to device positioning, as they both maintain an up-to-date database which contains access points locations and signal strengths, which the devices can query in order to figure out their position. The rest of the positioning solutions do the analysis and data gathering on the access point side.

Another highlighted finding is that most companies do not rely on proprietary hardware, but can instead be deployed on existing infrastructure. It is also possible with these companies to add other sensors to improve the accuracy, but it is not needed. This finding is possible since all companies use RSSI fingerprinting, either active or passive, in order to track a device within space. This type of technology is usually computational heavy but does not require speciality components in order to work.

4.7 Privacy aspect of RSSI fingerprinting

This section aims to give a brief overview of possible security and privacy concerns raised by RSSI fingerprinting.

After multiple high-level scandals, the ethics and privacy concerns of big data analysis concerning location and tracking has come under public scrutiny. A key issue of RSSI fingerprinting and similar technologies is that it operated in an unregulated sector. The companies developing such technologies do not provide insights into the type of data they

Company	Location	Used Radio Frequency	Hardware	Used for
Locatee	CH	WiFi Bluetooth	Enterprise WiFi	Real Estate management
Onway	CH	WiFi	Proprietary AP	Transportation WiFi System Transportation Tracking
Livealytics	CH	WiFi Bluetooth	Enterprise WiFi Beacons	Marketing research
Skyhook	USA	WiFi Bluetooth Cellular	Surveyed data	Marketing research Asset tracking Geopositioning
Cloud4wi	USA	WiFi	Enterprise Wifi	Marketing research
Comabain	SE	WiFi Bluetooth Cellular	Crowd sourced data	Asset tracking Geopositioning

Table 4.1: Summary of discussed companies and their positioning solutions

collect and can infer from it, as it is a company secret. Even though most of the companies mentioned in this report gather data which is then anonymised, as a report from the New York Times [49] highlighted, the sheer amount of data which is collected is enough to pinpoint a device's movements throughout the year. As visible on figure 4.10 by analysing this data, they were able to trace multiple persons daily lives from home to the workplace and their free-time activities. This data is legally collected, and under American law, it is also legal to resell this data. While in western countries, the trust in the government is high, there are multiple instances where law enforcement in less democratic countries used such databases for retroactively tracking activists and monitoring their movements. In the last couple of years, the world has been shaken up by major pro-privacy and pro-freedom of speech protests, and more research is being done in order to find a way to continue developing technologies such as indoor positioning without eroding personal rights of its users [50].

The European Union has already started implementing more strict guidelines regarding data privacy; under the GDPR, users now have to opt-in in order for companies to track the device. Furthermore, further rules and guidelines are being developed in order to create a stricter regulation on how companies can gather, utilise and store the data. This only represents the first step towards more strict guidelines on data ownership and elaboration, and therefore multiple researchers have focused their attention towards making more privacy-oriented tracking systems [51],[52]

4.8 Future Trends in Radio based Positioning

In this section, this report aims to give an overview of the current developments in indoor localisation and their advantages or disadvantages.

4.8.1 Machine learning improvements

Recently, several machine learning-based positioning techniques have been proposed; these learning algorithms only use the Received Signal Strength Indication [53], and most do not require special equipment. A research conducted by Baccar & Bouallegue [54], presented a new architecture of a geolocation artificial intelligence-based application. The system required one week to be trained on data gathered within the building and the definition



Figure 4.10: Inferred movement pattern from one device sent ping over 1 year time [49]

of internal zones, once the model was trained, the AI-model was able to define the correct zone with 93% accuracy without data pre-processing. This was then compared to other learning algorithms and proved that their algorithm was able to achieve more than 12% increase in accuracy.

Subsequently, in 2017, Ahmadi et al. [55]. were developed a learning-based algorithm which was able to furtherly decrease the positioning error, even when the device was moving with variable speeds. The proposed approach used a Neural Network classification in order to determine the best anchor node combined with an ensemble algorithm based on RT and RSSI measurements. The proposed algorithm proved that it was possible to decrease the computational complexity while also decreasing the error in localisation. Such developments mean that Learning-based localisation algorithms would be more suitable for large scale networks, but also highlight the importance of proper training selection strategies [55].

The main advantages of using Machine Learning in indoor positioning is the reduction of complexity and computational power needed to get an accurate position, with the continuous development of more Neural-network centric processors, more and more devices are starting to be capable of performing such calculation locally on-device instead of relying on the cloud. This also offers a further advantage: data privacy, by removing the need to send the data in order to process the device and its user are in complete control over the process. A disadvantage that has been highlighted by Ahmadi et al. [55], is the fact that training the model requires a significant amount of high-quality data which contained both RSSI and precise location. Their approach is not scalable beyond buildings as it would need to retrain the model in each new location.

Nonetheless, as mentioned in section 4.7, a great deal of research is currently being done in order to decrease inherent data privacy issues. This is mainly done by using learning-algorithms, which are less computationally complex, and can be run locally on a device, without it needing to connect to databases. This removes the possibility of companies to track a person movements by snooping on the localisation requests the server received.

4.8.2 IoT development

Cisco annual Internet report of 2020 [1] forecasted that the number of WiFi hotspots will grow four-fold from 2018 to 2023, totalling over 628 million public WiFi hotspots and that by 2023 there will be 8.7 billion handheld or personal mobile-ready devices. The demand and usage of WiFi and Bluetooth enabled fitness tracking devices, such as Apple Watch, FitBit or Samsung Gear, is also growing and in 2018 reached a 10% adoption rate in the US. Globally the number of connected wearable devices worldwide has more than doubled in the space of three years, increasing from 325 million in 2016 to 722 million in 2019. The number of devices is projected to reach more than one billion by 2022. Internet-of-Things market is also expected to grow sixfold within the next five years [56], and one of the causes is the increase of need of real-time data to be segregated into real-time analytics, security and remote monitoring [57].

The growth in IoT devices' ubiquity also means that there will be a higher density of WiFi signals senders and receivers, these can be used as signal sending beacons or as nodes capturing RSSI [58] and uploading them to a datacenter. One such example of IoT beacons was proposed by Spachos & Plataniotis in 2020[58]; the beacons were used inside a museum in order to estimate a device's distance from an artwork and give contextual content. At the same time, the application could collect useful data about the visitor permanence in the museum.

The main advantage of IoT is the ubiquity of such devices; nowadays, it is easy to add IoT capabilities to any device, such as lighting fixtures, air ventilation systems and other electronic appliances. By incrementing the number of nodes retrieving RSSI data, the higher the accuracy becomes [59]. A direct disadvantage of this is that by deploying a large number of nodes, the technical overhead increases, and so does the complexity of the network.

4.8.3 UWB

Ultra-WideBand is one of the fastest-growing technologies adopted in Real-Time Location Systems (RTLS), with considerable growth potential due to its accuracy in tracking assets [60]. One significant advantage of using UWB in position approximation is that it has a minimal theoretical position area of just around 2 to 7 centimetres [5]. A research conducted in 2017 by Xu et al.[61] were able to achieve a 0.2m average error over by using just four nodes and walking in a circumscribed area.

This technology is also touted as a game-changing development that will give to future smart devices a more accurate spatial awareness and is currently already in most flagship phones, such as Apple's iPhone 11 onwards and Samsung S20 onwards. As the technology is currently still not widespread, the only use in these phones is that they can understand which of the phone's contact is closest and therefore prioritise message sending to the respective phone.

UWB main advantages are the lower power consumption, the natively more precise data they can gather [5] and now that mainstream companies are investing in it, the price per chip is rapidly decreasing. The fact that it is still such an emerging technology also brings its downsides, such as its availability outside of selected devices is lacklustre and that its use has not been proven on a consumer-level scale. Another critical issue is related to the frequency that UWB use, 3.1 to 10.6 GHz, which overlaps WiFi 802.11a which uses the 5GHz band, in order to mitigate any interferences UWB is operated at lower power (-41.3dBm/MHz) [5]. The performance of UWB also quickly deteriorate if the sender and receiver are in a Non-Line-Of-Sight conditions[62].

4.8.4 Combination of the three

UWB and Machine Learning can be used in conjunction in order to improve Non-Line-Of-Sight detection [62], by combining UWB with learning-algorithms the researchers were able to increase the position accuracy while also decreasing the computational performance needed to calculate the device's position. Together with the ubiquity of **IoT** devices and their low cost, it will be possible to effortless create a vast network of nodes, which will further reduce the error as they can then combine WiFi, BLE and UWB RSSI positioning [63].

4.9 Survey analysis

This section discusses the results of the survey conducted during the lecture, which results are depicted in Appendix I, pp. 135-136. 25 Master answered the survey, or Bachelor students enrolled in the seminar *2740/2741 - Internet Economics* at the University of Zurich and was handed out at the end of the presentation based on the first revision of this report.

The survey showed that the majority of the students was unaware of the existence of implementations (Q1) of such technologies discussed in this report. This is particularly interesting when combined with the fact that most students try to diminish their data footprint (Q4). As discussed in previous sections, many companies' only target is to gather granular details about the movements of the device in order to sell the data for marketing research purposes. This companies thrive on devices having their WiFi and/or Bluetooth always enabled, which also in such a more technical versatile audience, applies to roughly 75% of the surveyed students (Q2).

Overall, the surveyed audience did not express the desire to change their habits regarding their device utilisation's (Q5), even though the vast majority identify in such implementations a major privacy issues (Q6) if they do not trust the provider/the uses of the data (Q7).

4.10 Final Considerations

This report aimed to investigate the current use of Wireless Sensing technologies and their market. This was done by analysing the current state of the technologies and their advantages and disadvantages. Followed by an overview of current market trends and application, in order to then conclude with an overview of selected companies and their approaches to the analysed technology.

With the current development and the increase of general interest regarding privacy and the ethics of information systems, it is increasingly important to understand how technologies work in order to mitigate their negative effect on individuals. Overall the technology is well perceived by a knowledgeable audience, and the use cases do also provide important insights and data for the general population. One example of such use is provided by the contact-tracing protocol created by *Apple* and *Google* in 2020, in order to limit the outspread of COVID-19.

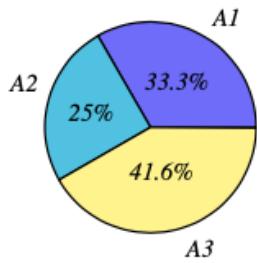
The lack of knowledge and understanding of the discussed technologies and their applications can undermine the trust of individuals, as visible by the low adoption rate of the various contact-tracing applications provided by governments. The general audience is unwilling to change their habits in order to limit their data footprint in databases such as those created by *Skyhook* or *Combain*. When combined with the increase in overall use of devices, the rise in IoT products and new network standards such as 5G and WiFi6in

the upcoming years, it means that companies providing services based on Wireless Localisation will be able to provide much more granular data about device movements and interactions.

Overall the technology discussed in this report will further develop and disrupt the workplace environment, marketing and public safety. With the constant technological developments both for devices and wireless networks, the ubiquity of Wireless sensing will only increase in uses and interest.

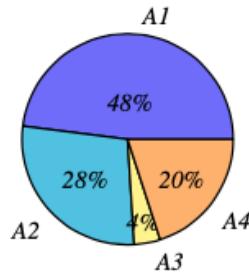
4.11 Appendix I: Survey results

Q1:Do you know of any application in which you experienced indoor localization?



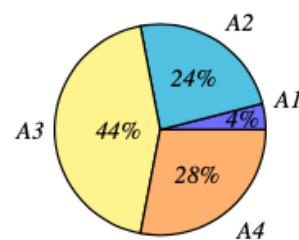
1. Yes
2. No
3. I wasn't aware of it before

Q2:How much do you leave Bluetooth/WiFi on?



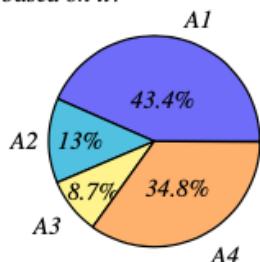
1. Always
2. Most of the time
3. Only in trusted locations
4. Only when needed

Q3:Do you connect to public hotspots (SBB/Starbucks/City FreeLan/...)?



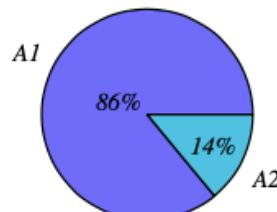
1. Yes, always
2. Yes, rarely
3. Yes, only when there is no cellular connection
4. Never

Q4:Are you conscious of which data you make available of yourself, or you just accept that companies tailor your content based on it?



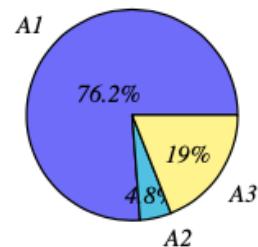
1. Yes, I try to hide my data as much as possible
2. Yes, I do whatever it takes to make myself invisible
3. No, better a tailored experience than one filled with useless stuff
4. No, It's impossible to escape every data collection method

Q5:Will you change anything to your wifi/bluetooth utilization now that you are aware of WiFi tracking?



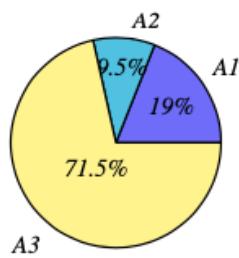
1. No
2. Yes

Q6:Do you think that such technology can pose major privacy issues?



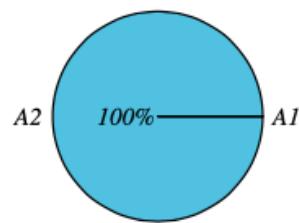
1. Yes
2. No
3. It depends

Q7:Do you think that such technology can pose major privacy issues?



1. Yes
2. No
3. Only if I know I can trust the provider

Q8:Would you prefer the use of Video cameras and Computer Vision compared to WiFi Sensing?



1. Yes
2. No

Bibliography

- [1] Cisco: Cisco Annual Internet Report - Cisco Annual Internet Report (2018-2023); White Paper, 2020 <https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html> Retrieved 6 October 2020.
- [2] Wireless sensors market - Future Market Insights, <https://www.futuremarketinsights.com/reports/wireless-sensors-market>, retrieved on 12.10.2020
- [3] Giuliano, Cardarilli, Cesarini, Di Nunzio, Fallucchi, Fazzolari, Mazzenga, Re, Vizzari: Indoor Localization System Based on Bluetooth Low Energy for Museum Applications, Electronics ,(Basel, Vol. 9, No. 6), June 2020, p. 1055-, doi:10.3390/electronics9061055.
- [4] Skyhook - Home Page <https://skyhook.com/> retrieved on 1.10.2020
- [5] Mautz, Rainer: Indoor positioning technologies. Habilitation Thesis, ETH Zurich, Department of Civil, Environmental and Geomatic Engineering, Institute of Geodesy and Photogrammetry (Zurich, Switzerland) 2012
- [6] Gioia, Sermi, Tarchi, and Vespe: On Cleaning Strategies for WiFi Positioning to Monitor Dynamic Crowds, Applied Geomatics, (Vol. 11, No. 4), (Springer Berlin Heidelberg), Dec. 2019, pp. 381-99, doi:10.1007/s12518-019-00260-z.
- [7] P. Davidson and R. Piche, "A Survey of Selected Indoor Positioning Methods for Smartphones, IEEE Communications Surveys & Tutorials, (Vol. 19, No. 2), Q2 2017, pp. 1347-1370, doi: 10.1109/COMST.2016.2637663.
- [8] Zafari, Gkelias, and Leung: A Survey of Indoor localisation Systems and Technologies, IEEE Communications Surveys & Tutorials, (Vol. 21, No. 3), Q3 2019, pp. 2568-2599, doi: 10.1109/COMST.2019.2911558.
- [9] S. He and S. -. G. Chan: Wi-Fi Fingerprint-Based Indoor Positioning: Recent Advances and Comparisons, IEEE Communications Surveys & Tutorials, (Vol. 18, No. 1), Q1 2016, pp. 466-490, doi: 10.1109/COMST.2015.2464084
- [10] ITWissen - Beacon Periode <https://www.itwissen.info/Beacon-Periode-beacon-period-BP.html>, retrieved on 10.10.2020
- [11] Maghdid, HS., Lami, IA., Ghafoor, KZ., Lloret, J.: Seamless Outdoors-Indoors Localization Solutions on Smartphones: Implementation and Challenges. ACM Computing Surveys. (Vol. 48, No. 4), 2012, pp. 1-34. <https://doi.org/10.1145/2871166>
- [12] Rosa F. D., Pelosi M. and Nurmi J.: Human-induced effects on RSS ranging measurements for cooperative positioning, Hindawi Int. J. Navig. Observation, (Vol. 2012), 2012, pp. 1-13

- [13] Youssef M. A., Agrawala A., Shankar A. U.: "WLAN location determination via clustering and probability distributions", Proc. 1st IEEE Int. Conf. Pervasive Comput. Commun. (PerCom), (Fort Worth, TX, USA), 2003, pp. 143-150
- [14] Mueller P., Raitoharju M. and Piche R.: A field test of parametric WLAN-fingerprint-positioning methods", in Proc. 17th Int Conf. Inf. Fusion, (Salamanca, Spain), 2014, pp. 1-8
- [15] Bargh, de Groote: Indoor Localization Based on Response Rate of Bluetooth Inquiries, Proceedings of the First ACM International Workshop on Mobile Entity Localization and Tracking in Gps-Less Environments, 2008, pp. 49-54, doi:10.1145/1410012.1410024.
- [16] Zonith - Products, <http://www.zonith.com/products/ips/>, last accessed 21. October 2011.
- [17] Povalac, A. and Sebesta, J.: Phase of Arrival Ranging Method for UHF RFID Tags Using Instantaneous Frequency Measurement, Proceedings of the 20th International Conference on Applied Electromagnetics and Communications (ICECom 2010), (Vol. 1), 2010, pp. 1-4
- [18] Larranaga, J., Muguira, L., Lopez-Garde, J.M. and Vazquez, J.I.: An Environment Adaptive ZigBee Based Indoor Positioning Algorithm, Proceedings of the 2010 International Conference on Indoor Positioning and Indoor Navigation (IPIN), (Zurich), September, 2010
- [19] Mautz, R. and Ochieng, W.Y.: A Robust Indoor Positioning and Auto Localisation Algorithm, Journal of Global Positioning Systems, (Vol. 6, No. 1), 2017, pp. 38-46.
- [20] Indoor Location Market, <https://www.marketsandmarkets.com/Market-Reports/indoor-location-market-989.html>, retrieved on 1.10.2020
- [21] Amin, Zhang: Radar Signal Processing for Elderly Fall Detection: The Future for in-Home Monitoring, IEEE Signal Processing Magazine, (Vol. 33, No. 2), March, 2016, pp. 71-80, doi:10.1109/MSP.2015.2502784.
- [22] A. Khalili, A. Soliman, M. Asaduzzaman and A. Griffiths: Wi-Fi sensing: applications and challenges, in The Journal of Engineering, (Vol. 2020, No. 3), March, 2020, pp. 87-97, doi: 10.1049/joe.2019.0790.
- [23] Filonenko, V., Cullen, C. and Carswell, J.: Investigating Ultrasonic Positioning on Mobile Phones, Proceedings of the 2010 International Conference on Indoor Positioning and Indoor Navigation (IPIN), (Zurich) September, 2010
- [24] Y. Wang, K. Wu, and L.M. Ni: Wifall: Device-free fall detection by wireless networks, IEEE Transactions on Mobile Computing, (Vol. 16, No. 2), 2017, pp. 581-594
- [25] Reschke, Schwarzl: Situation Awareness Based on Channel Measurements, 2011 IEEE 73rd Vehicular Technology Conference (VTC Spring), IEEE, 2011, pp. 1-5, doi:10.1109/VETECS.2011.5956453.
- [26] C. Xu, B. Firner, R. S. Moore, Y. Zhang, W. Trappe, R. Howard, F. Zhang, and N. An: Scpl: Indoor device-free multi-subject counting and localization using radio signal strength, The 12th ACM/IEEE Conference on Information Processing in Sensor Networks (ACM/IEEE IPSN), 2013, pp. 79-90, doi:10.1145/2461381.2461394.

- [27] S. Depatla, A. Muralidharan, and Y. Mostofi: Occupancy estimation using only WiFi power measurements, IEEE Journal on Selected Areas in Communications, (Vol. 33, No. 7), 2015, pp. 1381-1393
- [28] S. Depatla, A. Muralidharan, and Y. Mostofi: Occupancy estimation using only WiFi power measurements, IEEE Journal on Selected Areas in Communications, (Vol. 33, No. 7), 2015, pp. 1381-1393
- [29] A. Vertiy, S. Gavrilov, V. Stepanyuk, and I. Voynovskyy: Through-wall and wall microwave tomography imaging, IEEE Antennas and Propagation Society International Symposium (APS'04), (Vol. 3), June, 2004, pp. 3087-3090
- [30] K. Chetty, G.E. Smith, and K. Woodbridge: Through-the-wall sensing of personnel using passive bistatic Wi-Fi radar at standoff distances, IEEE Transactions on Geoscience and Remote Sensing, (Vol. 50, No. 4), 2012, pp.1218-1226
- [31] Wilson, Patwari.: See-Through Walls: Motion Tracking Using Variance-Based Radio Tomography Networks., IEEE Transactions on Mobile Computing, (Vol. 10, No. 5), May 2011, pp. 612-21, doi:10.1109/TMC.2010.175.
- [32] P. E. Bull: Posture and gesture, International Series in experimental social psychology - Pergamon press (Vol. 16), 1987.
- [33] M. Zhao, F. Adib, and D. Katabi: Emotion recognition using wireless signals, Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking, October, 2016, pp. 95-108
- [34] A. Kamran, A. X. Liu, W. Wang, and M. Shahzad: Recognizing Keystrokes Using Wi-Fi Devices., IEEE Journal on Selected Areas in Communications, (Vol. 35, No. 5), 2017, pp. 1175-1190
- [35] Karpagavalli P, Ramprasad AV: Estimating the density of the people and counting the number of people in a crowd environment for human safety, Proceedings 2013 IEEE conference on communication and signal processing. 2013, p. 663-7.
- [36] Locatee - Home Page, <https://locatee.com> retrieved on 1.10.2020
- [37] Locatee - Crunchbase Company Profile & funding, <https://www.crunchbase.com/organization/locatee>, retrieved on 1.10.2020
- [38] Onway - Home Page <https://onway.ch/> retrieved on 1.10.2020
- [39] Livealytics - Home Page <https://livealytics.com/> retrieved on 1.10.2020
- [40] Screenshot taken from video demonstration, on <https://www.youtube.com/watch?v=GI1bJXUq04I> , retrieved on 1.10.2020
- [41] Gallagher, T., Li, B., Kealy, A., & Dempster, A. G.: Trials of commercial WiFi positioning systems for indoor and urban canyons, IGNSS 2009 Symposium on GPS/GNSS, 2009, December.
- [42] Skyhook - Crunchbase Company Profile & funding, <https://www.crunchbase.com/organization/skyhook-wireless> , retrieved on 1.10.2020
- [43] Combain - Home Page <https://combain.com/> retrieved on 1.10.2020
- [44] Combain - Crunchbase Company Profile & funding, <https://www.crunchbase.com/organization/combain> , retrieved on 1.10.2020

- [45] Combain - API documentation, <https://combain.com/api/#combain-location-api> retrieved on 1.10.2020
- [46] Cloud4wi - Home Page <https://cloud4wi.com/> retrieved on 1.10.2020
- [47] Cloud4wi - Crunchbase Company Profile & funding, <https://www.crunchbase.com/organization/witech> , retrieved on 1.10.2020
- [48] Cloud4wi - How to profit from wifi analytics, <https://cloud4wi.com/resources/how-to-profit-from-wifi-analytics/> retrieved on 1.10.2020
- [49] Thompson, S., & Warzel, C.: Twelve Million Phones, One Dataset, Zero Privacy. <https://www.nytimes.com/interactive/2019/12/19/opinion/location-tracking-cell-phone.html> Published on December 19, 2019, Retrieved November 12, 2020
- [50] M. Smith, C. Szongott, B. Henne and G. von Voigt: Big data privacy issues in public social media, 2012 6th IEEE International Conference on Digital Ecosystems and Technologies (DEST), (Campione d'Italia), 2012, pp. 1-6, doi: 10.1109/DEST.2012.6227909.
- [51] L. Schauer, F. Dorfmeister and F. Wirth: Analyzing passive Wi-Fi fingerprinting for privacy-preserving indoor-positioning, 2016 International Conference on localisation and GNSS (ICL-GNSS), (Barcelona), 2016, pp. 1-6, doi: 10.1109/ICL-GNSS.2016.7533851.
- [52] Leping Huang, K. Matsuura, H. Yamane and K. Sezaki: Enhancing wireless location privacy using silent period, IEEE Wireless Communications and Networking Conference, (New Orleans, Vol. 2) 2005, pp. 1187-1192, doi: 10.1109/WCNC.2005.1424677.
- [53] Zheng J, Dehghani A.: Range-Free localisation in Wireless Sensor Networks with Neural Network Ensembles. Journal of Sensor and Actuator Networks. (Socorro, Vol. 1, No. 3), November, 2012, pp. 254-271.
- [54] N. Baccar and R. Bouallegue: Intelligent type 2 fuzzy-based mobile application for indoor geolocalisation, 2015 23rd International Conference on Software, Telecommunications and Computer Networks (SoftCOM), (Split), 2015, pp. 165-169, doi: 10.1109/SOFTCOM.2015.7314101.
- [55] Ahmadi, H, Viani, F, Polo, A, Bouallegue, R.: Learning ensemble strategy for static and dynamic localisation in wireless sensor networks. Int J Network Mgmt, (Vol. 27, No. 4) July, 2017, p. e1979-n/a <https://doi.org/10.1002/nem.1979>
- [56] H. Tankovska: Sep 23, 2020 Connected wearable devices worldwide 2016-2022. In Statista - The Statistics Portal. <https://www.statista.com/statistics/487291/global-connected-wearable-devices/> Retrieved October 1, 2020
- [57] Fortunebusinessinsights.com: Internet Of Things Market Size, Growth | Iot Industry Report [2020-2027]; 2020 <https://www.fortunebusinessinsights.com/industry-reports/internet-of-things-iot-market-100307>, retrieved on 1 November 2020
- [58] S. Sadowski and P. Spachos: RSSI-Based Indoor localisation With the Internet of Things, in IEEE Access, (vol. 6), 2018, pp. 30149-30161, doi: 10.1109/ACCESS.2018.2843325.

- [59] O. Kaltiokallio and M. Bocca: Real-Time Intrusion Detection and Tracking in Indoor Environment through Distributed RSSI Processing, 2011 IEEE 17th International Conference on Embedded and Real-Time Computing Systems and Applications, (Toyama), 2011, pp. 61-70, doi: 10.1109/RTCSA.2011.38.
- [60] Researchandmarkets.com: Ultra-Wideband Market By Application (RTLS, Imaging, Communication), Positioning System (Indoor, Outdoor), Vertical (Health-care, Automotive & Transportation, Manufacturing, Consumer Electronics, Residential, Retail), & Geography - Global Forecast To 2025; 2020 <https://www.researchandmarkets.com/reports/5025111/ultra-wideband-market-by-application-rtls> retrieved on 2 November 2020
- [61] Y. Xu, Y. S. Shmaliy, Y. Li and X. Chen: UWB-Based Indoor Human Localization With Time-Delayed Data Using EFIR Filtering, in IEEE Access, (vol. 5), 2017, 16676-16683, doi: 10.1109/ACCESS.2017.2743213.
- [62] S. Krishnan, R. Xenia Mendoza Santos, E. Ranier Yap and M. Thu Zin: Improving UWB Based Indoor Positioning in Industrial Environments Through Machine Learning, 2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV), (Singapore), 2018, pp. 1484-1488, doi: 10.1109/ICARCV.2018.8581305.
- [63] R. Ijaz, M. A. Pasha, N. U. Hassan and C. Yuen: A novel fusion methodology for indoor positioning in IoT-based mobile applications, 2018 IEEE 4th World Forum on Internet of Things (WF-IoT), (Singapore), 2018, pp. 742-747, doi: 10.1109/WF-IoT.2018.8355168.

Chapter 5

State of Blockchain-Based Banking Services

Marion Dübendorfer, Ramon Solo de Zaldivar, Raphael Imfeld

Blockchain (BC) has been designated to be the next revolutionary catalyst of the global financial industry. The Swiss financial sector, in particular the banking sector, which is one of the leading banking sectors worldwide [1], has welcomed new players that aim to bring BC to the forefront. But how true is this forecast? Is BC really as good as it seems? Do banks really need this technology? What is the regulatory situation in Switzerland? This report aims to answer these questions. More specifically, it aims to discuss to what extent BC is being implemented by the Swiss banks as of today and what laws are being made to facilitate the use of BC technology in the financial sector. In order to dive deep into the subject we first establish how BC works at its basic level, why its technology is considered revolutionary, and what advantages and challenges it can present with respect to banking. Then we look at different scenarios in which Swiss banks can or are implementing BC. We analyze a couple of key banks in Switzerland and discuss who is implementing which BC service and who is looking to expand their offering. Furthermore, we assess the structural, technical and legal challenges for the Swiss banking sector with regards to BC. The results of the assessment show that BC is indeed the banking technology of the future, however not without its drawbacks. Although some are hesitant, a lot of established banks are trying to broaden their offering with regards to BC services, some partly because of the fear of missing out on the potential benefits that BC-based services could bring on a operational level, others because of the positive market-share effects of being a first mover. Furthermore, the legal landscape in Switzerland is shaping an even and transparent playground for the whole financial sector, which helps instill trust in the technology, allowing banks and other entities to confidently invest into new BC ventures.

Contents

5.1	Introduction to Blockchain	144
5.1.1	The Blockchain Architecture	144
5.1.2	Consensus Algorithms	146
5.1.3	Advantages and Challenges of Blockchain	149
5.2	Blockchain-based Banking Services	151
5.2.1	Tokenization	151
5.2.2	Lending	152
5.2.3	Smart Contracts	153
5.2.4	Over the counter Trading	153
5.2.5	Custody	153
5.2.6	Storage	154
5.3	Swiss Banking Landscape	154
5.3.1	Julius Baer	154
5.3.2	Sygnum Bank	154
5.3.3	Seba Bank AG	155
5.3.4	Swissquote	155
5.3.5	Incore Bank	155
5.3.6	Falcon Private Bank	155
5.3.7	BEKB	155
5.3.8	UBS	155
5.4	Challenges for Swiss Banking	156
5.4.1	Structural	156
5.4.2	Technical	158
5.4.3	Legal	160
5.5	Conclusion	161
5.6	Limitations	162

5.1 Introduction to Blockchain

This section describes the basic principles of BC and addresses the following topics: The architecture underlying the BC technology, the essential security procedures which involve consensus algorithms, as well as the primary advantages and disadvantages of the BC technology.

5.1.1 The Blockchain Architecture

At its most basic level, a BC is just a sequence of blocks, where a block can store any sort of digital information and the chain is a public database that stores said blocks. Furthermore, two blocks are linked through their hashes. Each block stores its own hash as well as the hash of the previous block, which is also called the parent block. The first block of a BC is called genesis block and has no parent block.

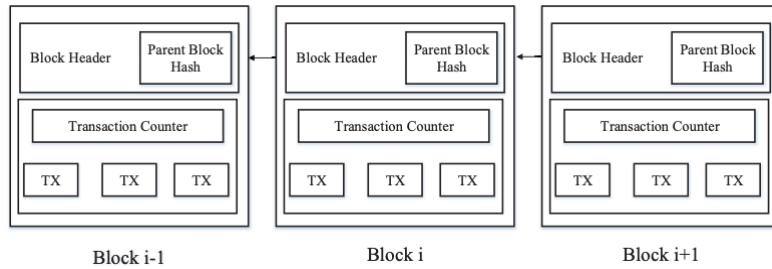


Figure 5.1: An Example of a Blockchain [2]

Figure 5.1 depicts an illustration of a BC. As we can see, each block consists of a block header and a block body. As mentioned above, the blocks are linked through their headers which store the parent block's hash. It is important to note that if any data inside the block body changes, its hash value is computed anew. If this is the case, the block's hash does not match anymore with the hash that is stored on the child block's header and which links the two blocks together. Thus, if the data inside a block's body changes, the BC starting from that block either becomes invalid, or all hash values starting from that block have to be computed anew and be approved by the network.

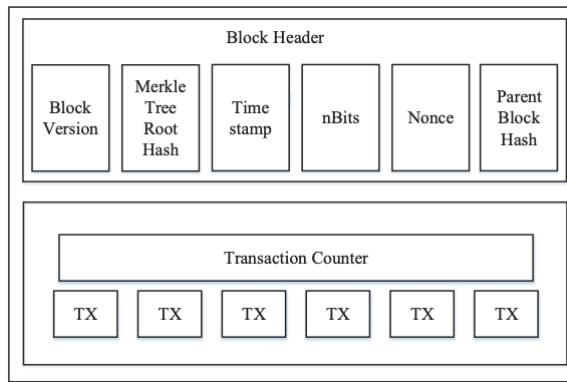


Figure 5.2: An Example of a Block in a Blockchain [2]

As we can see in Figure 5.2, the block header includes the block's version number, the hash value of all the transactions in the block, a timestamp, and the parent block's hash [2]. The block header may also contain a nonce, the concept of which will be addressed in section 1.2.1. The block body contains a transaction counter and information about the transaction the block embodies. For example, in the case of Bitcoin, the information

Property	Public	Consortium	Private
Centralization	No	Partial	Yes
Consensus determination	All miners	Selected set of nodes	One organization
Read permission	Public	Public/restricted	Public/restricted
Immutability	Difficult to manipulate	Mutable	Mutable
Efficiency	Low	High	High
Consensus process	Permissionless	Permissioned	Permissioned

Table 5.1: Types of Blockchain and their Properties[2]

about a transaction would consist of the sender, the receiver, and the amount of coin of the transaction.

Moreover, BCs are managed in a decentralized, distributed peer-to-peer network. More specifically, every node that joins a BC network receives a full copy of said BC. Whenever a new potential block gets added to the BC, each node in the network receives a copy of that block. The network, or rather the nodes in the network, then decide whether the data in that block is valid or not. If the network agrees on the authenticity of a block, it gets added to the BC, or rejected otherwise. This creates *consensus* among the nodes and ensures that the BC is in a valid state at all times. The different mechanisms that can be used in order to achieve consensus are discussed in section 1.2. As of today, BC systems can be categorized roughly into three types of BC: public blockchain, private blockchain and consortium blockchain (also called permissioned blockchain)[3]. In a public BC, anyone can take part in transactions as well as participate in the consensus mechanism. Public BCs are secured by the combination of economic incentives and cryptographic verification [3] using consensus mechanisms such as Proof of Work (PoW) and Proof of Stake (PoS), which will be discussed in section 1.2.

In a private BC, write permissions are limited to a specific group of participants, while reading access is either public or otherwise restricted to some extent. Furthermore, private BCs are commonly managed and controlled by a single organization or authority. In a consortium BC, a subset of an available pool of nodes is selected to create consensus. A comparison among these three types of BC is depicted in Table 5.1.

Centralization: The main difference between public and consortium or private BCs is that a public BC is decentralized, with each node independent from all other nodes. With consortium and private BCs, this is not the case. While a consortium BC is only partially centralized, a private BC is fully centralized since it is fully controlled by a single authority or organization[2].

Consensus determination: In a public BC, any computer is allowed to join the consensus mechanism. Meanwhile, in a consortium BC, only a set of nodes is selected to validate a new block. Finally, in the case of a private BC, only the organization that administers the BC is able to regulate the final consensus.

Read permission: In a public BC, transactions are always accessible by the public. However, in consortium and private BCs, reading access to transactions can either be public or restricted to a certain extent[2].

Immutability: In the case of a public BC, every node in a network stores a copy of the BC. This makes it very difficult to tamper with data since manipulation would not go unnoticed. Meanwhile, in a consortium or private BC, it is easier to tamper with data,

since the number of nodes is limited and, in the case of a private BC, even originate from the same organization or central authority.

Efficiency: Since there usually is a large number of nodes in a public BC, it takes longer to validate a transaction against every single node in a network. This leads to lower efficiency and less transactions per second (TPS). Therefore, with fewer validators as is the case with consortium and private BCs, BCs are more efficient in terms of throughput.

Consensus process: While anyone can join the consensus process of a public BC, a new node has to be permitted to join the network in the case of a consortium or private BC[2].

5.1.2 Consensus Algorithms

In this section, we will investigate which procedures can be used to reach consensus in a BC network. A BC is a distributed database which holds a continuously growing list of records of transactions. The BC is controlled by a certain amount of nodes that may not trust each other. New blocks are added to the BC through a distributed protocol executed by the controlling nodes. The nodes communicate over a network and agree on a new state of the BC without relying on a central authority[4].

Naturally, individual nodes might crash, behave maliciously, or the network communication could be interrupted [4]. In order to guarantee a continuous service, the network runs a fault-tolerant consensus protocol to ensure that all transactions occurring are valid and all nodes agree on the current state of the BC[4]. The fundamental consensus approaches are Proof of Work (PoW) and Proof of Stake (PoS). However, there are many newer, less traditional consensus protocols like Stellar Consensus Protocol (SCP) [5] and Ripple Protocol Consensus Algorithm (RPCA) [4] of the Ripple cryptocurrency.

5.1.2.1 Proof of Work (PoW)

In a decentralized network, a node has to be selected to record a transaction. Instead of choosing a random node, a node has to prove that it is trustworthy, or in other words, not likely to attack a network. For a node to be considered trustworthy, it needs to do a lot of work, which in this case means solving a cryptographic puzzle. In Proof of Work, every node of a network calculates the hash value of a new block. The new block contains a nonce (a 4-byte field, which usually starts with 0 and increases for every hash calculated) which changes frequently[2]. In order to reach consensus, the calculated hash must be equal to or smaller than a certain given value. As soon as one node reaches the target value, it broadcasts the block to other nodes, which in turn must confirm the correctness of the hash value. After the block is validated, each node in the network appends the new block to their copy of the BC[2]. Once the proof of work has been satisfied, the block cannot be changed without redoing the work for that specific block and all blocks that come after it [5]. The procedure of calculating a hash value and adding new blocks to the BC is called mining and the nodes that try to solve these cryptographic puzzles are called miners.

Notably, the mining process requires a vast amount of computing power and thus results in high electricity consumption as well as electronic waste. A more sustainable alternative in terms of energy consumption to Proof of Work is Proof of Stake.

5.1.2.2 Proof of Stake (PoS)

In the Proof of Stake approach, miners have to prove the ownership of a certain amount of currency, or in other words, provide proof of their stake. This approach is mainly based on the assumption that people with more currency involved are less likely to attack the network since they hold more at stake. The problem with this approach by itself is that

it is plutocratic, meaning that the few richest participants dominate the network. To mitigate this, apart from the stake that miners contribute, each coin's age is taken into consideration. When spending a coin in a transaction, said age is dissolved and reset to zero.

Under the Proof of Stake system, a miner pays himself (and thereby consumes his coin age) in order to gain the privilege to mine a new block for the BC. The following condition determines the amount that a miner needs to contribute in order to mine a new block:

$$\text{proofhash} < \text{coins} \times \text{age} \times \text{target},$$

where *coins* are the number of coins a miner has spent for the mining privilege, *age* is the age of the coins that have been spent, and *target* is the required amount of coin specified by the network through a difficulty adjustment process similar to the Proof of Work implementation[5]. *Proofhash* is the sum that depends on a stake, the unspent output, and the current time[5].

While Proof of Stake is more sustainable and more efficient, since mining cost is nearly zero, it is more susceptible to hacker attacks.

5.1.2.3 Stellar Consensus Protocol (SCP) and Ripple Protocol Consensus Algorithm (RPCA)

As defined in [6], the Stellar Consensus Protocol (SCP) is a decentralized consensus protocol where nodes in a network do not have to trust the entire network. Instead, each node can choose which nodes to trust. A group of nodes that trust each other is called a *quorum slice*. Furthermore, a quorum is a set of nodes sufficient to reach an agreement, and a quorum slice is a subset of a quorum[5].

SCP starts with a *nomination protocol* by proposing potential values for agreement. Every node that receives these values votes for one of these values. This of course eventually results in one value winning the majority. After the nomination protocol, the *ballot protocol* is executed[7]. In that protocol, nodes vote on whether the values selected in the nomination protocol should be committed or not. If a set of nodes can not reach consensus, the values are moved to a higher ranked ballot which then vote on these values again[7].

The Ripple Protocol Consensus Algorithm (RPCA) works somewhat similar to SCP. As the name implies, RPCA is used by the Ripple cryptocurrency[8].

The process of adding blocks is controlled by validating nodes and works as follows: These nodes periodically start to create a new block and vote in rounds on its content. Each node takes all valid transactions it has seen prior to a new consensus round and puts them in a list[9]. Each node then votes on the correctness of each transaction in one or multiple rounds[9]. Each node accepts a proposed transaction if 50%, ..., 80% (increasing by 10% in each round) of the signed updates that it receives match[4]. Accordingly, all transactions that meet a minimum of 80% approval in the final round are written to the BC. This consensus algorithm was designed to address latency issues present within other algorithms[5].

5.1.2.4 Comparison of Consensus Algorithms

In this section, a comparison of the consensus algorithms discussed above with regards to their popularity in cryptocurrencies, energy consumption and efficiency will be performed.

Currency	Consensus Algorithm	Market Cap
Bitcoin	Proof of Work	\$ 639.9 B
Ethereum	Proof of Stake	\$ 121 B
Ripple	Ripple Protocol Consensus Algorithm	\$ 13.4 B
Bitcoin Cash	Proof of Work	\$ 8.8 B
Stellar	Stellar Consensus Protocol	\$ 6.4 B

Table 5.2: Top Five Cryptocurrencies by Market Cap (as of 01/2021)[10]

As can be seen in Table 5.2, apart from Ripple with its own consensus algorithm, Proof of Work is the most popular consensus protocol, with popularity in this context defined by the market cap of a cryptocurrency. It should be noted that only as of December 2021, Ethereum has switched to a Proof of Stake protocol[11] and was using Proof of Work before. Meanwhile, currencies like Ripple or Stellar show that different approaches from Proof of Work or Proof of Stake are likely to gain popularity in the future.

Property	PoW	PoS	RPCA	SCP
Energy-saving	No	Partial	Yes	Yes
Tolerated power of adversary	<25% computing power	<51% stake	<20% faulty nodes	Variable

Table 5.3: Consensus Algorithm Properties [2][5]

Since it is impossible to provide precise numbers about how much energy each consensus mechanism consumes, the binary energy-saving property in Table 5.3 is presented in comparison to other mechanisms rather than as an absolute number.

In the case of PoW, miners are continuously computing new hashes in order to match a certain value (as discussed in section 1.2.1). Accordingly, the electricity consumption of the mining process is vast. Regarding PoS, miners still have to calculate hashes in order to match the proofhash (see secteion 1.2.2), but the mechanism is designed to reduce computing effort on the basis of the stake involved. Lastly, in the case of SCP and RPCA, there is no mining involved in the consensus mechanism and thus huge amounts of energy can be saved.

The *tolerated power of adversary* relates to the security of a consensus mechanism. More specifically, it refers to how much control an attacker would need to gain over a network in order to be able to overwhelm it [5]. In PoW and PoS, an adversary group would need control over 51% of the computing power of a network in the case of PoW and 51% of the stake in the case of PoS to be able to overwhelm it. However, in PoW, miners could gain more revenue with only 25% of the computing power by using selfish mining strategies [2]. A selfish mining strategy stands for increasing profit of certain miners by holding back blocks they themselves created and postponing their inclusion into the BC [12]. Ripple is proved to maintain correctness if the share of faulty nodes in the network is less than 20% [2]. In the case of SCP, it is impossible to calculate an absolute tolerance, since each node can choose which other nodes to trust. Hypothetically, an attacker could control a significant share of the network but not be trusted by any of the other nodes. On the other hand, the attacker could just manipulate separate quorum slices instead of the whole quorum, thus creating false trust [5].

5.1.3 Advantages and Challenges of Blockchain

While BC is one of the most revolutionary, attention-grabbing technologies in the internet economics of today, it is also a disputed technology that is far from being widely accepted. On the one hand, its potential to modernize and optimize not only the financial sector as well as the banking landscape, but also other domains like healthcare, supply chain management, identity management and many more is undeniable. On the other hand, some valid concerns have been raised with regard to the sustainability and efficiency of BC. In this section, we will discuss the main advantages and disadvantages of the distributed ledger technology that is BC.

5.1.3.1 Advantages

BC is often referred to as a *distributed, decentralized ledger*. This definition incorporates two of the main advantages of BC, namely decentralization and distribution.

Decentralization: As opposed to traditional services, BC does not store any of its digital information in a central database that is managed by a third party. Instead, a BC is spread across a network of many independent nodes. This decentralized approach, paired with the consensus algorithms discussed in section 1.2, makes data manipulation increasingly difficult. If a node would try to corrupt its copy of a BC, all other copies remain intact and the consensus algorithm would deem the corrupted copy invalid. Decentralization can also lead to cost savings and increased availability, which will be discussed later.

Distribution: Since a BC network is distributed onto a large set of nodes, it allows for distributing the computing power. In a public BC, any computer can join the network, thus reducing the risk for the network to be overwhelmed by a majority with malicious intent. Therefore, distribution reduces the risk of tampering, cyber crime and fraud [13]. It must be noted, however, that this only holds for a public or partly permissioned BC, but not necessarily for a private one.

Cost reductions: BC eliminates the need for third-party-verification costs as well as transaction fees[14]. That being said, financial services that implement BC offer potential cost reductions.

Availability: In traditional banking services, where transactions are processed through a central authority, it can take several days for a transaction to settle. While traditional institutions only operate during business hours, BC technologies are available at every hour of the day, every day[14]. This can make transactions much faster, which in turn makes them safer.

Open Source: Most consensus protocols are open source software, as are most of the public BC systems. This is an advantage since open source technologies pose low entry barriers into said technology, which in turn lead to a strong developer base. Moreover, open source code can be an advantage because it is produced collaboratively and transparently[15]. Also, open source software is extended and refined for the sake of the community, not the profit of a single institution. In summary, the fact that the BC technology is open source enhances the transparency and sincerity of public BCs.

Transparency: In public BCs, anyone can access, verify, and track transactions. For instance, using so-called *Blockchain Explorers*, anyone has permission to read the entire transaction history of a public BC. Thanks to the transparency displayed in BC networks, attempts to corrupt a BC are unlikely to go unnoticed.

5.1.3.2 Challenges

Scalability: BCs only ever grow larger, not smaller, since information about transactions in the form of blocks are added continuously and are never removed. Of course, each node has to store the whole chain to maintain consensus. Possible measures to mitigate the

BC growing too fast or getting too big in terms of storage, is restricting the block size or the amount of transactions per time unit. Thus, the more popular a BC gets, the more inefficient it gets. In other words, BCs scale badly.

Efficiency: BCs can be very inefficient. While this does not hold for every BC, it is almost certainly true for BCs that reach a certain size. For example, the PoW mechanism of Bitcoin allows for a new block to be added to the BC only every 10 minutes. This results in a maximum capacity of 7 transactions per second (TPS) [16]. In comparison to the 24'000 TPS that Visa can handle, this is very inefficient[17].

Sustainability: BCs that use mining as consensus mechanism are very unsustainable in terms of carbon footprint, electrical energy consumption and carbon footprint. In the case of Bitcoin, the annualized carbon footprint is comparable to that of New Zealand. Similarly, the carbon footprint of one single transaction is comparable to that of 756'605 Visa transactions[18]. What makes the lack of sustainability worse is that it is quite ironic; if only a small proportion of miners existed, only a small proportion of computing power would be needed, however, the transaction rate would not decrease and the BC would still operate at the same speed. Simultaneously, a critical mass of miners has to be reached in order to guarantee consensus. Thus it can be said that with mining mechanisms, sustainability is sacrificed for the sake of security.

5.2 Blockchain-based Banking Services

BC has the potential to thrive in the financial sector and have a positive impact for banks, i.e increase efficiency and lower costs. One of these positive impacts is in the area of transactions. In order to understand what the benefits are, we first need to understand how payments work today in Switzerland.

On a high level, payments in Switzerland mostly revolve around the firm SIX, which is a financial services company owned by Swiss financial institutions. Inter bank payments go through the Swiss National Bank, which uses the SIX Interbank Clearing system (SIC) to process these transactions. SIC acts as an intermediary between banks and regulates the payments. Although the Swiss transactions system is mostly automated, it still needs manual control and supervision, specially in the area of payment settlement, the process of fully crediting the owed amount minus transaction fees to the payee [19].

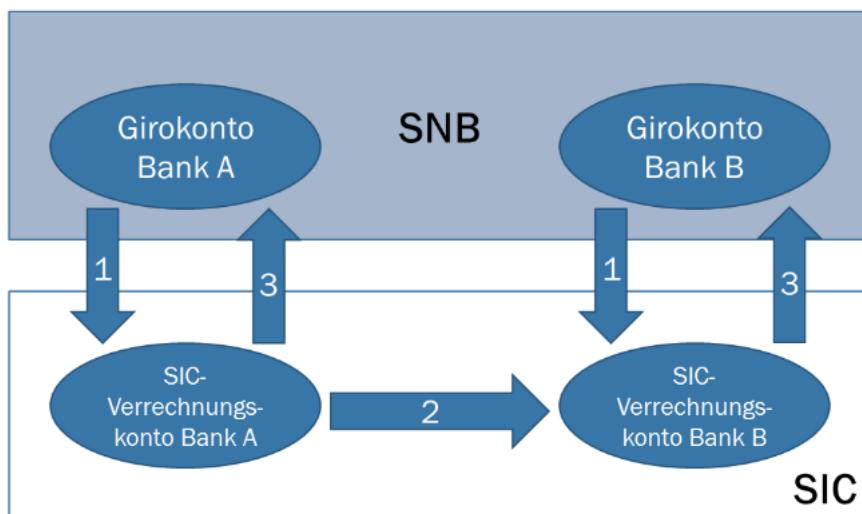


Figure 5.3: Illustration of SIC Payment Settlement Workflow

Source: Bakbasel

Figure 5.3 provides basic understanding of how the SIC system works. For example, if person A sends person B 100 Swiss francs, the bank of person A sends the transaction instruction to the SIC, which in turn gets in contact with person B's bank and then settles the transaction. This means that the settlement of the transaction and the involved accounts happens internally at the Swiss National Bank using the SIC system. As BC is based on distributed ledgers, it provides real-time verification and automated settlement of payments, thus reducing the administrative costs, minimizing settlement latency and has a positive impact on efficiency.

5.2.1 Tokenization

Tokenization of assets is the process of issuing a security token, which is a random string of characters, by an initial security token offering (STO) that digitally represents a given asset [20]. A security token offering is the process in which the investor receives a token that represents their investment, that is also asset backed [21], meaning that if the borrower would fail to return the investment, the ownership of this asset would then be transferred to the investor. A security token, which after being issued, is recorded on a BC, making the information immutable and can represent any asset, i.e. equities, bonds real estate, metals, and even to copyrights of ownership, digitally. This brings multiple benefits.

For one, as tokens are highly divisible, it is possible to buy or sell fractions of an asset. This causes the minimum investment fees and investment period to lessen and in turn

makes a specific market more attractive to invest for less wealthy investors, as they won't have to invest high amounts over a long period of time, thus broadening the amount of potential investors. Furthermore, transparency is also guaranteed, as a token has the capability of having the owner's rights and legal responsibilities recorded on itself along with an immutable record of ownership, allowing investors to know exactly from whom they're buying and who owned the token previously since it was issued [22].

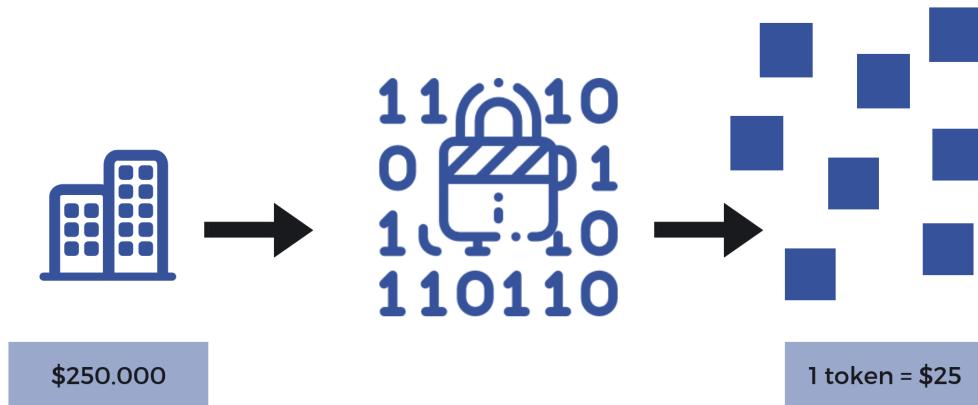


Figure 5.4: Asset Tokenization [22]

Figure 5.4 shows how an assets can be converted, through STO, to multiple tokens that together amount to the total price of the asset. These tokens can then be traded independently.

Tokenization has the potential of reducing trading transaction times and administrative costs by permitting automated transaction settlements, as no manual verification would have to be completed. Trading settlement latency would also decrease together with smart contracts. Smart contracts are programmable contracts with predefined parameters that carry themselves out automatically as soon as these predefined terms are met. Smart contracts are also stored on a BC which makes it almost impossible to tamper with them. Issuing a security token for a given asset also minimizes physical paperwork, making the transactions more cost and time efficient.

5.2.2 Lending

Another advantage of BC can be found in the area of loans and mortgages. On the one hand, crypto-asset collateralized lending enables borrowers to provide their crypto-assets, which can be any cryptocurrency or tokenized assets, as collateral in order to get a loan. The loan can be either in fiat money or cryptocurrencies. Crypto-asset collateralized lending highly resembles a lombard loan, where the loan amount is backed by a deposit of for example, fiat money, stocks, bonds, life insurance policies and other assets [23]. However, instead of having these aforementioned type of assets as collaterals, crypto-assets are used instead. These crypto-assets are then stored in a crypto vault for the duration of the repayment of the loan[24]. A crypto vault is a crypto wallet, which is software developed to store transact with crypto-assets, with added security and encryption layers, that is stored on a piece of hardware that are not connected to the internet.

On the other hand, mortgages can also be simplified with BC. A BC service that is not yet available in Switzerland [25] is the automation of the mortgage application and acceptance process. Instead of spending multiple months and money on fees to get a mortgage, the borrower can complete the application online. The application is then added to a block on the BC upon submission. Then every third party involved in the process will have access to this immutable application, stored securely on the BC, with their own private keys. They will then be able to verify the details given by the borrower, such as ID, asset

to debt ration, among other [26]. After each party verifies the application, the loan offer is generated and the user can access it and sign it. As soon as they sign it the funds get transferred to them and this transaction is then stored inside of a block on the BC as well, keeping an immutable record of the transaction and of the whole application process [27].

5.2.3 Smart Contracts

Unlike paper contracts which can be destroyed, tampered with or have multiple versions, a smart contract, as mentioned before, is an automated contract that has its terms and conditions written as code parameters and is stored on a BC, that will only be executed once if and only if all the encoded parameters are met. The actual set up of the contract needs human interaction, as the contract has to be programmed. The execution, however, is an automated process which fires as soon as the defined parameters, as in the terms and conditions of the agreement, are met[28].

The advantages of smart contracts lie in its immutability, meaning no party can change the terms and conditions, in its reusability, once a standard contract is developed it can be tweaked and reused in other scenarios. However, disadvantages such as programmed back doors that can be "exploited after the contract has been signed." [29] and faulty code can compromise smart contracts and create financial loss [29].

5.2.4 Over the counter Trading

In contrast to a cryptocurrency exchange, over the counter (OTC) cryptocurrency trading desks are a peer to peer trading platform. It is a key player for investors and funds that are looking to buy large amounts of a given cryptocurrency. The amount and price for the cryptocurrency are negotiated beforehand, the OTC desk then gathers the currency and upon the customers payment, the cryptocurrency is sent to their crypto wallet. If a fund or investor would try to secure large amounts of a cryptocurrency through an exchange, they would incur in slippage, which "[...] occurs when you run out of people selling at your desired price causing you to 'slip' further from the original market price" [30], meaning that they would self induce a price spike in the currency and thus buy it at a higher market rate.

5.2.5 Custody

Another BC based banking service is custody, where a cryptocurrency wallet and its respective private keys are held by a third party. This third party has complete authority over the users funds, the users only has to dictate which transactions should be executed his behalf. As the amount of digital assets is ever growing, the need to securely store and insure these assets is growing rapidly.

Custody solutions are applicable for institutional investors, hedge funds and cryptocurrency exchanges. These cryptocurrency custody solutions normally consist of a mixture of hot and cold storage to maximize security but also grant a certain degree of flexibility. Hot storage means that the users cryptocurrency wallet is stored on the third parties online portal. Hot storage of cryptocurrencies and digital assets offers easier liquidity because it is connected to the internet, but this in turn makes it vulnerable unauthorized third party data breaches. Cold storage enables the entity to have greater security, but limits the chance to generate liquidity quickly because it is stored offline[31].

5.2.6 Storage

The biggest difference to custody, is that in cryptocurrency storage the private keys reside with the end user. A complex string combination of an alphanumeric sequence, these keys are used to access digital assets and conduct transactions with them. Because private keys are complex and hard for any human to remember, they are stored either online or offline, which makes private keys inherently susceptible to being stolen or lost. Online wallets are an online solution for storing these private keys but have proven to be susceptible to hacker attacks [27]. Offline solutions, for example writing the private keys down on paper or storing it on a hard drive, are efficient to avoid hacker attacks but the risk of losing them is much greater[32].

5.3 Swiss Banking Landscape

Recently, as BC has become evermore present and trusted, the Swiss financial sector has increased its adoption of the technology in order to strategically benefit from it [33]. There has also been a shift in regulations, as "Seba Crypto and Sygnum both clinched banking licenses from Swiss financial regulator Finma[...]" in 2019 [34]. Even further, the newly adjusted Anti-Money-Laundry (AML) regulations laid out by FINMA also apply: "Institutions supervised by FINMA are only permitted to send cryptocurrencies or other tokens to external wallets belonging to their own customers whose identity has already been verified and are only allowed to receive cryptocurrencies or tokens from such customers." [35].

5.3.1 Julius Baer

The Zurich headquartered private bank partnered up with Seba Bank AG, one of the first two FINMA approved Crypto-Banks in Switzerland, in an attempt to meet the demand of their customers on digital assets. Julius Baer allows it's users to store, transact and invest into digital assets and also tokenize illiquid assets, assets that normally do not have an active market, as most of their pricing model is subjective. The tokens created through tokenization of illiquid assets, such as luxury cars, real estate and fine art, can then be traded on a secondary market, consequently increasing the liquidity of these illiquid assets, as the access to investors increases.

5.3.2 Sygnum Bank

The worlds first digital asset bank, Sygnum Bank, was one of the two crypto banks to get a banking licence in Switzerland, next to Seba Bank. They offer crypto collateralized loans. They allow the customer to scale liquidity according to their needs and agree on a flexible maximum amount of available fiat liquidity against crypto assets in order to get better interest rates on the loan. In addition to loans, they also offer custody services, making it easier for customers to handle their private keys, as Sygnum takes up the administrative step to secure them. Furthermore, they have a top of the line Anti Money Laundry (AML) checking software to ensure that the customers transactions are fully compliant and legal. Sygnum Bank also offers tokenization of existing and new assets, together with a secondary trading platform for those tokens, however, this service is still in development. Furthermore, Sygnum offers faster and automated transaction settlements through smart contracts. Through Sygnum's brokerage and asset management offerings, customers can trade their digital assets with fast settlements and high levels of security, as well as get help from digital assets experts to further expand their portfolio.

5.3.3 Seba Bank AG

Seba Swiss, which as mentioned before partnered up with Julius Baer offers, apart from crypto storage, transaction and investment services, also crypto-collateralized lending services and digital custody services, which users can access through the Seba E-banking environment.

5.3.4 Swissquote

Swissquote, which also has an official swiss banking licence from FINMA, offers amongst other services crypto OTC trading and access to a cryptocurrency exchange platform. This allows users to invest into crypto assets like Theme Trading Certificates, which enables them to invest into a market as a whole and not just in one company or commodity, as well as in exchange traded funds and exchange traded products and products like Mini-Futures, which have an unlimited term.

5.3.5 Incore Bank

Incore Bank is a business to business transaction focused bank. They offer Digital Asset Banking as a Service for cryptocurrencies and tokenized assets. They take care of custody, brokerage and transfer of digital assets for their customers. In addition to the aforementioned services, Incore Bank also provides tokenization for previously non-bankable or illiquid assets.

5.3.6 Falcon Private Bank

A private investment bank that now offers it's customers the option to invest into BC assets via e-banking or with an advisor. Through direct BC asset transfers system, customers can make their BC assets bankable by transferring them to a crypto wallet within Falcon's environment and convert that cryptocurrency into fiat money or transfer them into an external wallet.

5.3.7 BEKB

The Berner Kantonalbank grants digital asset custody services, but mainly focuses on providing a market for unlisted assets of small and medium-sized enterprises (SMEs) that are then tokenized, listed and traded on the OTC-X platform through the BEKB environment.

5.3.8 UBS

UBS, one of the biggest banks of the world, has not fully committed to BC for various reasons. They did however launch a trading platform called *we.trade* designed for SMEs in Europe. The platform looks to help the growth of trust in between SMEs in Europe. It includes services like bank payment guarantees and invoice financing.

5.4 Challenges for Swiss Banking

The provided blockchain-based solutions of Swiss banks are either done by a new bank, which is specialized on these services or still work-in-progress when being done internally, as seen in the chapter before. In a contested market, this behavior seems rather untypical, since banks try to find new sources of income due to the decreasing margin in traditional businesses. It arises the thesis of other existing challenges causing the stiffness of the once worldwide-known innovative Swiss Banking. From those challenges this report focuses on three main challenges for banks which are challenged by the introduction of BC services, namely: Structure, Technology and Legal affairs. Using a bottom-up approach, challenges on the lowest structural level for each individual are reflected on higher layers as well, since the management will be confronted with those. But even different areas like technology and legal will have to react to the changes in the operative segment.

5.4.1 Structural

Instead of investigating the hierarchy in a horizontal manner first, this subsection divides it vertically in personnel and organization, as the impact of BC is potentially fundamental. Obviously, these two areas are closely interconnected in both directions, because employees are affected by organizational changes but can force the organization to change as well.

5.4.1.1 Personnel changes

Due to the technical nature of this change, employees will face changes in their daily work, because of different and even automated processes. Therefore, new training is needed in order to keep the affected workers up-to-date.

Not every worker will accept these changes and might refuse to continue working in such a different environment [36]. This in turn can lead to higher fluctuation, which denotes another challenge since the amount of needed training has already been increasing in the last years [37]. This implies that not all of the positions could be replaced but rather should be filled with the right person. First BC appliances appeared to a broader audience through Bitcoin already in the late 2000's and more than 10 years later, we're still speaking of how to implement it for basic financial services [38]. It shows that the technology was lacking acceptance in economy and therefore didn't have the needed support to convince society of it. Consequently there hasn't been a need to include BC in usual education, which leads to a smaller number of capable workers in the market [39]. As supply of a needed special good is getting short, its price will increase - BC developers are in demand and therefore are in a good position to negotiate their wage.

Apart from the financial impacts, each employee will cope with personal consequences in such a far-reaching change. Depending on the company's communication and support for affected employees, the severeness can be eased, however the productivity will suffer during the phase of change along with the company's internal mood [40]. The uncertainty about the future of the job will be further amplified when considering the expected middle-man businesses. Since banks nowadays carry out those middle-man tasks like transactions, such jobs could be either replaced or dropped[41].

However, all of these challenges are tightly linked to strategic decisions on the organization's structure and extent of implementing technology, since each employee will have to carry out what has been discussed on this.

5.4.1.2 Organizational changes

On a higher organizational level the challenges of lower hierarchical levels apply, since managers cope with fundamental changes as well. Additionally, they have to consider the

effects in both directions: How can the intentions of higher levels be fulfilled while also solving most of the concerns of other employees?

As already stated before, one approach could be involving employees closely and trying to train them for the upcoming changes. Due to the needed training of workers, the company will create and assess new training sessions for the required knowledge about BC, and the new processes respectively. This leads to a more technology-driven approach in the training of the trainers, as this is a shift in the core business from mere financial to a more technological knowledge of the underlying processes. As not all of the employees will get along with these changes, there will be supplemental costs in order to find either a new position within the company or let them leave [42]. Additionally, costs for a new employment increase as well due to the lack of BC developers on the market. Wage structures will then have to be re-evaluated, particularly in a country like Switzerland with a high average wage in the financial sector, where banks are already trying to be as cost-efficient in terms of personnel expenses [43].

Apart from all expenses, BC is highly related to decentralization, which is a significant change in power of control. As banks nowadays take care of their own business and only a few shared infrastructures like Swiss Infrastructure and Exchange, SIX or TWINT exist, two of which are settled in the area of inter-bank transactions. Why should a bank reorganize itself when considering the above mentioned costs and challenges to implement a technology, which leads to a loss of control and is based on technology rather than the well-known financial knowledge?

Since BC is a change that affects every bank, there are possibilities of using synergies. The workload is not only on one party, so costs could be held low [44]. Bonds created in those committees lead to a closer global collaboration as there are major banks involved. Thanks to these joint efforts, systems could be more likely to work together and therefore lower transaction costs as well as development costs, both of which are a major issue for banks nowadays [45].

However innovation does have its price, due to the working hours invested into the whole process from the first idea to an actual implementation. If an added benefit is visible to customers, they're more willing to pay for it. Additionally, being a first mover, a bank will be able to gain a bigger market share, because the market is lacking alternatives [46]. It provides valuable information about how the market is reacting to the new technology and could lead to a closer interaction with market players like investors.

Being in such a position a bank will be the driver of new standards, since the implementation will draw the industry's attention and, in the case of success, force the other players in the market to come up with similar solutions. Hence the first mover sets the bar, also in legal aspects, because the new technology has to be approved by the legislator. BC will disrupt the current understanding of security and privacy, since data could now be handled decentralized, as blocks are at least shared between all banks in a distributed ledger, depending on the chosen type of BC. Each participant has to ensure that existing security and privacy standards are either held or even improved with the new technology, although the system is not (only) being maintained internally.

Those systems rely on a certain critical mass to work efficiently, which is seen when evaluating the accumulated costs which were mentioned above. Without a network of nodes, which are located outside of the own company as well, the costs of calculation power, power consumption and system maintenance (backups) are enormous. Therefore, banks will have to make sure that the system will be used by peers, since a small network leads to the productivity paradox [47].

With the introduction of BC organizational structures will change, affecting employee demographics of each company. Careful planning of the procedure during the implementation of BC, along with involvement of the employees could ease the effect on productivity and costs. However, these expenses won't be put on one bank only as the whole financial

sector will be affected by the technology change. Hence, the costs can be split up, since every player on the market will have to implement a system that interoperates with all other parties, because of the decentralized aspect. It will be challenging for first movers to establish a system which is used by everyone in order to hit the efficiency target, despite the advantages they have like setting standards and being ahead on the learning curve of the new technologies.

5.4.2 Technical

With the implementation of BC, fundamental processes will be changed. Information has to be stored and secured in blocks as well as be distributed to all other nodes, which is handled in the back-end. Consequently, the front-end won't be changed if the underlying process still persists and needs human interaction. Developers have to ensure that interoperability between existing and new services is still given after the implementation.

5.4.2.1 IT Legacy

It seems to be a daunting task as banks are already struggling with digitizing currently manual processes [48]. This is due to the old legacy systems being used in Swiss Banks. They're considered "backbones" of the main tasks like carrying out transactions or book-building. BC changes this, instead of matching it centralized after checking plausibility internally, BC could be a solution to speed up with its (partial) decentralized approach. However, this example shows the complexity of the issue: To implement BC, most sensitive processes are being altered or revamped. It implies that there is a serious operational risk along with this implementation, as it touches the core of the current IT system. Meticulous testing of functionality will be key as well as performance testing, regarding time sensitivity of the actions to be carried out like stock market transactions. Performing those tests is challenging, since a network with several nodes with its calculation power can't be mocked easily. [49]

5.4.2.2 Security

Implementing such a network including all other partners is a project of a bigger extent, which will require different techniques, since it could be a peer-to-peer solution. In order to achieve such a network, the information of the now centralized approach has to be migrated to the decentralized BC, assuming that a full decentralized system is chosen by banks. Preventing any loss of information will be key as this is a major operational risk when migrating big chunks of stored data [38].

Integrity and immutability of the blocks with its stored data is a main security characteristic of the technology, however this is achieved through mathematical premises based on the current possible computing power for individuals. If Quantum Computers will be accessible for a broader audience, it could be easier to harm the system by brute force approaches. Compared to currently used security algorithms, this isn't a disadvantage, since most security systems are not using post-quantum security algorithms [50].

5.4.2.3 Scalability

A secure system is a main requirement to be used, but as soon as it is used by an increasing amount of users (nodes) it should be scalable. In terms of transaction speed it shouldn't get noticeably slower with an increasing amount of transactions. With the current technology used for BC, it grows in size as time advances. This is causing longer load times for synchronizing the blocks in the system [51]. There are several approaches which ease this flaw, so banks will have to think of this as they make their decision on

how to implement. Customers won't accept lower transaction speeds as they're used to now - actually the trends point into other directions when looking at the popularity of modern digital solutions like Revolut or Transferwise. Accessing information and doing transactions is just a touch away, without any added costs, this is what the implementation of BC should achieve or even more.

Ensuring a fast system with a technology which isn't quite as scalable as the current one used, while being able to interoperate with old legacy systems sums up the complexity of technological challenges for Swiss banks. As they're already struggling with current projects to automate manual processes, they now face the task of implementing a new technology, which isn't quite ready yet for handling the same amount of transactions as current systems do. So they should find solutions for a new technology, which potentially will simplify the handling and reduce costs, when it's a fully fledged alternative. However, as the system will be decentralized, these challenges will most likely not be handled by a single bank but rather in the group of participating parties. This also applies to security aspects, even though every participant will have to look after its own interface. Regulatory inquiries will be easier handled when acting in a group, since these rules will apply on every party in the same way.

5.4.3 Legal

All the above mentioned challenges are depending on how BC will be treated by regulators as they will decide on what is legally accepted. This includes the treating of structural changes in companies since there might be negotiations about redundancy programmes which will affect costs for each bank. Even the internal choice of choosing which organizational structure will fit best to the new technology depends on the regulators, because they decide if the chosen degree of decentralization is feasible to respect existing rules. Last but not least the implementation will be subject due to auditing, this has to be ensured, especially in terms of traceability of data.

5.4.3.1 General issues

This data is being distributed in so-called tokens as already shown above. How are these handled legally? There were several attempts to categorize it with the existing jurisdiction, since it's only the implementation that differs, but not the intention of having a proof of being owner of a certain good.

It could be considered as a res, since res is not defined by law. However, in current jurisdiction this term is used very specifically for non-digital data, which contradicts this classification. Such legal claims in digital form exist in the financial market. However, it is not even clear if with current jurisdiction a token could qualify for a monetary claim, due to the different applications of a token. Therefore, it isn't possible to set it equal with cash.

Bonds are closer to a token, since they denote a contract between a debtor and a creditor. In a decentralized system it's not that easy to name a particular debtor, but the needed counterparty could be seen as each node in the network. There are already contracts in which several parties could be involved with the possibility to join and leave continuously as long as there is at least one.

Securities denote a securitization of a right being given through a certificate. However the claim to be owner of this right is linked with a physical ownership, which is being transferred from the former to the future owner. Such a physical transfer is not given with the digital distribution of tokens through a BC based system. Hence a token can't be considered as security [52]. These insecurities of classifying a token is a major challenge for every company which tries to use tokens as a mean of determining ownership (also in a legal way). Customers are used to regulated securities and won't accept any drawbacks, since it's their money they invest.

5.4.3.2 Blockchain law in Switzerland

As seen above, there are already a lot of new banks and companies using BC as their main business model. Therefore, Swiss politicians are already active to handle the BC topic. In autumn 2020 the parliament of Switzerland discussed BC and the regulations and came up with some solutions, which are now assessed with cantons until February 2021. A major subtopic was the handling of tokens in case of bankruptcy, since it is not possible to classify them and therefore they aren't regulated yet.

They decided on some changes, which clarify still open legal aspects, along with other measures, which should minimize the risk of abuse and other barriers of BC implementations. This is achieved through licensing of exchanges and especially introducing a concept to resolve the above mentioned issues about handling tokens legally.

There will be a new type of digital security "Uncertificated Register Securities", which can be treated separately in case of bankruptcy. However they didn't invent anything new here, it's more an adaptation of the already existing laws for securities applied on the tokenization aspects. Moreover they clarified how information have to be made accessible to the owners and the counterparties, while preventing them to be mutated by a third party[53].

Switzerland tries to improve the legal framework for using BC in order to support innovative start-ups and existing companies in the financial sector like banks and insurances. This is needed, since new technologies won't be considered in the economy if there is no legal certainty about how it will be treated in key topics.

5.5 Conclusion

BC certainly has the potential to disrupt the financial sector. The technology is breaking new grounds in terms of decentralization of data, but also saving costs and enhancing security. Currently, each bank maintains its own individual IT system, while the processes being handled are very similar. Moreover, every opportunity to automate and increase efficiency is highly welcome since margins are shrinking due to new players in the banking landscape providing full digital solutions.

Several banks decided to use BC to achieve this increase in efficiency, but instead of trying to implement it themselves, they use already existing services in order to be faster and save costs. This approach is a way to avoid hiring new people with proficiency in the BC technology, which are rare in the current labour market, which in turn leads to higher wages being offered. However, existing Swiss banks are still hesitating to declare a BC strategy. Most of them are active in research, but as of today are often collaborating with newly formed BC banks instead of presenting own solutions. This could be due to the high effort that has to be done in order to implement BC, as it will change the company's structure. Such a disruptive transition has to be well-planned in order to ensure a flawless functionality of the daily business during and after the transition. Going through this transition as a first mover can increase costs but will bring benefits like being able to decide on standards as well as reputational improvements. This can be used to acquire new customers and therefore lead to a long-term revenue source. Paired with future-ready technology this could be a possibility to consolidate their position in an already competitive market.

Although legacy systems are still present in Swiss banks, the implementation of BC could lead to a needed update of core functionalities. Scalability of BC at the current stage isn't given, but there are already approaches to tackle this and make BC needing less resources, which increases efficiency. Together with the recent plan to change current jurisdiction in order to facilitate innovations concerning BC, Switzerland provides favorable conditions

for companies in the financial sector to maintain a leader role in the global financial market.

However, apart from scalability and efficiency, there still remains one big issue to tackle, which is sustainability. Especially the traditional, established banks but also the banking sector as a whole have come to play a leading role when it comes to embodying sustainability in new technologies. In other words, the public expects the banking sector to prioritize on sustainable innovation rather than profit-based innovation. That being said, the issue of BC being potentially very unsustainable could stall the acceptance of banks moving to a BC infrastructure or even result in a damage of reputation. But whether banks will accommodate to the demand of sustainability or whether they are powerful enough to go through with the transformation to BC in spite of it remains an open question which will likely be answered in the next few years.

In conclusion, since the financial crisis in 2008 and the following tax law suits, Swiss banks had to reorient themselves in a fast changing financial market, shaken by negative interests and the Euro crisis. With accepting above challenges and mastering them, a pioneer position could be established. Swiss banks are carefully approaching the transition to BC-based services, while new banks based on BC are already formed. However, existing and even more the established banks have yet to undertake the core transition, which is the migration of their legacy-based transaction systems to a BC based approach, as it seems costs and uncertainty are still too high to make the final step. The next few years will be crucial in determining whether BC will indeed be the revolutionary technology that is designated to change the Swiss banking landscape.

5.6 Limitations

As the Swiss banking sector is inhomogeneous in terms of types of banks (private or public banks, varying in sizes), our findings can't be applied on the whole banking sector, as we were considering mostly big and established banks. As we focussed on finding challenges through a rather theoretical approach, we did not take into account any practical view, which other papers already did. By interviewing affected developers and employees, they discovered that most BC developers also might not know a lot about currently present automated software-solutions such as ERP (Entreprise-Resource-Planning), which are not BC based[54]. This leads to a biased view on the technology. When discussing the legal situation in Switzerland we did not consider international jurisdiction. Since BC will be used worldwide for international transfers, there are still some open topics and unknown factors in cross-border laws concerning BC.

Bibliography

- [1] Z. Ali, “The world’s 100 largest banks, 2020,” Apr 2020, visited on 01.12.2020. [Online]. Available: <https://www.spglobal.com/marketintelligence/en/news-insights/latest-news-headlines/the-world-s-100-largest-banks-2020-57854079>
- [2] Z. Zheng, S. Xie, H. Dai, X. Chen, and H. Wang, “An overview of blockchain technology: Architecture, consensus, and future trends,” in *2017 IEEE International Congress on Big Data (BigData Congress)*, 2017, pp. 557–564, (visited on 03.11.2020).
- [3] V. Buterin, “On Public and Private Blockchains,” Aug 2015, (visited on 30.10.2020). [Online]. Available: <https://blog.ethereum.org/2015/08/07/on-public-and-private-blockchains/>
- [4] C. Cachin and M. Vukolic, “Blockchain consensus protocols in the wild,” *CoRR*, vol. abs/1707.01873, 2017. [Online]. Available: <http://arxiv.org/abs/1707.01873>
- [5] L. M. Bach, B. Mihaljevic, and M. Zagar, “Comparative analysis of blockchain consensus algorithms,” in *2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, 2018, pp. 1545–1550.
- [6] D. Mazières, “The Stellar Consensus Protocol: A Federated Model for Internet-Level Consensus,” *Stellar Development Foundation*, 2016, visited on 02.11.2020. [Online]. Available: <https://www.stellar.org/papers/stellar-consensus-protocol>
- [7] L. Sankar, M. Sindhu, and M. Sethumadhavan, “Survey of consensus protocols on blockchain applications.” 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS), 01 2017, pp. 1–5.
- [8] B. Chase and E. MacBrough, “Analysis of the XRP ledger consensus protocol,” *CoRR*, vol. abs/1802.07242, 2018. [Online]. Available: <http://arxiv.org/abs/1802.07242>
- [9] D. Schwartz, N. Youngs, A. Britto *et al.*, “The ripple protocol consensus algorithm,” *Ripple Labs Inc White Paper*, vol. 5, no. 8, 2014.
- [10] “Cryptocurrency prices, charts and market capitalizations.” [Online]. Available: <https://coinmarketcap.com/>
- [11] C. Kim, “The ‘hot swap’ plan to switch ethereum to proof-of-stake explained,” Sep 2020, visited on 04.11.2020. [Online]. Available: <https://www.coindesk.com/the-hot-swap-plan-to-switch-ethereum-to-proof-of-stake-explained>
- [12] A. Sapirshtein, Y. Sompolinsky, and A. Zohar, “Optimal selfish mining strategies in bitcoin,” in *Financial Cryptography and Data Security*, J. Grossklags and B. Preneel, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2017, pp. 515–532.

- [13] W. Fauvel, “Blockchain advantage and disadvantages,” Aug 2017, visited on 03.11.2020. [Online]. Available: <https://medium.com/nudged/blockchain-advantage-and-disadvantages-e76dfde3bbc0>
- [14] N. Reiff, “Blockchain explained,” Feb 2020, visited on 10.10.2020. [Online]. Available: <https://www.investopedia.com/terms/b/blockchain.asp>
- [15] M. J. Heron, V. L. Hanson, and I. Ricketts, “Open source and accessibility: Advantages and limitations,” *Journal of Interaction Science*, vol. 1, no. 1, p. 2, 2013.
- [16] Blockchain.com, “Transaction rate per second,” visited on 10.10.2020. [Online]. Available: <https://www.blockchain.com/charts/transactions-per-second>
- [17] Visa, “Visa acceptance for retailers,” visited on 10.10.2020. [Online]. Available: <https://usa.visa.com/run-your-business/small-business-tools/retail.html>
- [18] Digiconomist, “Bitcoin energy consumption index,” visited on 10.10.2020. [Online]. Available: <https://digiconomist.net/bitcoin-energy-consumption>
- [19] L. Fuchs, “Warum werden banküberweisungen am wochenende nicht verarbeitet?” 2018. [Online]. Available: <https://www.srf.ch/sendungen/kassensturz-espresso/warum-werden-bankueberweisungen-am-wochenende-nicht-verarbeitet>
- [20] P. Laurent, T. Chollet, M. Burke, and T. Seers, “The tokenization of assets is disrupting the financial industry. are you ready?”
- [21] C. Pauw, “What is an sto, explained,” 2019. [Online]. Available: <https://cointelegraph.com/explained/what-is-an-sto-explained>
- [22] B. Mellon, “Tokenization: Opening illiquid assets to investors,” 2019. [Online]. Available: <https://www.bnymellon.com/us/en/insights/all-insights/tokenization-opening-illiquid-assets-to-investors.html>
- [23] L. SAS, “What is crypto lending?” 2020. [Online]. Available: <https://www.ledger.com/academy/what-is-crypto-lending>
- [24] B. Crowell, “Crypto lending, explained,” *finews.com*, 2020. [Online]. Available: <https://cointelegraph.com/explained/crypto-lending-explained>
- [25] “Swiss bank widens crypto services,” *finews.com*, 2019. [Online]. Available: <https://www.finews.com/news/english-news/35847-crypto-swiss-banks-hypotheekbank-lenzburg-hypi-tokesuisse-cooperation>
- [26] H. L. Experts, “Blockchain mortgage | a cheaper and faster home loan,” 2017. [Online]. Available: <https://www.homeloanexperts.com.au/home-loan-articles/blockchain-mortgage/>
- [27] B. Patel, “Will blockchain revolutionize mortgage lending?” *Forbes*, 2019. [Online]. Available: <https://www.forbes.com/sites/forbesfinancecouncil/2019/08/19/will-blockchain-revolutionize-mortgage-lending/?sh=9145a791cab7>
- [28] T. Ko, J. Lee, and D. Ryu, “Blockchain technology and manufacturing industry: Real-time transparency and cost savings,” *Sustainability*, vol. 10, no. 11, p. 4274, 2018.
- [29] R. Knecht, “Are smart contracts the future?” translated to English by Veronica Bielawski. [Online]. Available: <https://www.digitec.ch/en/page/are-smart-contracts-the-future-17659>

- [30] C. Dempsey, “How does crypto otc actually work?” 2019. [Online]. Available: <https://medium.com/circle-research/how-does-crypto-otc-actually-work-e2215c4bb13>
- [31] R. Sharma, “What are cryptocurrency custody solutions?” 2020. [Online]. Available: <https://www.investopedia.com/news/what-are-cryptocurrency-custody-solutions/>
- [32] N. Zahed Benisi, M. Aminian, and B. Javadi, “Blockchain-based decentralized storage networks: A survey,” *Journal of Network and Computer Applications*, vol. 162, p. 102656, 2020. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1084804520301302>
- [33] E. Ede-Obarogie, “The importance of blockchain in the swiss financial sector,” *Geneva Business News*, 2019. [Online]. Available: <https://www.gbnews.ch/the-importance-of-blockchain-in-the-swiss-financial-sector/>
- [34] “Switzerland approves crypto banks,” *finews.com*, 2019. [Online]. Available: <https://www.finews.com/news/english-news/37740-seba-sygnum-switzerland-banking-license-swiss>
- [35] V. Mathys, “Finma guidance: stringent approach to combating money laundering on the blockchain,” *Finma*, 2019. [Online]. Available: <https://www.finma.ch/en/news/2019/08/20190826-mm-kryptogwg/>
- [36] A. Agboola and R. Salawu, “Managing deviant behavior and resistance to change,” *International Journal of Business and Management*, vol. 6, 12 2010.
- [37] Schweizer Bankiervereinigung, “Sbvg bankenbarometer 2019,” Aug 2019, (visited on 02.11.2020). [Online]. Available: https://www.swissbanking.org/finanzplatz-in-zahlen/wp-content/uploads/2019/08/SBVg_Bankenbarometer_2019_DE_2S.pdf
- [38] M. Crosby, P. Pattanayak, S. Verma, V. Kalyanaraman *et al.*, “Blockchain technology: Beyond bitcoin,” *Applied Innovation*, vol. 2, no. 6-10, p. 71, 2016.
- [39] Mearian, “Blockchain jobs remain unfilled, while skilled workers are being poached,” (visited on 03.11.2020). [Online]. Available: <https://www.computerworld.com/article/3387441/blockchain-jobs-remain-unfilled-while-skilled-workers-are-being-poached.html>
- [40] W. A. Weeks, J. Roberts, L. B. Chonko, and E. Jones, “Organizational readiness for change, individual fear of change, and sales manager performance: An empirical investigation,” *Journal of Personal Selling & Sales Management*, vol. 24, no. 1, pp. 7–17, 2004. [Online]. Available: <https://doi.org/10.1080/08853134.2004.10749012>
- [41] P. Witzig and V. Salomon, “Cutting out the middleman: a case study of blockchain-induced reconfigurations in the swiss financial services industry,” 03 2018.
- [42] J. Duda, L. Žúrková *et al.*, “Costs of employee turnover,” *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, vol. 61, no. 7, pp. 2071–2075, 2013.
- [43] S. Gerber, “Soviel verdient man im swiss banking wirklich,” (visited on 03.11.2020). [Online]. Available: <https://www.finews.ch/themen/finews-life/36288-bonus-lohn-studie-michael-page-swiss-banking>
- [44] V. Morabito, “Business innovation through blockchain,” *Cham: Springer International Publishing*, 2017.

- [45] Ernst Young AG, “Bankenbarometer 2019,” (visited on 05.11.2020). [Online]. Available: <https://www.eycom.ch/de/Publications/20190110-Bankenbarometer-2019-Zeichen-der-Zeit/download>
- [46] P. Tufano, “Financial innovation and first-mover advantages,” *Journal of Financial Economics*, vol. 25, no. 2, pp. 213 – 240, 1989. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/0304405X89900822>
- [47] Grewal-Carr and Marshall, “Blockchain enigma.paradox.opportunity,” (visited on 02.11.2020). [Online]. Available: <https://www2.deloitte.com/content/dam/Deloitte/ch/Documents/innovation/ch-en-deloitte-innovation-blockchain-full-report.pdf>
- [48] O. Schneider, “Jörg Gasser: ”Die Banken waren gut vorbereitet,”” (visited on 07.11.2020). [Online]. Available: <https://www.netzwoche.ch/news/2020-06-19/joerg-gasser-die-banken-waren-gut-vorbereitet>
- [49] B. Koteska, E. Karafiloski, and A. Mishev, “Blockchain implementation quality challenges: a literature,” in *SQAMIA 2017: 6th Workshop of Software Quality, Analysis, Monitoring, Improvement, and Applications*, 2017, pp. 11–13.
- [50] B. S. Haney, “Blockchain: Post-quantum security & legal economics,” *NC Banking Inst.*, vol. 24, p. 117, 2020.
- [51] I.-C. Lin and T.-C. Liao, “A survey of blockchain security issues and challenges.” *IJ Network Security*, vol. 19, no. 5, pp. 653–659, 2017.
- [52] H. C. von der Crone, F. J. Kessler, and L. Angstmann, “Token in der blockchain–privatrechtliche aspekte der distributed ledger technologie,” *Schweizerische Juristen-Zeitung*, vol. 114, pp. 337–345, 2018.
- [53] MME, “Schweizer parlament genehmigt neues dlt gesetz,” (visited on 10.11.2020). [Online]. Available: https://www.mme.ch/de/magazin/schweizer_parlament_genehmigt_neue_dlt_verordnungen/
- [54] R. Esmander, P. Lafourcade, M. Lombard-Platet, and C. N. Ribalta, “A silver bullet? a comparison of accountants and developers mental models in the raise of blockchain,” in *Proceedings of the 15th International Conference on Availability, Reliability and Security*, ser. ARES ’20. New York, NY, USA: Association for Computing Machinery, 2020. [Online]. Available: <https://doi.org/10.1145/3407023.3409193>

Chapter 6

An Analysis and Comparison of Blockchain-as-a-Service (BaaS) Providers

Felix Hoffmann, Charlotte Eder, Alain Küng

Both Blockchain and Cloud Computing are technologies that have seen a big increase in popularity during the past years. In order to meet the needs of the current market, more and more platforms have emerged that offer a combination of Cloud Computing services and Blockchain technology under the name Blockchain-as-a-Service (BaaS). Since the field of BaaS is rapidly changing, an overview over the current market situation and trends in BaaS is missing. Thus, this paper aims to give an overview over the current BaaS market by first presenting six BaaS platforms and their services from the providers Microsoft, IBM, Amazon, Dragonchain, Alibaba and Oracle. Then, a comparison is drawn between the BaaS providers based on various criteria. The findings of the comparison show that mostly major Cloud Computing providers have adopted a Blockchain service in order to complement their other Cloud Computing services. They use two different pricing models, such as pay-as-you-go or a subscription based payment system. Further, Hyperledger Fabric is the Blockchain framework that is supported by most of BaaS providers.

Contents

6.1	Introduction	.	.	.	169
6.2	Blockchain	.	.	.	169
6.2.1	History	.	.	.	169
6.2.2	Technology	.	.	.	170
6.2.3	Uses	.	.	.	173
6.3	Cloud Computing	.	.	.	175
6.3.1	Service Models	.	.	.	176
6.3.2	Advantages and Challenges of Cloud Computing	.	.	.	177
6.4	Blockchain-as-a-Service (BaaS)	.	.	.	178
6.4.1	Advantages of BaaS	.	.	.	179
6.4.2	Challenges and Risks of BaaS	.	.	.	180
6.5	Providers	.	.	.	181
6.5.1	Microsoft Azure	.	.	.	181
6.5.2	IBM	.	.	.	184
6.5.3	Amazon's AWS Blockchain	.	.	.	186
6.5.4	Dragonchain	.	.	.	188
6.5.5	Alibaba Group	.	.	.	189
6.5.6	Oracle	.	.	.	190
6.6	Discussion	.	.	.	192
6.6.1	Products and services	.	.	.	192
6.6.2	Cloud Revenue	.	.	.	193
6.6.3	Known Supported Frameworks	.	.	.	193
6.6.4	Pricing model	.	.	.	193
6.6.5	Service Kits	.	.	.	194
6.7	Conclusion	.	.	.	194
6.8	Future Work	.	.	.	194

6.1 Introduction

In the past years, there has been a steep rise in the usage of Blockchain technology first in the domain of cryptocurrencies and now also for more diverse applications, such as to manage supply chains, create more secure systems to store patients records in health care or to validate candidates application credentials [1]. At the same time, Cloud Computing is getting more and more popular as a business model for providing dynamic, scalable and pay-per-use services [2]. Unsurprisingly, a fusion of these two technologies has recently started to emerge under the name BaaS. Several well-established cloud providers such as Microsoft Azure, Amazon Web Services (AWS) and IBM have already included BaaS in their services, but also smaller providers such as Dragonchain try to get a foot into the market of BaaS [3]. However, since BaaS is still a new and rapidly changing field, a detailed, up-to-date overview over the different BaaS providers is missing in literature.

This is why this paper aims to give an overview over the current BaaS market by comparing established and niche BaaS providers. This paper focuses on answering the questions of who the current BaaS providers are, what their service consists of and how their offers differ from each other.

6.2 Blockchain

Blockchain is a technology in the Information Technology (IT) sector with the aim to prevent information fraud or tempering with data [4]. In the last years it turned into a popular tool for which by now a wide range of use cases has been found.

6.2.1 History

Half a century ago the digital revolution changed our society permanently. One of its biggest advantages is that information can be shared and multiplied at basically no cost. This enables us to share information and data with a huge community. The big disadvantage is that we do not know where the information originated from and if it is even true. Not knowing if a information is valid or tempered with might loose its whole value.

6.2.1.1 Merkle Tree

There are many structures in which digital data can be stored. However, a particularly time efficient way is to store data in a Merkle Tree. Here we use nodes that, unless its the root node, each has one predecessor, and, unless its the last node of a branch, each has one or more successors. If we only have one successor for each node the terminology shifts to the linked list and this benefit of speed is lost. Ralph Merkle was still able to still see a potential in this kind of tree or linked list and the idea of immutably chaining blocks was born [5].

6.2.1.2 David Chaum

In 1982, only a few years after Merkles idea of immutably changing blocks, cryptographer David Chaum introduced a vault system in his dissertation at Berkeley University. According to Chaum the system allows mutually suspicious users to have a save transaction of data using a distributed network. In this vault system, the chaining process exists as every new transaction signs, records and broadcasts the done transaction through the whole network. The broadcast is necessary as every user of this virtual community needs

to agree with it, should there be discrepancies the node can not be added. This closely aligns with the usage of Blockchain as we know it today [5].

However, Chaums work was not getting attention, probably due to the fact that he did not publish it in any prestigious venue. There is also the issue that Chaums vault system is based on a distributed, not fully decentralised, network system and which, depending on the definition, might therefore interfere with the term Blockchain today [5]. Next chapters present how a Blockchain is build and the differences of the networks will be presented in a later chapter.

6.2.1.3 Bitcoin

Even though the idea and technology of Blockchain dates back many decades it has found very little application in practice. It took until 2009, when Bitcoin started, that the general public heard for the first time about the term Blockchain. Bitcoin is a digital currency, called cryptocurrency, which users can buy with real money [6]. But what exactly is Bitcoin, and what is it not?

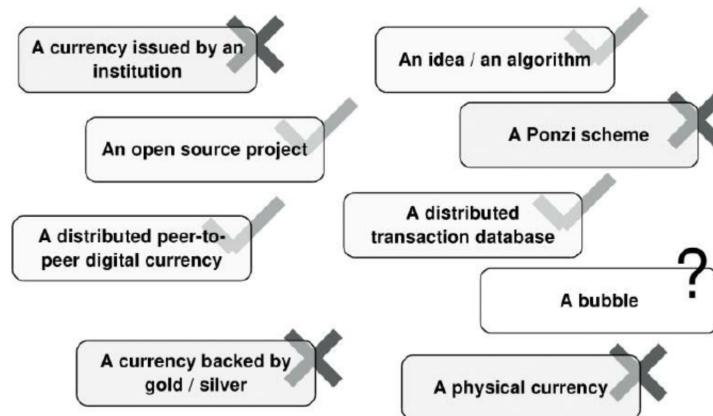


Figure 6.1: What Bitcoin is, and what not. [6]

In Figure 6.1 we have a display of what Bitcoin actually is, and why it might be irritating to some. In the end, bitcoins are just bits, digital data. It does not hold any physical value, and one could argue that the physical currency we handle every day and store in our wallets also is just a piece of paper that does not hold any actual value. But when our made up currency originated it used to stand for an amount of actual gold or silver stored at a bank. With bitcoin that is not the case. A bitcoin does not have physical properties we can grasp, nor is it issued by any institution. Bitcoin can be described as an idea, an algorithm, and a distributed database. It is realized in form of an open source project. With its success story, Bitcoin created huge media attention. This created a hype and led to other innovative uses of Blockchain. After some years the idea of BaaS was born and many tech companies provide Blockchain as a solution to their customers [16]. The aspects of Bitcoin potentially being a bubble are analyzed in the chapter 2.2.3.1 with a focus on cryptocurrencies in general.

6.2.2 Technology

The technology of Blockchain has been constantly evolving. An innovation during this process has brought out the permissionless Public Blockchain as we know it from Bitcoin [5]. This section will have a look into what a Blockchain consists of, the different types and methods being used.

6.2.2.1 Structure

In Figure 6.2, an example of a Blockchain is displayed. A Blockchain starts with a genesis block which consists of a unique hash number, a time stamp, the transactions (in Figure 6.2 displayed as TX 1 .. TX n), and a nonce, which is a random number to verify the hash. An added block will contain the parental hash number as well as a new, own hash number, created out of the nonce and the parental hash. As new blocks are generated with the information about new transactions and the blocks extend a Blockchain as a whole the Blockchain can serve as a ledger of the complete transaction history [4].

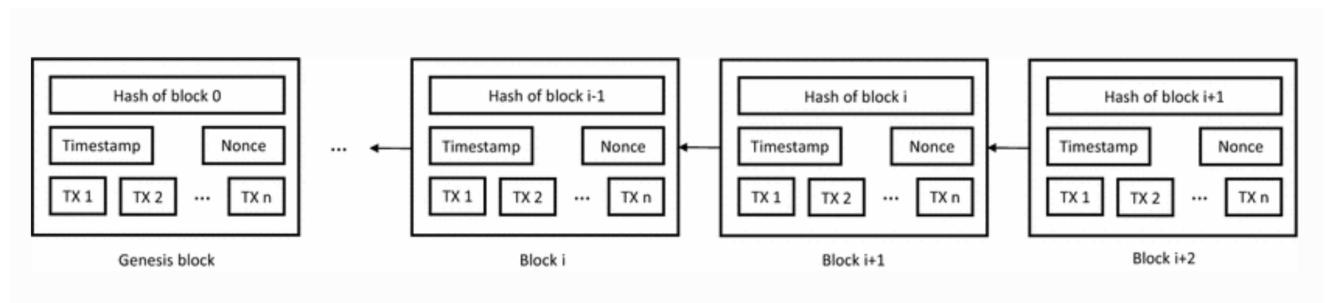


Figure 6.2: Example of a Blockchain [4]

Blockchains deployed as open, permissionless, public networks do not directly add a block to the chain. The information is first broadcasted throughout the whole network to see if there is no interference and, depending on the system, for example more than half of the networks members agree with the new block, it can be added to the chain [4].

Most Blockchain systems add an additional measure of security. A very popular method is to demand a so called Proof-of-Work (PoW), which was introduced by Satoshi Nakamoto, inventor of Bitcoin [5].

6.2.2.2 Types of Network

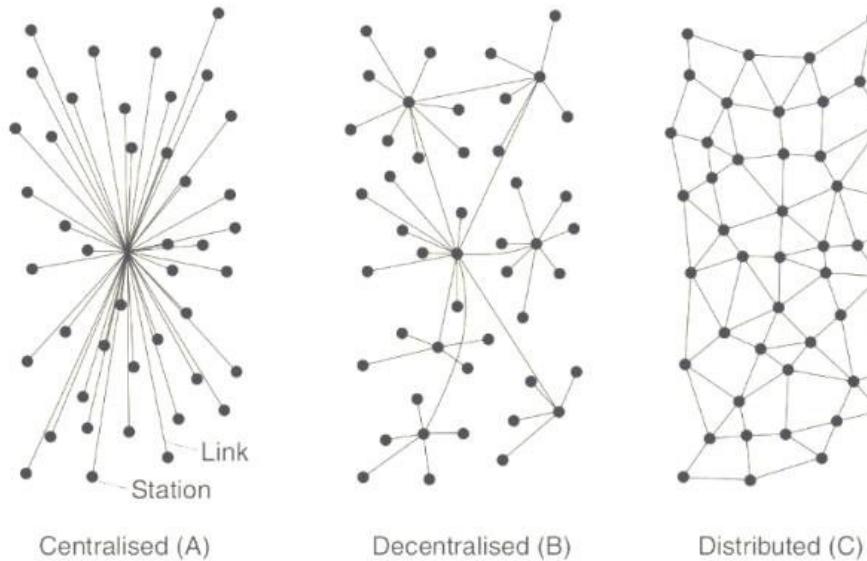


Figure 6.3: Network Methods [7]

There are different methods how a network can be build. Figure 6.3 displays a diagram of how these methods might look like.

- **Centralised (A):** In a centralised network all nodes are connected to a center node which could for example be a managing server. Many local area networks are build in this way, connected through one point. The big disadvantage of this method is that in case the managing center node should break down the whole network would be unusable [4]. This can be very expensive in terms of time and money.
- **Decentralised (B):** To avoid the problematic in which the whole network relies on one node there is the option to use a decentralised system. This method also increases the trust among users as they are not dependent on a single node [4].
- **Distributed (C):** When Blockchain first was introduced it was aiming for a distributed network method. As mentioned in the history of Bitcoin the idea was to broadcast the new information, the block to be added, through the whole network in which each node has the same amount of power. The majority of the network needs to agree with the authenticity of the chain and if the block can be added [5]. This allows for a maximum level of trust as there would not be a single party that can steer this vote to a personal gain.

While the distributed method is advertised to gain users trust it remains a fact that in a few Blockchain systems, even the one of Bitcoin, a number as low as 20 mining pools are in control of roughly 90 percent of the whole computing power in the network [5]. Therefore, it is more comparable to a decentralised method. Should some of the holders of these powerful nodes decide to work together and temper with the system they might be able to do that successfully. However, there are many rules and even punishments for owners of nodes. If they do not follow their duties, such as being online to receive the broadcasted information [4], or avoiding double spending [5], they have to deal with consequences from the system. Tempering with the system is also not in their interest as it would lose the users trust, might cause the bubble to burst and the digital property would lose its value.

6.2.2.3 Types

Blockchain systems can occur in different types. We differentiate between permissioned and permissionless. In a permissionless Blockchain system there is no administrator that allows who can hold a node. With permissioned Blockchains there is one or more trustees that are responsible for the census [5].

- **Public:** In a public Blockchain system anybody can become a user and participate. Bitcoin is a good example for this. Here the permissionless method is quite popular as it creates more trust. Chaums Vault system was relying on 2nd level trustees to validate the nodes and is therefore in some way using a permissioned method [5].
- **Private:** In context of our report the private Blockchain systems are more interesting. Here only invited users can participate in the network and if somebody uses Software as a Service the users are usually the employees or partners of a company [30]. One can assume that the company trusts itself and the own Blockchain system so there is no need to have a public, permissionless system. Private, permissioned systems are much more efficient as they do not need security measures such as PoW in order to be completely secure [5].

6.2.2.4 Controversy

There is a debate about what exactly is considered a Blockchain. Some argue that adding blocks of information to a linked list does not suffice to deserve the name Blockchain. In

this case, Satoshi Nakamoto is considered the inventor of Blockchain with the creation of Bitcion, a public and permissionless system. Other systems would be labeled a ledger, hyper ledger, distributed ledger [8].

The topic of this report is the comparison of BaaS providers. These providers label their products as Blockchains, regardless of them being permissioned, private, public etc. For this reason we consider all types of Blockchains in this report as true Blockchains.

6.2.3 Uses

The technology of Blockchain has various uses. This chapter will have an insight into some of the most dominant ones.

6.2.3.1 Cryptocurrencies

The most famous use of Blockchain is clearly in cryptocurrencies. Since Bitcoin was introduced thousands of new digital currencies have been developed that work in the same public, permissionless way. As displayed in Figure 6.4, Bitcoin still holds the highest Market capitalization. Even though there is this huge amount of digital currencies only the most valuable 14 hold 78 percent of the total digital currency's market capitalization [9].

No.	Cryptocurrency	Code	Exchange rate	Number of coins	Market capitalization
1	Bitcoin	BTC	\$ 8951.6394	17 180 188	\$ 153,790,855,757
2	Ethereum	ETH	\$ 646.5789	100 154 753	\$ 64,757,950,249
3	Ripple	XRP	\$ 0.80371665	39 541 619 593	\$ 31,780,258,035
4	Bitcoin Cash	XBC	\$ 1281.7728	17 275 946	\$ 22,143,838,920
5	EOS	EOS	\$ 16.7904	835 329 772	\$ 14,025,521,010
6	Cardano	ADA	\$ 0.3254823	26 188 960 137	\$ 8,524,042,980
7	Litecoin	LTC	\$ 143.9658	56 898 395	\$ 8,191,423,098
8	Stellar Lumes	XLM	\$ 0.39935016	18 759 309 869	\$ 7,491,533,398
9	Tronix	TRX	\$ 0.09088398	66 412 089 292	\$ 6,035,794,995
10	NEO	NEO	\$ 80.4078	65 659 718	\$ 5,279,553,500
11	IOTA	IOT	\$ 1.8513	2 800 940 157	\$ 5,185,380,514
12	Monero	XMR	\$ 230.3433	16 146 465	\$ 3,719,230,081
13	Dash	DASH	\$ 454.2615	8 121 006	\$ 3,689,060,743
14	Nem	XEM	\$ 0.39074508	9 090 909 088	\$ 3,552,227,999
15	Tether	USDT	\$ 0.99	2 450 101 824	\$ 2,425,600,806
16	Vechain	VEN	\$ 4.3362	530 773 566	\$ 2,301,540,339
17	Etherum Classic	ETC	\$ 20.7207	102 514 915	\$ 2,124,180,807
18	Qtum	QTUM	\$ 21.4731	89 476 920	\$ 1,921,346,857
19	OmiseGO	OMG	\$ 16.1964	103 074 544	\$ 1,669,436,555
20	Binance Coin	BNB	\$ 13.7511	115 221 419	\$ 1,584,421,260
21	Lisk	LSK	\$ 12.2463	106 472 521	\$ 1,303,894,445
22	RaiBlocks	XRB	\$ 9.4347	134 639 002	\$ 1,270,278,593
23	Bitcoin Gold	BTG	\$ 69.7455	17 146 505	\$ 1,195,891,577
24	Verge	XVG	\$ 0.07307982	15 092 890 527	\$ 1,102,985,723
25	Zcash	ZEC	\$ 272.6064	3 856 723	\$ 1,051,367,505

Figure 6.4: List of most valuable digital currencies sorted after highest market capitalization in \$ on 01/05/2018 [10]

Digital currencies consist of only bits or bytes. They do not stand for any physical value in any sort. The only value created is the artificial scarcity of a kind of digital data and participants gamble on its success [9]. The future will tell if this bubble will burst, either due to speculators loosing trust or interest, or maybe due to newly introduced laws.

6.2.3.2 Smart Contracts

Contracts are agreements that are binding by law. A smart contract is a contract that uses IT technology to assure that the agreed terms are being followed. This can create an autonomous process that gives every party involved confidence. An example could be the automated payment once a project has finished.

Smart contracts can be based on Blockchain technology. This will automatically increase the trust. A supervising authority that secures a accurate transition, for example a bank, would not be needed anymore. With real peer-to-peer contracts, administrative effort is reduced [11].

Using a Blockchain in a smart contract can also make it more flexible. For example, a project might be crowd founded and the investments of a big amount of founders will be immutably saved in the change. Should the project succeed they could receive their product, otherwise their money back, dependent on contract.

There are also other possible uses for smart contacts. Securing intellectual property can be made very easy. The process to create a patent for an invention is very time consuming and expensive. Using a smart contract based on a Blockchain we can easily validate the origin and timing of something submitted and save a lot of administrative cost [11].

6.2.3.3 Supply Chain

The use of Blockchain technology to monitor the supply chain of a company brings valuable benefits and makes the process much more straight forward. This, again, will save administrative work. In Figure 6.5 we can see an example from an agriculture business.

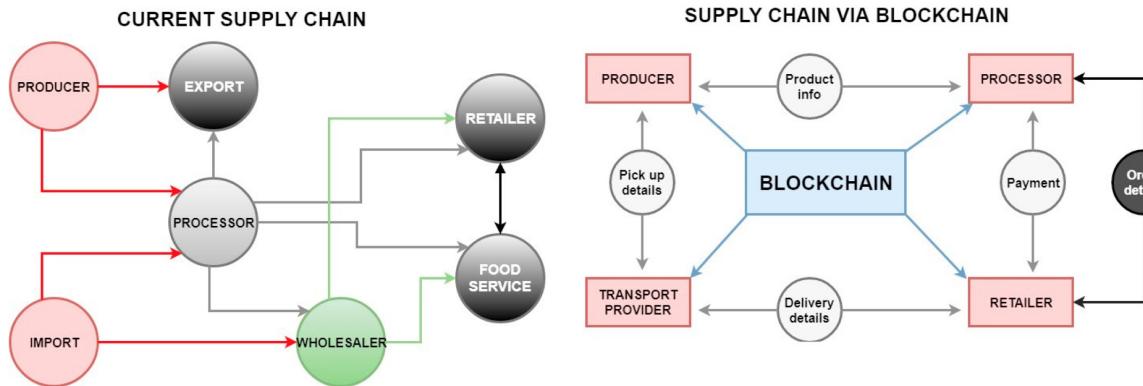


Figure 6.5: Conventional agriculture supply chain and one with Blockchain [12]

Using a Blockchain, the supply chains are, in this example, decentralized. Every member of the supply chain can document their transactions into the chain. Members can also read the blocks of the Blockchains that are in direct contact with them [12]. This leads to more efficiency and control through the whole supply chain.

One of the companies using Blockchain technology to monitor their supply chain is Starbucks. Starbucks even goes one step further and uses this advanced supply chain for their marketing. Coffee production has been a controversy topic as people are worried about the circumstances how and where it is sourced. With this sophisticated, Blockchain based supply chain it was easy for Starbucks to print QR codes onto their products with which customers can easily track their beans to the source [13]. The solution of Starbucks supply chain is provided by one of the BaaS providers, Azure by Microsoft, which we will introduce later in this report.

6.2.3.4 Medical Records

Medical records are a very sensitive topic. They must be private, protected, and if they are complete and accurate they can be very useful for patients and medical staff. This is a concern that exists as long as professional medical care and Blockchain technology brings a new solution.

Using Blockchain we have a medical history that is immutable. Nothing gets forgotten or lost. Malpractice can not be erased as the Blockchain does not allow to be tampered with. Allergies, previous procedures and medication are all traceable in a secure and protected way. It cuts down on administrative effort and optimizes a crucial and outdated process that is long overdue [14].

6.2.3.5 Energy Trading

Renewable energies are in high demand. While in the past it was reserved to major corporations to generate and provide electrical energy it is now possible for most home owners to invest into solar panels and generate their own energy. To store this energy in house in form batteries has many disadvantages. A big amount of the energy will get lost and the batteries are involved with a high cost in acquiring and maintaining. The environmental impact of the batteries with our technological standard is also a cause for concern. Since most houses were already connected to the national power network prior the investment into solar panels it is an easy procedure to feed the generated solar energy during the day into the national power network and source it back during the night, when needed. Its in both, the home owners and energy providers interest to monitor precisely how much energy went from one place to the other. An investment into solar panels can be at very high financial cost and by today's standard often produces more than the household uses. Home owners want to be compensated for the excess energy generated and Blockchain can be an elegant solution to monitor the flow. This can even include automated payments as described in the smart contracts. Using a Blockchain system creates trust and might convince some investors to cooperate with the energy providers. The only concern is that a permissionless Blockchain system with PoW uses a large amount of energy itself [15].

6.3 Cloud Computing

During the late 2000's, Cloud Computing emerged as a new paradigm for hosting and delivering services over the internet [18]. Cloud Computing can be described as set of network enabled services that provide scalable and inexpensive computing platforms. These platforms need to be accessible in a simple and persuasive way [19].

One of the reoccurring characteristics of Cloud Computing is its on-demand self-service structure. Consumer can instantaneously get access to computing resources without any human interaction. Secondly, these computing resources are available through broad network access, normally via the internet. Also, many Cloud Computing providers group their resources into pools so that they can use multi-tenancy and virtualization to use these resources more efficiently, profiting from economies of scale. However, even though resources are pooled together, the cloud infrastructure still provides a way of measuring the resource usage for each individual customer so that it is possible to bill them whenever they are using any Cloud Computing services. Through this, computing resources are not an up-front commitment for consumers anymore but rather an on demand service that can scale up immediately when a lot of computing power or storage is needed [20].

6.3.1 Service Models

Cloud Computing service models are often described through expressions ending with “-as-a-Service”. The three most used classifications are:

- **Software-as-a-Service (SaaS):** The Cloud Computing provider offers ready to use applications running on the cloud infrastructure. The applications are accessible through a thin client interface like a web browser. An example of a SaaS application are Google’s Gmail, Dropbox or Salesforce [2].
- **Platform-as-a-Service (PaaS):** Customers can deploy their own applications onto the cloud infrastructure. These applications can be created using programming languages, libraries, services and tools that are supported by the provider [21]. Examples of PaaS providers are Microsoft Azure and Amazon Web Services [22].
- **Infrastructure-as-a-Service (IaaS):** Storage, processing power, networks and other fundamental computing resources are provided to the customer. The customer can then run arbitrary software on the infrastructure which can include operating systems and other applications [21]. Examples for IaaS providers are NTT Communications and OpSource [22].

Further service models can be distinguished such as Hardware-as-a-Service, Safety-as-a-Service or Data-as-a-Service [19]. However, since the line between low level infrastructure and high-level platform is already difficult to draw, only the three categories BaaS, PaaS and IaaS will be used in this paper. These three categories are also visually explained in Figure 6.6.

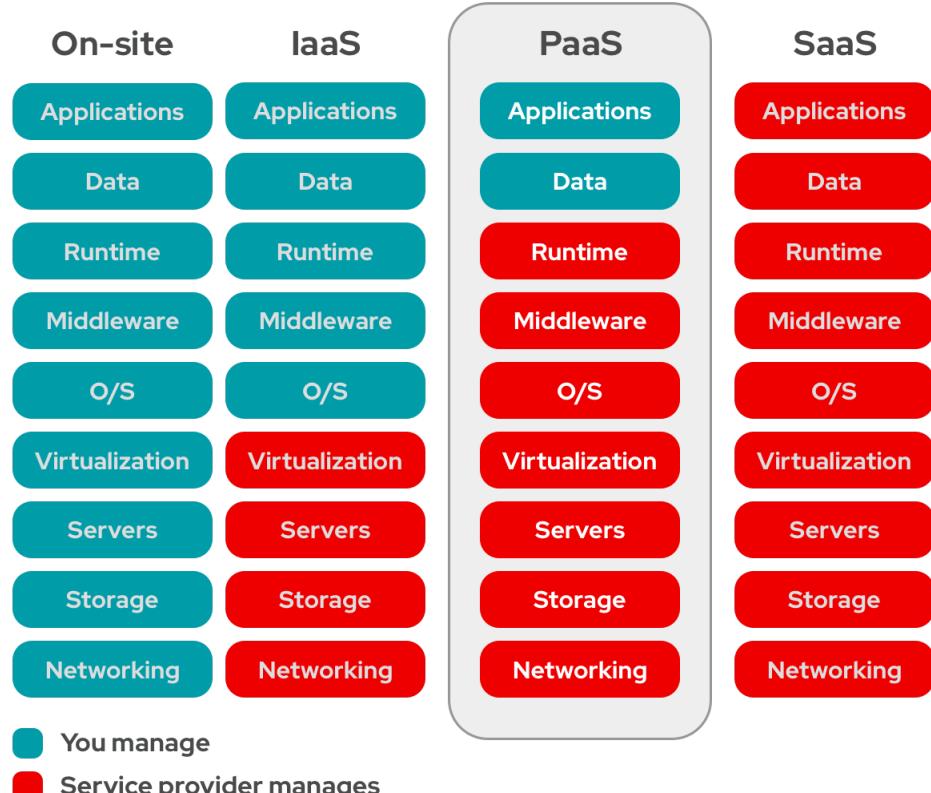


Figure 6.6: Management of SaaS, PaaS and IaaS [24]

6.3.2 Advantages and Challenges of Cloud Computing

Cloud Computing services are highly appreciated for their ability to immediately provide on-demand computational power to customers. This gives companies the flexibility to get instantaneous access to resources when needed leading for example, to a faster time to market in many businesses. This reduces the upfront capital investment into hardware for users, making it possible for smaller companies to primarily focus on their core strengths instead of having to also deal with infrastructure, leading to more innovation and promising startups. Also, this allows small companies to benefit from compute-intensive services such as business analytics that were only available to large corporations before [18].

From the providers point of view, it is possible with Cloud Computing to allocate resources much more efficiently than single companies could ever do on their own. The same infrastructure can be used by different end users that are also using the infrastructure in different ways. Since many companies are subject to seasonal or abrupt raises and declines in demand, Cloud Computing gives these companies the opportunity to react accordingly [23]. This seasonal demand can also be seen in Figure 6.7. Because of this, the user likely pays less for Cloud Computing services than when maintaining the infrastructure by themselves. Lastly, Cloud Computing makes it possible to develop new types of application. For instance, interactive mobile applications have been developed that respond in real-time to locational, environmental and contextual data provided by humans and sensors. They use external batch processing to analyze this data quickly which would not have been possible without Cloud Computing [18].

Figure 2. (a) Even if peak load can be correctly anticipated, without elasticity we waste resources (shaded area) during nonpeak times. (b) Underprovisioning case 1: potential revenue from users not served (shaded area) is sacrificed. (c) Underprovisioning case 2: some users desert the site permanently after experiencing poor service; this attrition and possible negative press result in a permanent loss of a portion of the revenue stream.

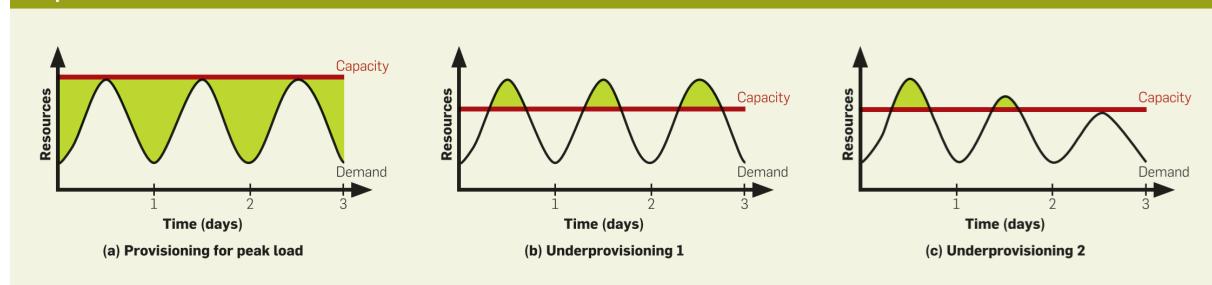


Figure 6.7: Inefficient resource allocation for a single company [23]

Although there are many benefits to Cloud Computing, challenges arise as well. One of the big issues that keep companies from using Cloud Computing services are security concerns. Security issues such as data loss, phishing and botnets propose real threats to Cloud Computing users [20]. For example in 2019, data of 100'000 customers stored on the cloud storage service S3 of AWS by the Capital One bank was leaked by using a vulnerability in the bank's cloud infrastructure by hacker who previously worked at AWS [25]. This is why some companies are not willing to entrust Cloud Computing providers with their sensitive data. When data is stored and used on a companies internal hardware, it is clear that the company itself is also responsible for securing that data. With cloud computing however, there is a middle ground between the surface layers where the customer has to ensure security and the bottom layers where the provider is responsible for data security where both of them share the responsibility to protect the data, which can lead to conflict [23]. Also, Cloud Computing providers have to adapt to compliance requirements of various countries. For example, Cloud Computing providers have to maintain business legal documents that comply with various laws of specific countries [18].

Another issue with Cloud Computing is the fear of inconsistent service availability. Even though big Cloud Computing providers have proven to be very reliable when it comes

to server availability, the few instances when disruptions occur are quickly picked up by major news sources [23]. For example, the two outages of AWS S3 during the August of 2008 (2 hours and 8 hours) have been highly reported on [26; 27]. Another concern that makes companies reluctant to switch to Cloud Computing is the fear of vendor locking and interoperability. Many cloud providers have already made considerable progress to standardize their systems. However, there are still many interface points that make it difficult to integrate the existing software with Cloud Computing services [18]. Moreover, once a system is tightly integrated into the Cloud Computing services of a provider, switching to another provider becomes causes new integration points which leads to vendor lock-in [23].

Moreover, another challenge that companies encounter when migrating their applications to Cloud Computing provider infrastructure is the lack of support. When buying on-premises software, the support is usually included in the price of the software. With Cloud Computing however, the providers have to employ and train support adequate staff that the customers then often have to pay for [28]. All in all, these hidden costs such as costs for support, disaster recovery, a data loss insurance, application modifications and integration costs diminish the economic value of Cloud Computing and have to be weight against the benefits of Cloud Computing both short and long term [18]. All the previously mentioned advantages and challenges are summarized in Table 6.1.

Table 6.1: Advantages and Challenges of Cloud Computing

Advantages	Challenges
Scalability	Security and Data privacy
On-demand, immediate access	Reliability concerns
Reduces upfront costs	Integration difficulties
Efficient resource allocation	Vendor locking and interoperability
Low costs	Compliance
Supports innovation	High support costs

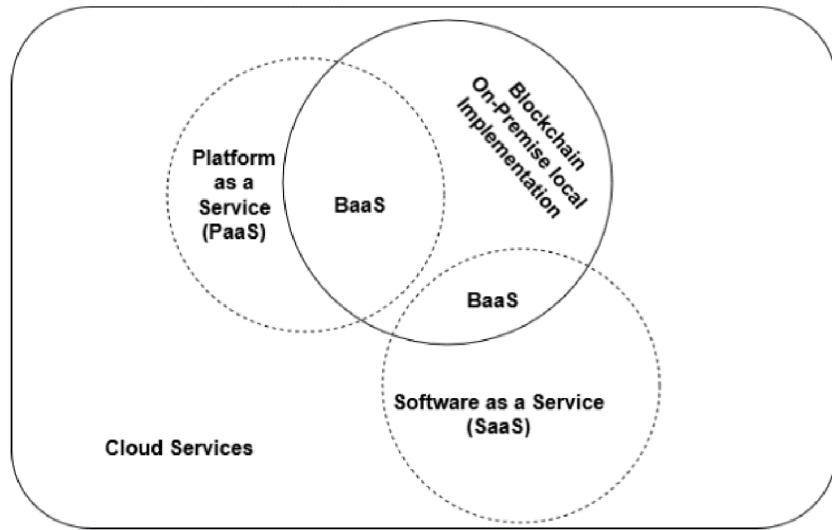
6.4 Blockchain-as-a-Service (BaaS)

BaaS is a term used to describe a service provider offering a combination of Blockchain technology and Cloud Computing [29]. This encompasses the management of the infrastructure, identity management services, middleware and other components that provide a foundation for creating, deploying and managing Blockchain solutions [30].

BaaS providers can be grouped into two types. Firstly, there are providers that work together with clients to develop a state-of-the-art Blockchain application. This approach is strongly application oriented and therefore similar to SaaS. In this case, clients only need a minimum understanding of Blockchain technologies.

The second approach allows clients select, use, combine, integrate and customise different components that are provided by the platform to create their own, individualised Blockchain solution. These providers give clients the possibility to select and manage the chain, define ledger records or configure the access regime. This approach corresponds to the idea of PaaS [30]. These two service model types are also visualized in Figure 6.8. Blockchain dedicated services that are equivalent to IaaS are still very sparse, but in the future there might be the option for renting Blockchain-tailored hardware [32].

Since BaaS providers are mainly focused on the application of Blockchain, they themselves rely on other Blockchain platforms and communities to create the foundation for

**Figure 6.8:** Types of BaaS [31]

their BaaS solutions. Two of the most used platforms whose code base and services are integrated into the applications of Amazon's AWS BaaS, Microsoft's BaaS and IBM's BaaS [30], [33] are the Hyperledger Consortium and the Enterprise Ethereum Alliance (EEA). The Hyperledger Consortium is an open source community which is focused on developing and providing tools, libraries and frameworks for Blockchain deployment on an enterprise level, often for private Blockchains [34]. One of the services that they are providing is for example Hyperledger Fabric, a framework which allows to write Blockchain applications with access control and permissions for Blockchain data [33].

The EEA on the other hand is an organisation which aims to customize Ethereum for companies in the industry [35]. The EEA is governed by the Ethereum foundation, a non-profit organisation designated to support Ethereum services [36]. For example, the EEA provides an Ethereum framework with whom it is possible to write Blockchain applications with very little downtime, censorship and fraud. Ethereum's framework is mainly used if users want to perform peer transactions on the Ethereum public network or use Ethereum's Solidity smart contract language [33]. Another functionality that is provided by Ethereum's Blockchain platform is a complete programming language for writing smart contracts. These contracts can be used to create a decentralized uncensored computer that is automatically maintained. The functionalities of the programming language are mainly focused on digital currency transactions [37].

6.4.1 Advantages of BaaS

Since BaaS is the combination of Cloud Computing and Blockchain, it also shares a lot of the advantages with Cloud Computing. One of the advantages of BaaS is the ability to integrate the Blockchain technology into other tools and services that are already provided by the owner of the BaaS platform. This way, Blockchain applications can profit from other applications that are part of the repertoire of the BaaS provider such as analysis, access tools or survey tools [33]. A second advantage of BaaS for its customers is that the infrastructure is managed by the BaaS provider [29]. This keeps the infrastructure agile and lets the provider perform efficient resource allocation which helps saving energy and costs [37]. A lot of this is possible because most Blockchain related services are automatically measured and monitored, which lets the system automatically optimize the usage of its resources. The constant monitoring also gives transparency to the user about the availability, performance and pricing of the BaaS application [33]. Moreover, the

environment can grow with the Blockchain application by designating more storage space and computational power to specific applications [33].

Another advantage of BaaS is the fact that especially with application-focused providers, nearly all knowledge about Blockchain technology is being provided by the platform owner. Since Blockchain is still an emerging technology, profound domain-knowledge is limited in many companies and experts are in high demand. With BaaS, the knowledge of few Blockchain experts can be made available to a wide range of companies, optimally distributing the expertise of the present Blockchain specialists. This lets companies with minimum Blockchain knowledge explore Blockchain technologies by being provided abstractions of lower-level details [29].

Further, companies can focus on their core products and think of new ways to build Blockchain applications that suit their needs. For example, there is a need for supporting applications around the development of industrial Internet of Things technologies. One of the needs of industrial Internet of Things consists of a way to secure transactions. This can be achieved either by an in-house Blockchain solution or a BaaS provider [38]. Such a system can help with the tracking of suppliers identity and reputation, the traceability of products and smart diagnostics for machine maintenance [39]. Other use cases for BaaS are the tracking of goods inside an industrial supply chain, a way to help sending funds abroad or the tracking of less expensive real estate transactions [16].

The third big advantage which is often promoted in advertising about BaaS is its potential to fasten the development and deployment process of Blockchain applications. Applications that would have taken months to deploy by using an in-house Blockchain solution can be deployed in weeks or days when using a BaaS solution and are by that also often cheaper than in-house solutions [16]. Lastly, most BaaS provider help standardizing the Blockchain field by building their services on the Blockchain services of big consortiums like the EEA and the Hyperledger Consortium [33]. This leads to faster and more efficient Blockchain applications [40].

6.4.2 Challenges and Risks of BaaS

Even though there come lot of benefits with the use of BaaS, certain challenges and risks arise as well. The most prominent challenge of BaaS involves data privacy. Traditional Blockchain applications are designed to be accessible by all users [3]. However, there are multiple application scenarios where not every participant of a Blockchain application should have access to the full system. For example, a banking consortium wants to use Blockchain application to track assets between departments. However, only pre-vetted participants are allowed to access the asset tracking. Even though this situation is realisable with a BaaS provider that offers identity management services and digital signatures, the same task is much easier implementable with a secure append-only database [30]. In the same way, traditional Blockchain applications do only provide one level of permission control which provides full access to all users. However, with BaaS, this might lead to an inexperienced user triggering a function that becomes a vulnerability of the Blockchain-based application [3].

Another challenge that arises from BaaS concerns trust considerations around the provider and the tenants of BaaS. On the one hand, application-focused owners of BaaS platforms often provide the whole infrastructure underlying Blockchain services. However, this undermines the trust mechanism that Blockchain aims to address. Normally, one of the goals of Blockchain is that a single participant does not manage all the nodes so that not all copies of the ledger are held by just one participant. However, this is the case if the foundation of a Blockchain application is offered by just one provider. This issue can be addressed by either offering a more PaaS-like setting where the tenant has more control over the application or by giving a higher level of operational visibility to the tenant so

that they would see if something about the Blockchain gets changed. This is especially important for applications that are dealing with sensitive data [30]. News reports about data breaches on cryptocurrency trading platforms also raise management concerns once managers consider a BaaS solution instead of building an in-house Blockchain application [16].

On the other hand, for traditional SaaS applications there is a contract between a legal tenant and a provider. In the context of BaaS however, tenancy can mean that some participants have more power to control the infrastructure than others. For example, a company that wants to trace its products with the help of a Blockchain application might be the single tenant of the BaaS provider. The suppliers and customers of this company however might not interact with provider directly, even though they are using the Blockchain application as well. This can lead to an over-concentration of power for some tenants, especially with more open BaaS environments. A solution for this would be that each participant should manage their node through their own BaaS tenancy arrangement with the BaaS provider [30].

Governance can also cause issues in the context of BaaS. Design decisions of the BaaS application have direct implications for security and trust considerations. For example, Bitcoin has technical mechanisms to enforce the validity of transactions so that the same coin cannot be spent twice. These mechanisms are developed and negotiated by both BaaS provider and tenant. However, the degree to which a single party of the BaaS application can influence these design decisions are often not clearly defined. Therefore, provider and client should arrange contractual terms where governance issues are clearly agreed upon to reduce future negotiations [3].

Lastly, BaaS also raises the number of technologies used for a Blockchain application. This leads to more integration points, requiring a lot of attention during the development and deployment of BaaS applications [33]. Also, most BaaS solutions lock customers into a specific Blockchain or cloud platform, making the customer dependant on that specific technology [3]. The Table 6.2 summarizes the previously discussed advantages and challenges.

Table 6.2: Advantages and Challenges of BaaS

Advantages	Challenges
Survey tools, access management and analysis tools	Data privacy
Efficient resource allocation	Complicated access management
Lower energy needs	Permission control
Provider offers Blockchain know-how	Trust considerations tenant
Faster development	Trust considerations provider
Faster deployment	Governance
Helps standardizing Blockchain	Many integration points
New Blockchain technology application ideas	Locks in to one provider

6.5 Providers

6.5.1 Microsoft Azure

Microsoft is a worldwide company employing over 160'000 employees all over the world. Over half (58.5%) of their employees are engineering, working on their service products. It was founded 1975 and as of 2020, its headquarters is located in Redmond, USA

[41]. Microsoft has a long story of successfully deploying new products and software services in the last decades [42]. From developing a widely spread operating system Windows to Office products like Word, PowerPoint, Excel [41]. On the 1. of February 2010, Microsoft released its new platform as a service offering called Microsoft Azure [42]. As of today, the Microsoft cloud platform contains over 200 products and cloud-services, constantly expanding worldwide [43].

Microsoft is owning a large global network containing multiple data centers interconnected through different edge-nodes spread among the different regions in the world. As shown in the Figure 6.9, Microsoft has strategically positioned servers such that they can ensure quality experience for billions of Application Programming Interface (API) requests [44].



Figure 6.9: Microsoft Global Network [45]

Through their distributed data centers and servers they can manage services such as Microsoft Azure, Bing, Dynamics 365, Microsoft 365 and XBox. [44]

In the fiscal year 2020, ending June 30, Microsoft Coperation had a total revenue of over \$140 billion, increasing by 13% compared to last year. About \$48 billion is solely due to their Intelligent Cloud. This includes server products and cloud services. This segment by its own increased its revenue by almost \$10 billion compared to 2019 [46].

6.5.1.1 Product and Services

Over the last two years Microsoft built up its own Blockchain technology on Microsoft Azure. The service comes along with a Blockchain Development Kit, an Azure Blockchain Workbench and Azure Blockchain services [47]. These three parts are discussed in the next sections.

The Azure Blockchain Workbench is an easy-to-use platform to configure and define Blockchain applications. It provides a web application with a REST API and a message-based API. This should let the customer focus their work more on business logic and writing smart contract code [48]. The Workbench also allows customers to integrate into existing systems as well as signing and routing transactions to the appropriate Blockchain. Required components to run a consortium Blockchain network are provided by the tool and it also supports Ethereum. The whole package comes along with a lot of tutorials and documentation about managing Blockchain applications [48].

The Blockchain Development Kit is a Github Repository providing different content about Blockchain. The Repository is publicly available with various sample codes for connecting data to or from Blockchain, integrating tools, systems and general information about DevOps [49].

Along these products Microsoft released their Azure Blockchain Service. With this service customers can fully manage a Blockchain network, particularly creating and configuring a consortium Blockchain infrastructure. It includes full node management, scalable data streaming, monitoring smart contracts, transactions and stream on-chain data to off-chain data [50].

Microsoft also presents this services with high quality security certificates including ISO/IEC, CSA/CCM, ITAR, CJIS, HIPAA, and IRS 1075 [50]. They also support their Azure cloud with the Azure DDoS Protection, Key Vault for keeping track of keys and secrets, Security Center to unify security management and many more products and services to ensure multi-layered security [51]. They also state that they yearly invest over \$1 billions into cybersecurity and have over 3500 security experts working on security and data privacy [50].

In the following section, well-known enterprises that are using Microsoft Blockchain Service are listed.

- Starbucks, a global chain of coffeehouses, is using the Microsoft Azure Blockchain to trace their path of a coffee bean to the coffee cup. Stakeholders can therefore see where the bean were grown and roasted, and how they landed in their cup [52].
- Singapore Airlines has another use case for their Azure Blockchain application. In their current loyalty program for customers they use Blockchain to convert their KrisFlyer air miles into KrisPay miles that they can spend on other servers. The Microsoft Blockchain is considered as an optimal technology by Singapore Airlines to have a highly secure ledger of transactions [53].
- GE Aviation's Digital Group is large company with over 10000 employees. Their goal is to deliver data and domain-driven solutions for the aviation industry, this includes manufacturing. GE Aviation partnered with Microsoft to reduce costs and raise efficiency by better keeping track of part and repairs traced over their lifetime with the Blockchain service [54].

6.5.1.2 Pricing

Microsoft Azure Blockchain Service has no upfront costs. As a customer one will pay for the amount of validator nodes, transaction nodes, Blockchain storage and the data manager for the transactions. In the Figure 6.10 the pricing details for the East US in US Dollars can be seen. The prices will vary depending on the region where the customer will use the infrastructure [55]. It is also possible to decide between the Basic or the Standard pricing, decided by using it either as a test environment or to run a production workload respectively.

It also has to be mentioned that the Blockchain service is or can be coupled to other Microsoft Azure products and services, which are not included in the monthly service. It is very likely that a customer would use other services like Azure Cosmos DB, Azure SQL Database or other management tools which increases the cost [56].

	Basic	Standard
Compute	Environment for dev/test 1 vCore	Run production workloads 2 vCores
Consortium Governance	✓	✓
Transaction node price (per hour)	\$0.0996/hour	\$0.318/hour
Validator node price (per hour)	\$0.0996/hour	\$0.318/hour
Blockchain storage price GB per month	\$0.05/month	\$0.05/month
Blockchain data manager	\$0.0001/transaction Included transactions – 50/day	\$0.0001/transaction Included transactions – 50/day

Figure 6.10: Pricing Microsoft [55]

6.5.2 IBM

6.5.2.1 IBM Corporation

IBMs roots goes back to 1911, where three companies in the manufacturing industry merged together to a new one called Computing- Tabulating- Recording Company (C-T-R). At this time C-T-R produced a variety of products such as punch cards, meat and cheese slicers and tabulators [57]. In the 1920s C-T-R change their name to International Business Machines Corporation, short IBM, due to the expansion to Europe and their current products [58]. During the second world war IBM began to develop along with Harvard University the first machine that could execute long calculation tasks automatically called Mark I [59]. This set the step for the future development of many successful stories like helping NASA to land on the moon 1969 by providing computers and complex software programs, IBM Floppy Discs invention 1971 and today's supercomputer with computing power over 200 petaflops [60]. It can be stated that IBM is a huge internationally established company. It has been stated on their website that their total revenue is around \$77.1 billion, including a total cloud revenue of \$21.2 Billion [61].

6.5.2.2 Product and Services

Before introducing IBM's BaaS, first IBM cloud must be presented. IBM cloud is a cloud platform that contains over 170 products and services. This includes products such as databases, security and network services [62]. One of these products is the IBM Blockchain Platform. The Blockchain platform comes along with a lot of helper tools such as technical brief about the platform itself, a free Ebook from O'Reilly about developing and designing Blockchain, such that a potential customer has an easy time to get familiar with their Blockchain platform [63]. They also provide customers an E-mail address that can be used for asking questions on how to start your own Blockchain product. IBM also provides an VS-Code extension with that a customer can experience a easier way to integrate and test the project onto the network that is used [63]. At the beginning customer has to chose between using the IBM Blockchain Platform deployed on the IBM cloud or deployed on an other system, also supporting Hyperledger Fabrics [69]. This means that a customer either has its own kubernetes clusters where he can deploy his platform for the computing or the customers uses IBM cloud kubernetes clusters to compute. This of course leads to an other pricing model, in this case we will look at deploying it on the IBM cloud since it has a clear price table that can be used to compare it with providers [64].

IBM has some successful stories with their provided Blockchain platform. In the next section there is a description of different clients using the service.

- Nordea Bank Abp is a financial company that has build up a trade financing platform (we.trade) that is using IBM Blockchain Platform running on IBM Cloud. The

platform is used for international trading with a heavy mark on cross-border trading. They service allows them to have a protected ecosystem for that [65].

- Kroger, a large grocery retailer uses the IBM Blockchain Platform to have better traceability. In order to have a better understanding of their supply chain they apply the service to trace their products, ultimately to identify possible drop in quality upon their suppliers[66].
- Plastic Bank is a institution that plans to minimize the ocean plastic pollution by monetizing ocean plastic. They setup recycling systems in second and third world countries to help local citizens to gain money by collecting plastic. This deployed network is used to have a secure and scalable reward system for users [67].

6.5.2.3 Pricing

The pricing model of IBM is based on CPU allocation. Different to per node per hour price model, a customer has to pay for cores used per hour [68]. IBM uses for their kubernetes clusters on IBM cloud to allocate the CPU the Blockchain platform needs. An additionally to the CPU pricing a customer is also charged with a storage price that is needed for a Blockchain. There are no upfront costs such as member fee and it is stated that developers can have clear cost estimation by using a cost estimator tool such that a customer always has comprehensible billing. Additionally, it is stated on their website that a customer has always the option to scale to extend their Blockchain or downscale to minimize their expense [68]. Another benefit of the CPU allocation is that a customer can choose on how much CPU a node needs. Therefore, a developer can decide on how many nodes they need with the needed CPU allocation. The Figure 6.11 below shows the pricing model.

Pricing options** (1 VPC = 1 CPU = 1 vCPU)	Test Network	Join a Network
CPU allocation	1.25 vCPU Includes: - 1 peer (0.7 vCPU) - 2 CAs (0.1 vCPU x 2) - 1 ordering node (0.35 vCPU)	2.9 vCPU Includes: - 2 peers (for HA) (2x default compute = 2 x 0.7 x 2) - 1 CA (0.1)
Hourly cost: IBM Blockchain Platform	\$0.46 USD (1.25 vCPU x \$0.29 USD/VPC-hr)	\$0.81 USD (2.9 vCPU x \$0.29 USD/VPC-hr)
Hourly cost: IBM Cloud Kubernetes cluster	\$0.29 USD (Compute: 4 x 16 lowest tier; 1 worker node; 1 zone)	\$0.29 USD (Compute: 4 x 16 lowest tier; 1 worker node; 1 zone)
Hourly cost: Storage	\$0.06 USD 340GB Bronze ↗ 2 IOPS/GB	\$0.09 USD 420GB Silver ↗ 4 IOPS/GB
Total hourly cost	\$0.71 USD	\$1.19 USD

Figure 6.11: IBM Pricing BaaS [71]

The hourly cost on storage is calculated on IOPS/GB which means input/output operations per second per GB which costs \$0.05, \$0.12, \$0.16. and \$0.48 for 0.25, 2, 4, 10 IOPS/GB respectively per month [70]. It has to be mentioned that there is an hourly cost

charged for the platform itself for allocating clusters as it can be seen on the Figure 6.11. In a demo scenario where a customer wants to host a network including 1.65 CPU (\$0.29), 240 GB storage (\$0.05 per hour) the customer will get a approximate total cost of \$606 per month [71].

6.5.3 Amazon's AWS Blockchain

6.5.3.1 Amazon.com, Inc

Amazon.com, Inc. is an internet-based company that sells directly or as a middleman books, electronics, movies, music and other goods. Under the name Amazon Web Services (AWS), the company also offers data storage and computing resources for renting. The company was formed by Jeff Bezos in 1994 and started as an online book-seller [72].

In 2002, the company launched Amazon Web Services, which offers today more than 175 services that are focused around but not restricted to Cloud Computing such as databases, developer tools, analytics, compute, storage and security [73]. Amazon leases and owns more than 25'000 m² of fulfilment and data center space around the world [74]. A big part of this space is used to provide AWS services in 24 geographic regions on every continent (See Figure 6.12). Amazon employed 876'800 people as of June 30th and reported sales of \$88.912 million and an income of \$5,851 million for its second quartal of 2020. For AWS, Amazon reported net sales of \$10,808 million and an income of \$3,357 [75]. The numbers of employees working for AWS is not getting published by Amazon but it is estimated that between 40'000 and 62'000 people are currently working for AWS [76].

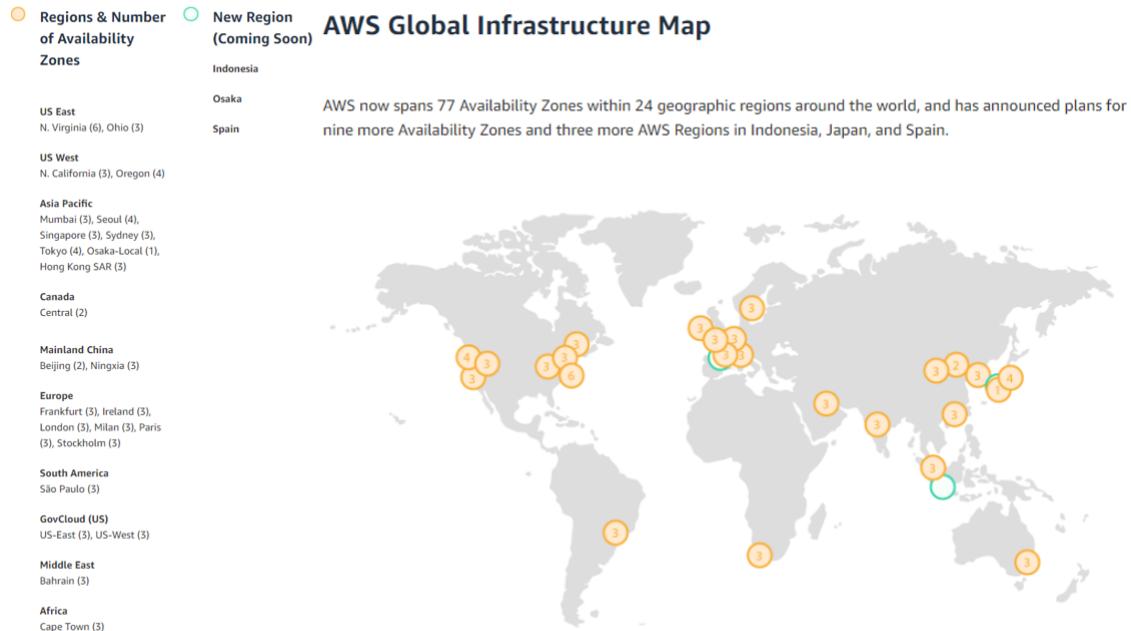


Figure 6.12: AWS Global Infrastructure [85]

6.5.3.2 Product and Services

In December of 2018, Amazon announced its own fully managed Blockchain service, that supports both Ethereum and Hyperledger Fabric frameworks [77]. Previously, Amazon only provided Blockchain options by forging partnerships with other companies such as R3 or Kaleido [79].

One service that Amazon provides in relation to Blockchain is the Amazon Managed Blockchain, where the user is provided with a fully managed Blockchain service that lets him or her create Blockchain networks based on the open source framework Hyperledger Fabric. Even though the support of Ethereum networks was already announced in 2018, it is still not available to this point [80]. Included in the Amazon Managed Blockchain are tools to set up and manage a Blockchain network. Also, Amazon provides the infrastructure for running the Blockchain network and automatically scales it with increased storage and computational needs. Moreover, there is a Blockchain voting API so that network participants can add or remove members to the Blockchain network. In order to make it as easy as possible for the users to learn the Amazon Managed Blockchain, Amazon provides an extensive Blockchain management guide where it is explained in detail how to set up and manage Blockchain networks. Further, Amazon hosts loads of Webinars and Seminars about the Amazon Managed Blockchain [81]. Lastly, it also secures Hyperledger Fabric's certificate authority with AWS Key Management Service technology to meet security needs of sensitive applications [82].

For companies that are in need of a way to track and maintain data completely and accurately with an immutable and verifiable history of all data changes, Amazon also provides the Amazon Quantum Ledger Database (QLDB). This service is especially useful for companies that have a lot of data and security needs. For them, traditional databases with audit logs are not suitable since they are prone to errors, not easily scalable and mutable. A fully fledged Blockchain network on the other hand produces a large overhead and adds unnecessary complexity for an application that is only used by a single user. QLDB is a mix between these two solutions. It is immutable and provides a verifiable history of all data changes, managed by a central authoring to scale as needed [83]. The third service that Amazon is providing in relation to Blockchain is a platform on which Amazon brings companies in need of Blockchain technology with trusted third party Blockchain partners together. These partners provide validated solutions for leveraging Blockchain and DLT services on AWS [84].

6.5.3.3 Pricing

Amazon's managed Blockchain systems are based on a pay-as-you-go pricing system with no upfront costs. It uses per-second billing and charges each member for their own resources and data it writes to the network. Based on the type of membership, service pricing rates are different and provide higher or lower transaction throughput and availability. For the Amazon managed Blockchain, there is a starter edition membership which is suited for test and small production networks. It has a lower transaction throughput and availability than the Standard Edition network but is also priced lower. The standard edition membership is best suited for production networks and has higher transaction throughput and availability than the Starter edition. However, pricing rates are also higher [85]. As a pricing example, following comparison of two two node networks is presented. For a two member test network, there is first of all the membership cost of \$0.30 per hour per user, (US East), the cost per node of \$0.034 per hour, the storage cost \$0.1 per GB-month and the data written cost 0.1 per GB.

For a two member test network, this leaves the following calculation that gives roughly an hourly cost of (2 Starter Edition members) x (\$0.30 per hour) + (2 members) x (1 bc.t3.small peer node per member) x (0.034 per hour) + (2 members) x (1 peer node per member) x (20 GiB storage per peer node) x (\$0.10 per GB-month) + (2 members) x (9 MB per hour) x (0.10 per GB) = \$0.676 per hour or \$487 per month. A production network running on a standard license has a member cost of \$0.55 per hour per member, \$1.23 per hour for two peer nodes, \$0.139 per hour storage cost for two nodes and \$0.01 data written cost. In total, a network like this costs its company 1 Standard Edition

member) x (\$0.55 per hour) + 1 member) x (2 m5.2xlarge peer nodes per member) x (0.615 per hour) + (1 member) x (2 peer node per member) x (500 GiB storage per peer node) x (\$0.10 per GB-month) + (1 member) x (100 MB per hour) x (0.10 per GB) = \$1.93 per hour or \$1390 per month. The prices do change depending on the region of the infrastructure. For example, the membership rate in the US East (N. Virginia) region is \$0.55 per hour, whereas in the Europe (Ireland) region it is \$0.62 per hour, a 11.3% difference. Also, this is just the pricing for the blockchain system. Often other Amazon services such as Amazon's Cloud9 or CloudFormation are used together with the Blockchain system and are also adding to its costs [86]. A summary of all pricing components can be seen in Figure 6.13.

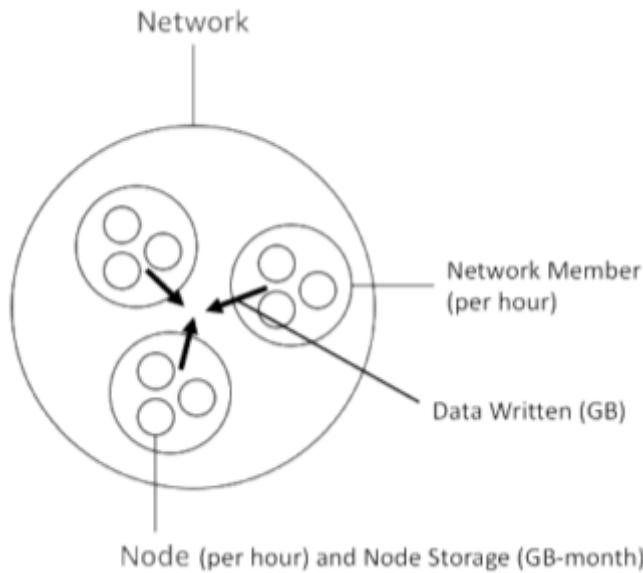


Figure 6.13: Pricing dimensions [85]

6.5.4 Dragonchain

6.5.4.1 Dragonchain

Dragonchain was developed at the Walt Disney company in 2014 for securing Walt Disney's internal data. The first multi-chain hybrid was created in 2016. In the same year, the platform was released as open-source software under Disney. Joe Roets founded in 2017 the commercial Dragonchain, Inc. in Seattle. On October 30th 2018, Dragonchain managed for the first time to connect a private and a public Blockchain through their patented Interchain technology and Dragonchain since then steadily improved the Interchain technology [95]. Today, about 56'000 transactions per month are performed on the Dragonchain infrastructure [88]. On LinkedIn, Dragonchain has 26 registered employees. Their headquarters is located in Bellevue, Washington [87]. On Growjo, Dragonchains revenue has been estimated to be around \$6 millions annually [89].

6.5.4.2 Product and Services

Dragonchain is an open-use Blockchain platform that provides developers and companies with a public/private Blockchain hybrid (see Figure 6.14) [90]. As services, Dragonchain provides a console with whom developers can set up decentralized applications, create smart contracts and manage their Blockchain networks [91]. Also, the infrastructure for running a Blockchain network and smart contracts is provided by Dragonchain [92]. In order to help developers to learn working with Dragonchain, they provide a complete

documentation about all important facets of the framework [96]. Lastly, Dragonchain also has a consulting team that helps users with product access, capabilities, methods, and tools [94].

One unique feature of Dragonchain is its patent on the Interchain technology, a tool that lets different Blockchains and traditional systems connect to each other in a secure manner. For example, Ethereum's Ethers can be transferred to a Dragonchain network [95].

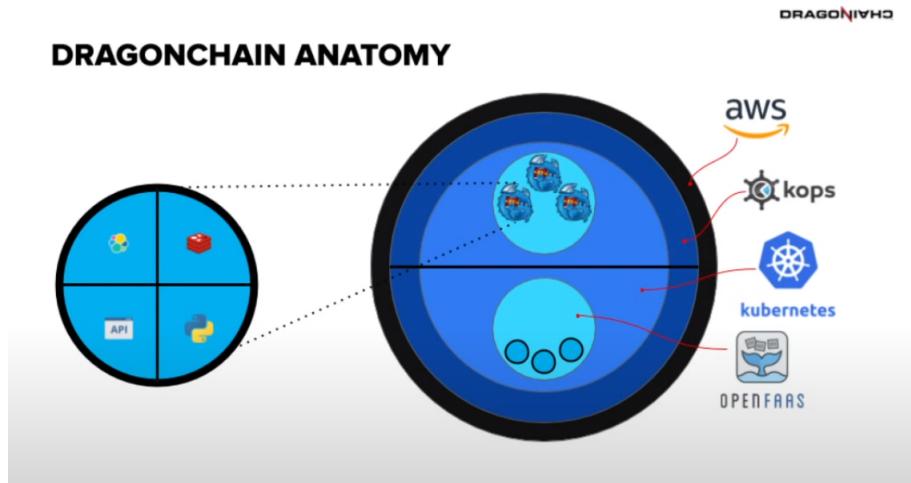


Figure 6.14: Dragonchains Architecture [92]

6.5.4.3 Pricing

One pillar of Dragonchains pricing is a subscription system. Buying a managed business node of type L1 for example costs \$99 per month. In addition, there are fees for each transaction of at least \$0.0000025 per transaction. Other costs include consultation fees or fees for services of third parties. A user can also create an unmanaged L1 node with no monthly costs. The services are billed monthly [96].

6.5.5 Alibaba Group

Alibaba Group is a corporation with different business models in the technology sectors. It was founded 1999 by Jack Ma in China and had an enormous growth since then [97]. This digital economy includes businesses such as AliExpress, Alibaba Cloud, ANTGROUP, Taobao, DingTalk and many more. Alibaba Cloud is the basis of their digital service and product successes. It was setup 2009 and celebrated together with their 10th anniversary [99]. Alibaba Group can be looked at as one of the largest IaaS provider by revenue as claimed by Gartner's April 2020 report [98]. Their annual revenue of 2019 is about \$56 Billion and \$3.6 Billion is solely due to their Cloud Computing services and products [100].

6.5.5.1 Products and services

Since Alibaba Cloud is the business sector that is responsible for their BaaS it has to be looked at more closely. The Alibaba Cloud offers a wide amount of cloud services worldwide such as elastic computing, database storage, network, machine learning platform and many more. Having a large range of 140 different products and services it comes that one of them is the BaaS [101]. The BaaS support three different types of Blockchain: Hyperledger Fabric, AntChain (Blockchain technologies of ANTGROUP) and Quorum. It

allows customers to quickly deploy Blockchains and manage nodes and smart contracts. It is mentioned that the service provides a CA certificate and is underlined by cryptographic algorithms that are suggested by China's authorities [102]. On their website the service comes along with user and development guides to have an easy time to deploy the preferred Blockchain [103]. One of the customer that is using Alibabas Baas is Trusple. In this case they are using Antchain, the Blockchain solution of ANT GROUP (member of the Alibaba digital economy) [105]. They are using it to have more efficient and secure transactions within the trading system across borders. In order to help developers to build up their Blockchain they provide them with a development guide in their website [106].

6.5.5.2 Pricing

The pricing method of Alibaba Cloud is subscription based, depending on what instance type will be used for the Blockchain (Hyperledger Fabric or Ant Blockchain). The Hyperledger Fabric is divided into five different categories as shown in the Figure 6.15. This table includes the monthly subscription cost among six different region supported such as China, Singapore, Australia, Japan, Germany and US [104]. It is indicated that the more storage space a customer needs the more the customer will pay as it is a fixed price for each storage package. Additionally to the monthly subscription, a customer has a pay-

Instance type	Instance specification	Free storage space	Monthly subscription
Test network	Starter Edition	50GB	US\$623
Consortium	Basic Edition	100GB	US\$1,078
Consortium	Enterprise Edition	500GB	US\$5,254
Organization	Basic Edition	100GB	US\$735
Organization	Enterprise Edition	500GB	US\$3,645

Figure 6.15: Alibabas BaaS pricing [104]

as-you-go billing that will charge him \$0.000472 per GB per hour if he needs any further storage to the limited free space included in the subscription. Alongside the Hyperledger Fabric there is also the Ant Blockchain option which is charged with a yearly subscription billing. There is only the Enterprise Edition specified which includes 4 nodes as well as 6 TB free storage space each. This sums up to a yearly cost of \$160'000 [104].

6.5.6 Oracle

Oracle was founded 1977 in Silicon Valley by Larry Ellison, Bob Miner and Ed Oates. It has concentrated on software development and later in 1987 it was one of the largest database management company with sales of over \$100 Million. Meanwhile Oracle products do not only include huge databases but ranges from infrastructure services, software and application for industry or cloud [107]. Oracles infrastructure data center are located all over the world as shown in Figure 6.16. It allows them to have an optimized support for any customer, supporting them with high-bandwidth interconnections and high security [110]. Having a total revenue of \$39.1 Billion in the fiscal 2020 it can be said that the company is well-established. Cloud services and license support had a total revenue



Figure 6.16: Oracle Server location [109]

of \$27.4 billion which indicates that they put in a lot of effort into Cloud Computing as it is one of their main revenue sources [111].

6.5.6.1 Services and products

Oracle Blockchain Platform Cloud Service is embedded in the Oracle Cloud infrastructure which is a collection of services managing various parts such storage, computation, networking, Cloud Database and many more [112]. Their cloud infrastructure provides more than 80 products [108]. According to their website, the BaaS is supporting Hyperledger Fabric platforms which either can be deployed on their cloud system or to join on another Hyperledger Fabric. The service comes with dynamic scaling of nodes and storage, an easy setup of the Blockchain network tutorial, Identity management and manage access permissions [112]. Alongside the services there is a developer community with a lot of resources for constructing such a Blockchain app and its possible to explore their Oracle Cloud Marketplace to observe other pre-built Blockchain solutions. As many other providers Oracle has numerous large customers using Blockchain for supply chains, trading or managing large amounts of transactions for their Blockchain network [112]. The following section provides some use cases of the Oracle Blockchain Platform.

- Circulor is a corporation that uses the Oracle Blockchain to support industries with new Industrie 4.0 technology. They concentrate on Traceability as a Service to enhance supply chains of various customers with Blockchain technology to better trace and secure customers material sources and traffic [113].
- INDETAIL is a software company that originates from Japan [114]. One of their main activities is to build up privat Blockchain solutions for customers. For their Platform they are using the Oracle Cloud Blockchain. Their goal is to integrate new technology such as Blockchain, AI and IoT into customers enterprises[115].
- Another supply chain tracing software developer is Retraced. They build for their customers supply chain solutions such that companies have an improved data connectivity and more manageable overview of all activities in the customers supply chain. They work with the Oracle Cloud infrastructure, including Oracle's Blockchain [116].

6.5.6.2 Pricing

The Oracle Blockchain Platform Cloud Service has an OCPU-based pricing model. OCPU or short OCU means having one physical core Intel Xeon. It is equal to two vCPUs or two threads[117]. It is stated that a user could use the provided Control Panel to either scale up or scale down the number of OCPU to reduce costs. As shown in Figure 6.17 the amount of storage a customer need will also be charged. The downscaling is limited to certained minimum threshhold that can not be passed. Oracle mentioned that the advantage of their BaaS is the dynamiycally scaling of OCU per hour which lead to a higher price-performance and information governance [112]. The Figure 6.17 shows that a customer either uses the Standard - Blockchain Platform Cloud Service or the Enterprise - Blockchain Platform Cloud Service product for either \$0.215 or \$0.4301 respectively [112]. Additionally, there is a fee for storage that gets calculated per TB storage per month for \$79.40.

Product	Unit price	Metric
Oracle Cloud Infrastructure - Blockchain Platform Cloud Service - Standard	\$0.215	OCU per hour
Oracle Cloud Infrastructure - Blockchain Platform Cloud Service - Enterprise	\$0.4301	OCU per hour
Oracle Cloud Infrastructure - Blockchain Platform Cloud Service - Storage	\$70.40	TB Storage Capacity per month

Figure 6.17: Oracle Server location [112]

6.6 Discussion

The Table 6.3 presents the collected data of the analysed providers in this report in a side-by-side view. It is divided into five sections that are considered to be important for the comparison. In the following section each property of the providers will be analysed in a respective manner.

6.6.1 Products and services

All providers, except Dragonchain, offer, besides their BaaS, various cloud products and services ranging from security, storage and networking to management systems and computation. These providers have a natural advantage over Dragonchain since they already have a huge amount of customers before they provided BaaS. The customers are more likely to use their BaaS since they are already using their cloud framework in the respective companies. It is more convenient to have all services work with each other in one interconnected framework. Of course all newly acquired customers of BaaS have the advantage of using other products and services of their respective provider. Dragonchains main focus is BaaS, they are certainly in their domain and do not have consider to interconnect or develop other products and services of their own. One drawback of the multi-service provider is also that if a customer once uses the infrastructure cloud it is not easy to change your provider without losing the benefit of the connected services.

6.6.2 Cloud Revenue

The leading cloud provider by revenue is Microsoft Azure with over \$48 billion followed by Oracle and IBM with \$27.4 and \$21.1 billion. In the midrange we have AWS and Alibaba with \$10.8 and \$3.6 Billion and at the end Dragonchain with \$6 million. It has to be considered that the revenues of the provider do not reflect the direct revenue of the BaaS. Except Dragonchain, it only gives us an indication on how well the cloud providers sell their cloud products. As it was stated in the report all BaaS providers have customers in the supply chain or trading sectors where it can be assumed that they need a lot of computation and storage power which leads to an increased revenue. Dragonchain with \$6 millions also shows that they may be one of the more niche and small providers among all.

6.6.3 Known Supported Frameworks

Hyperledger Fabric is almost supported by all providers as it seems to be one of the more popular frameworks. A vast amount of their customers are supply-chains as it seems to be the solution for many to have a private permissioned Blockchain network for tracing supplies. One can say that no BaaS provider stands out if we look at the use-cases. Having Ethereum also seems to be the state of art for almost all providers since it provides increasing value to support an open-source that manages data-values. As one of the smaller companies compared to its competitors, Dragonchain offers its Interchain technology that lets different Blockchain systems and traditional system connect to each other, which gives them an advantage over the other providers. Alibaba also stands out with its own Antchain network which of course get only supported by them. That gives them an unique lead over the others.

6.6.4 Pricing model

In a BaaS a customer pays for the infrastructure he gets from each provider. Only the way of pricing seems to differ between the providers. There seems to be 3 dominant types of pricing models. First off the pay-as-you-go per node and storage model from Microsoft Azure and AWS. Secondly the pay-as-you-go per CPU and storage used by IBM and Oracle. Thirdly the subscription-based model where a customer pays monthly for using their infrastructure from Dragonchain and Alibaba. In the next sections the different models get more deeply discussed.

6.6.4.1 Per node and storage model

The advantage of this model is that a customer will only pay for their nodes and storage they need. If a customer needs less nodes or more storage he can dynamically lower or upper the amount used such that a customer only has to pay for what they really need. The drawback is the following: if you need many nodes it will increase drastically even if he does not need that much of computing power (CPU) per node.

6.6.4.2 Per CPU and storage model

The per CPU and storage model has the advantage that a customer can increase the amount of nodes used without paying more. It depends on how much computing power each node needs. For example a customer can divide a CPU cluster to 4 nodes instead of one. As it comes to scaling it certainly can be that a node may need even more computation power over times, thus increasing the amount a customer has to pay where a per node model could come in more handy.

6.6.4.3 Subscription-based model

The advantage of a subscription-based model is that a customer might have a better overview on how much he will pay per month. A customer can pay for extra amounts used over the allocated subscribed amount in the case of Alibaba such it has some kind of dynamics to it. Speaking of dynamic, the obvious drawback of having a subscription-based model is that a customer may be paying too much for the amount he is actually using. It may also hold a customer back as the pricing will not dynamically grow with his Blockchain successes.

6.6.5 Service Kits

As it comes to Blockchain being a fairly new discovered technology, the knowledge of developers about the implementation of such may be limited. Because of that all selected providers have some sort of service kits that help customer to develop their Blockchain on the providers infrastructure. Microsoft Azure BaaS seems to have one of the more elaborated Service Kits among all providers with free source code and easy-to-use work environment to have a nice user experience. IBM has an interesting E-Book and tutorials which give the developers a nice introduction to Blockchain. Dragonchain and Alibaba have a slightly less developed service Kit but it serves its purpose.

6.7 Conclusion

BaaS is a new branch in the economy with tremendous potential. Blockchains give us trust, security, transparency or privacy when needed, and are a benefit to all honest parties involved. This did not remain unnoticed by major technology companies like Microsoft, IBM, Oracle or Amazon and they quickly introduced BaaS into their line of products. It also inspired new companies like Dragonchain to enter the market. In our report we provided an overview of Blockchain, Software as a System, introduced some of the most relevant providers and put them in comparison. With Dragonchain being the only exception all our providers have been in the technology market for a long time. The differences of the providers are displayed in revenue they generate, service kits, and pricing models which are differently calculated. Some providers charge per nodes, CPU usage, storage or a combination. Some providers offer a pay-as-you-go solution, others only offer a subscription package. In our report we came to the conclusion that for every use case and customer we could think of there is a suitable provider on the market.

6.8 Future Work

Since this paper can only provide a snapshot of the current situation of the BaaS market, it is important that the market situation also gets monitored in the future, so that new trends and novel applications of Blockchain technology can be detected and analysed in a timely manner. Secondly, a lot of Blockchain applications are being developed for the usage in Asian and African countries. Therefore, further research about Blockchain applications in these countries and providers that are operating in these regions would help to give a more holistic view of the current BaaS market. Thirdly, a lot of unseen use cases of Blockchain technology are developed in connection with BaaS. Hence, further research in connection to Blockchain related applications that have been specifically developed with BaaS might provide many novel applications use cases for Blockchain technology. Lastly, security and data privacy concerns seem to be a big issue when working with

Table 6.3: Provider Comparison

Providers\Properties	Microsoft Azure	AWS	IBM	Dragonchain	Alibaba	Oracle
Products & Services	200+ products, worldwide network	175+ products, worldwide network	170+ products, worldwide network	Only Blockchain	~140 products worldwide network	~80 cloud infrastructure products, worldwide network
Cloud Revenue	\$48 billion	\$10.8 billion	\$21.1 billion	\$6 million	\$3.6 billion	\$27.4 billion
Known Supported Frameworks	Consortium, Etherum	Hyperledger Fabrics, Ethereum	Hyperledger Fabrics	Interchain technology, Ethereum	Antchain, Hyperledger Fabrics, Quorum	Hyperledger Fabrics
Pricing model	pay-as-you-go: per validator/ transaction node	pay-as-you-go: per node, per storage, per data written	pay-as-you-go: per CPU hour, per storage	Subscription, transaction fee	Subscription, per storage	pas-as-you-go: per CPU hour
Service Kits	Blockchain Workbench, Blockchain Development Kit	Management Guide	E-Book, tutorials, technical brief	Documentation	Development Guide	Developer Community, Oracle Cloud Market place

Cloud Computing. Therefore it should be examined as well if BaaS suffers from the same data privacy and security challenges as Cloud Computing.

Bibliography

- [1] Forbes Technology Council, "13 Evolving And Emerging Uses For Blockchain Technology," *Forbes*, June 10, 2020. [Online]. Available: <https://www.forbes.com/sites/forbestechcouncil/2020/06/10/13-evolving-and-emerging-uses-for-Blockchain-technology/?sh=2f39fc90162e> [Accessed: Oct. 31, 2020]
- [2] I. H. Chuang, S. H. Li, K. C. Huang, Y. H. Kuo, "An effective privacy protection scheme for Cloud Computing," in Proc. 13th International Conference on Advanced Communication Technology (ICACT2011), pp. 260-265, 2011.
- [3] Q. Lu, X. Xu, Y. Liu, I. Weber, L. Zhu, W. Zhang "uBaaS: A unified BaaS platform. Future Generation Computer Systems," in *uBaaS: A unified BaaS platform. Future Generation Computer Systems*, vol. 101, pp. 564-575, 2019.
- [4] M. Nofer, P. Gomber, O. Hinz, et al. "Blockchain," in *Bus Inf Eng* 59, 183-187 (2017).
- [5] A. T. Sherman, F. Javani, H. Zhang and E. Golaszewski, "On the Origins and Variations of Blockchain Technologies," in *IEEE Security & Privacy*, vol. 17, no. 1 pp. 72-77, Jan.-Feb. 2019.
- [6] P. Franco, "Understanding Bitcoin," in *Cryptography, Engineering and Economics*, (John Wiley & Sons, Incorporated), 2014.
- [7] Researchgate, "Centralized vs Decentralized vs Distributed Networks," *Researchgate*, 2020. [Online] Available:[Accessed Nov. 5, 2020] https://www.researchgate.net/figure/Centralized-vs-Decentralized-vs-Distributed-Networks_fig1_316042146.
- [8] O. Konashevych, "Why 'Permissioned' and 'Private' are not Blockchains," in *SSRN Electronic Journal*, 2019.
- [9] E. A. Coskun, C. K. M. Lau, H. Kahyaoglu, "Uncertainty and herding behavior: evidence from cryptocurrencies," in *Research in International Business and Finance*, vol. 54, 2020.
- [10] I. Miciula, K. Kazojc, "The global development of cryptocurrencies," in *Research Papers of Wroclaw University of Economics*, vol. 63, 2, pp. 183-196, 2019.
- [11] T. Meitinger, "Smart Contracts," in *Informatik Spektrum* vol. 40, pp. 371-375, 2017.
- [12] R. Casado-Vara, J. Prieto, F. De la Prieta, J. M. Corchado, "How Blockchain improves the supply chain: case study alimentary supply chain," in *Procedia Computer Science* vol. 134, 2018.
- [13] V. Thiruchelvam, A. Shaka Mugisha, M. Shahpasand, M. Bamiah, "Journal of Telecommunication," in *Electronic and Computer Engineering*, (Asia Pacific University of Technology & Innovation Prince Sultan University), vol. 10, No. 3-2, 2018

- [14] Y. Chen, S. Ding, Z. Xu, et al, "Blockchain-Based Medical Records Secure Storage and Medical Service Framework," in *J Med Syst* vol. 43, 5, 2019.
- [15] K. Brousmichc, A. Anoaica, O. Dib, T. Abdellatif, G. Deleuze, "Blockchain Energy Market Place Evaluation: An Agent-Based Approach," in *IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, (Vancouver, BC), pp. 321-327, 2018.
- [16] V. J. Morkunas, J. Paschen, E. Boon, "How Blockchain technologies impact your business model," in *Business Horizons*, vol. 62, no. 3, pp. 295-306, 2019.
- [17] M. M. H. Onik and M. H. Miraz, "Performance Analytical Comparison of Blockchain-as-a-Service (BaaS) Platforms," in *Emerging Technologies in Computing* (M. H. Miraz, P. S. Excell, A. Ware, S. Soomro, and M. Ali, eds.), (Cham), pp. 3–18, Springer International Publishing, July 2019.
- [18] M. G. Avram, "Advantages and challenges of adopting Cloud Computing from an enterprise perspective," in *Procedia Technology*, vol. 12, pp. 529-534, 2014.
- [19] L. Wang, J. Tao, M. Kunze, A. C. Castellanos, D. Kramer, W. Karl: "Scientific Cloud Computing: Early definition and experience," in Proc. 2008 10th IEEE international conference on high performance computing and communications, pp. 825-830, 2008.
- [20] t. Dillon, C. Wu, E. Chang, E., "Cloud Computing: issues and challenges," in Proc. 2010 24th IEEE international conference on advanced information networking and applications, pp. 27-33. 2010.
- [21] P. Mell, T. Grance, "The NIST definition of Cloud Computing," NIST, NIST Special Publication 800-145, 2011.
- [22] D. Rani, R. K. Ranjan, "A comparative study of SaaS, PaaS and IaaS in Cloud Computing," in *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 4, no.6, 458-461. 2014.
- [23] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, M. Zaharia, "A view of Cloud Computing," in *Communications of the ACM*, vol. 53, no. 4, pp. 55-58, 2010.
- [24] Red Hat, "What is PaaS?," *Red Hat*, 2020. [Online]. Available at: <https://www.redhat.com/de/topics/cloud-computing/what-is-paas> [Accessed Oct. 15, 2020]
- [25] Jeb Su, "More Than 100 Million Consumer Personal Data Leaked After A Massive Cloud Breach At Capital One," *Forbes*, July 30, 2019. [Online]. Available: <https://www.forbes.com/sites/jeanbaptiste/2019/07/30/more-than-100-million-consumer-personal-data-leaked-after-a-massive-cloud-breach-at-capital-one/?sh=711641b54caa> [Accessed: Oct. 31, 2020]
- [26] J. Nicholas Hoover, "Outages Force Cloud Computing Users To Rethink Tactics," *InformationWeek*, August 16, 2008.
- [27] A. Stern, "Update from amazon regarding Friday's s3 downtime," *Centernetworks*, February 2008. [Online], Available at: <http://www.centernetworks.com/amazon-s3-downtime-update>. [Accessed Oct. 15, 2020]
- [28] W. Kim, "Cloud Computing: Today and tomorrow," in *J. Object Technol.*, vol. 8, no. 1, pp. 65-72, 2009.

- [29] W. Zheng, Z. Zheng, X. Chen, K. Dai, P. Li, R. Chen, "Nutbaas: A Blockchain-as-a-service platform," in *Ieee Access*, vol. 7, pp. 134422-134433, 2019.
- [30] J. Singh, J. D. Michels, "BaaS (BaaS): providers and trust," in Proc. 2018 IEEE Symposium on Security and Privacy Workshops, 2018, pp.67-74.
- [31] M. Onik, M. Miraz, "Performance Analytical Comparison of Blockchain-as-a-Service (BaaS) Platforms," in Proc. International Conference for Emerging Technologies in Computing, 2019, pp. 3-18.
- [32] M. Taylor, "The Evolution of Bitcoin Hardware," in *Computer*, vol. 50, no. 9, pp. 58-66, 2017.
- [33] E. F. Coutinho, D. E. Paulo, A. W. Abreu, I. B. Carla. "Towards Cloud Computing and Blockchain Integrated Applications," in Proc. 2020 IEEE International Conference on Software Architecture Companion, 2020, pp. 139-142.
- [34] Hyperledger, "Hyperledger - Open Source Blockchain Technologies," *The Linux Foundation*, 2020. [Online]. Available: <https://www.hyperledger.org/> [Accessed Oct. 15, 2020]
- [35] Enterprise Ethereum Alliance, "Enterprise Ethereum Alliance: Home," *Enterprise Ethereum Alliance (EEA)*, 2020. [Online]. Available: <https://entethalliance.org/> [Accessed Oct. 15, 2020]
- [36] Ethereum Foundation, "About the Ethereum Foundation," *Ethereum Foundation*, 2020. [Online]. Available: <https://ethereum.org/en/foundation/> [Accessed Oct. 15, 2020]
- [37] Y. Lu, "The Blockchain: State-of-the-art and research challenges," in *Journal of Industrial Information Integration*, vol. 15, pp. 80-90, 2019.
- [38] C. Roy, S. Misra, S. Pal, "Blockchain-Enabled Safety-as-a-Service for Industrial IoT Applications," in *IEEE Internet of Things Magazine*, vol. 3, no. 2, pp. 19-23, 2020.
- [39] A. Bahga, V. K. Madisetti, "Blockchain platform for industrial internet of things," in *Journal of Software Engineering and Applications*, vol. 9, no. 10, pp. 533-546, 2016.
- [40] M. Vukolic, "Behind the architecture of Hyperledger Fabric," *IBM*, 2020. [Online]. Available at <https://www.ibm.com/blogs/research/2018/02/architecture-hyperledger-fabric/> [Accessed Oct. 15, 2020]
- [41] Microsoft, "Facts about Microsoft," *Microsoft*, 2020. [Online]. Available: <https://news.microsoft.com/facts-about-microsoft/> [Accessed Oct. 25, 2020]
- [42] Janakiram MSV, "A Look Back At Ten Years Of Microsoft Azure," *Forbes*, 3. Feb. 2020. [Online]. Available at <https://www.forbes.com/sites/janakirammsv/2020/02/03/a-look-back-at-ten-years-of-microsoft-azure/#5aded56d4929> [Accessed Nov. 1, 2020]
- [43] Microsoft Azure, "What is Azure?," *Microsoft*, 2020. [Online]. Available at <https://azure.microsoft.com/en-gb/overview/what-is-azure/> [Accessed Nov. 1, 2020]
- [44] Microsoft Azure, "Microsoft global network," *Microsoft*, Jun. 13. 2019. [Online]. Available at <https://docs.microsoft.com/en-gb/azure/networking/microsoft-global-network> [Accessed Nov. 2, 2020]

- [45] Microsoft Azure, "microsoft-global-wan," *Microsoft*, 2020. [Online]. Available at: <https://docs.microsoft.com/de-de/azure/networking/media/microsoft-global-network/microsoft-global-wan.png> [Accessed Nov. 2, 2020]
- [46] Microsoft, "Earnings Release FY20 Q4," *Microsoft*, Jul. 22. 2020. [Online]. Available at <https://www.microsoft.com/en-us/Investor/earnings/FY-2020-Q4/press-release-webcast> [Accessed Nov. 2, 2020]
- [47] Frederic Lardinois, "Microsoft launches a fully managed Blockchain service," *techcrunch*, May 2. 2019. [Online]. Available at <https://techcrunch.com/2019/05/02/microsoft-launches-a-fully-managed-Blockchain-service/> [Accessed Nov. 4, 2020]
- [48] Microsoft, "What is Azure Blockchain Workbench?," *Microsoft*, May 22. 2020. [Online]. Available at <https://docs.microsoft.com/en-gb/azure/Blockchain/workbench/overview> [Accessed Nov. 5, 2020]
- [49] Microsoft, "Azure Blockchain Development Kit," *Microsoft*, Jul. 5. 2020. [Online]. Available: <https://docs.microsoft.com/en-gb/samples/azure-samples/Blockchain-devkit/azure-Blockchain-development-kit/> [Accessed Oct. 25, 2020]
- [50] Microsoft Azure, "Azure Blockchain Service," *Microsoft*, 2020. [Online]. Available: <https://azure.microsoft.com/en-gb/services/Blockchain-service/> [Accessed Oct. 28, 2020]
- [51] Microsoft Azure, "Strengthen your security posture with Azure," *Microsoft*, 2020. [Online]. Available: <https://azure.microsoft.com/en-gb/overview/security/> [Accessed Oct. 27, 2020]
- [52] Jennifer Warnick, "Knowledge is valuable: Coffee journey going digital for customers, farmers," *Starbucks*, Mar. 20. 2019. [Online]. Available: <https://stories.starbucks.com/stories/2019/knowledge-is-valuable-coffee-journey-going-digital-for-customers-farmers/> [Accessed Oct. 28, 2020]
- [53] Microsoft, "Singapore Airlines transforms customer loyalty with Blockchain on Azure," *Microsoft*, May 2. 2019. [Online]. Available: <https://customers.microsoft.com/en-us/story/singapore-airlines-travel-transportation-azure> [Accessed Oct. 27, 2020]
- [54] Microsoft, "GE Aviation's Digital Group streamlines tracking of aircraft parts, reduces inefficiencies with Azure Blockchain technologies," *Microsoft*, Nov. 2. 2019. [Online]. Available: <https://customers.microsoft.com/en-gb/story/755328-ge-aviation-manufacturing-azure> [Accessed Oct. 29, 2020]
- [55] Microsoft Azure, "Azure Blockchain Service pricing," *Microsoft*, 2020. [Online]. Available: <https://azure.microsoft.com/en-us/pricing/details/Blockchain-service/> [Accessed Nov. 2, 2020]
- [56] Microsoft Azure, "Pricing calculator," *Microsoft*, 2020. [Online]. Available: <https://azure.microsoft.com/en-gb/pricing/calculator/?service=Blockchain-service> [Accessed Nov. 2, 2020]
- [57] IBM, "Chronological History of IBM 1910s," *IBM*, 2020. [Online]. Available: https://www.ibm.com/ibm/history/history/decade_1910.html [Accessed Nov. 4, 2020]
- [58] IBM, "Chronological History of IBM 1920s," *IBM*, 2020. [Online]. Available: https://www.ibm.com/ibm/history/history/decade_1920.html [Accessed Nov. 4, 2020]

- [59] IBM, "Chronological History of IBM 1940s," *IBM*, 2020. [Online]. Available: https://www.ibm.com/ibm/history/history/decade_1940.html [Accessed Nov. 4, 2020]
- [60] IBM, "About IBM," *IBM*, 2020. [Online]. Available: <https://www.ibm.com/ibm/us/en/> [Accessed Nov. 5, 2020]
- [61] IBM, "IBM Reports 2019 Fourth-Quarter and Full-Year Results," *IBM*, 2020. [Online]. Available: <https://newsroom.ibm.com/2020-01-21-IBM-Reports-2019-Fourth-Quarter-and-Full-Year-Results> [Accessed Nov. 5, 2020]
- [62] IBM, "IBM Cloud products," *IBM*, 2020. [Online]. Available: <https://www.ibm.com/uk-en/cloud/products> [Accessed Nov. 1, 2020]
- [63] IBM, "IBM Blockchain Platform - Developer," *IBM*, 2020. [Online]. Available: <https://www.ibm.com/uk-en/cloud/Blockchain-platform/developer> [Accessed Nov. 2, 2020]
- [64] IBM, "Getting started with IBM Blockchain Platform," *IBM*, Oct. 15. 2020. [Online]. Available: <https://cloud.ibm.com/docs/Blockchain> [Accessed Nov. 3, 2020]
- [65] IBM, "Nordea leverages IBM technology to enable smooth and secure trading for its customers," *IBM*, 2020. [Online]. Available: <https://www.ibm.com/Blockchain/use-cases/success-stories/#section-1> [Accessed Nov. 6, 2020]
- [66] IBM, "Kroger uses IBM Blockchain technology for farm to fork food traceability," *IBM*, 2020. [Online]. Available: <https://www.ibm.com/Blockchain/use-cases/success-stories/#section-3> [Accessed Nov. 6, 2020]
- [67] IBM, "Revolutionizing recycling by creating an ecosystem for plastic," *IBM*, 2020. [Online]. Available: <https://www.ibm.com/Blockchain/use-cases/success-stories/#section-5> [Accessed Nov. 6, 2020]
- [68] IBM, "Pricing for IBM Blockchain Platform for IBM Cloud," *IBM*, Feb. 11. 2020. [Online]. Available: <https://cloud.ibm.com/docs/Blockchain?topic=Blockchain-ipb-saas-pricing> [Accessed Nov. 6, 2020]
- [69] IBM, "Hyperledger Fabric: the flexible Blockchain framework that's changing the business world," *IBM*, 2020. [Online]. Available: <https://www.ibm.com/Blockchain/hyperledger> [Accessed Nov. 11, 2020]
- [70] IBM, "IBM Cloud File Storage: Pricing," *IBM*, 2020. [Online]. Available: <https://www.ibm.com/cloud/file-storage/pricing> [Accessed Nov. 7, 2020]
- [71] IBM, "Detailed pricing scenarios," *IBM*, Feb. 10. 2020. [Online]. Available: <https://cloud.ibm.com/docs/Blockchain?topic=Blockchain-ipb-detailed-pricing> [Accessed Nov. 7, 2020]
- [72] M. Hall, "Amazon.com," in /em Encyclopaedia Britannica, 9. Apr. 2020. [Online]. Available: <https://www.britannica.com/topic/Amazoncom> [Accessed Oct. 15, 2020]
- [73] Amazon Web Services, "Cloud Products," *Amazon Web Services*, 2020. [Online]. Available: <https://aws.amazon.com/products/?pg=WIAWS-mstf> [Accessed Oct. 15, 2020]
- [74] Amazon, "Annual Report 2019," *Amazon.com, Inc.*, 30. Jan. 2020. [Online]. Available: <https://ir.aboutamazon.com/annual-reports-proxies-and-shareholder-letters/default.aspx> [Accessed Oct. 15, 2020]

- [75] Amazon, "Amazon.com announces second quarter results," *Amazon.com, Inc*, 30. Jul. 2020. [Online]. Available: <https://ir.aboutamazon.com/annual-reports-proxies-and-shareholder-letters/default.aspx> [Accessed Oct. 15, 2020]
- [76] K. Stalcup, "Inside Amazon Web Services: AWS By The Numbers," *Emerald X*, 25. Jul. 2019. [Online]. Available: <https://mytechdecisions.com/it-infrastructure/inside-amazon-web-services-aws-by-the-numbers/> [Accessed Oct. 15, 2020]
- [77] Amazon Web Services, "AWS re:Invent 2018 - Announcement of Amazon Managed Blockchain," *Youtube.com*, para. 13, Dec. 2018. [Online]. Available: <https://www.youtube.com/watch?v=gCZfJWkfqLo&feature=youtu.be> [Accessed Oct. 15, 2020]
- [78] Amazon Web Services, "AWS Partner Network," *Amazon Web Services*, 2020. [Online]. Available: https://aws.amazon.com/partners/?nc1=h_ls [Accessed Oct. 15, 2020]
- [79] M. Nystrom, "Everything Enterprises Need to Know About Amazon's Blockchain as a Service," *ConsenSys*, para. 7, Jan. 2019. [Online]. Available: <https://media.consensys.net/everything-enterprises-need-to-know-about-amazons-Blockchain-as-a-service-9bd740e09276> [Accessed Oct. 15, 2020]
- [80] Amazon Web Services, "Amazon Managed Blockchain," *Amazon Web Services*, 2020. [Online]. Available: https://aws.amazon.com/managed-Blockchain/?nc1=h_ls [Accessed Oct. 15, 2020]
- [81] Amazon Web Services, "Amazon Managed Blockchain," *Amazon Web Services*, 2020. [Online]. Available: <https://aws.amazon.com/de/managed-Blockchain/resources/?nc=sn&loc=5> [Accessed Oct. 15, 2020]
- [82] Amazon Web Services, "Amazon Managed Blockchain Features," *Amazon Web Services*, 2020. [Online]. Available: <https://aws.amazon.com/managed-Blockchain/features/> [Accessed Oct. 15, 2020]
- [83] Amazon Web Services, "AWS QLDB Overview: Animated Ex- plainer Video," *Youtube.com*, para. 13, Dec. 2018. [Online]. Available: https://www.youtube.com/watch?v=jcZ<_rsLJrqk&feature=youtu.be [Accessed Oct. 15, 2020]
- [84] Amazon Web Services, "APN Spotlight: Blockchain Parnters," *Amazon Web Services*, 2020. [Online]. Available: https://aws.amazon.com/partners/spotlights/Blockchain-partner-spotlight/?nc1=h_ls [Accessed Oct. 15, 2020]
- [85] N. Modi, S. Rao, K. Alusi, "Amazon Managed Blockchain - Deep Dive," *AWS re:Invent*, 2019. [Online]. Available: https://d1.awsstatic.com/events/reinvent/2019/Enterprise_blockchain_AWS%27s_open-source_approach_OPN217.pdf [Accessed Oct. 15, 2020]
- [86] Amazon Web Services, "Amazon Managed Blockchain pricing," *Amazon Web Services*, 2020. [Online]. Available: <https://aws.amazon.com/managed-Blockchain/pricing/> [Accessed Oct. 15, 2020]
- [87] Dragonchain, "Dragonchain," *LinkedIn*, 2020. [Online]. Available: <https://www.linkedin.com/company/dragonchain/about/> [Accessed Oct. 15, 2020]
- [88] Dragonchain, "Metrics," *Dragonchain*, 18. Aug. 2020. [Online]. Available: <https://metrics.dragonchain.com/> [Accessed Oct. 15, 2020]

- [89] Growjon, "Dragonchain Competitors, Revenue and Alternatives," *Growjo*, 2020. [Online]. Available: <https://growjo.com/company/Dragonchain> [Accessed Oct. 15, 2020]
- [90] M. Cavicchioli, "What is Dragonchain," *the Cryptonomist*, 10. Jun. 2020. [Online]. Available: <https://en.cryptonomist.ch/2020/06/10/what-is-dragonchain/> [Accessed Oct. 15, 2020]
- [91] Dragonchain, "A flexible API along with official SDKs," *Dragonchain*, 2020. [Online]. Available: <https://dragonchain.com/developers> [Accessed Oct. 15, 2020]
- [92] Dragonchain, "Blockchain as a Service at Scale for Enterprise," *Youtube.com*, para. 14, Jun. 2019. [Online]. Available: <https://www.youtube.com/watch?v=ef4SF25v9ys> [Accessed Oct. 15, 2020]
- [93] Dragonchain, "Dragonchain Documentation," *Dragonchain*, 2020. [Online]. Available: <https://dragonchain-core-docs.dragonchain.com/latest/index.html> [Accessed Oct. 15, 2020]
- [94] Dragonchain, "Enterprise Ready, Startup Friendly," *Dragonchain*, 2020. [Online]. Available: <https://dragonchain.com/business> [Accessed Oct. 15, 2020]
- [95] Dragonchain, "Interchain Between Blockchains and Traditional Systems," *Dragonchain*, 18. Aug. 2020. [Online]. Available: <https://dragonchain.com/blog/interchain> [Accessed Oct. 15, 2020]
- [96] Dragonchain, "Add Business Node Subscription," *Dragonchain*, 2020. [Online]. Available: <https://console.dragonchain.com/chains/new/managed> [Accessed Oct. 15, 2020]
- [97] Alibaba Group, "COMPANY OVERVIEW," *Alibaba Group*, 2020. [Online]. Available: <https://www.alibabagroup.com/en/about/overview> [Accessed Nov. 8, 2020]
- [98] Alibaba Group, "OUR BUSINESSES," *Alibaba Group*, 2020. [Online]. Available: <https://www.alibabagroup.com/en/about/businesses> [Accessed Nov. 8, 2020]
- [99] Alibaba Group, "HISTORY AND MILESTONES," *Alibaba Group*, 2020. [Online]. Available: <https://www.alibabagroup.com/en/about/history?year=2009> [Accessed Nov. 8, 2020]
- [100] Alibaba Group, "Alibaba Group Announces March Quarter and Full Fiscal Year 2019 Results," *Alibaba Group*, May 15. 2019. [Online]. Available: <https://www.alibabagroup.com/en/news/press-pdf/p190515.pdf> [Accessed Nov. 8, 2020]
- [101] Alibaba Cloud, "Alibaba Cloud Products & Services," *Alibaba Cloud*, 2020. [Online]. Available: <https://www.alibabacloud.com/product> [Accessed Nov. 9, 2020]
- [102] Alibaba Cloud, "BaaS," *Alibaba Cloud*, 2020. [Online]. Available: <https://www.alibabacloud.com/products/baas> [Accessed Nov. 9, 2020]
- [103] Alibaba Cloud, "Quick Start," *Alibaba Cloud*, Aug. 7. 2020. [Online]. Available: <https://www.alibabacloud.com/help/doc-detail/85649.htm> [Accessed Nov. 9, 2020]
- [104] Alibaba Cloud, "BaaS," *Alibaba Cloud*, 2020. [Online]. Available: <https://www.alibabacloud.com/product/baas/pricing> [Accessed Nov. 9, 2020]
- [105] Trusple, "Trade with Payment Security and Inclusive Financial Services," *Trusple*, 2020. [Online]. Available: <https://www.trusple.com/> [Accessed Nov. 9, 2020]

- [106] Alibaba Cloud, "BaaS," *Alibaba Cloud*, 2020. [Online]. Available: <https://www.alibabacloud.com/help/product/84950.htm> [Accessed Nov. 10, 2020]
- [107] Oracle, "A history of possibilities," *Oracle*, 2020. [Online]. Available: <https://www.oracle.com/uk/corporate/> [Accessed Nov. 9, 2020]
- [108] Oracle, "Oracle Cloud Infrastructure products," *Oracle*, 2020. [Online]. Available: <https://www.oracle.com/uk/cloud/> [Accessed Nov. 10, 2020]
- [109] Oracle, "oci-region-map," *Oracle*, 2020. [Online]. Available: <https://www.oracle.com/a/ocom/img/cc01-oci-region-map-1075x450.png> [Accessed Nov. 9, 2020]
- [110] Oracle, "Oracle Cloud Infrastructure Data Center Regions," *Oracle*, 2020. [Online]. Available: <https://www.oracle.com/uk/cloud/architecture-and-regions.html> [Accessed Nov. 8, 2020]
- [111] Oracle, "Oracle Announces Fiscal 2020 Fourth Quarter and Fiscal Full Year Financial Results," *Oracle*, Jun. 16. 2020. [Online]. Available: <https://investor.oracle.com/investor-news/news-details/2020/Oracle-Announces-Fiscal-2020-Fourth-Quarter-and-Fiscal-Full-Year-Financial-Results/default.aspx> [Accessed Nov. 7, 2020]
- [112] Oracle, "Oracle Blockchain Platform Cloud Service," *Oracle*, 2020. [Online]. Available: <https://www.oracle.com/application-development/cloud-services/Blockchain-platform/> [Accessed Nov. 6, 2020]
- [113] Circulor, "Circulor is transforming the supply chains of various industries with Industry 4.0 tech," *Circulor*, 2020. [Online]. Available: <https://www.circulor.com/technology> [Accessed Nov. 6, 2020]
- [114] INDETAIL, "Company Information," *INDETAIL*, 2020. [Online]. Available: <https://www.indetail.io/company/> [Accessed Nov. 5, 2020]
- [115] INDETAIL, "Technology," *INDETAIL*, 2020. [Online]. Available: <https://www.indetail.io/technology/> [Accessed Nov. 5, 2020]
- [116] Retraced, Start Your Transparency Journey., "Retraced, 2020. [Online]. Available: <https://retraced.co/en> [Accessed Nov. 5, 2020]
- [117] Retraced, "Oracle Cloud Infrastructure (OCI), IaaS from Oracle," *Retraced*, Jan. 17. 2018. [Online]. Available: <https://acevedoapps.com/en/oracle-cloud-infrastructure-oci-iaas-from-oracle/> [Accessed Nov. 6, 2020]
- [118] E. Bellini, P. Ceravolo, and E. Damiani, "Blockchain-Based E-Vote-as-a-Service," in *IEEE 12th International Conference on Cloud Computing (CLOUD 2019)*, (Milan, Italy), pp. 484–486, July 2019.

Chapter 7

Public Proof-of-Stake Blockchains From an Operational Perspective

Bill Bosshard, Simon Bachmann

Blockchain (BC) has been an emerging technology in recent years. With the rise of Bitcoin there has been a lot of critique for the energy hungry Proof-of-Work (PoW) consensus algorithm. There has been a trend in distributed ledger technology (DLT) where existing BC either transition to Proof-of-Stake (PoS) or new BC are developed with the environment in mind. This report highlights the intended improvements over PoW. Furthermore, the requirements and the security risks for operating a PoS validator node is stated based on the analysis of six different PoS protocols. Multiple strategies for managing the private keys are discussed including on-premise and cloud-based hardware security modules (HSM). Key management services and smart contract wallets are cheaper alternatives to HSM. Also, different architectural setups are proposed, highlighting their strengths and weaknesses. The economics of each setup is analyzed and evaluated. There is no perfect fit for every validator. The recommended validator setup depends on the characteristics of the consensus mechanism such as the slashing penalties, the initial staking amount and if the node functions as a delegator node, where the server up-time is of importance.

Contents

7.1	Introduction	.	.	.	207
7.1.1	Consensus Mechanisms	.	.	.	207
7.2	Comparison of Different PoS Architectures	.	.	.	211
7.3	Security Considerations	.	.	.	214
7.3.1	Key Management	.	.	.	214
7.3.2	Supply Chain Attacks	.	.	.	214
7.3.3	Distributed Denial-Of-Service (DDoS) Attacks	.	.	.	216
7.4	Architectural Considerations	.	.	.	216
7.4.1	Sentry Node Architecture (SNA)	.	.	.	216
7.4.2	Possible Validator Configurations	.	.	.	217
7.4.3	Cryptographic Key Management	.	.	.	219
7.5	Economics	.	.	.	221
7.6	Centralization	.	.	.	224
7.7	Summary	.	.	.	225

7.1 Introduction

With the rise of Bitcoin in 2009 there has been a lot of critique for the energy intensive Proof-of-Work (PoW) consensus algorithm. There has been a trend in distributed ledger technology (DLT) where existing BC either transition to Proof-of-Stake (PoS) or new BC are developed with the environment in mind. This report highlights the intended improvements over PoW, such as the scalability of PoS and the environmental friendliness. These improvements are achieved by replacing the cryptographic puzzle of PoW with other consensus mechanisms, which are computationally less expensive [48]. This report gives an overview over different PoS implementations and highlights their key features.

PoS introduces the concept of validator, which are nodes in the network that fulfill the same role as the miners in PoW. They validate transactions and append new blocks. This report investigates different PoS ecosystems and compares various metrics such as minimum stake, slashing risks and expected annual yield (APY). The comparison highlights potential risks and benefits for validators for each BC ecosystem.

Furthermore, the requirements and the security risks for operating a PoS validator node are explained. One big risk for the operators of validator nodes, are Distributed-Denial-Of-Service Attacks (DDoS), since unresponsiveness can lead to punishments [53]. The threat of a DDoS attack is not unique to network participants in PoS. However, the financial penalty that has to be paid when being offline makes it more pressing in a PoS environment. This report provides an introduction to the sentry node architecture, which is a common hardware and software setup to mitigate DDoS attacks. Furthermore, a range of different installations is covered. From a basic single validator node to a more complex multi-active node setups and their possible advantages and disadvantages are part of this report. Further, the economics of each setup are analyzed showcasing the estimated costs and expected revenue.

An important topic in BC technology is the key management. The private key of a validator is the most sensitive part of the installation and requires additional attention [52]. Four key management strategies are covered in this report. HSM execute cryptographic operations, such as encrypting, decrypting data without exposing sensitive keys, even with physical access to the device [40]. HSM have become a standard for financial service providers. However, adding an HSM to a setup also means having an additional interface that may be malfunctioning. HSM are often integrated as black boxes with only knowing the interface specification. [54]

In recent years there has been a rise in cloud solutions. Since some of the controls and responsibilities are managed by cloud provider, setting up a cloud based validator installation requires additional attention with regards to the key management. This report discusses possible solutions in cloud such as cloud HSM and key management services (KMS) including their strength and weaknesses. On-premise as well as cloud HSM are expensive. Therefore, cheaper solutions for protecting a validator such as smart contract wallets are explained.

7.1.1 Consensus Mechanisms

This section covers the fundamentals for understanding the basic concepts behind the different kinds of consensus algorithms that are used in BC architectures. Furthermore, the advantages and challenges of each type is discussed.

A consensus mechanism is a fault-tolerant algorithm to achieve a mutually agreed state among the majority of nodes in the network [1]. For crypto-currencies such as Bitcoin, the data that is stored in all the nodes is a ledger that keeps track of all the transactions. However, it is not limited to a ledger. Second-generation BC such as Ethereum allow storing computer code in the nodes, and effectively enable more complex data structures

and programmatic rules. [2] These so-called Smart Contracts (SC) event-driven scripts that modify the state of the BC according to the terms defined in the contract. If the majority of the network agrees upon the execution, the SC can be executed. The objectives of SC are to reduce the need of trusted third parties, to automate processes with less enforcement costs and less opportunities for fraud. [45]

Unlike in centralized client-server systems, there is no authority with special privileges for accessing or manipulating the data in a BC. Every node in the network maintains its own copy of the ledger. Updates to the ledger can only be made if the rules of the consensus mechanism are followed. For example the consensus mechanism determines under which circumstances new data can be appended to the ledger.

The problem in a multi-agent network is that a node may fail or may be unreliable. Thus, a consensus mechanism in a decentralized system must be fault tolerant. In order to achieve consensus without the need of a trusted third party, cryptoeconomics play an important role [3]. By incentivizing nodes to participate in the consensus mechanism and help securing the network, they are financially compensated. Cryptoeconomics combine cryptography and economics to establish consensus and create robust decentralized P2P networks. The benefits of a cryptoeconomics-driven P2P network is the fault-tolerance (1), the attack resistance (2) and collusion resilience (3). Such systems are less likely to accidentally fail because they the execution and verification is replicated among many nodes (1). Also, there is no central point of failure making an attack more expensive to execute (2). Lastly, decentralized governance of a protocol helps to reduce the collisions of individuals that only benefit a minority (3).

7.1.1.1 Proof of Work (PoW)

PoW was the first consensus mechanism in BC technology that was introduced with Bitcoin in 2009 [46]. It requires a proof that the work performed by a node qualifies it to receive the authority to add new transactions to the ledger. In order to qualify, one must come up with a number in a cryptographic puzzle which can only be found by brute-forcing. With more computing power a node can increase its chances to find this number and be compensated with the mining reward. The difficulty of the cryptographic puzzle is constantly updated based on the amount of combined hashing power in the network. The nodes that actively participate in the transaction validation process are referred to miners. The advantage of PoW is its simplistic design of finding a random node in the network. Also, it has proven to be secure for more than eleven years for Bitcoin. However, the environmental harm of Bitcoin is tremendous. By the time of writing this paper, Digiconomist [8] reports that the Bitcoin BC consumes more than 74 TWh annually which is more than the entire population and industry of Austria.

The total hashing power has risen to a level such that an individual miner may never mine a block. Nodes that collaboratively solve the cryptographic puzzle, share the block reward and the transaction fees form so-called mining pools. These formations of collaborative mining is considered as a threat of centralization. More than 80% of the hashing power of the 20 largest mining pools comes from China. Furthermore, the three largest mining pools have combined more than 51% of the total hashing power. [19] These pools together can take over the network since they could always produce a longer chain. This is the so-called 51% attack.

Economies of scale are also one of the reason why PoW system become more centralized. For an individual miner it is difficult to compete with large mining farms. These mining farms either develop proprietary hardware or buy specialized hardware in large quantities. This effectively forces individual miners out of the market as they cannot operate at a similar price level [47].

7.1.1.2 Proof of Stake (PoS)

The nodes that validate transactions in a PoS system are called validators. Most PoS implementations require a node to deposit or lock some minimum amount of the underlying crypto currency in order to become a validator. Proportionally to the stake that is locked up, a node will be elected to propose a new block. Other validators are requested to check the authenticity of the proposed block and communicate the result with the other peers in the network. If the majority agrees the new block is added to the ledger [48].

An advantage of PoS is that the validation process does not include the cryptographic puzzle solving as it does in PoW. Therefore, validators must no be compensated for the burned electricity as the miners do. Lower costs for operating the network result in lower transaction fees. Also, the more energy efficient nature of PoS reduces the environmental impact compared to PoW. No specialized hardware is needed in the validation process. Thus, Economies of scale do not give large validator the same advantage as in PoW.

However, validators must make sure to be online during most of the time and actively participate in the consensus algorithm. Otherwise, the deposit may be at risk. Thus, a different hardware setup is required for operating a node in PoS. Ensuring that a node is online most of the time is a challenge and requires technical know how. How such a setup may look like and the economics behind it are discussed later in this paper. There exist many different types of PoS implementations. These different concepts are described in the following paragraphs and some of implementations are discussed in more detail in Section 7.2.

Delegated Proof of Stake (DPoS) In this type of consensus algorithm the token holders are able to cast votes for electing a subset of trustworthy nodes. These selected nodes are responsible for securing the network, validate transactions and establishing consensus for the entire network. The number of votes that a node is allowed to cast depends on the stake in the network. The token holders are responsible to elect trustworthy validators [49].

The benefit of such a delegation system is that the system can run more efficiently since the consensus is established among fewer nodes. However, the disadvantage is that token holders have to trust that the elected validator act truthfully and can setup a system that is resilient against attacks and system outages. Some implementations of DPoS also punish nodes that delegate their stake as shown in Section 7.2.

Pure PoS (PPoS) In this type of consensus protocol every token has the same chance of being selected as a validator. Most other flavours of PoS require tokens to be deposited before they are considered as stake in the network. Other PoS implementations also have different lengths of withdrawal periods when a validator wants to leave the set of validators. Both of these characteristics do not apply to PPoS since no stake is ever held hostage. [10] PPoS is constructed under the assumption that a super majority is always online for establishing consensus.

Liquid PoS (LPoS) The LPoS consensus functions similar to the previously discussed DPoS. The system of LPoS allows token owner to optionally delegate their voting and endorsing rights to other participants, but still keep the ownership of the tokens. Furthermore stakeholders are not punished for mistakes or malicious behaviour of their delegates. The punishment should incentives the delegates to honest behaviour. The major difference between LPoS and DPoS is that delegation is optional in LPoS the consensus algorithm does not require delegation to run. It has a dynamic validator set and allows token holders to take part in the governance. LPoS was first introduced with the Tezos BC [5].

Bonded Proof-of-Stake (BPoS) BPoS is similar to LPoS. The consensus algorithm allows delegation but it is optional and the system could run without it. The stakeholder benefits from voting rights and the possibility to delegate their stake to one or more multiple delegates. Inactivity and malicious behaviour gets punished by slashing a part of the bonded stake, this affects not only the offending validators, all his delegators are slashed too. Delegators need to make a careful choice, which validator they trust and want to delegate their stake to. BPoS was first introduced in projects such as Cosmos and IRISnet [6].

7.1.1.3 Liveness vs. Security

Every consensus protocol is designed to either favour liveness (availability) or security (finality). If the network favours security, it may halt at some point when a majority of the validators are offline. However, it guarantees that newly added block is approved by the majority of validators. This results in instant finality as soon as the block is published since the network never soft-forks. It can be considered final since otherwise a majority of the validators' deposits is lost.

On the other hand, if the system favours liveness, the system can continue to run as long as there is at least one validator online. Thus, the network remains available. However, with only one validator online an attack on the system becomes more feasible. Furthermore, it is possible that multiple valid chains (soft-forks) coexist and the protocol must contain a rule on how the network comes to consensus in such a scenario. Therefore, a block cannot be considered as final at the moment when it is published [50].

7.1.1.4 Nothing at Stake Problem

One common challenge that liveness favoured PoS consensus algorithms must solve is the nothing at stake problem. In liveness-favoured protocols 7.1.1.3, it is possible that two blocks are added to the chain before both are fully propagated through the network. This scenario is called a soft-fork. The protocol must define how the network comes to consensus even if there are multiple valid chains proposed.

Unlike in PoW, there are no resources burned in the process of adding a new block. This allows the validator to append a new block on any of the competing chains without additional costs. Therefore, it is in the validators' incentive to build on top of every competing chain. This strategy assures that the suggested block is included in any of the forks. Figure 7.1 illustrates that adding a block to both competing chains results in the highest expected value EV . In the context of PoS, probability p is determined by the stake that voted for that particular chain.

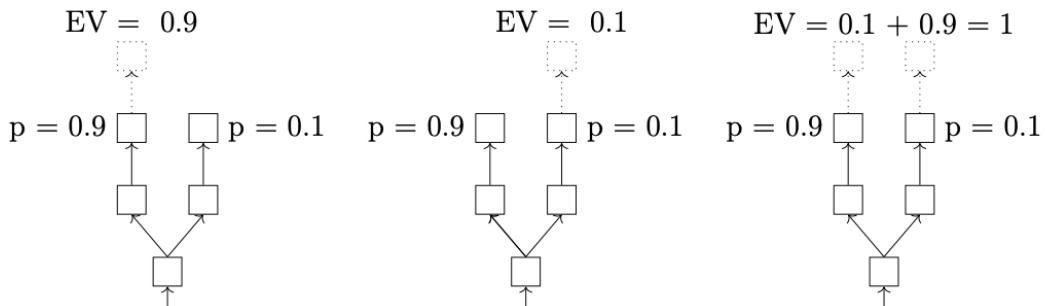


Figure 7.1: PoS: Nothing at stake problem in a soft fork [27]

Since validators want to maximize their returns, they append new block to any of the forks. By doing so, the BC will never reach consensus, even if there are only honest validators. Thus, validators must be forced by the protocol to only propose new blocks

on one of the forks. The mechanism of punishing validators that sign blocks on multiple chains is called slashing and is discussed in Section 7.1.1.5.

Unlike in PoS, miners must dedicate their resources to one chain, since the previous hash is included in the cryptographic mining puzzle. These resources are burned and cannot be used for solving the puzzle on both competing chains as illustrated in Figure 7.2. In the context of PoW, the probability p of a chain depends on the fraction of the total hashing power that is put into each chain. If the chains are of the same length, the percentage of mining node The the percentage of mining nodes that received this chain prior to the any other chain influences the likelihood p as well. [51]

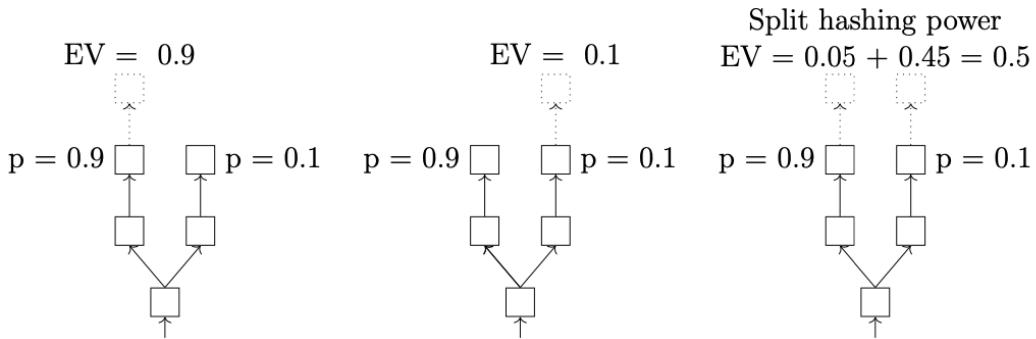


Figure 7.2: PoW: resource allocation in a soft fork [27]

7.1.1.5 Slashing

PoS-validators are punished if they are offline, attack the network or run modified software. [9] The process of subtracting a fraction of their deposited coins for such behavior is called slashing. It differs from protocol to protocol how much of the deposit is slashed. Some DPoS protocol also slash the token holders that delegated their coins, holding them responsible for delegating their coins to a untrusted actor. Thus, Table 7.2 gives an overview among the different protocols and their slashing penalties.

7.2 Comparison of Different PoS Architectures

As part of this report multiple PoS implementations were analyzed in order to understand the magnitude of rewards, penalties and other characteristics. In the following subsections the most important differences of each implementation is highlighted. Specifically, the following architectural characteristics were identified as important for analyzing the economics for running a full staking node.

Type As described in Section 7.1.1.2 there exists different flavours of PoS.

Minimum stake This is the minimum stake that a validator must possess directly or indirectly through delegations. For better comparison it is converted to USD with the exchange rates from Coinmarketcap¹ on the 25th of October 2020.

Validators Some protocols limit the amount of active validators. These restriction are may be imposed by the number of votes or stake in the network.

Penalty unresponsiveness This is the slashing penalty that is imposed for unresponsiveness. This can occur due to connectivity issues, unreliable hardware, software issues or attacks the the running node.

¹<https://coinmarketcap.com>

Penalty bad behaviour This is the slashing penalty that is imposed for deliberately attacking the network. These penalties occur when running modified software or voting on multiple chains in the same epoch.

Penalizes delegator In DPoS systems delegating nodes do not run a node but instead delegate their stake to a representative. This characteristic defines if the delegator is held responsible in a case of responsiveness or bad behavior of the representative.

APY The annual yield percentage is defined the expected return that is paid out for securing the network.

Slashing destination When a node is slashed, the removed stake can either be burned or be credited somewhere else.

7.2.0.1 Cosmos Hub (Tendermint)

Cosmos is a ecosystem of independent parallel BC that can interpolate with each other. Cosmos Hub is the first BC in the Cosmos Network, it is powered by the Tendermint consensus [14]. Tendermint is a byzantine fault tolerant (BFT) consensus algorithm. BFT means it can tolerate up to 1/3 nodes failing or acting maliciously, without the system crashing [15]. The Tendermint consensus has instant finality, therefore forks are not created as long as 2/3 nodes are honest. Blocks are finalized as soon as the block is created [14]. Cosmos hub uses BPoS as described in 7.1.1.2. It has a fix set of active validators, which consist of the 125 validators with the most total stake (self delegated and delegated). Both validator and delegator can be slashed for misbehavior of the validator, i.e double signing or downtime [16].

7.2.0.2 Cardano (Ouroborous)

Cardano uses Ouroboros as consensus algorithm, which is an implementation of PoS. In Ouroboros time periods are defined as epochs. Each epoch is a time frame of 5 days, which is divided into 21.600 slots of 20 seconds. For each slot there is a slot leader is randomly selected, where the chances are distributed equally to the stake of each stakeholder. The slot leader can create a block and receives the reward for doing so. Anyone that holds any amount of stake can participate in this lottery, but the more stake, the higher the chances of being elected. As a stakeholder you could also delegate your stake to a third party to act on your behalf and get some revenue without the need of being always online or creating blocks.

7.2.0.3 Tezos

Tezos uses a LPoS consensus algorithm. Stakeholders which create blocks are called bakers. To be a backer, participants have to own at least one roll (10'000 XTZ). The chance to be elected as validator is proportional to the number of rolls owned. Additionally 32 random participants are select as endorser. Endorser have the responsibility to vote for the block that they think will receive the most endorsements. The number of endorsement is used to determine which chain is the canonical one. Backers are rewarded for creating a block or endorsing a block that becomes a canonical one. For both action they have to deposit a fix amount of XTZ that can be slashed in case of misbehaviour. Stakeholders can delegate their stake to other participants, if they do not want to be backers and still get some rewards, they will not be slashed if the validator they chose misbehave [18].

7.2.0.4 Algorand

As explained in Section 7.1.1.3 Algorand is an example that favours security over liveness meaning the produced blocks are considered finalized as soon as they are published. The networks assumes that at least 2/3 of the validators are honest. With that assumption in place, there is no need for locking up coins or a sanction mechanism that punishes malicious nodes. [38] The reasoning for such an assumption is that an economy can only function if at least 2/3 of the total stake is honest. Algorand claims that under these assumptions the network forks in less than $1/10^9$ blocks [39].

7.2.0.5 Polkadot

The Polkadot consensus algorithm attempts to prevent the network from punishing the validators if they go offline. What makes the network stand out from others is that the actions of the other validators influence the penalties on each individual validator. The protocol differentiate between four levels of offences. [9]

level 1 If a node goes offline and at least 90% of the other validators remain online, the unresponsive node is put into so-called chilling mode, where no more rewards are earned but not punished either.

level 2 If more than 10% of the nodes go offline, the unresponsive nodes and its delegators are slashed. The fraction of the stake that is slashed depends on the number of other offline nodes in the same epoch. To calculate the penalty they use the following formula where x is the number of offenders and n the total number of validators.

$$\min((3 * (x - (n/10 + 1))) / n, 1) * 0.07$$

It is intended that there are 1000 nodes securing the network. Therefore, level 2 is being executed when at least 11 validators go offline. Thus, it is expected to loose 7% of its stake in such a scenario and more then 18% if all nodes are unresponsive.

level 3 This level contains all the misconducts that do not harm the security of the network to any large extend such as accidental double signing from one node. The formula used for these offences is defined below.

$$\min((3 * x/n)^2, 1)$$

This aggregates to less than one 0.1% if there is only one offender and 99 honest validators.

level 4 Any misconduct that poses a serious risk on the security of the network. It uses the same formula as in level 3. As soon as a third of the nodes is compromised, the bad actors are slashed with their entire stake.

Level 3 and 4 may also occur due to faulty software. The penalties collected from staking inefficiencies/attacks are credited towards the treasury. In case of a software bug, the community may accept this error and may refund the validators with the lost money. However, this is not guaranteed.

The Polkadot wiki claims to have 20% APY with 50% of all the tokens staked. However, currently there are more tokens locked. With the BC explorer the reward payouts were analyzed. At the time of writing this thesis the APY add up to 13%.

Validator pools with larger total stake backing them will get slashed more harshly than less popular ones. The rational behind this is to encourage nominators to shift their nominations to less popular validators.

7.2.0.6 Ethereum 2.0

The Ethereum network has proposed a plan to upgrade its architecture in terms of scalability. One of the proposals is the transition from a PoW to a PoS consensus mechanism. However, it is important to know that the described design is still in development. It is scheduled for 1st of December to go live on the mainnet. Nevertheless, Ethereum is the largest crypto-currency in terms of usage (cumulative transaction fees) [11] and second largest in terms of market capitalization. The economics for the PoS algorithm in Ethereum 2.0 will be influential and therefore, the design of Ethereum 2.0 is studied as part of this report.

The Network will roll out in three phases. PoS will be part of the fist phase. To incentives early adopters, staking rewards will be higher at launch and will decrease with more node securing the network. Table 7.1 shows how the staking rewards are calculated. [22]

2/3 of the validators must be online in order to run the network. No information could be found on how large the slashing penalty is, if a validator is offline. However, if more than 1/3 of the validators is offline, the slashing penalties can be as high as 50% of the deposited amount. [21]

ETH validating	ETH validating of total supply	Max annual network issuance (inflation)	Staking reward
1'000'000	0.88%	0.17%	18.10%
3'000'000	2.65%	0.30%	10.45%
10'000'000	8.83%	0.54%	5.72%
30'000'000	26.50%	0.94%	3.30%
100'000'000	83.33%	1.71%	1.81%

Table 7.1: Staking rewards in Ethereum 2.0

7.3 Security Considerations

7.3.1 Key Management

In the context of BC and cryptography the private key management is highly critical. Running a validator node requires that the validator appliance must be connected to the internet. This offers a range of possible attacks where an attacker attempts find the physical location of the node or event tries to steal the private keys of the validators remotely [29]. In Section 7.4.3.3 possible key management solutions are discussed in more detail.

7.3.2 Supply Chain Attacks

When buying hardware or using other services there is a risk that the product is compromised. It is important when purchasing hardware or services the source can be trusted. For example in case of a HSM, it is important to use certified and directly from the source, to avoid potential tampered or faked hardware. If the operator, wants to run the node on his own hardware, it is essential that the hardware is secure, so it should not be bought from an untrusted source. Using a service provider requires to trust in the security and honesty of that provider. [30].

	type	minimum stake	#validators	penalty offline behaviour	penalty bad behaviour	penalizes delegator	APY	slashing destination
Cosmos Hub	BPos	<5 USD	125	0.01%	5%	true	7 - 20%	burned
Cardano	PPoS	<1 USD	restricted	no slashing	false		6.93%	no slashing
Tezos	Liquid PoS	16000 USD	restricted	no slashing	1060 USD	false	6%	50% burned, 50% accuser
Algorand	PPoS	<1 USD	restricted	not	0%	0%	7%	no slashing
Polkadot	DPoS	1500 USD	50 (max 1000)	7 - 18%	0 - 100%	true	13%	treasury
Ethereum 2.0	BPos	13000 USD	not restricted	up to 50%	min 1 ETH	true	2 - 18%	accuser attestor

Table 7.2: PoS Protocol Comparison

7.3.3 Distributed Denial-Of-Service (DDoS) Attacks

A DDoS attack targets a single system. The attacker uses a network of compromised computer called bots or zombies to flood the target system with network requests. The general objective of the attack is to deplete the resource of the target to hinder or stop the service, which leads to the service being unavailable [12]. In the context of BC and cryptocurrencies, there are further incentives due to their popularity and the capital involved. Attacks can cause devaluation or loss of mining/staking rewards of a cryptocurrency[13]. In the context of PoS a DDoS attack could be used to hurt a competitors of a validators, by making the service of a validator unavailable and damage its reputation.

7.4 Architectural Considerations

In this section we will discuss different architectural considerations such as network topology, validator setup and key management. The architecture can help mitigate threats such as DDoS attacks and have an impact on the economics for the operator. There is a trade off between security, performance and costs.

7.4.1 Sentry Node Architecture (SNA)

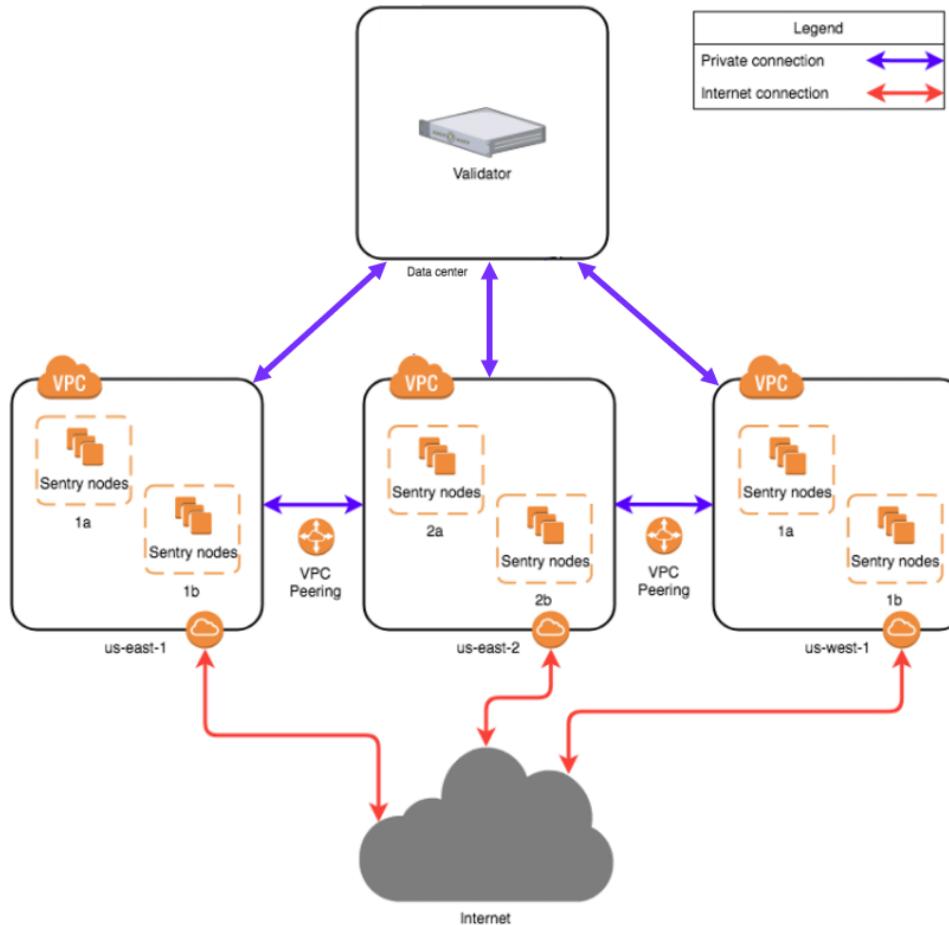


Figure 7.3: Sentry node architecture overview [31]

Validators are responsible for the security of their network to ensure availability. A key risk is the previously discussed attack angle of DDoS 7.3.3. The sentry node architecture is a recommended way to structure the network topology to provide DDoS protection. The problem of running a validator, is that the node has a fixed IP address and it opens

a RESTful API port facing the internet[31]. Having a fixed IP address opens up a vulnerability to DDoS attacks. To overcome this problem the SNA applies a solution similar to the classical backend/fronted separation of services in cooperate environments. The validator node represents the backend and is hidden in a private network of the validator in a data center. The network of the data center might use different strategies to setup the validator involving firewalls, multiple subnets and redundant devices [29]. Figure 7.3 illustrates a simplified version of a possible SNA setup based on Amazon AWS services. AWS services allows direct connection to a VPC, furthermore VPC can have private connections between each other, although these connections are not mandatory but can be useful. The validator node should only connect to trusted sentry nodes, avoiding exposing it to the public internet. The sentry nodes build the frontend, they are exposed to the internet and communicate with the outside. Sentry nodes can be hosted in virtual private clouds (VPC), communication between validator and the sentry nodes are done in private connections. Amazon offers VPC in different regions, which enables to have sentry nodes in different regions too. By having multiple sentry nodes, a DDoS attack would require to take down all nodes, but because of the flexible scaling of the cloud environments, new sentry nodes can be spawned, allowing the validator to continue working [31].

7.4.2 Possible Validator Configurations

There are different ways to setup a validator, with different advantages and disadvantages, in this section we will briefly discuss the different possibilities. Going from the simple single-node to the complex active-active validator setups, highlighting strength and weaknesses of each setup.

7.4.2.1 Single-Node Validator

A single-node validator is the simplest solution. Running a single-node has the advantage of avoiding the danger of inconsistency [32]. An example of how a inconsistent setup can lead to dramatic outcome can be seen in the example of the validator CosmosPool.org in the Comsmos network. The potential danger of inconsistency is illustrated in figure 7.4. The validator ran two different nodes with the same validator key, to ensure availability, by having a backup validator. The primary validator failed and the backup validator started working, which is how it supposed to work. But for unknown reasons the primary node came back online and started working too, which lead to the validator proposing two different blocks for the same heights. The double signing got punished by slashing 5% of the validators and all delegated stake, showcasing how important a sophisticated setup is [33]. With a single-node validator this risk can be avoided, since there are never multiple validator nodes running. Furthermore in a modern data centre network there should be a redundant network. The network should be designed by an expert with a very low failure rate, expecting high availability. This solution can be used by low-stakes validators, since the setup is sufficient and cheaper than other solutions. But the hardware can potentially still have a catastrophic failure, which is not acceptable for commercial validators, even if the risk is low the damage can be as high as 50% as described in section 7.1.1.5 not even factoring in the damage from reputation damage. For this reason commercial validator should use a different setup. [32]

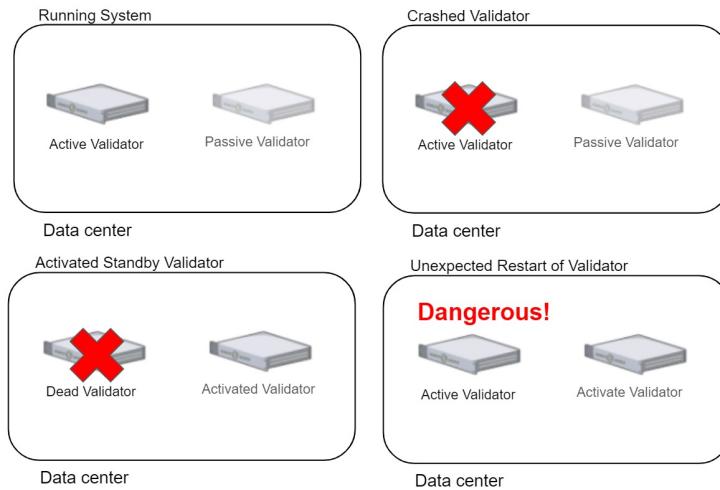


Figure 7.4: An active-standby validator setup.

7.4.2.2 Active-Standby Validator

The Active-Standby setup increases the up-time, by adding another validator node with the same configuration as shown in figure 7.4. As seen in the case of CosmosPool.org [33] the added complexity can lead to the risk of inconsistency. One possibility is to manually promote slaves, if the active node goes down the system administrator gets notified and is required to promote the standby node. However, a manual routine is susceptible to human errors and there are different tools to automate this task. The setup for automating the process is complex, especially ensuring that a node that went down does not switch back on for example through spontaneously resetting or by an operator mistake. The Active-Standby validator can lead to high availability but the setup can be complex and needs good testing before running it [32].

7.4.2.3 Active-Active Validator

The Acitve-Active has multiple validators with the same key running at the same time as shown in figure 7.5. To run multiple validators a the same time, it is required to have another consensus layer on top of the validator base. Although the consensus algorithm does not have to be byzantine fault tolerant, since it can be assumed that all validator nodes are honest, since they are all controlled by the validator operator. Having an active-active setup allows for consistency and high availability, without the unpredictability of validator nodes going down. However it requires the most complex setup, having to implement another layer of consensus on top of the validators, is not a trivial task [32].

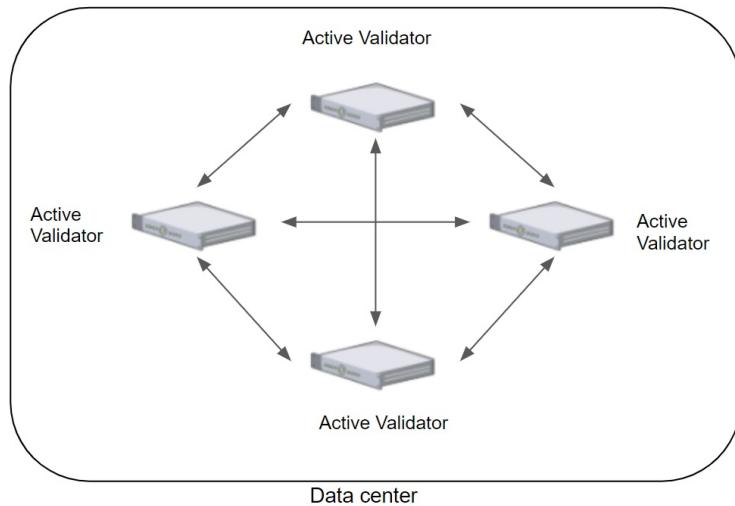


Figure 7.5: An active-active validator setup.

7.4.3 Cryptographic Key Management

One of the most insecure links in a distributed ledger system is the wallet, which is located on application level. Thus, it is crucial to store the cryptographic keys in a secure location. Validator nodes must be connected to the internet and thus, the keys are susceptible to a wider range of attacks. [52]

7.4.3.1 Hardware Security Module (HSM)

HSMs are dedicated processors designed to protect sensitive cryptographic keys that are used in an application. HSMs make it harder to extract the secret keys. It prevents the exposure of the keys even with physical access. All operations related to encryption and decryption are performed within the HSM. The hardware is well-tested and certified by specialized laboratories. The interface to a HSM is strictly controlled by the rules defined by the module to hide and protect the keys [40].

Keys can be generated within the module itself. It uses a physical process to create a source of randomness which is essential when creating secure encryption keys. Other life-cycle methods such as import, export, usage, rotation, destruction and auditing are key features of a HSM. Due to the custom hardware architecture, cryptographic operations are executed more efficiently with a HSM than a general purpose CPU. [20] Using a HSM does not fully secure a validator node setup. Although it increases the difficulty for an attacker to gain access to the private key, with access to the validator node the attacker can still perform slashable actions. As explained in Section 7.3.3 there are also other incentives for an attacker to compromise a validator node.

7.4.3.2 Cloud HSM

As more companies transition to cloud deployments, on-premise HSM can become a bottleneck in terms of latency. Thus, cloud providers offer HSM in the cloud. System architects must not concern with the HSM appliance selection or provisioning. Furthermore, an application can make use of the high availability of cloud providers and scale the application more easily. Some cloud providers allow to complement their infrastructure with an option of using an external HSM for storing the master keys for an additional level of security. [41]

The main advantage of (cloud) HSMs is that the key is (physically) separated from the rest of the application. The cloud provider can host the HSM in a physically more secure environment, monitor it individually and make an exploit less likely. However, adding a

HSM to a setup also means adding at least one trusted third party. The HSM is treated as a black box assumed to work as advertised. If the HSM of the cloud provider is compromised or out-of-service, many other validator nodes may be affected as well. As explained on the example implementation of Polkadot in Section 7.2.0.5, validators are slashed more drastically if many other validators go offline. Thus, with many validators on the same cloud provider it makes an attack on the cloud provider more lucrative.

7.4.3.3 Key Management Services (KMS)

Storing the private key in the file system means that an attacker who gains access to the validator node also get access to the private key, risking all the funds in that account. Thus, it is recommended to encrypt this key. KMS allows the creation of a key hierarchy. As shown in Figure 7.6 a key is encrypted with a master key and stored with the encrypted data. The master key can then be further re-encrypted in order to create a hierarchy. This makes it possible to be very specific on who has access to which data of the application. Furthermore, it makes it possible to manage all keys centrally from one place. However, the key on the top of the hierarchy must be stored in a very secure location as it has access to all the data. [42]

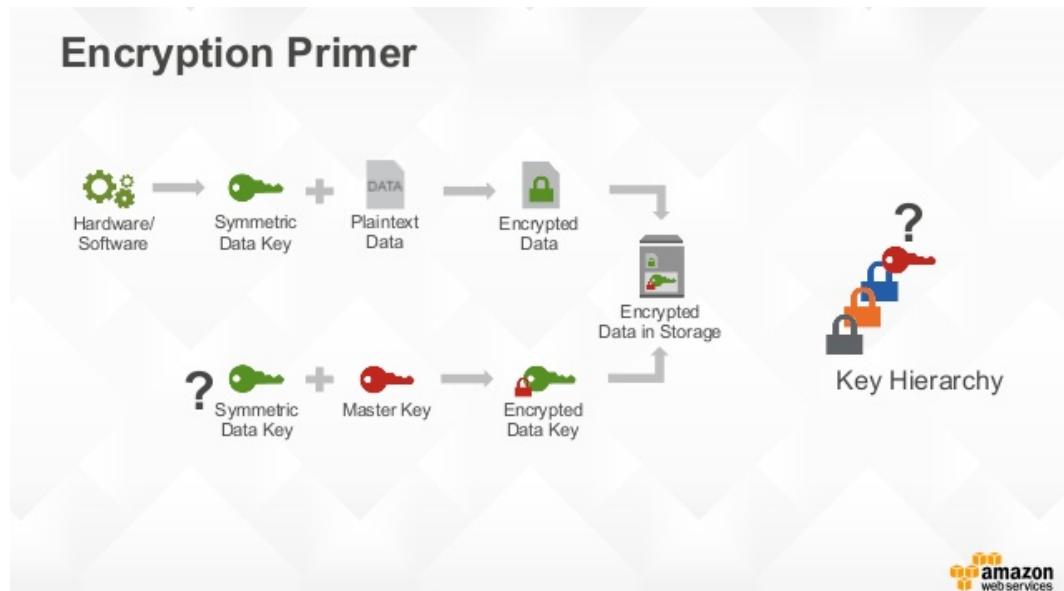


Figure 7.6: KMS re-encryption schema to create a key hierarchy²

If it is desired to run one or multiple validator nodes in the cloud, a KMS can be utilized to encrypt the private key and store the master key securely in one location. The computationally expensive tasks of validating transactions and executing SC is made in the cloud but signing the transaction still requires access to the master key. The master key can be store on a computationally less powerful device on-premise with a VPC to the validator node.

One can argue that storing a master key instead of a private key on the on-premise device does not add any additional security benefits. This is correct, however, depending on the cloud provider and the APIs made available, integrating the KMS with other products of the cloud provider may be simpler than signing transactions with the private key across multiple devices.

²<https://www.slideshare.net/AmazonWebServices/encryption-and-key-management-in-aws>

7.4.3.4 Smart Contract Wallets

SC-enabled BC such as Ethereum allow a SC to be operate as a wallet. The access control is split among multiple key where some of them should be store as a cold wallet (offline). These SC wallets can be configured such that funds cannot be drained if an attacker gets access to one of the private key. Two-factor authentication (2FA) can be enabled for example with a hardware wallet and an additional mobile wallet. Furthermore, by whitelisting certain addresses or set a daily spend limit can protect the owner. [24] Mobile wallets such as Gnosis³, Argent⁴ and Authereum⁵ have implemented SC wallets. SC wallets can protect an address from being liquidated to the attacker's wallet when one of the keys is compromised. However, slashable behavior can still be executed. Also, no information was found if any of the protocols allow SC wallets participate in the validation process.

7.5 Economics

Setting up a validator node in PoS requires a different hardware setup than a mining node in PoW. Many of the security considerations must be made for both types of consensus mechanisms. However, as explained in Section 7.3.3 an DDoS attack on a validator node can be more severe than an attack on a mining node. If a miner goes offline and cannot publish a successfully mined block, it misses out on the block reward and the transaction fees. If a validator node is unresponsive, besides the lost block reward and transaction fees the network actively punishes the node. Thus, more resources should be allocated when setting up the validator node. This section covers the economical factors that play an important role and visualize the expected revenues for different scenarios.

For all the following calculations, the system requirements for running a validator node was estimated according the recommendations from Polkadot [9]. The minimum suggested hardware requirements are shown in Table 7.3. Running such a machine in the cloud is estimated to cost around 200 USD\$/year. For the calculation the AWS (EC2: t4g.micro instance)⁶ and the Azure (A0 instance) pricing calculators⁷ are used. Buying these components for setting up such a machine on-premise costs around 1500 USD. For the ongoing costs 10\$/month is calculated for the internet access and 400\$/month for the electricity. The electricity was estimated with a power consumption of 200W running 24h/day with a kilowatt-hour price of 0.23\$.

CPU	Intel(R) Core(TM) i7-7700K CPU @ 4.20GHz
Memory	64GB
Storage	160GB SSD

Table 7.3: Minimum hardware requirements for running a validator node

As explained in Section 7.4, a sentry node only forwards the request such that the IP address of the validator node can be hidden. Thus, these machines need less computing power and barley any storage. It is estimated that the described workload can be carried out be a machine that is similar to a Raspberry Pi 4 model B which costs 50 USD [43] and consumes 3.4 W on an average load [44]. Including 10\$/month for internet access and a kilowatt-hour cost of 23 cents, this adds up to 130\$/year for ongoing costs. A comparable device on AWS and Azure costs around 10\$/year.

³<https://docs.gnosis.io/safe/>

⁴<https://www.argent.xyz/security/>

⁵<https://authereum.com/>

⁶<https://calculator.aws/#/createCalculator>

⁷<https://azure.microsoft.com/en-us/pricing/calculator/>

Section 7.4.3 covers the reasons to include an HSM in the setup. The prices for on-premise HSM were investigated. Most HSM providers do not specify the prices online. Securosyst⁸ is a Zurich based company that also builds BC specific HSM e.g. supporting the required hashing operations used in BC. Their products range from 10000\$ up to 34000\$. The following scenarios estimate the expected revenue based on the average HSM price which is 22000\$. A dedicated cloud HSM on Microsoft Azure costs 4.85\$/hour. This adds up to more than 3500\$/month or 42000\$/year. AWS does not offer cloud HSMs in Switzerland. An HSM in AWS located in Germany costs yearly 16800 USD. Thus, for the calculation the average price of 29400\$/year is used.

Additional parameters are introduced to make a more accurate prediction for the calculations. The average **API** among the investigated protocols in Section 7.2 is used, which aggregates to 9.3%. The **slashing penalties** are also included and are multiplied with the **server up-time** to estimate the slashing penalty that for every year.

Creating a hybrid setup, where the validator node is on-premise and the layer of sentry nodes is in the cloud, gives full control over the validator node, while still having the scalability of the cloud. In such a setup it is crucial to have a secure connection to the cloud service provider. The cost for a VPN connection are estimated based on the pricing of AWS, which charges \$0.05 per hour and \$0.09 per GB [25]. Since the connection has to be maintained the entire time, it leads to yearly base cost of 438\$. The traffic that runs over the connection depends on the BC. To make an estimation of the growth of a BC, ETH is used, which grew 48GB in 2019 [26]. This growth rate indicates that the cost for the traffic is rather neglectable. For the cost estimation model the traffic is estimated to 144GB, which results in costs of 12.96\$, giving a total cost of 450\$ yearly cost per VPN connection.

Scenario 1 The slashing penalties vary a lot between each protocol. If there is no penalty for being offline, additional DDoS prevention helps a validator to maintain the revenue stream from the block rewards as well as the public reputation of a validator. The reputation is relevant if this node functions as a delegator node. Therefore, without slashing penalties the server up-time only affects the expected revenue indirectly from the lost of block rewards. The simplest setup consists of one validator node run on-premise. With a total investment (before hardware acquisition) of 20000\$ it is possible to break even after one year. After ten years the stake is expected to grow up to 38000\$. However, this setup exposes the IP of the validator and is not recommended.

Scenario 2 This scenario is based on scenario 1 but adding another layer of security with three additional sentry nodes. Figure 7.7 illustrates the expected revenue over ten years with an initial investment of 20000\$ and one validator node and three sentry nodes. Compared to scenario 1 the expected stake after 10 years is about 10000\$ less. Therefore, as shown in this scenario without slashing penalties, the additional hardware has a great impact on the expected revenue. If the validator does not function as a delegator node and the reputation is not of importance, the number of sentry nodes should be selected with the financial effect in mind.

⁸<https://www.securosyst.com/en/>



Figure 7.7: One validator and three sentry nodes without slashing penalties

Scenario 3 Extending the setup from scenario 2 with an HSM, the validator does not operate on a profitable level. A higher initial stake is necessary to cover the high costs of the HSM. If the initial stake is increased to 100000\$ the validator breaks even after more than three years of operation as shown in Figure 7.8. Therefore, only commercial or large capital investors should consider adding an HSM to the setup.

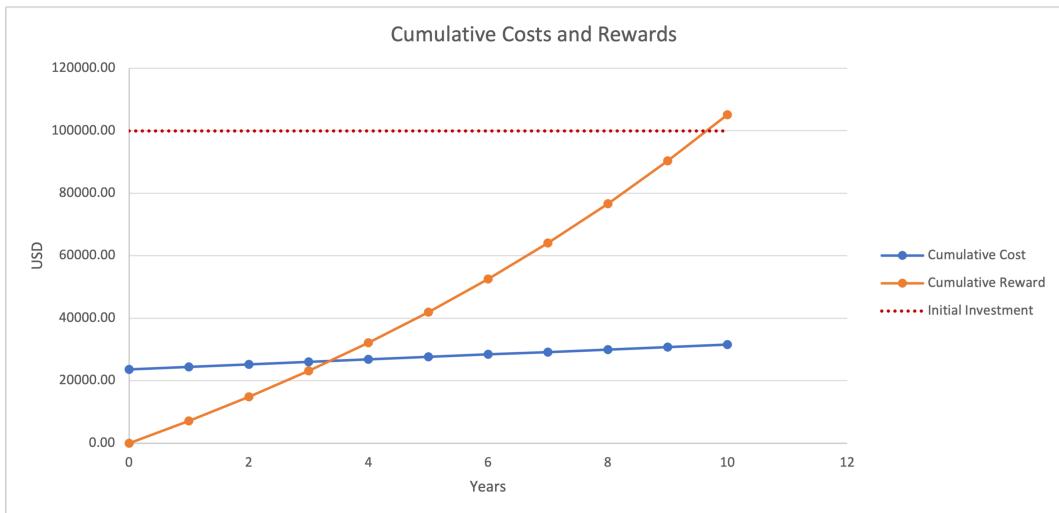


Figure 7.8: 100'000\$ initial investment and securing private key with an HSM

Scenario 4 Having the same setup as in scenario 2 but considering a slashing penalty of 7% (as in Polkadot 7.2.0.5), the server up-time plays an important role. Assuming that a validator node is slashed every two years (50% slashing probability for each year) the validator breaks even after about 5 years and only expects to have less than 10% profit from the initial investment after 10 years as shown in Figure 7.9. This scenario shows the impact of slashing penalty on the expected return. Therefore, high availability is desired such that the slashing probability can be kept low.



Figure 7.9: Expected rewards with slashing penalties imposed every two years

Scenario 5 Setting up a cloud-based architecture has the benefit that there are no initial costs. Furthermore, it is easier to scale the architecture up or down, depending on the requirements. However, the ongoing costs are higher. Just like in scenario 2, having one validator node and a layer with three sentry nodes in a cloud environment allow a validator to be profitable from the first moment. With the accumulated higher rewards, after ten years the expected return is also higher 7.10. With the proposed setup the validator can double his investment of 20000\$ after ten years. This multiplication factor is increased with a higher initial stake.



Figure 7.10: Cloud-based setup immediately yield profits with the right setup

As shown in these scenarios, it is important to analyze the slashing policy of the protocol. If this slashing penalties are imposed high availability is crucial to be profitable. Also, if no hardware is available to the validator and therefore first has to be purchased, setting up a cloud-based setup is more profitable.

7.6 Centralization

In this section the degree of centralization of different BC networks will be compared, to determine whether PoS or PoW is leading to more centralization. Table 7.4 compares four different BC (two based on PoW and two based on PoS). The table compares the percentage of power the 5 and 10 largest validators/mining pools have in their network.

	Consensus Algorithm	Sum Top 5	Sum Top 10
Bitcoin	PoW	63%	87%
Ethereum	PoW	74%	83%
Cosmos Hub	PoS	27%	42%
Tezos	PoS	28%	45%

Table 7.4: Comparison between different BC networks and their degree of centralisation

For PoW-based BC this is measured by the number of blocks created from a mining pool in the last year. This gives an overview of the current state of centralization in these BC. For PoS-based BC the stake is used to measure the power of validators in a network [37][36]. In both PoS BC (Ethereum and Bitcoin) the top mining pools are controlling more than 60% of the total power, and the situation for the top 10 is even more drastic with over 80%. Both PoS BC are showing drastically lower number with the top five holding less than 30% and top 10 around 45%. Although the PoS BC are not as extreme as the PoW one, there is already a large percentage of the total stake accumulated in the top 10. Furthermore, the PoS-based BC are not as established as the PoW ones. If the newer BC are proven to be stable, it could attract more investors to stake larger amount of money into these ecosystems. Neither consensus algorithm seems to be leading to a truly decentralized system regarding the creation and validation of blocks.

7.7 Summary

From a operational perspective, participating in a PoS consensus algorithm shares many similarities with PoW. In both cases, the node has to protect itself against DDoS attacks. A miner misses out on the block reward while the validator is slashed. In some of the PoS implementation the slashing penalties can be more severe than missing out on a block reward. Thus, it is essential that the validator has a dedicated hardware setup. The sentry node architecture and the use of HSMs are recommended across different PoS communities. However, on-premise as well as cloud-HSM are expensive. High capital investors or commercial validators should consider setting up a dedicated hardware device for managing the private keys. For low-capital investors buying specialized hardware with price tag of figures does not yield to any profits. There are cheaper alternatives such as the KMS and the use of multi-signature and SC wallets that can protect an enthusiast validator from being liquidated by an attacker.

To achieve a high level decentralization, the setup process must be simplified in order to make it accessible for more people. If centralized exchanges can offer highly competitive staking rewards on their platform, there is a risk of centralization. Blockchain-as-a-service (BaaS) has the potential to offer validator packages to simplify deploying the suggested setups and help paving the way for more decentralized networks.

When comparing the different PoS implementations, there were different approaches on handling inactivity or misbehaviour of validators from not slashing at all to up to 50%. Thus, becoming familiar with the financial penalties, that are imposed when going offline, must be studied carefully. Furthermore, it is difficult to manifest the economical performance of each protocol since many of them recently launched their mainnet or are still in development. Most protocols offer attractive APY to attract users to their platform. However, a trend of decreasing staking rewards for the future was observed in most of the protocols since many of them attempt.

The topic of PoS and operating a validator node is still in its infancy. There has not been a lot of research in this areas, further it is hard to identify standards for operating a validator. However, the goal of a well-functioning distributed network is to not only have

one standardized architectural setup for all its validators. Instead, having a diverse set of possible architectures makes the entire ecosystem more resistant to a potential exploit. With only one common standard, there is a risk for central point of failures such as the validator software, cloud provider, HSM manufacturer or even the ISP.

Bibliography

- [1] Jake Frankenfield *Consensus Mechanism (Cryptocurrency)*, <https://www.investopedia.com/terms/c/consensus-mechanism-cryptocurrency.asp>, (Accessed on 25th October 2020).
- [2] Nathan Reiff *Blockchain Technology's Three Generations*, <https://www.investopedia.com/terms/c/consensus-mechanism-cryptocurrency.asp>, (Accessed on 25th October 2020).
- [3] *Cryptoeconomics*, <https://blockchainhub.net/cryptoeconomics/>, (Accessed on 25th October 2020).
- [4] Jake Frankenfield, Julius Mansa: *Consensus Mechanism (Cryptocurrency)*, July 2020. <https://www.investopedia.com/terms/c/consensus-mechanism-cryptocurrency.asp>, (Accessed on 20th October 2020).
- [5] Michael Neuder, Daniel J.Moro, Rithvik Rao, David C. Parkes: *Selfish Behavior in the Tezos Proof-of-Stake Protocol*, CoRR January 2020. <http://arxiv.org/abs/1912.02954>.
- [6] Stakin: *The Proof-of-Stake Guidebook* <https://medium.com/stakin/proof-of-stake-guide-dpos-vs-lpos-vs-bpos-vs-hybrid-1393a33e849c> (accessed 20th October 2020).
- [7] Nathan Reiff, Somer Anderson: *Bitcoin vs. Ethereum: What's the Difference?*, July 2020. <https://www.investopedia.com/articles/investing/031416/bitcoin-vs-ethereum-driven-different-purposes.asp>, (Accessed on 20th October 2020).
- [8] *Bitcoin Energy Consumption Index*, November 2020, <https://digiconomist.net/bitcoin-energy-consumption>, (accessed 20th October 2020).
- [9] Logan Saether: *Polkadot Wiki*, July 2020, <https://wiki.polkadot.network/docs/en/learn-staking>, (accessed 20th October 2020).
- [10] *ALGORAND'S PURE PROOF-OF-STAKE APPROACH*, <https://www.algorand.com/what-we-do/technology/pure-proof-of-stake>, (accessed 20th October 2020).
- [11] Jamie Redman: *Cumulative Ethereum Transaction Fees in 2020 Supersede Bitcoin's by a Long Shot*, September 2020. <https://news.bitcoin.com/cumulative-ethereum-transaction-fees-in-2020-supersede-bitcoins-by-a-long-shot/>, (Accessed on 20th October 2020).
- [12] S.Dong and K.Abbas and R.Jain: *A Survey on Distributed Denial of Service (DDoS) Attacks in SDN and Cloud Computing Environments*, June 2019 <https://ieeexplore.ieee.org/document/8735686>

- [13] M.Saad and J.Spaulding and L.Njilla: *Exploring the Attack Surface of Blockchain: A Comprehensive Survey*, March 2020 <https://ieeexplore.ieee.org/abstract/document/9019870>
- [14] *What is Cosmos?*, <https://cosmos.network/intro#what-is-tendermint-core-and-the-abci>, (Accessed on 25th October 2020).
- [15] *What is Tendermint*, <https://docs.tendermint.com/master/introduction/what-is-tendermint.html>, (Accessed on 25th October 2020).
- [16] *Cosmos Hub*, <https://hub.cosmos.network/master/hub-overview/overview.html>, (Accessed on 25th October 2020).
- [17] Aggelos Kiayias, Alexander Russell, Bernardo David,Roman Oliynykov *Ouroboros: A Provably Secure Proof-of-Stake Blockchain Protocol* Advances in Cryptology - CRYPTO 2017. https://link.springer.com/chapter/10.1007/978-3-319-63688-7_12
- [18] *What consensus algorithm does Tezos use?* <https://learn.tqtezos.com/files/proofofstake.html#consensus>, (Accessed on 25th October 2020).
- [19] Jordan Tuwiner: *Bitcoin Mining Pools*, October 2020, <https://www.buybitcoinworldwide.com/mining/pools/>, (Accessed on 25th October 2020).
- [20] Peter Smirnoff: *Understanding Hardware Security Modules (HSMs)*, September 2017, <https://www.cryptomathic.com/news-events/blog/understanding-hardware-security-modules-hsms>, (Accessed on 25th October 2020).
- [21] *Ethereum 2.0: Where the rubber meets the road*, February 2020, <https://www.seba.swiss/research/Ethereum-2-Where-the-rubber-meets-the-road>, (Accessed on 25th October 2020).
- [22] Carl Beekhuizen: *Validated staking on eth2: #1 - Incentives*, January 2020, <https://blog.ethereum.org/2020/01/13/validated-staking-on-eth2-1-incentives/>, (Accessed on 25th October 2020).
- [23] Carl Beekhuizen: *Polkadot Wiki*, last updated June 2020, <https://wiki.polkadot.network/docs/en/maintain-guides-secure-validator>, (Accessed on 25th October 2020).
- [24] *Ethereum Documentation*, last updated June 2020, <https://docs.ethhub.io/using-ethereum/wallets/smart-contract-wallets/>, (Accessed on 25th October 2020).
- [25] *AWS VPN pricing*, <https://aws.amazon.com/vpn/pricing/>, (Accessed on 25th October 2020).
- [26] *Blockchain size*, <https://blockchair.com/ethereum/charts/blockchain-size?granularity=year>, (Accessed on 25th October 2020).
- [27] Simon Bachmann: *Proof of Stake for Bazo*, February 2018, https://files_ifi_uzh_ch_CSG_staff_bocek_extern_theses_BA-Simon-Bachmann.pdf, (Accessed on 25th October 2020).
- [28] : *Sentry Nodes*, <https://www.irisnet.org/docs/concepts/sentry-nodes.html>, (Accessed on 25th October 2020).

- [29] : *Validator*, <https://docs.tendermint.com/master/tendermint-core/validators.html>, (Accessed on 25th October 2020).
- [30] : *Delegators*, <https://hub.cosmos.network/master/delegators/delegate-security.html>, (Accessed on 25th October 2020).
- [31] : *Sentry Node Architecture Overview*, <https://forum.cosmos.network/t/sentry-node-architecture-overview/454>, (Accessed on 25th October 2020).
- [32] : *Validator High-Availability*, https://kb.certus.one/validator_ha.html, (Accessed on 25th October 2020).
- [33] : *Slashing Risks and Validator Diligence*, <https://medium.com/@staked/slashing-risks-and-validator-diligence-f6901cc9622a>, (Accessed on 25th October 2020).
- [34] : *Bitcoin mining pool distribution*, https://btc.com/stats/pool?pool_mode=year, (Accessed on 25th October 2020).
- [35] : *Ethereum mining pool distribution*, https://eth.btc.com/?_ga=2.132512654.1894298181.1605196655-367420409.1605080623, (Accessed on 25th October 2020).
- [36] : *Tezos validators*, <https://stake.fish/de/tezos/leaderboard/>, (Accessed on 25th October 2020).
- [37] : *Cosmos validators*, <https://stake.fish/de/cosmos/leaderboard/>, (Accessed on 25th October 2020).
- [38] Silvio Micali *Algorand's Core Technology (in a nutshell)*, <https://www.algorand.com/resources/blog/algorands-core-technology-in-a-nutshell>, (Accessed on 25th October 2020).
- [39] Yongge Wang: *Another Look at ALGORAND*, UNC Charlotte, February 2013. <https://arxiv.org/pdf/1905.04463.pdf>, (Accessed on 25th October 2020).
- [40] Peter Smirnoff: *Understanding Hardware Security Modules (HSMs)*, September 2017. <https://www.cryptomathic.com/news-events/blog/understanding-hardware-security-modules-hsms>, (Accessed on 25th October 2020).
- [41] Margaret Rouse: *AWS CloudHSM*, September 2016. <https://searchaws.techtarget.com/definition/AWS-CloudHSM>, (Accessed on 25th October 2020).
- [42] *AWS Key Management Service Wiki* <https://www.amazonaws.cn/en/kms/features/#Overview>, (Accessed on 25th October 2020).
- [43] *Raspberry Pi Website* <https://www.raspberrypi.org/products/raspberry-pi-4-model-b/>, (Accessed on 25th October 2020).
- [44] Avram Piltch: *Raspberry Pi 4: Review, Buying Guide and How to Use*, June 2020. <https://www.tomshardware.com/reviews/raspberry-pi-4>, (Accessed on 25th October 2020).
- [45] Ameer Rosic: *Smart Contracts: The Blockchain Technology That Will Replace Lawyers* <https://blockgeeks.com/guides/smart-contracts/>, (Accessed on 25th October 2020).

- [46] Jake Frankenfield: *What is Bitcoin?*, May 2020. <https://www.investopedia.com/terms/b/bitcoin.asp>, (Accessed on 25th October 2020).
- [47] *Understanding Economies of Scale*, July 2014. https://digiconomist.net/understanding_economies_of_scale/, (Accessed on 17th December 2020).
- [48] Jake Frankenfield: *Proof of Stake (PoS)*, August 2019. https://digiconomist.net/understanding_economies_of_scale/, (Accessed on 17th December 2020).
- [49] *Explain Delegated Proof of Stake Like I'm 5*, September 2017. <https://hackernoon.com/explain-delegated-proof-of-stake-like-im-5-888b2a74897d>, (Accessed on 17th December 2020).
- [50] *Safety and Liveness — Blockchain in the Point of View of FLP Impossibility*, May 2018. <https://medium.com/codechain/safety-and-liveness-blockchain-in-the-point-of-view-of-flp-impossibility-182e33927ce6>, (Accessed on 17th December 2020).
- [51] Julian Martinez *Understanding Proof of Stake: The Nothing at Stake Theory*, June 2018. <https://medium.com/coinmonks/understanding-proof-of-stake-the-nothing-at-stake-theory-1f0d71bc027>, (Accessed on 17th December 2020).
- [52] Julian Martinez *Blockchain only as strong as its weakest link*, August 2018. <https://www.csosonline.com/article/3295013/blockchain-only-as-strong-as-its-weakest-link.html>, (Accessed on 17th December 2020).
- [53] *Slashing* <https://bisontrails.co/slashing/>, (Accessed on 17th December 2020).
- [54] Julian Martinez *HARDWARE SECURITY MODULES: THE ULTIMATE BLACK BOXES*, January 2019. <https://www.usenix.org/conference/enigma2019/presentation/lackey>, (Accessed on 17th December 2020).

Chapter 8

Economic Advances of Blockchain-based Trading Platform

Saiteja Reddy Pottanigari, Ankan Ghosh

After the evolution, we had a problem of handling large group without relying on anybody. We are using monarchy , anarchy or democracy for solving but we always found a centralized method only. Blockchain handle this consensus by rewards or contributions like either by work or stake or capacity. Earlier we used to isolate things from external access to keep it safe but blockchain does it completely opposite of the procedure providing independent verification rather isolation. Enthusiasts say that blockchain can be a transformative as internet due to its repercussions in diverse fields. In recent years, Blockchain technology has gained tremendous interest. This research focuses on the commercial component of Blockchain-based trade networks and explores how Blockchain leads to economic development in multiple business sectors and the applications of blockchain technology. Blockchain can extend similar or different corporate logic across multiple stakeholders in a single industry.

Contents

8.1	Introduction	233
8.1.1	Overview of Blockchain	233
8.1.2	Why would Blockchain Contribute to Economy	233
8.1.3	Rise of Blockchain Ecosystem	234
8.1.4	Adaption of Blockchain Technology	236
8.2	Traditional Trade Finance Methods	236
8.2.1	Open Account	237
8.2.2	Letter of Credit	237
8.3	Blockchain-based Trading platform	238
8.3.1	Trade Finance	238
8.3.2	Delivery	239
8.3.3	Asset Exchange	241
8.4	Blockchain based various Use Cases	244
8.4.1	Economic significance and direct improvement in economic value	244
8.4.2	Marketplace	245
8.4.3	RealEstate	245
8.4.4	Finance	245
8.4.5	Energy	249
8.4.6	Supply Chain	251
8.4.7	Agriculture	253
8.4.8	HealthCare	255
8.5	Summary and Conclusions	258

8.1 Introduction

8.1.1 Overview of Blockchain

Blockchain is characterized by consensus, computing and authentication. The network can be managed by all network players. Without the approval of the whole network, the BC network consists of various computers linked together and cannot modify the block. It will replace certain conventional procedures including intermediaries or third-party reliance. BC is an ever-expanding block chain, which by cryptographic functions is interconnected and secured. A collection of protocols and approval from any network member is required to validate new blocks. The data is contained in a linear sequence. In BC for block representation, points and connected list data structures are used.

8.1.2 Why would Blockchain Contribute to Economy

BC is one of the most promising, innovative and creative distributed leader in the industry today. It is the solution for many of the common problem in digital solutions. Some are below:

1. Decentralization
2. Transparency
3. Anonymity
4. Security and Privacy

Decentralization The information is not stored by one single entity. All in the network always have all the details. The ledger is shared with everyone and is continually revised and synced. Public BCs are completely distributed and decentralized, where the validation of the transaction is trusted upon members of the network, decentralized meaning it does not have any governing authority or a single person looking after the framework. Rather a group of nodes maintains the network making it decentralized.

Transparency Transactions stored in BC network are accessible to all members hence creating transparency. This feature increases the trust of people in the system and allow more people to get involved and enhance a network for all in true sense.

Anonymity With the emergence of Turing complete smart contracts, developing an system which can handle the governance part of the system in decentralized infrastructure by implementing consensus rules and making the BC transactions reliable, accurate and available on every nodes on the network. However, anonymity in true sense is very difficult to achieve on public BCs.

Security and Privacy The use of private and public keys is a crucial feature of BC's privacy. Every consumer has a public and private key in such networks. Blockchain platforms use asymmetrical encryption to protect transfers amongst users. This increases protection and defends against hackers. Public keys should be exchanged with other network users and they do not have sensitive information. Users place their public keys in published blocks and exchange information with prior and subsequent blocks through the participation nodes. This makes the public key unchangeable and makes publishing false keys more difficult for the attackers. Tamper tolerant transfers across all network nodes make the BC safer. This is perhaps the most promising privacy study for BC currently has zkSNARKs.

8.1.3 Rise of Blockchain Ecosystem

In recent years, several conceptual implementations of research topics from distributed and parallel processing systems have progressed in BC ecosystem that have driven the BC adoption into efficient, decentrally focused applications of infrastructure.

Below are few components which helps in huge impact in the BC ecosystem:

1. Smart Contract
2. Initial Coin Offerings
3. DAO
4. DApps
5. Smart Wallets
6. Blockchain Oracles
7. Non-Fungible Token
8. OFF-Chain Operations

8.1.3.1 Smart Contract

Smart Contract (SC) is a contract that is self-performing. As in physical paper written contracts written in any natural human understandable language, SCs are drafted in a computer-interpretable language. With the emergence of Ethereum, the turing-complete smart contracts came into existence. SCs are a type of BC based account.[18] This means they have a balance and can send transactions over the network. Money can be stored on the SCs' specific address. Additionally Every token in blockchain is just a piece of code in a SC. every node in the network is allowed to write a smart contract and deploy it to the network. The tokenisation was created due to the use of ERC token specifications in smart contract came into existence.

8.1.3.2 Initial Coin Offerings

Initial Coin Offerings is a workaround for a pricey affair involving marketing banks (ie) Initial Public Offerings that make it impossible for SMEs to collect capital by these approaches. ICO is originally used by and for BC related ventures as a fundraising tool. This is first utility token and also security token offering.

8.1.3.3 DAO(Decentralized Autonomous Organization)

An autonomous organization based on the SC code including the framework for governance. DAOs run by configured smart contract rules on decentralized infrastructure. By considering how an entity might do without its administrators, Vitalik Buterin expanded the notion. Holders of a network governance token can control their parameters of DAO, MakerDAO is a protocol to construct a synthesis stablecoin (DAI).

8.1.3.4 DApps(Decentralized Applications)

Dapps are apps that run on a BC-built P2P network, often enabled by smart contracts. These distributed, robust, transparent and incentivized applications will prove themselves to the world by remapping the technological landscape. Dapp had crossed \$20 billion by Q3 2020.[20] Some of the key benefits of using Decentralized Applications in the real world:

1. No individual node will be able to control the network (Censorship-resistant).
2. No Downtime because no single point of failure
3. Open Source

8.1.3.5 Smart Wallets

Smart wallet is a SC-based wallet. In addition, the smart wallet may be renamed DApp for the control of different BC cryptographic currency. Smart wallets will provide functionality like the below:

1. Better recovery process for Private key
2. Multi Signature feature
3. Limit of transactions

Smart wallets will help us to carry out gas-less transactions. Meta mask is not a smart wallet but it is non-custodial wallet(No need of trust our private key with such wallets). Coinbase is a custodial wallet(We have to give our private key with such wallets).

8.1.3.6 Non-Fungible Token

Only fungible (non differentiable) token were available up to ERC 721 standard. ERC (Ethereum Request for Comment) defines a common list of rules that all Ethereum tokens must adhere to Non-Fungible Tokens(NFTs) allows tagging real world distinguishable assets and they allow digital representation tied to a physical object. Token generated using ERC-721 standard additionally stores the meta description or any meta tags for non-fungible property. Token generated using ERC-1155 standard can be both fungible and non-fungible. Cryptokitties and many games have been developed using these both standards.

8.1.3.7 Blockchain Oracles

Protocol or utilities linking off-chain data to the deterministic BC. Oracles are often referred to as middleware blockchain. Using API calls or other non-deterministic sources in the Blockchain infrastructure. There is a good opportunity that the source might depreciated, compromised, or otherwise disabled , and we would not be able to validate transactions. smart contracts also depend on off-chain inputs and we would have the issue with Oracle. There are currently in the existing market no effective Oracles' decentralized designs. Chainlink provides some necessary services to Oracles.

8.1.3.8 OFF-Chain Operations

In order to conclude an on-chain transaction, there has to be an agreed number of confirmations by miners. This sounds fantastic but poses a minor challenge that sending every single change to every single participant in the network and every single participant requires to record every single change which need both storage, transaction speed and computation resources. It also relies on network congestion to complete a chain transaction. Downloading those transactions in chain and supplying users with the means to connect outside of chain. IPFS is a offchain hypermedia P2P protocol to improve the network speed, protection and transparency, like a shared cloud called Filecoin. While Bitcoin increased scalability of the base layer by ten times, it will still not meet the needs of the system of payment of the world economy.

OFF CHAIN Operations offers :

1. Lower Transaction fees.
2. Lower Storage resources used.
3. More Integrability.

8.1.4 Adaption of Blockchain Technology

The majority of BC services are offered on a pilot scale and with limited revenues. Technavio predicts that adoption in the banking, financial services and insurance sectors will increase 63% by 2022.

Government exploration of BC for national currencies continues to increase globally, Technavio expects to see paper or national currency replacement BC increased the spread of the technology in all sectors. The use of national currencies is expected to improve the efficiency of the services offered and create new sources of income in various sectors.[36] With the high prevalence of the Internet and associated technology, a number of Industry 4.0-based applications have been used around the world to register, measure and communicate data for industrial automation through sensors and actuators. In order to resolve problems such as data heterogeneity, data confidentiality and data redundancy, as well as security and privacy concerns, such systems process data in large amounts and this is therefore essential. In addition, different applications require records from different domains in different formats. Therefore, the data format must also be standardized so that several applications based on Industry 4.0 can use this. The use of smartphones and smart apps for family, technical and social activities is rising exponentially worldwide, contributing both to an increase in network traffic and to an increase in overall expenditure (in billions of dollars).

According to this report, by 2020, smart industry would spend \$ 40 billion on the Internet of Things across sectors including transportation and manufacturing. However, due to the large number of data exchanges over the Internet, maintaining confidentiality, data protection and integrity becomes a major problem in Industry 4.0. In addition, according to surveys carried out by various authorities, almost 60 million people are affected by identity theft. There were 12 billion people registered in 2018 and it is expected to increase to 33 billion by 2023. After the development of BC, people believe in the expectation of the changes in the world economy have increased dramatically because of its decentralized and immutable property that helps overcome a big problem in a country's economy, the middleman, broker or intermediary and (word for light) trusted the world of smart contracts. A human code for monetary purposes began and this started investments in various applications creating an ecosystem for BC to flourish. DAO (Decentralized Autonomous Organization) and many other were discussed in this report about the impact on P2P trading and its effect directly in economic.[35]

8.2 Traditional Trade Finance Methods

Trade finance may be understood as a transaction or financial assistance that takes place during international trade. In international trade, an importer typically has to have collateral or to pay for goods that are either shipped or will be shipped. While the world monitored decline in the global manufacturing output and global export compared to 2019 during the pandemic and lock downs. Since the beginning of 2020, no less than six large international commodity traders have gone out of business due to the traditional trade finance methodologies[31]. Trade finance has a huge influence on the world economy. There are many different traditional trade finance methods but this report covers two major methods.

1. Open Account
2. Letter of Credit

8.2.1 Open Account

Open Account is the traditional method started before Barter system¹. An open account transaction is a sale where the goods are shipped and delivered before payment is due. Obviously, this option is the most advantageous for the importer in terms of cash flow and cost, but it is consequently the highest risk option for an exporter. Upon an agreement between exporter(seller) and importer(buyer), The goods, together with all the necessary documents, are shipped directly to the importer who has agreed to pay the exporter's invoice at a specified date. The exporter should be absolutely confident that the importer will accept shipment and pay at the agreed time and that the importing country is commercially and politically secure. As the exporter here waits for payment from the importer.



Figure 8.1: Open Account Trade Finance

8.2.2 Letter of Credit

With the advent of electronic message transmission banks started using SWIFT (Society of Worldwide Interbank Financial Telecommunication) and established a standard format for Letter of Credit to communicate. A typical Letter of Credit transaction is as shown in figure[1.2] involved stakeholders and sequential steps involved to carry out a complete Letter of Credit transaction. This is extensively used in international trade. Letter of Credit process usually takes more than 7 days of time only for the documentation process other than the shipment delivery time. Here are the detailed steps of how the letter of credit occurs.

1. Initially After an agreement between seller and buyer
2. Buyer approaches his bank and request for a letter of credit to the amount he is requesting
3. Buyer's Issues the letter of credit based on Buyers deposit in the bank or Loan on him and send this letter to Seller's Bank
4. Seller banks check and confirms him that he can proceed to ship.
5. Seller ships the product and gets the documents to Seller's bank.
6. Seller's Bank sends the documents to Buyers Banks and receive the payment and pays to Seller

¹A barter system is an old method of exchange. This system has been used for centuries and long before money was invented. People traded services and products in return for other services and goods.

7. Buyers will receive his products meanwhile

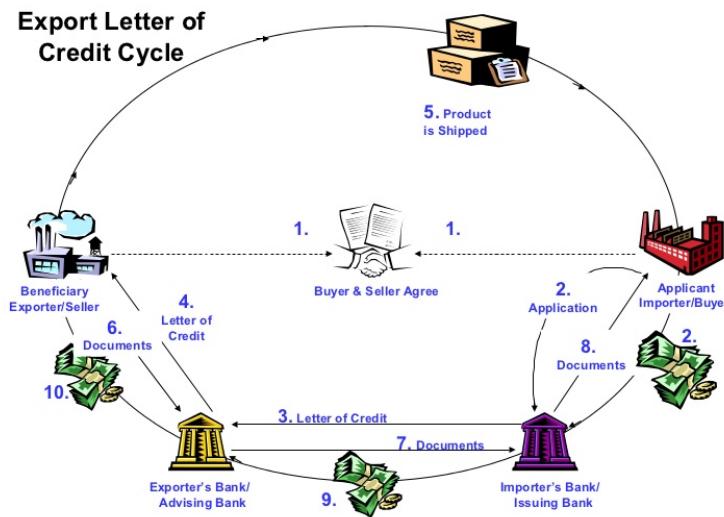


Figure 8.2: Letter of Credit Trade Finance

Usually the providers of trade finances are Banks, syndicates, trade finance houses. The users of trade finance are manufacturer, importer, trader. The main disadvantages observed from above two methods is trusting a third party and need of documentation. There are other disadvantages like poor control over funds, poor control over the goods, tough chances of refund or repayment, not better visibility and monitoring over the whole trade cycle through the entire transaction, poor security for the data tampering. Since trade finance is crucial for the global economy, better methodologies are required.

8.3 Blockchain-based Trading platform

As several problems have not been addressed by conventional approaches. BC may have an effect on the framework with its assets. This report covers different stages where BC can impact in an entire trade cycle.

1. Trade Finance.
2. Delivery.
3. Asset exchange between blockchain.

8.3.1 Trade Finance

Although credit certificates are now stored online through digital processing, they are not always accessible or visible to at least all the participating parties in the trade. In Blockchain based trade finance methodologies, SC plays a role in the authentication of goods to be delivered or shipped for both a deposit escrow and a validation escrow. We can lay down few rules in the SC and in these trading systems for a seller with buyer or seller with transporter or buyer. While using SC's in blockchain based solution, In order to facilitate the deal, the seller and the buying firm shall sign the terms and conditions or may also deposit the collateral. The physical letter of credit documentation by UBS will weigh approximately 500 grams of 36 papers.[34].

UBS, IBM and several others are collaborating on Batavia[37], a global trading financial platform based on the BC. It just takes about 1 hour to record the whole workflow

of the Batavia project with several systems involved. Whereas, The project Batavia helps simplify the entire market chain by offering interactive and streamlined solutions for foreign trading deals, security and funding. Since the trade finance platform is based on BC, the sharing of all documents by all members is carried out in a single distributed version. This makes everyone aware of the exchange. This includes trade arrangements and the implementation of smart payments. And "smart" it means that transactions are automatically triggered and immutably reported in the BC by some events in the supply chain.

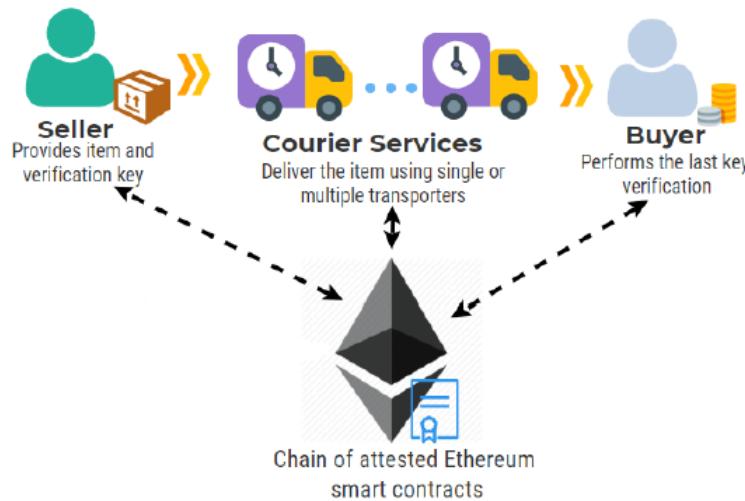


Figure 8.3: Letter of Credit Trade Finance

8.3.2 Delivery

There are numerous Blockchain-based Delivery Solutions due to the emergence of a turing-complete smart contract this advent was possible.

1. Key Signature
2. Multi Signature escrow
3. Double Deposited escrow(Double collateralized escrow)
4. Proof of Delivery with Arbitrator

Physical Asset Delivery BC solutions use the best peer-to-peer security features to create a distribution solution between a vendor, a customer and carriers who might any miles apart or also can be in separate countries. These strategies can be expanded into as many carriers (or delivery services) as possible with efficient and safe manner. This solution allow both parties to behave sincerely by making any involved party to deposit a collateral amount. Any individual party in the following figure has an address in Ethereum and interacts on permission with the smart contracts created during an asset delivery process. When the package is shipped by two parties, a contract chain for this particular package delivery arrangement is established and dependent on the number of courier services, the number of smart contracts. However, between a seller and a buyer, at least two contracts are needed. Before the contracts are initiated they should sign and agree on the terms and conditions of the arrangement between the seller, the buyer and the courier services(s).[31]

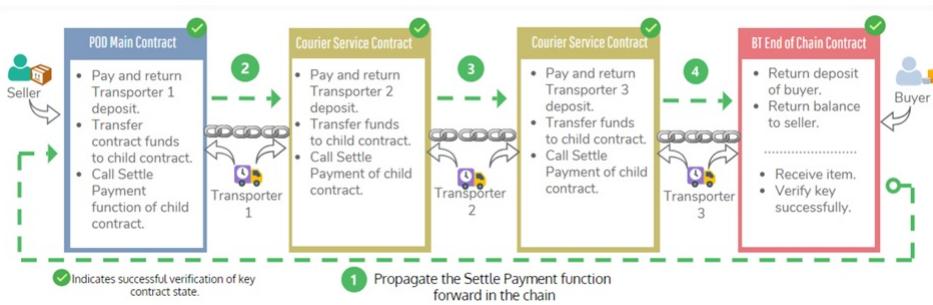


Figure 8.4: Smart Contract of Proof of Delivery with Multiple Transporter

Digital Asset Delivery Now-a-days, Digital assets delivery involves a bunch of illegal transactions. So , BC based solution is much needed and this report addresses a crucial BC file sharing, which will allow a vast range of businesses to provide an autonomous and decentralized data sharing framework with arbitrator.[32]

1. Firstly, customer requests the digital content which could be a file, a book, an image, a video or music which can be streamed or downloaded.
2. Internally, A contract receiving a request by the customer indicates that the customer agreed to the terms and conditions of the contract.
3. Customer will be asked for a double deposit collateral and it takes the deposit.
4. The file server would also then have to deposit the same amount as collateral.(This helps to encourage all organisations to behave honestly and to ensure that all participating organisations are equal in their positions of authority.)
5. The contract then generates automatically a specific token for the client which is legitimate according to the specified form of digital content.
6. The server will allow consumers to retrieve the content safely using the token generated.(The token is accessible to anyone listening to the network.)
7. The file server checks the download by running a smart contract function. every node in the network is informed of the customer's update from the file server of the content.
8. The customer will also conduct a feature with the effect of the download validation.
9. The bill shall be paid and all collaterals shall be reimbursed until the client is pleased.
10. Item price will be shared between the owner and the processor.

However, often there may be a discrepancies in delivery, then the arbitrator takes the token of the customer and uses it to download the same file. The results of the arbitrator will decide whether or not the customer had the right to a refund. In addition, the framework requires consumers to seek refunds before the file server has yet to deposit its collateral.

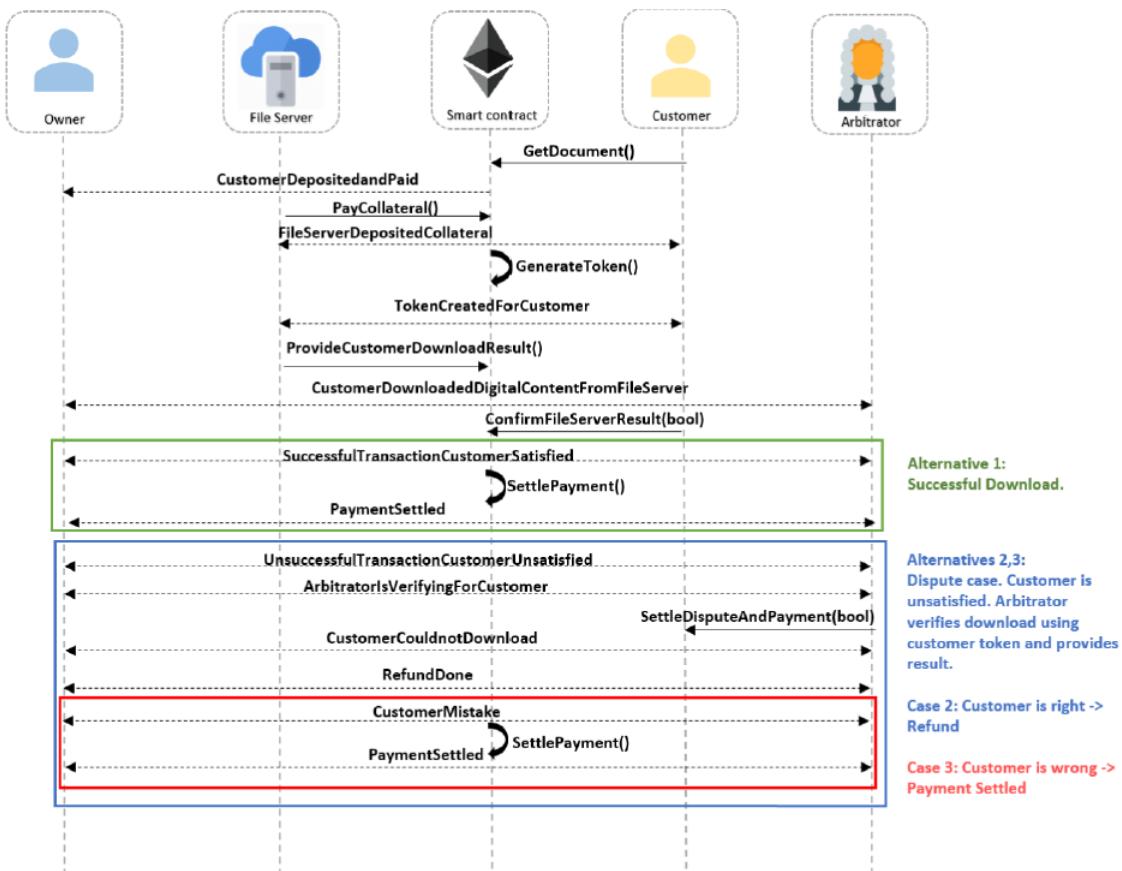


Figure 8.5: Smart Contract of Proof of Delivery with Multiple Transporter

8.3.3 Asset Exchange

As BC numbers on the market are rising. The interoperability of BC will be an important future problem for genuinely complex architecture, but we have little decentralized interoperability solutions that bring more power to mutual sharing or trade on any two blockchains of heterogeneity.

Two Accepted decentralized interoperability approaches between two separate chains:

1. Sidechains
2. Atomic swaps(Hashed Time-locking Approaches)

Sidechains originally meant a chain that validates events in outside chains as part of the local validation process. Later, though, it is also a fantastic choice to link two BC's or add an additional BC to two-way connection mechanism to improve BC scalability. One of the key reasons that they have been introduced are to congested big networks and increase performance. The same methods as a new BC are required to make the wait time even longer, but you are not out of storage.

Atomic Swaps Allows the exchange of assets without Trustworthy Third Parties between separate chains. Which allows transparent interactions between two nodes/users in different BC's. Alice wants to exchange 10α tokens with 100β tokens of Bob. They exchanged using Atomic swaps as below steps:

1. Firstly, Alice and Bob must set up their account in each other's blockchain.
2. Alice creates a smart contract of tx1 locked with a secret(x) using hash function $H(x)$ (Secret Hash)¹ In transaction tx1 in smart contract, Alice locked her 10α

¹The Hash function $H(x)$ can be considered as a lock, and x is the key. The value of x can be used as the way to decrypt your contracts as tx1 and tx3 have both been generated with $H(x)$. where x is a random value(Secret).where x is a random value (Secret).

tokens with few conditions to pay 10α tokens from Alice to Bob's public key if (x is submitted and signed by Bob) or (signed by both Alice and Bob).

3. Since Alice α tokens is locked in tx1 she needs to generate another smart contract, tx2, to retrieve her α tokens back if Bob defaults. In transaction tx2 in smart contract with few conditions to pay 10α tokens from tx1 to Alice's public key after 48 hours and signed by Bob. Usually tx1 and tx2 are often referred to as Hashed Timelock agreements (HTLCs).
4. Later, Alice then sends tx2 to Bob.
5. When Bob signs tx2, Alice transmits tx1 to the network or sends it directly to Bob.
6. Bob creates smart contracts of tx3 with same secret hash $H(x)$. In transaction tx3 in smart contract, Bob locked her 100β tokens with few conditions to pay from Bob to Alice's public key if (x is submitted and signed by Alice) or (signed by both Alice and Bob).
7. Bob also creates transaction tx4 with few conditions to pay 100α tokens from tx3 to Bob's public key after 24 hours and signed by Alice.
8. Then Bob sends tx4, signed and returned to Bob by Alice and After that Bob transmits tx3 to the entire network or sends it directly to Alice.
9. Alice then signed tx3 and issued 100β tokens at her address from the contract and disclosed to Bob the value of x simultaneously.
10. With the x value, Bob will now sign tx1 and collect tokens of α .

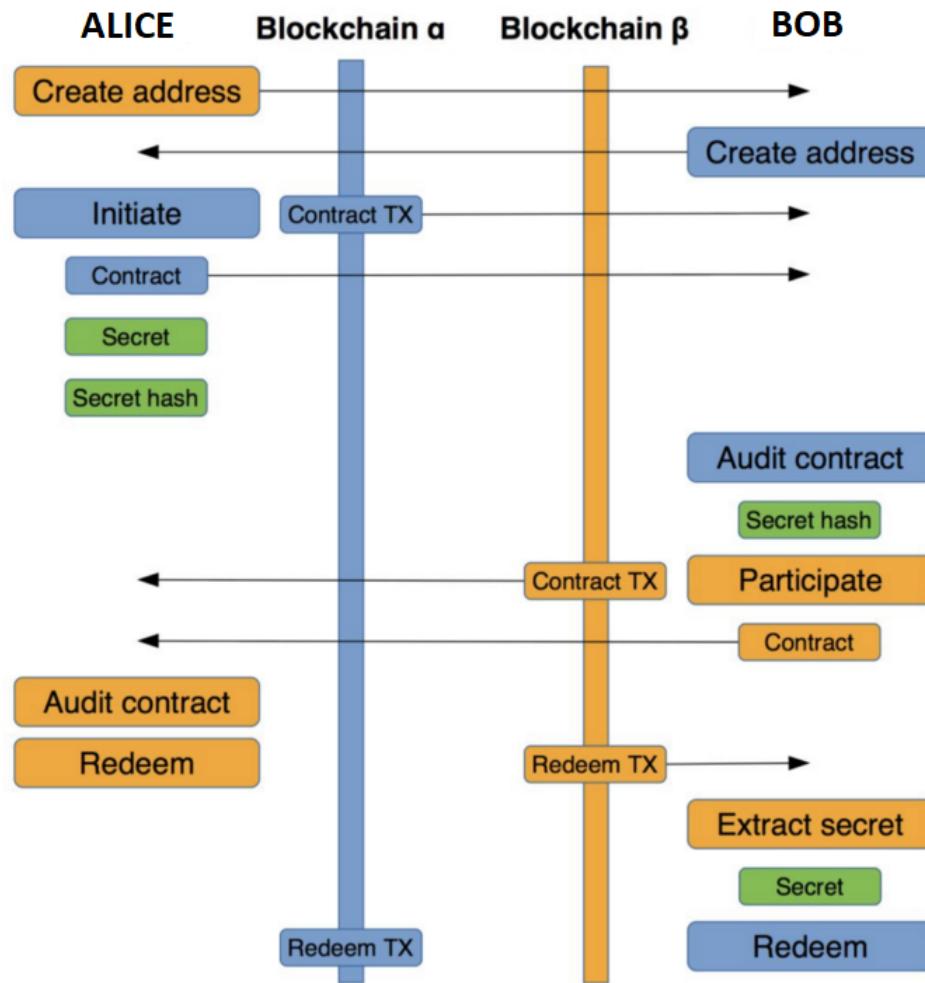


Figure 8.6: Atomic Swaps

8.4 Blockchain based various Use Cases

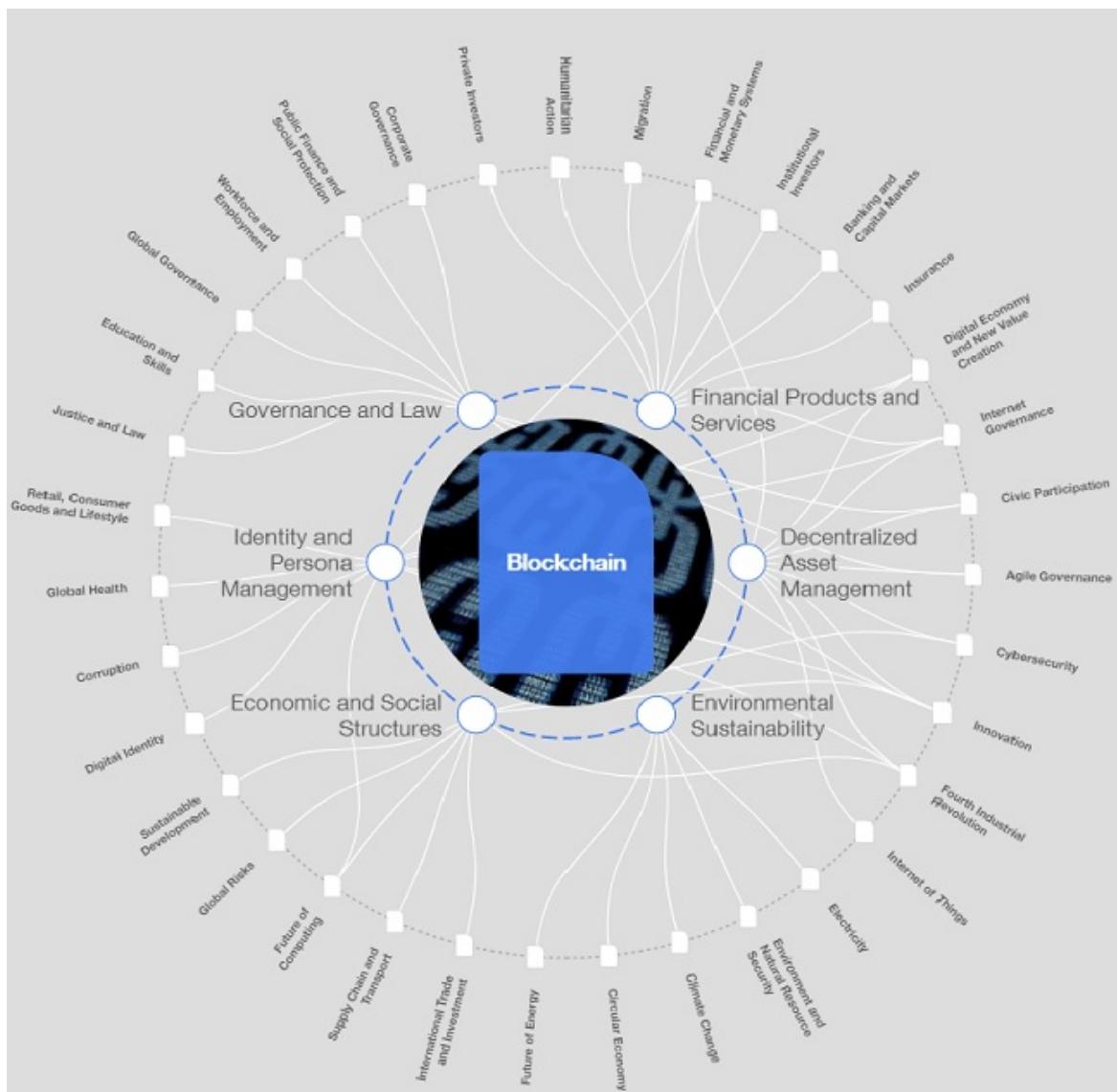


Figure 8.7: Blockchain based use-cases

8.4.1 Economic significance and direct improvement in economic value

BC has attracted great interest not only from the IT industry, but also from other industries and governments around the world. Over the past decade, significant investments have been made in research and development, testing and implementation of BC in multiple sectors. Increasingly, large companies are forming alliances and consortiums around BC ecosystem investing significantly in its research and development. As an example, more than 145 companies including IBM, Microsoft, Cisco, Intel, JP Morgan and Toyota form the Hyperledger consortium. Market research report from Technavio forecasts adoption in the banking, financial services and insurance sectors to increase 63% by 2022. World Economic Forum expects at least 10 percent of the global GDP being stored on BC platforms by 2025. Through BC technologies, value creators such as artists, composers, and designers could transfer value to their clients or consumers directly. BC make it possible to track and trace intellectual transfers of property thereby protecting equally producers and consumers of products and services.

8.4.2 Marketplace

In Marketplace sector, there are currently several upcoming DApps in a single setup for each product. This study only covers a few markets which already support the e-commerce economy.

Opensea is the first decentralized, P2P marketplace for BC-based assets. Opensea has collectibles, gaming items, domain names, digital art traded over its platform and OpenSea could let 20,000 ETH pass through their market.

District0x is a network of decentralized markets and communities. it believe Create, operate, and govern powered by Ethereum, Aragon, and IPFS. Since the ERC721 and 1155 launch of NFTs, the e-commerce industry has opened up the prospect of marking real world items. District0x is a community of market places and societies that function as decentralized autonomous organisation. The Bazaar name is an ENS trading application. ETHlance is a DApp, just like another freelance website which has better secure payments are taken place in eth rather than defaulter-tolerant traditional website.

Provenance is a DApp BC platform that facilitates safe traceability in the supply chain for certifications and other relevant details. Provence allows a physical product to have an auditing record of their path behind all physical goods. They provide an application that provides how they are manufactured or mined and sent while they are monitored at all times.

Decentraland is a virtual reality platform powered by the Ethereum BC. Users can create, experience, and monetize content and applications. Land in Decentraland is permanently owned by the community, giving them full control over their creations. This app is the marketplace for gaming and dealing of bits of property, carrying and exclusive names in the virtual universe by setting the MANA own price and an expiry date. It also presents a land mortgage concept which opens the virtual experience to any investor. Land is a non-fungible, transferable, scarce digital asset stored in an Ethereum SC. It can be acquired by spending an ERC20 token called MANA. MANA can also be used to make in-world purchases of digital goods and services.

8.4.3 RealEstate

BlockIMMO is a swiss based Decentralized Application aiming to create simple, secure real estate investment platform. It is marketplace for real estate tokenization. It is great platform for real estate portolio management. In this platform, dividend payments often take place for commercial properties for owners and also make P2P exchange tokens of any property including building shares. BlockIMMO have to abide by essential anti-money laundering (AML) and know-your-customer (KYC) requirements, which are not only costly but also time-consuming. BlockIMMO obeys regulatory rules by third party integrations like BLOCKID integration for KYC. it a great platform for real estate portfollio. Investors can buy IMMO tokens that are designed like securities, thereby benefiting on the one hand from increases in value and, on the other hand, from dividend distributions

8.4.4 Finance

Blockchain in Finance is now-a-days called DeFi(Decentralized Finance) and a popular field where BC based solutions for most of the problems like anonymity, immutable transaction, global in nature both by public/private permissioned/permissionless BC with smart contract with help of Oracle.

Advantages:

1. Instant Settlements

2. Improve Capital Optimisation
3. Reduced Counterparty Risks
4. Improved Contractual Performance due to Smart Contracts
5. Increased Transparency
6. Increased Financial Solutions in terms of Crisis
7. Reduced Error Handling and Reconciliation

Some of the diverse applications in finance:

1. Uniswap, the most popular peer to peer Decentralized Exchange platform which let you swap ERC20 tokens and liquidity pool.
2. Compound is a decentralized protocol which establishes money markets with algorithmically set interest rates based on supply and demand, allowing users to frictionlessly exchange the time value of Ethereum assets.

8.4.4.1 P2P DEX

Uniswap, Curve, Balance are some of the popular Decentralized Exchange platforms. With just a smaller and efficient methodological changes these three popular DEX platforms differ. This report cover about Uniswap the most used DEX with \$2 Billion locked in the platform and \$0.5-1 Billion daily trading volume.

Uniswap: is a protocol to trustlessly exchange ERC20 tokens. Uniswap uses liquidity pools concept with automatic market maker logic rather than the standard order book paradigm for trading tokens. Uniswap pools tokens to smart contracts and users exchange tokens against these liquidity pools. Anyone can swap tokens, add tokens for fees to a pool or can also list their token in Uniswap. Uniswap contains two main components (i.e) liquidity pools² and Automatic Market makers³. Everyone can establish an interface that connects to these contracts in uniswap and automatically initiate an exchange with someone who using the same service. The quantity and prices in Uniswap are deterimed based on an automated market maker formula. Uniswap has two separate kinds of contracts. First is the Exchange contract which holds a pool of a specific token and Ethereum against which the user can swap. Second is the Factory contract which is responsible for establishing new Exchange contracts and registering the ERC20 token address to its Exchange contract address.

This report discusses three usecases of Uniswap:

1. Liquidity Providing
2. Exchanging ETH to Any ERC20 Token
3. Swapping ERC20 Tokens

Liquidity Providing Initially when an Exchange contract is first created for a specific ERC20 token, both the token and Ethereum pools are empty. The first party to deposit in a contract determines the token's ratio to the ether. If they deposit a ratio other than the current market ratio, then there is an arbitrage opportunity. When liquidity providers are adding to tokens to an already established pool, they should add a proportional amount

²Liquidity pools is pool of tokens locked in smart contract to facilitate trading

³Market maker facilitate always buying and always selling option without waiting for counter party

of token and Ether to the pool. If they do not, they would also face arbitration of the liquidity that they attached.

Uniswap convinces users to introduce liquidity by rewarding liquidity providers with some fees. A 0.3% fee is charged for exchanging Ether with a ERC20 token and approximately 0.6% fee is applied to the ERC20 token to ERC20 token swaps. The Liquidity provider will be given unique ERC20 tokens known as liquidity tokens proportionate to the amount of liquidity they added to the pool. The token is burnt as the customer decides to obtain liquidity contributed plus receives the payment accumulated during the lock of liquidity.

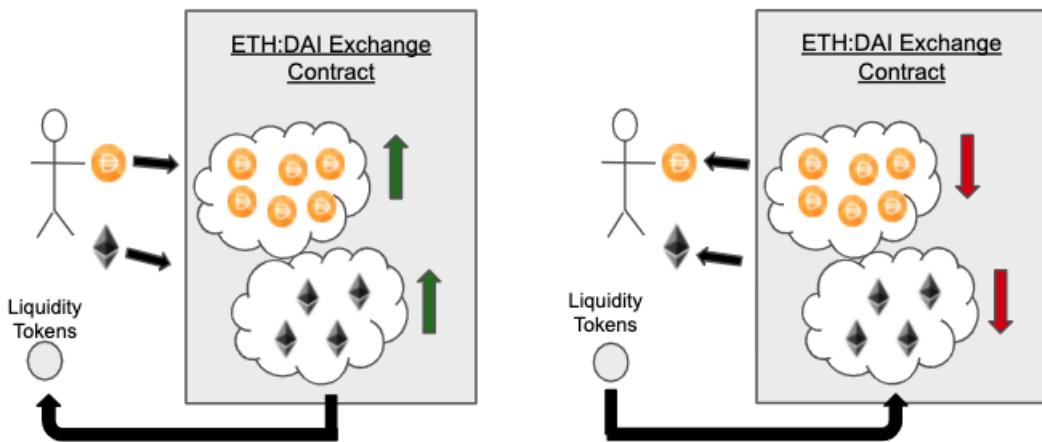


Figure 8.8: Liquidity Providing

Exchanging ETH to Any ERC20 Token The Ether is sent to the contract pool as the token is exchanged, and the token is returned to the customer. The customer does not need to wait for a counterparty to swap or about the price setting. As anybody can list/liquidate the tokens and users don't have to think about matching with another user.

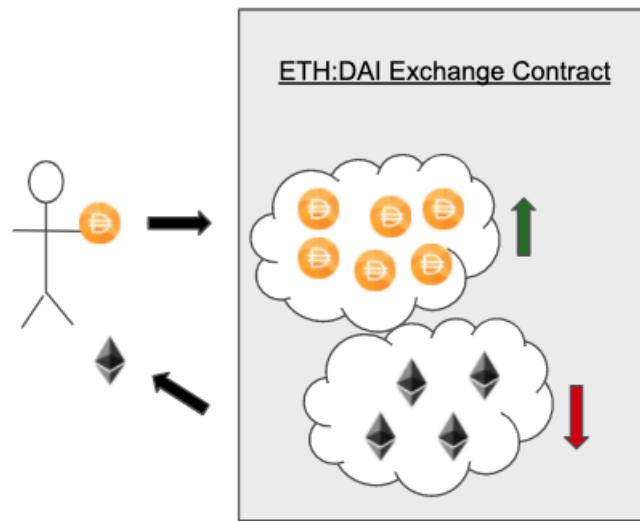


Figure 8.9: Liquidity Providing

Swapping ERC20 Tokens In a single atomic transaction, Uniswap helps users to exchange an ERC20 token directly for another ERC20 token. The consumer has DAI and wants to substitute for MKR in the figure 1.11. The user then calls `tokenToTokenSwap` function and adds DAI to the DAI pool and kicks ETH to the MKR pool and returns the `tokenToTokenSwap` to the address that the transaction initially sent.

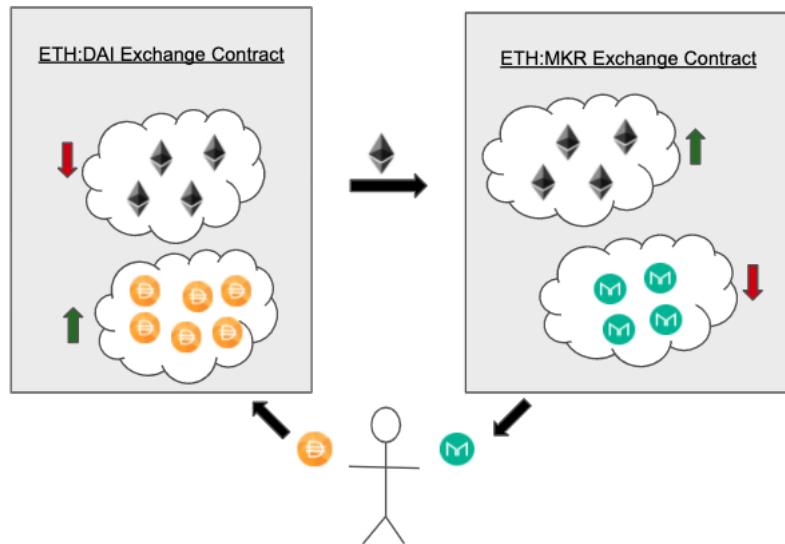


Figure 8.10: Liquidity Providing

8.4.4.2 P2P Lending

MAKER, Compound , AAVE : are the popular Decentralized Lending platforms. This platform makes the full ownership of both the lender and the borrowers' funds with no hassle over the interest rate and the locking time frame. Compound is the largest lending platform in the market with \$1.66 billion, 2.8 million ETH and 91.4 thousand BTC locked into the platform and \$0.5–1 billion dollars daily.

Compound is an algorithmic money market protocol on Ethereum which helps users to acquire interest or borrow assets against collateral. Anyone can supply assets to the liquidity pools of Compound and gain interest automatically. Tariffs for the supplied amount are dynamically adapted according to supply and demand. The lending mechanism in the crypto industry is primarily over-collateralized, i.e. the debt value of the token is less than the loan amount of the token \times collateral factor ⁴ In exchange for loans issued ctokens. This report discusses three use cases of Compund:

1. Supplying tokens
2. Borrowing tokens

Supplying Tokens

When anyone supplies an asset to the Compound protocol the liquidity token "cToken" is given. For example, if you lock Dai, you will obtain a cDai in exchange. cTokens is a deposit partition which demonstrates that you have added liquidity to the protocol and can also be used to recover the assets that are given, plus the accumulated interest, for each token that is locked in the protocol.

⁴The maximum amount users can borrow is limited by the collateral factors of the assets they have supplied.

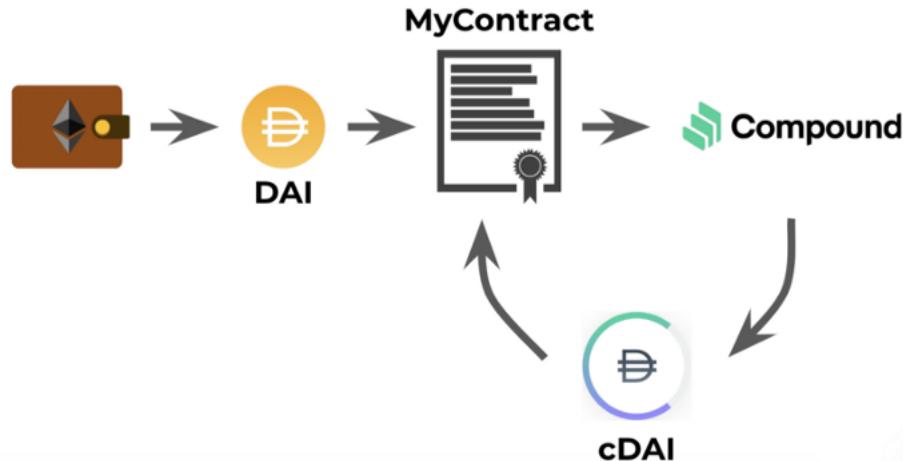


Figure 8.11: Liquidity Providing

Borrowing tokens Upon locking up assets and holding cTokens, users may borrow various other ERC20 tokens from the protocol. For instance, a user who locks Dai gets cDai, can then borrow BAT from this cDai as a collateral. It is important to remember that for each asset there are certain collateral ratios dependent on collateral efficiency. If the price of the asset you borrowed versus the one that you supplied diverges too far, then you have the possibility of having a portion of your collateral liquidated.

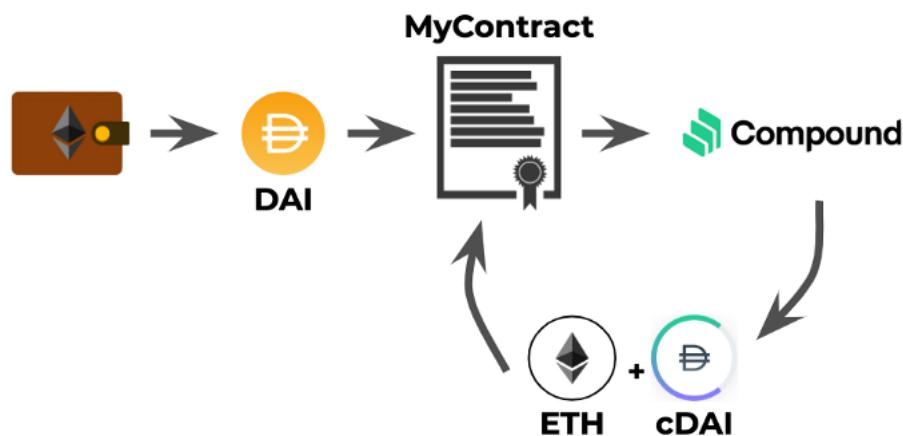


Figure 8.12: Liquidity Providing

8.4.5 Energy

The energy industry has been consistently adopting innovative solutions including rooftop solar, electric vehicles, and smart metering. Now, with the worldwide adoption of BC as it offers advanced features through its smart contracts and systems interoperability energy sector is poised to grow in terms of efficiency, innovative use cases and technology adoption in small communities. The World Economic Forum, Stanford Woods Institute for the Environment, and PwC released a joint report identifying more than 65 existing and emerging blockchain use-cases for the environment. BC offer unique solutions for renewable energy distribution and has the potential to improve efficiencies for the utility providers by tracking the chain of custody for grid material. [7]

Some of the reasons why BC is been considered by the energy industry to resolve some of its current problems:

1. Cost reduction for energy providers and consumers.
2. More innovative solutions for clean energy and environmental sustainability.
3. Increased transparency for everyone between regulators to end-users.
4. Enhanced data privacy.
5. Less dependencies on intermediaries.
6. Single points of failure reduces vulnerability.

The simulation result shows that electricity expenses of a household has reduced up to 44.73%.

8.4.5.1 P2P Energy Trading

A Blockchain in Energy report by Wood Mackenzie from a while back shows that 59% of blockchain energy projects are building P2P energy markets, shows an interesting adoption in inclusive and flexible energy market.

BC enabled P2P energy trading allows consumers to buy and sell excess energy from prosumers directly benefiting the masses as it reduces control from central authorities, plus it allows users to be truly in charge of their energy supply. BC makes the trade secure and real time so that users can track the usage ensuring the transactions are smooth and safe. With the increase in local energy generation from Renewable Energy Sources(RESs), the concept of decentralized P2P Local Energy Market (LEM) is becoming increasingly popular. This also allows the possibilities which allows end users to shift their load to off-peak hours and to use cheap energy from the LEM. Many applications that offer such solutions allow users to maximize their economic gains at both neighborhood and end-user levels in terms of energy use and power prices, and provide the LEM with adequate energy. Furthermore, smart applications of BC can unveil a surplus of unique, game changing solutions for energy distribution.[8]

8.4.5.2 Robotina ROX

Robotina founded in 1990 that focuses on innovation in automation and control technology. A bold idea of converting the energy systems that run your home into a full-fledged BC ecosystem is the foundation for the platform. RobotinaRox allows its users to trade and receive cryptos. It uses integrations of BC and IoT also allowing the cryptos to be converted to electricity afterwards. In addition to having full control over the energy expenditures, all energy sources in the RobotinaRox network are eco-friendly, so there is a good sense of protecting the natural environment that is in line with the BC initiative. Platform will be a vertical, high-tech all-inclusive enabling solution, consisting of connected IoT elements (things, processes, data, people) and the Power Platform (PP). It will use Smart Rules, Artificial Intelligence and BC technologies. Robotina Platform is a crypto-friendly token used in all internal transactions, and the Robotina Utility token (ROX) is an exclusive token. The client gets a special discount while paying with ROX (one of the ROX utility features), while the services and products sold on the Robotina site could also be bought with other cryptocurrencies, tokens or fiat currency.[9]



Smart Contract



IoT



Artificial Intelligence



Cloud computing



Battery Storage

Figure 8.13: RobotinaROX Platform Offerings

The IoT movement represents a true digital revolution involving the connection of items in the value chain that interact through the internet, people, processes and data. Considered as Internet-connected end-nodes to form communities, collaborate on usage and exchange data. Physical sensors, devices, actuators and other items generating data or receiving information from other sources. To make intelligent decisions and control mechanisms, raw data is analyzed and transformed into meaningful information. IoT, with local intelligence, senses and controls the properties. They are linked to cloud-based services that process data and customize procedures to better suit the needs and requirements of the customer.[9]

Robotina uses Ethereum to enforce trust among the consumers using a turing complete programming language. Robotina Utility Token(ROX) which integrates the whole process as token based in the life cycle of the transaction. AI and machine learning to build and constantly improve models, which allow us to predict consumption and we use cognitive optimization to achieve optimal results.[9]

8.4.6 Supply Chain

The supply chain is very crucial for a business to turn raw materials for the customer into finished goods and services. It begins with the raw material being harvested. Crops, poultry, wood, gold, or other natural resources might be the commodity. The commodity then goes to the manufacturer. That is when it transforms into a finished product. For instance, the contribution of the apple industry to the economy across the supply chain, US apple producers harvested 10.4 billion pounds in 2016, with approximately 61% of the crop produced in Washington, 10% in New York, and 9% produced in Michigan. The production of NYS was almost 1.2 billion pounds, estimated at over \$317 million. The aggregate NYS apple sector directly contributes \$1.3 billion in cumulative production, 8,033 jobs, and \$397.9 million in gross domestic product (GDP) to the New York State economy in 2016.[10]

The COVID-19 crisis has shaken the world's supply chains and raised fundamental concerns about the future of trade. The stabilization and reconstruction of economic growth

is crucial to regaining interest in these processes. Although digitization has dramatically lowered processing costs, most company domains are still working in silos, causing accounting inconsistencies that need to be aligned. For business performance, the need to process transactions rapidly and validate the development, delivery and receipt of a specific exchange of value is increasingly important. Transparency and credibility across disciplines would be required to render a supply chain robust, which can be strengthened by the implementation of BC technologies.

In January 2018, IBM and global shipping pioneer Maersk launched an effort to create a new BC-based global trading network. The aims were clear and optimistic at the same time, to reduce the expense of global distribution, increase visibility across supply chains and remove paper-based process inefficiencies. IBM projected that going fully digital could save around \$38 billion per year for shipping carriers.[11] A test case using avocados, shipped from Mombasa to Rotterdam was carried out by Maersk and IBM. IBM estimated that the cost of transporting the shipping container itself was approximately \$2000, with documentation contributing approximately \$300-between 15% and 20% of the overall cost. Therefore, this technology has the potential to, among many others, generate huge and scalable savings in this field. With input from various stakeholders throughout the global supply chain, TradeLens is built on open standards. It is a new approach that provides tremendous benefit around the community to all members.[11]

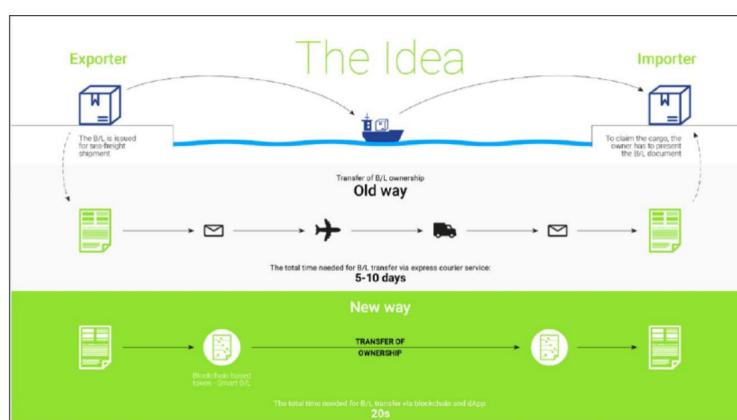


Figure 8.14: Comparison between solution provided and TradeLens

Around 100 organizations, including ports, customs authorities, container carriers, freight forwarders and cargo owners, have already agreed to participate in the TradeLens platform. The main issues addressed by the platform are boosting information exchange, improving communications between companies. parts and reduce shipping time.

1. **ECOSYSTEM:** The industry network of shippers, freight forwarders, ports and airports, ocean carriers, intermodal companies, government agencies, customs brokers and more form the base of TradeLens. Each agency exchanges data that can be tracked, processed and managed throughout the network during the trajectory of a shipment.
2. **PLATFORM:** The TradeLens Network is available through an open API and, via a series of open standards, ties together the ecosystem. The network, powered by Hyperledger Fabric blockchain technology and IBM Cloud, helps the industry to safely exchange knowledge and collaborate.
3. **MARKETPLACE:** Open Marketplace Software and Services helps both TradeLens and third parties to publish fit-for-purpose services on top of the TradeLens network, encouraging competition in the supply chain and building value. An transparent ecosystem for software and resources enables both TradeLens and third parties

to publish fit-for-purpose services on top of the TradeLens network, encouraging competition in the supply chain and value generation.

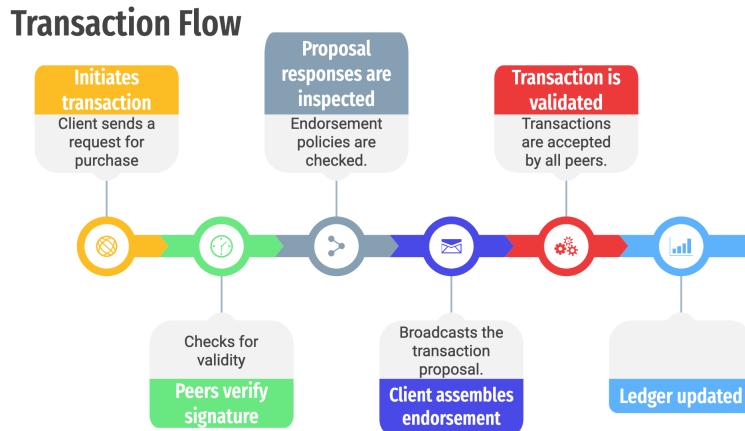


Figure 8.15: High Level transactional flow

An approved ledger called Hyperledger BC is used by Tradelens. Hyperledger Fabric is an open source distributed ledger technology (DLT) framework for enterprise-grade authorization. The platform is permissioned, meaning that the users are identifiable to each other, rather than anonymous and thus totally untrusted, unlike a public permissionless network. It supports pluggable consensus protocols that allow the platform to be more easily configured to meet individual usage cases and trusted models. Complete Byzantine fault tolerant consensus can be deemed needless and an unacceptable drag on efficiency and throughput while implemented within a single organization, or managed by a trustworthy authority. Fabric can leverage consensus protocols that do not require a native cryptocurrency to incentivize expensive mining or drive smart contract execution. Cryptocurrency avoidance eliminates certain major risk/attack vectors, and the absence of cryptographic mining activities ensures that the platform can be implemented at about the same running expense as any other distributed framework.[12]

8.4.7 Agriculture

The global food chain is highly multi-stakeholder and diversified, with a wide variety of stakeholders involved, such as farmers, shipping companies, wholesalers and retailers, traders and food. AgiDigital led the world's first ever agreement to sell 23.46 tonnes of grain on a BC (ICT4Ag, 2017). Since then, over 1,300 users and more than 1.6 million tons of grain have been processed through the cloud-based system, which included \$ 360 million in producer payments. [13] BC applications are increasingly used in supply chain management and are expected to increase from 45 million US dollars in 2018 to 3314.6 million US dollars by 2023 with an average growth rate of 87 percent. Current policies, programs, and economic impact were the focus of our interest. Based on paper[13], there 23 relevant papers identified and were divided into six main categories as follows:

1. food security (2 projects/initiatives, 4%)
2. food safety (3 projects/initiatives, 6%)
3. food integrity (24 projects/initiatives, 49.5%)
4. support of small farmers (8 projects/initiatives, 16%)
5. waste reduction and environmental awareness (5 projects/initiatives, 10%)

6. better supervision and management of the supply chain (7 projects/initiatives, 14.5%)

Goods and products, in relation to projects using blockchain technology and their overall objectives.		
Goods, Products, Resources	Initiative/Project/Company Involved	Objectives
Soybeans	LDC (Hoffman & Munsterman, 2018)	Financial, Faster Operations
Grains	AgriDigital (AgriDigital, 2017), GEBN study (Lucena et al., 2018)	Financial, Supervision and management
Olive oil	OlivaCoin (OlivaCoin, 2016)	Financial, Small farmers support
Turkeys	Cargill Inc. (Bunge, 2017), Hendrix Genetics (Hendrix Genetics, 2018)	Traceability, Animal welfare
Mangoes	Walmart, Kroger, IBM (CB Insights, 2017), (Kamath, 2018), Nestle (ITUnews, 2018)	Traceability
Canned pumpkin	Nestle (ITUnews, 2018)	Traceability
Pork	Walmart, Kroger, IBM (CB Insights, 2017), (Kamath, 2018)	Traceability
Sugar cane	Coca-Cola (Gertrude Chavez-Dreyfuss, Reuters 2018)	Supervision and Management
Beer	Downstream (Ireland Craft Beers, 2017)	Traceability
Beef	"Paddock to plate" project (Campbell, 2017), BeefLedger (BeefLedger Limited, 2017), JD.com (Adele Peter, Fast Company 2017)	Traceability
Cannabis	Medical Cannabis Tracking (MCT) system (Abelseth, 2018)	Traceability
Chicken	Gogochicken (Adele Peter, Fast Company 2017), Grass Roots Farmers Cooperative (Grass Roots Farmers' Cooperative 2017), OriginTrail (OriginTrail, 2018)	Traceability
Wood (Chestnut trees)	Infotracing (Frigorilli et al., 2018)	Traceability
Sea-food	Intel (Hyperledger, 2018), WWF (WWF, 2018), Balfegó (Balfegó Group, 2017)	Environmental impact, Traceability
Table grapes	"Blockchain for a grape" project (Ge et al., 2017 , pp. 2017–2112), Grape farm near the City of Skopje (Davcev et al., 2018)	Experimental feasibility study, Supervision and management
Organic food	Soil Association Certification (Soil Association Certification, 2018)	Financial, Traceability, Small farmers support
Food waste	Plastic Bank (Plastic Bank, 2019), Agora Tech Lab (Agora Tech Lab, 2018), SNCF (SNCF, 2017), Recereum (Recereum, 2017), Swachhcoin (Swachhcoin, 2018)	Waste reduction
Water	Global water assets (Poberezhna, 2018 , pp. 189–196)	Supervision and management
Rice	Quality of rice in transportation (Kumar & Iyengar, 2017)	Supervision and management
Food chain in general	AgriLedger (AgriLedger, 2017), FarmShare (FarmShare, 2017), Carrefour (Carrefour, 2018), ripe.io (Ripe.io, 2017), OriginTrail (OriginTrail, 2018), AgriBlockIoT (Caro et al., 2018a, b), Food supply chain prototypes enhanced with other technologies (Tian, 2017), (Kui et al., 2018), (Boehm et al., 2017)	Financial, Traceability, Food safety, Small farmers support, Waste reduction, Supervision and management

Figure 8.16: Blockchain initialives in Goods and Products

8.4.7.1 Ambrosus

Ambrosus is an optimized end-to-end solution that requires hardware, applications, a protocol layer, and resources from the developer. Ethereum BC and other distributed technologies are based on the Ambrosus protocol and the software layer on top of it, allowing information from IoT devices to be recorded on a decentralized network. To promote the development of applications on top of the network, we are also developing a developer toolkit. This provides an interactive API suited to the needs of participants in the food and pharmaceutical supply chain, one that is sufficiently scalable for use of any form of product or consumer good. On top of the BC framework, a JavaScript interface for Ambrosus enables users to build on our platform without any knowledge of BC programming.[14]

The Amber token, an ERC-20 compliant token that supports all transactions on the Ambrosus network, is the core of the Ambrosus network. Amber is the world's first data-linked token. It is used to constantly connect the modified and verified supply chain, logistical, environmental and biological data to the corresponding product as it moves between counterparts.[14]

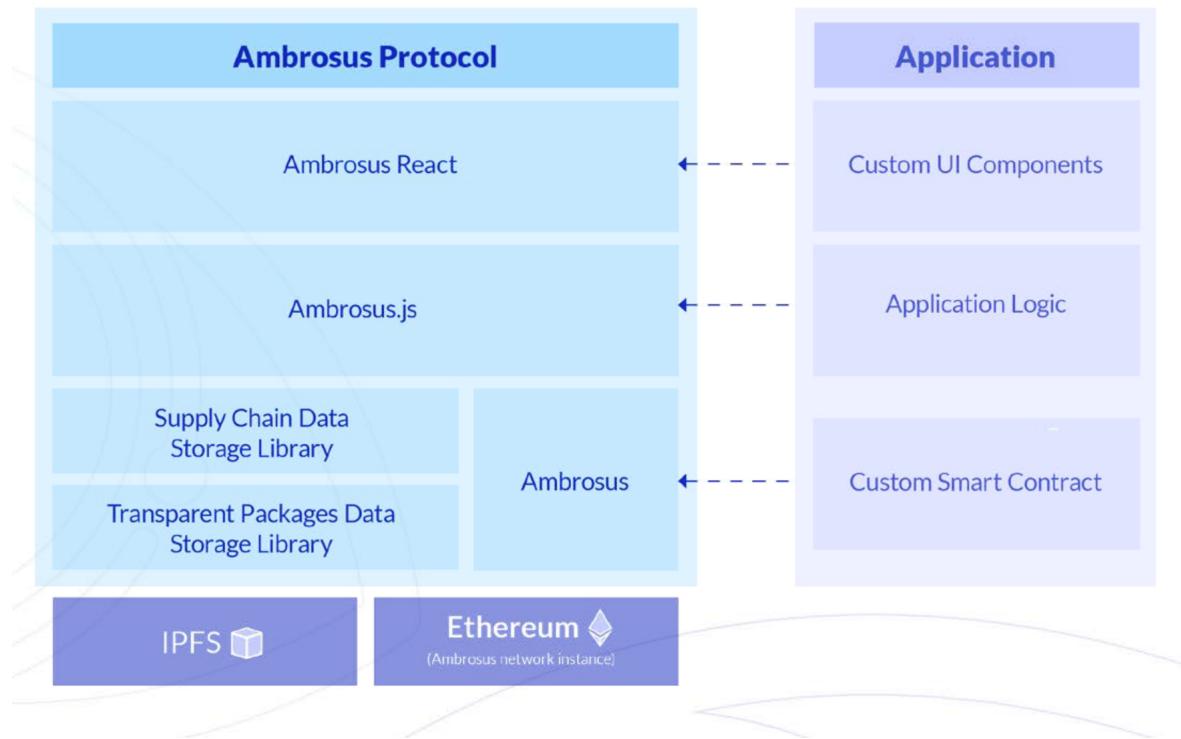


Figure 8.17: Dependency diagram for an application using Ambrosus

Ambrosus architecture uses an additional three-layer architecture for storing data:[14]

1. The first layer is a library used to store vast volumes of small data for blockchain and distributed file structures. In a transparent, permanent and immutable way, this layer can store arrays of individual data. Signed data and Merkle trees are central topics here.
2. The second layer is dedicated to supply chains. It employs concepts such as Measurements and Requirements Smart Contracts.
3. Ambrosus.js, a protocol devoted to food and pharmaceutical supply chains, with unique calculation forms and specifications applicable to those sectors, is the top layer.

8.4.8 HealthCare

BC is now gaining immense interest in healthcare. The global healthcare industry spend on BC is forecast to reach \$5.61 billion by 2025, according to a study by BIS Research. By 2025, the introduction of BC technologies will save the healthcare sector up to 100–150 billion a year in data breach costs, IT costs, operating costs, support feature costs and staffing costs, and by reducing fraud and counterfeit goods by 2025.[16]



Figure 8.18: Key problem area for Blockchain

To retain medical records today, most health-care providers still rely on obsolete programs. These services have the advantage of keeping geographic archives of patient data. This will make the doctor's diagnosis difficult, which is time-consuming for the doctor and also tedious for the patients. As a result, there is a significant rise in the expense of running a patient-oriented business. Issues that prevail in the healthcare industry are not limited. The need for a technically sophisticated computer should not be ignored. Consider the topic of opioid counterfeiting, resulting in damages of about \$200 million. Another time-consuming and boring method that results in high prices for the healthcare sector is the Patient Information Exchange. Since patients have no oversight of their documents, there are regular threats of identity breaches, spamming, and financial data crime.

8.4.8.1 MediLedger

The value of a blockchain-based framework is that, without exchanging private details, rivals will cooperate on a common network to, say, improve drug protection. That is precisely the concept behind the MediLedger Network, a group that includes pioneers including Gilead, Pfizer, Amgen, Genentech, AmerisourceBergen, and McKesson among its founders, focused on pharmaceutical supply chains. The underlying infrastructure is being provided by Chronicled, a startup.

MediLedger's first manufacturing approach is a medication authentication device that makes it easier to verify that a returned drug is genuine, a popular yet demanding practice. About 60 million units of saleable medications are returned yearly, according to the Healthcare Distribution Partnership. Wholesalers who must email suppliers to trace serial numbers, a procedure that may take up to 48 hours, must check that these drugs are genuine before they are resold. Using BC, wholesale dealers may use a barcode scanner to do the same verification in less than a second, meaning that goods can be placed back into commercial circulation easily, and producers retain ownership of their records. It is also possible to use this quick serial number verification to assist hospitals and pharmacies. With no infrastructure except a web browser and a barcode scanner, once it is put on the shelves, workers can check if a prescription is genuine. In an effort to pass drugs off as legal, counterfeiters may also copy barcodes, but the ledger would flag and permanently document suspicious behavior. MediLedger successfully completed a trial with 25 participants using a BC for track and trace, including Walgreens and Walmart stores, FedEx transport provider, GS1 standards association, bulk distributors including AmerisourceBergen and Cardinal Wellness, and suppliers ranging in scale from 100 to 125,000 employees. [15]

Transactions and life cycle of the medicine can be traced with just a web application:[17]

1. Solution provider will create a private node on your behalf to connect to the Mediledger Network. This is where your data will be stored.
2. Your products (GTIN's) are added to the lookup directory. The lookup directory is like a phone book that used for verification requests to the location of the repository where a manufacturer can respond to serial number's authenticity.
3. Directory updated and pushed in all the nodes after reaching an consensus.
4. Now any re-seller or distributors can look up the product with GTIN.

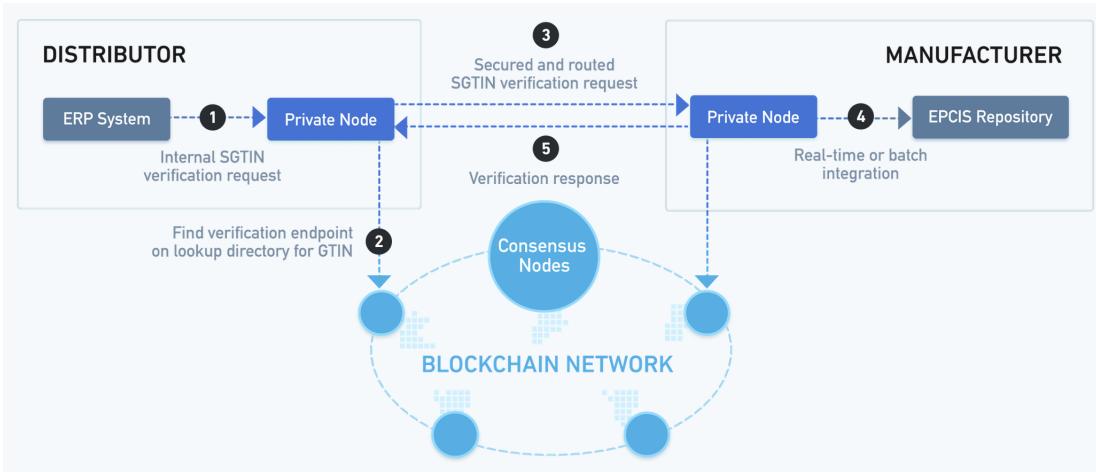


Figure 8.19: High level MediLedger Architecture

Industry	Use cases	Start-ups
Energy, utilities & mining	<ul style="list-style-type: none"> • Smart utility metering system • Decentralised energy data platform 	Bankymoon AutoGrid
Entertainment & media	<ul style="list-style-type: none"> • Control of ownership rights of digital media • Disintermediation of record labels 	Ascribe Mycelia
Financial services	<ul style="list-style-type: none"> • International P2P transactions • Anti-money laundering 	Bitcoin Coinfirm
Government & public services	<ul style="list-style-type: none"> • Land ownership records • Tamper-proof voting records • Digital identity of citizens 	Factom Follow My Vote Tradle
Healthcare	<ul style="list-style-type: none"> • Storage of healthcare records • Population health and clinical studies 	HealthNautica Tierion
Hospitality & leisure	<ul style="list-style-type: none"> • Loyalty programmes 	Loyyal
Insurance	<ul style="list-style-type: none"> • Peer-to-peer flight insurance policies • Micro-insurance 	InsurETH Stratum
Transportation & logistics (freight transport)	<ul style="list-style-type: none"> • Trade documentation (e.g. Bill of Lading) • Trade finance • Supply chain transparency 	Wave Skuchain Provenance
Transportation & logistics (aviation)	<ul style="list-style-type: none"> • Distribution of tickets and ancillary services • Loyalty programmes (cf. H&L) • Passenger identity management 	Loyyal

Figure 8.20: Other use-cases

8.5 Summary and Conclusions

The use of BC technology comes at a staggering cost. To create and manage a system that is based on BC technology, we have to consider the fact that it is relatively expensive and unfeasible it is to maintain such a system on a larger scale. First, developing on a system that uses BC technology is slow and restricted. This is because the idea of BC is fundamentally based on consistency, and that there is always a compromise between speed and consistency in development. Also, scaling becomes a major issue when dealing with a decentralized BC system, especially since the same piece of data in a BC database resides in many different places. As such, costs of scaling a single piece of data has to be incurred on the entire database, instead of on a single piece of data in a conventional database.

Bibliography

- [1] Yahaya, A.S.; Javaid, N.; Alzahrani, F.A.; Rehman, A.; Ullah, I.; Shahid, A.; Shafiq, M.: *Blockchain Based Sustainable Local Energy Trading Considering Home Energy Management and Demurrage Mechanism*, Energy Management in a "Smart Home" with Integrated Solar Cells, April 2020. <https://www.mdpi.com/2071-1050/12/8/3385>.
- [2] Edvard Tijan, Sasa Aksentijevi, Katarina Ivanic and Mladen Jardas : *Blockchain Technology Implementation in Logistics*, Faculty of Maritime Studies, University of Rijeka, February 2019. https://www.researchgate.net/publication/331325030_Blockchain_Technology_Implementation_in_Logistics.
- [3] Blockimmo *FACILITATING AN ACCESSIBLE & STREAMLINED REAL-ESTATE MARKET*, Blockimmo Whitepaper, November 2018. https://s3.eu-central-1.amazonaws.com/assets-prod-protected-jm01tirz25y4/blockimmo_whitepaper_web_201811.pdf.
- [4] Marija Jovic, Marko Filipovic, Edvard Tijan, Mladen Jardas : *A Review of Blockchain Technology Implementation in Shipping Industry*, Faculty of Maritime Studies, University of Rijeka, September 2019.
- [5] Oleksii Konashevych : *General Concept of Real Estate Tokenization on Blockchain*, Erasmus Mundus Joint International Doctoral Fellow in Law, Science and Technology (EU), June 2020. https://www.researchgate.net/publication/331325030_Blockchain_Technology_Implementation_in_Logistics.
- [6] Sundeep Tanvar, Umesh Bodkhe : *Blockchain for Industry 4.0: A Comprehensive Review*, Department of Corporate and Information Services, NTG of Australia. <https://ieeexplore.ieee.org/document/9069885>.
- [7] *The Transformative Impact of Blockchain in the Energy Sector*, <https://www.entrepreneur.com/article/353519>.
- [8] Adamu Sani Yahaya, Nadeem Javaid, Amjad Rehman, Fahad A. Alzahrani : *Blockchain Based Sustainable Local Energy Trading Considering Home Energy Management and Demurrage Mechanism*, <https://www.mdpi.com/2071-1050/12/8/3385>.
- [9] *Robotina ROX ICO whitepaper, IoT, Artificial Intelligence and blockchain empowering energy consumers*, https://robotinarox.io/wp-content/uploads/2018/07/Robotina_WP.pdf.
- [10] Todd M. Schmit, Roberta M. Severson, Jesse Strzok and Jose Barros : *Economic Contributions of the Apple Industry Supply Chain in New York State*, Charles H. Dyson School of Applied Economics and Management Cornell University. <https://dyson.cornell.edu/wp-content/uploads/sites/5/2019/02/Cornell-Dyson-eb1803.pdf>.

- [11] *Hyperledger Fabric release 2.2 doc*, <https://hyperledger-fabric.readthedocs.io/en/release-2.2/txflow.html>.
- [12] Andreas Kamaras, Agusti Fonts, Francesc X. Prenafeta-Bold : *The rise of blockchain technology in agriculture and food supply chain*, https://www.academia.edu/40259056/The_rise_of_blockchain_technology_in_agriculture_and_food_supply_chains.
- [13] *Ambrosus whitepaper*, <https://ambrosus.com/assets/en/-White-Paper-V8-1.pdf>.
- [14] *Harvard Business Review*, Why Big Pharma Is Betting on Blockchain <https://hbr.org/2020/05/why-big-pharma-is-betting-on-blockchain?ab=hero-main-text>.
- [15] *Healthcare Weekly*, The Global “Blockchain in Healthcare” Report: 2020 <https://healthcareweekly.com/blockchain-in-healthcare-guide/>.
- [16] *MediLedger Solutions*, MediLedger Solutions <https://www.mediledger.com/solution-protocols>.
- [17] *INTRODUCTION TO SMART CONTRACTS*:<https://ethereum.org/en/developers/docs/smart-contracts/>
- [18] *How ICO and STO powered by Blockchain platform is transforming modern day start-ups for crowdfunding — A Very Extensive Look*:<https://medium.com/@BangBitTech/ico-vs-ipo-44e1da262459>
- [19] *What are Dapps*:<https://decrypt.co/resources/dapps>
- [20] *What are Smart Contract Wallets*:<https://cryptotesters.com/blog/what-are-smart-contract-wallets>
- [21] *What Are NFTs and How Can They Be Used in DeFi* :<https://finematics.com/what-are-nfts-and-how-can-they-be-used-in-defi/>
- [22] *Atomic Swaps Explained*:<https://academy.binance.com/en/articles/atomic-swaps-explained>
- [23] *Understanding Uniswap*:<https://docs.ethhub.io/guides/graphical-guide-for-understanding-uniswap/>
- [24] *Guide to Compound*:<https://docs.ethhub.io/guides/graphical-guide-to-compound/>
- [25] *BLOCKIMMO Whitepaper*:https://s3.eu-central-1.amazonaws.com/assets-prod-protected-jm01tirz25y4/blockimmo_whitepaper_web_201811.pdf
- [26] *Decentraland Whitepaper*:<https://decentraland.org/whitepaper.pdf>
- [27] *Provenance Whitepaper*:<https://www.provenance.org/whitepaper>
- [28] *District0x White Paper*:<https://district0x.io/docs/district0x-whitepaper.pdf>
- [29] Sergei Chertishev *Can blockchain resolve the crisis of trust in commodity trade finance?*:<https://www.linkedin.com/pulse/can-blockchain-resolve-crisis-trust-commodity-trade-chertishev/>

- [30] Alan Llyod Paris, Sai Surendra kumar, Srinivasa Manikant Upadhyayula : *Blockchain in Trade finance: a new paradigm to fight trade-based money laundering*
- [31] HAYA R. HASAN AND KHALED SALAH : *Proof of Delivery of Physical Assets Using Blockchain and Smart Contracts*
- [32] <https://sxi.io/ubs-unveils-blockchain-for-trade-finance-at-sibos/>
- [33] <https://www.ijert.org/blockchain-the-power-of-supply-chain-40/>
- [34] <https://www.smart-energy.com/industry-sectors/policy-regulation/blockchain-feature>
- [35] <https://www.fintechfutures.com/2018/04/first-pilot-client-transactions-executed-on-batavia-trade-finance-platform/>

Chapter 9

Survey and Analysis of Models for Cybersecurity Economics

Jan Bauer and Yung-Hsin Chen

The goal of this seminar is to investigate and describe the main metrics and state-of-the-art approaches available for cybersecurity economics. Also, a technical discussion is provided regarding their possible application scenarios, followed by a comparison of its benefits and drawbacks. Furthermore, opportunities and challenges for cybersecurity economics have been investigated and discussed.

Contents

9.1	Introduction	264
9.1.1	History of Cybersecurity Economics	265
9.2	The Economics of Cybersecurity	265
9.3	Economical Models for Cybersecurity	268
9.3.1	Return on Security Investment (RoSI)	268
9.3.2	SEconomy Framework	270
9.3.3	SEconomy Framework Example	272
9.3.4	Gordon Loeb model	275
9.3.5	Conclusion	277
9.4	Case Studies: Economical Impacts of Cyberattacks	277
9.4.1	Security in Practice	278
9.4.2	WannaCry	280
9.4.3	Stuxnet	281
9.4.4	Sony Hack	282
9.5	Conclusion	282

9.1 Introduction

"In a digital era, data is the new gold [1]." This status makes data attractive to criminals and putting it at risk of cybercrime. Regardless of whether you are a private person, company or state organization, your sensitive data and IT systems need to be protected. The vulnerability to business disruption through Denial-of-Service (DoS) attacks, theft of proprietary data and intellectual property increases with the level of digitization. The economic damage caused by malicious cyber activity is immense. The cost for the U.S. economy total between \$57 billion and \$109 billion in 2016 alone [2] and the annual global spend on cybersecurity is approaching \$100 billion [3]. Cybercrime is mainly fueled by the financial profit expectations of the perpetrators, making it an economic issue. It has relatively low risk compared to other types of crime. Cybercriminals, being virtually invisible, are rarely caught and convicted. This makes cybercrime an attractive business opportunity. Since the problem is caused by economical thinking, it can be best solved by economical thinking [4][5]. IT security professionals must understand the tools as well as the attackers' motives in order to successfully protect their organization.

The incidence of cybercrime, cyberespionage and even acts of cyberwar is increasing. It is often difficult to estimate the associated costs and risks [6] because of different issues, such as information asymmetry and lack of investments in effective risk assessment. Society is spending more money to prevent bank robbery compared to the stolen amount, while spending less money on cybersecurity than the thieves and spammers steal [7]. Further economic biases will be discussed in the course of the paper. By better understanding the impact and risk associated with cybercrime, organizations and individuals can determine the right amount to invest. Models for cybersecurity economics help guide the way through this difficult terrain. The report aims to provide an overview of such models. Based on this, the effects on business and society are investigated. Economical models help executive leaders to estimate the right amount of money to invest in the protection of their IT systems.

9.1.1 History of Cybersecurity Economics

During the sixties and seventies, centralized computing using mainframe computers enabled the foundations for modern IT. However, these centralized systems were not yet connected to the outside world and were operated by trustworthy specialists. Therefore, it was enough to ensure the physical security of the mainframe and its connected user terminals [8].

Time sharing and dial-up connections made it possible for organizations to remotely use centralized computing resources. Companies used equipment to send the commands from their remote premise to the shared computing unit. The time on the computer was bought much the same way that households buy power and water from utility companies. Since the physical security of the command terminals was ensured by the participating organizations, it was well known who interacted with the computing resources. Therefore, ensuring physical security was enough to ensure overall system security [9].

The game changed with the emergence of the Internet. Individual parties did not know any longer who they were interacting with. The systemic vulnerability became evident when the Morris Worm, released in 1988, infected around 10% of the existing Internet within a couple of hours. Although, the Morris Worm was not designed with bad intent in mind, the incident made it clear that better security practices were necessary to protect critical data [10]. Since there were no formal laws regarding these types of incidents, the legislation had to be adapted to this new type of crime. In 1990, the United Kingdom's government passed the Computer Misuse Act, that criminalizes the act of accessing or modifying digital systems without appropriate consent or permission [11].

The first computer at the University of Minnesota suddenly went under attack from a network of 114 other computers infected with a malicious script called Trin00. This code caused the infected computers to send superfluous data packets to the university, overwhelming its computer and preventing it from handling legitimate requests. In this way, the attack knocked out the university computers for two days[12]. Operation Aurora in 2010 used Advanced Persistent Threats (ATP) to attack many US tech giants like Google, Adobe and Yahoo. This coordinated attack intended to steal intellectual properties like source codes [13]. The US election in 2016 was the biggest case of election hacking to date. Russian interference tried to damage the Clinton campaign, boost Trump's chances and spread distrust in American democracy overall [14]. The incident shows how extensive the influence of cybercriminals can be. In the future, cybercrime is expected to expand its reach by using tools like machine learning and artificial intelligence [15].

Ransomware like WannaCry, Petya, NotPetya wreaked havoc in 2017, targeting machines vulnerable to the EternalBlue and EternalRomance exploit. The ransomware encrypted all the data on the system and forced the user to pay a ransom in crypto-currency in order to regain the data. Since the ransomware was embedded within a worm, it was able to spread quickly within computer networks. Many hospitals were affected by the outbreak leading to surgeries and treatments being cancelled. This could have been prevented by following basic security principles like installing the available security updates. The incidents show how far reaching the effects of cyberattacks can be [16].

9.2 The Economics of Cybersecurity

Cyber security is not merely a technical issue, it is also subjected to many economical and political causes, making cybersecurity hard to achieve [17]. For instance, the confusing responsibility ascription, The outcome of Business Model, the Asymmetric Information and the Offense and Defense.

Negative Network Externalities

It is known from Metcalfe's law that the value of a network rapidly increases with the number of people using it [18]. The more shops accept credit cards as a payment method, the more usable it is for the credit card holders and vice versa. Furthermore, IT systems usually have high fixed costs and low marginal costs. This means it is expensive to produce the first copy of a software but then it is very cheap to produce an arbitrary number of copies. In such markets, the winner usually takes it all. The first company to acquire a big market share has a high chance of taking over the whole market in the end. Therefore, companies have to move fast and capture this critical market share as fast as possible in order to defeat their competitors. However, such negative network externality leads to a ship first and fix later mentality. This is dangerous in terms of cybersecurity [19].

The Confusing Responsibility Ascription

Wrong economic incentives can lead to unfavorable outcomes. One of the difficulties of making information security hard is that when the party which is responsible for the protection of the information will not be affected by the negative outcomes an attack. This can be illustrated by an example from the banking sector. If banking fraud happens to a customer of an American bank, the bank has to proof that the customer did something wrong. So the burden of proof is on the bank. If the same incident happens to the customer of a British bank, the customer has to proof the insecurity of the banking system. We see that the burden of proof is on the customer here. However, it is almost impossible for the customer to proof this. So the American banks have a high incentive to keep their systems secure since in case a mistake happens, they are in charge. On the other hand, putting the blame on the customers as the British system leads to a decline in the level of cybersecurity. This example shows how wrong economic incentives can shape the behaviour of organizations. Such problems occur when the party suffering from an attack is different from the party that is responsible for its avoidance [19].

A common sense among people is that they are willing to buy software to prevent their devices from being attacked. But on the other hand they are unwilling to buy software that could prevent their devices from being used as the slave of a DDoS attack. This is referred to as the Tragedy of Commons [19].

The Outcome of Business Models

Some business strategies may also cause the problems for information security. Firms tend to raise the cost of switching in order to lock the customer to their system and gain more profits out of them. However, this usually requires the use of proprietary and obscure architectures instead of relying on open, secure and tested standards [19].

Asymmetry of Information

An asymmetry of information between customers and suppliers may cause a high percentage of bad products. This can be explained by considering the following situation. Suppose there is a marketplace with 50% bad products and 50% good products. The good ones are to be sold for \$ 3000, while the bad ones are to be sold at \$ 1000. However only the merchant knows which products are good and which are bad. The customer will only be able to identify this after the purchase. The average market price would be \$ 2000. However, the merchant is not able to sell good products worth \$ 3000 at such a low price. He can only offer the bad products at this price. In this case, the customers suffer from the high price and the bad product quality due to the asymmetric information. Therefore, the customers know that they can only get bad products for the price at \$ 2000, making the price rapidly drop to the average price of bad products, which is \$ 1000. Due to the asymmetry of information the customer always has to assume the worst case. This tendency can drive down product quality in a regular marketplace [19].

Offense versus Defense

A very important detail is if a technology package favours offence or defense. In traditional armed conflicts, the defense has an advantage due to its consolidated position [20].

However it is the opposite with cybersecurity. In order to protect a system, a defender has to find and fix all vulnerabilities. The attacker on the other hand might be able to infiltrate a system through a single loophole. Although, a defender might be able to cover 99 of the 100 existing vulnerabilities of his system, an attacker might still get in through the last loophole [21]. Since large software systems like operating systems, webservers and applications comprise millions of lines of code, they certainly contain a high number of bugs. Usually the ratio between lines of code and bugs is around 3 to 5 percent. The ever increasing number of technologies applications and devices connected to the internet make it even harder for defenders to stay ahead. Therefore, the technical bias of IT systems is in favour of the attacker. This makes attacking much easier than defending [19].

9.3 Economical Models for Cybersecurity

Employees and organizations have become more and more concerned with cybersecurity issues. Carrying out a cybersecurity analysis is an initial step that can help improve the efficiency of cybersecurity investments. To characterize threats, and plan certain responses and investments, one should rely on proven methodologies. These methodologies are based on the quantification of system attributes according to metrics. In the following section, metrics and frameworks that help determine the efficiency of investments in cybersecurity will be introduced.

9.3.1 Return on Security Investment (RoSI)

Metrics are quantifying attributes, based on the measurements of several characteristics. It can be a number that tells whether an investment or action is effective or not. Security metrics measure whether the expected improvement is worth the investment, and whether it should be taken or not. The Return on Security Investment (RoSI) is a well-known metric and will be introduced in this section.

The concept of Return on Security Investment (RoSI) is pretty much the same as the Return on Investment (ROI)¹. But instead of measuring the profits gained from the investment, RoSI focuses on the loss prevented by the security investment. And thus, it is sufficient to measure the monetary value of risk. The monetary value of risk can be estimated by *quantitative risk assessment*, which is an analytical approach that can quantify the risk of a specific risk event. In the RoSI measurement, Single Loss Expectancy (SLE), Annual Rate of Occurrence (ARO), and Annual Loss Expectancy (ALE) are used.

- Single Loss Expectancy(SLE)

The SLE is the cost of loss due to a single incident. Since it is the summation of the loss, the value of the loss can be quite objective from company to company. The loss should include not only the direct loss, but also the indirect ones. For instance, the loss of a computer-lossing incident should include not only the computer itself, but also the loss of reputation, the investigation time, ...etc.

- Annual Rate of Occurrence (ARO)

The ARO is the frequency of an incident happening per year. Naturally, the probability of an incident occurring depends on some relevant factors. For instance, the ARO of a cyberattack depends on the implementation and effectiveness of anti-virus.

- Annual Loss Expectancy (ALE)

The ALE is determined by

$$ALE = ARO \times SLE \quad (9.1)$$

It describes the annual monetary loss from an expected incident. From the equation, it is obvious that it takes both the likelihood of the event happening, and the cost of it happening into account.

Since the purpose of RoSI is to measure whether it is worth the investment of security system, the concept is similar to the ROI calculation. But instead, the RoSI wishes to determine the effectiveness of the risk reduction. Thus, according to the quantitative risk assessments, the risk reduction can be determined as

$$risk\ reduction = ALE \times mitigation\ ratio \quad (9.2)$$

According to the concept of the ROI, the RoSI should follow

$$RoSI = \frac{risk\ reduction - solution\ cost}{solution\ cost} \quad (9.3)$$

¹ROI is the way to measure the benefits gained from a particular investment. It is calculated by (gain from investment - cost of investment) / cost of investment

If the RoSI is higher than 1, it means that the benefits gained from the cost of solution is more than the cost itself. However, if the RoSI is lower than 1, the decision of investment is better to be revisited.

The mathematical statements above might seem blur, and so the example below should help with the understanding of RoSI.

Example

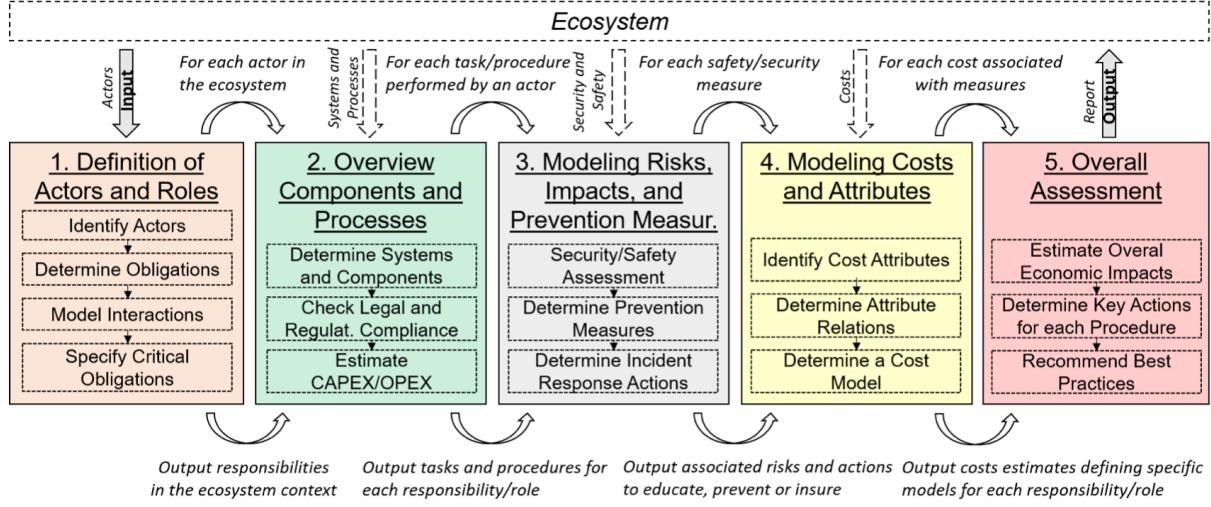
Suppose a company decides to invest a security system. Cyber attack strikes the company roughly 3 times a year, and the loss is approximately \$30000 including the data loss, the investigation fee, and so on. However, with the \$25000 anti-virus product, the probability of such incident can be reduced by 50%. Thus, the ROSI is

$$\begin{aligned} ROSI &= \frac{\text{risk reduction} - \text{solution cost}}{\text{solution cost}} \\ &= \frac{(ARO \times SLE) \times \text{mitigation ratio} - \text{solution cost}}{\text{solution cost}} \\ &= \frac{(3 \times 30000) \times 0.5 - 25000}{25000} = 0.8 \end{aligned} \tag{9.4}$$

According to what the ROSI suggests, the company should not invest in such a product.

9.3.2 SEconomy Framework

The SEconomy framework [22] is a step-based framework to measure economic impact of cybersecurity activities in a distributed ecosystem with several actors. The framework comprises five distinct stages as shown in the following figure.



Stage 1: Definition of Actors and Roles

In stage one, the analyst identifies relevant actors and their professional obligations. To be specific, relevant actors involve all those in the supply chain. Take the aircrafts for instance, the manufacturing of Airbus A380 involves manufacturers from 30 different countries. So in the stage one, the analyst should identify and list the duties and relevant regulations of all actors in the supply chains.

Stage 2: Overview Components and Processes

In stage two, the analyst specifies the ones with critical processes and components. The critical processes should not only be technically safe, but the actors operating them should also be able to be monitored. Meanwhile, the components should obey the security regulations. For instance, Boeing 737 MAX 8 were grounded immediately by the Federal Aviation Administration (FAA) after two fatal crashes, and are able to go back into services after the FAA allows [23].

Stage 3: Modeling Risks, Impacts, and Prevention Measures

In stage three, the costs are taken into consideration. This is done by assessing the risks, and then modeling costs according to the assessments, which will be done in stage four. To assess risks and threats, there are several risk assessments, such as STRIDE, LIND-DUN, and DREAD, that could categorize the threats and point out the corresponding mitigation measures.

However, since most systems do not operate independently, and thus, it is likely that the failure of a component might lead to the failure of others. In this case, *mapping dependencies (MD)* help calculate the covariance between failures.

$$MD(A, B) = \frac{p_A - p_A \times p_B}{\sqrt{p_A(1-p_A)p_B(1-p_B)}} \quad (9.5)$$

Stage 4: Modeling Costs and Attributes

The SEconomy is based on ROSI, and thus, the purpose here is to calculate the ROSI via cost modeling. According to ROSI, we need the risk reduction fee and the solution fee. Due to the fact that the costs of the mitigation do not necessarily form a linear relation, SEconomy follows it with a slight alternation, inputting the time dependency. For instance, a longer period mitigation procedure might not always mean the higher cost. To achieve this, *Reactive Mitigation Cost (RMC)* is introduced. It is a matrix describing the relation between time and cost as shown in the figure below.

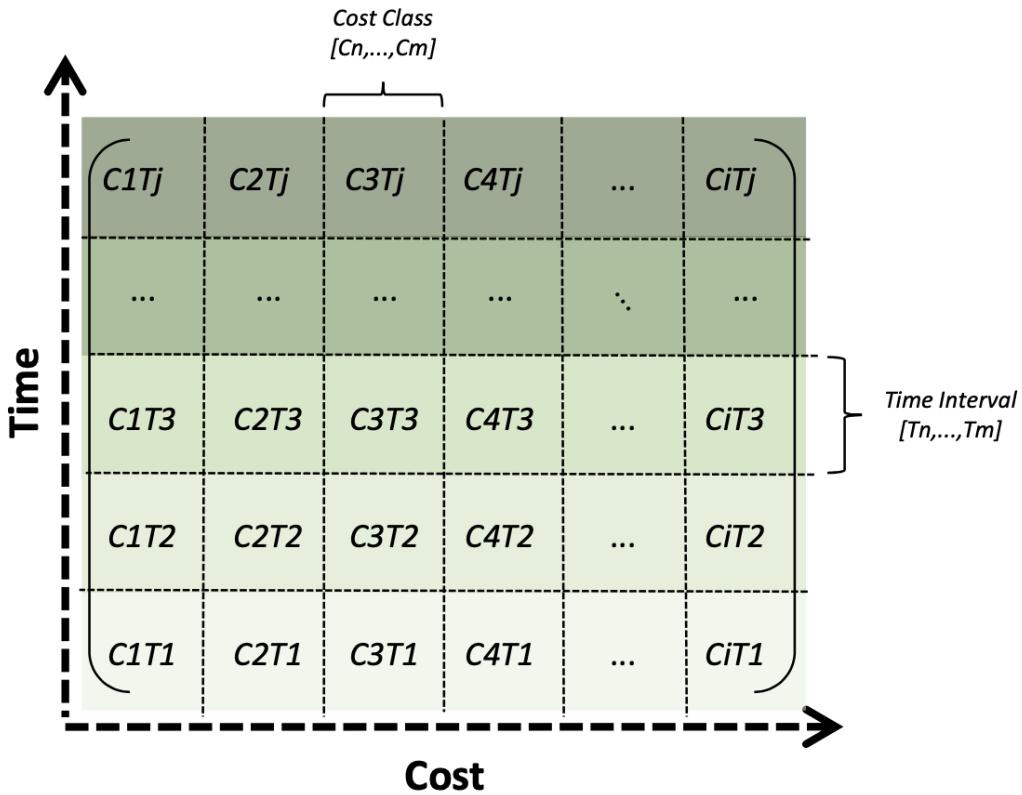


Figure 9.1: The MTC consists the cost function $f(C_i T_j)$ [22].

According to the alternation, the solution fee in SEconomy is transformed into the idea of *Proactive Mitigation Cost (PMC)*.

$$PMC(A) = \sum_{i=1}^{N_{threats}} \Delta T \times (\text{Proactive Cost} + \text{Insurance Cost}) \quad (9.6)$$

where A denote the system. Overall, PMC calculates sum of the proactive cost and the insurance cost in a period of time with N possible threats. Note that the proactive cost is the fee of precaution, while the insurance cost is to cover the unexpected fees.

Now, it is sufficient to calculate the corresponding term of the risk reduction term in ROSI. The risk reduction is composed of the ARO and the SLE, which are the similar idea of the MD and MTC in SEconomy. The term involving the MD and MTC is the *Reactive Mitigation Cost (RMC)*.

$$RMC(A, b) = \sum_{i=1}^{N_{system}} \left(\sum_{j=1}^{N_{threat}} MD(A, B) \times MTC[C_i][T_j] \right) \quad (9.7)$$

As for the threat vector, T_{cost} , it refers to the , which is the direct costs of each threat, for instance, the financial loss, the service downtime, business disruption and so on. Finally, the ROSI in SEconomy can be written as

$$ROSI = \Delta T \times \sum_{i=1}^{N_{system}} \frac{(T_{cost} \times RMC) - PMC}{PMC} \quad (9.8)$$

Stage 5: Overall Economic Assessment

As illustrated in the figure below, the steps in the previous four stages are summed up. By substituting all the terms into the calculations, SEconomy can define the overall estimated costs for the Ecosystem.

Algorithm 1: Overall Economic Assessment (OEA)

```

1 begin
2   for each Actor ∈ Ecosystem:
3     for each Role ∈ Actor:
4       for each System ∈ Role:
5         /* Correlation between linked systems in Equation 1 */
6         p(x) ← dependence(System, ∀ linkedSystems)
7         /* Estimate exposure costs in Equation 2 */
8         threatcosts ← Tcosts(A, p(x))
9         /* Estimate mitigation (Proactive and Reactive) costs
   in Equation 3 */
10        mitigationcosts ← PMCCosts(A)
11        mitigationcosts ← RMCCosts(A, p(x))
12        /* Get Overal Economic Assessment (OEA) in Equation 4
   */
13        OEA ← ROSI(threatcosts, mitigationcosts, InitSecCost)

```

Figure 9.2: The MTC consists the cost function $f(C_i T_j)$ [22].

9.3.3 SEconomy Framework Example

In order to understand the SEconomy framework, an example of the cybersecurity of Pfizer and BioNTech's COVID Vaccine is provided. Since COVID is causing serious disorder around the world, and has had a large impact on almost every aspects of our lives, being able to produce the vaccine not only leads to a large fortune, but also mitigate the negative impacts on economical and daily lives. In this case, cybersecurity for both companies are of high importance. Hence, we could apply this example to the SEconomy framework and see whether the investments are worth it.

Stage 1: Definition of Actors and Roles

In stage 1, the analyst should specify the duties of people involved in this whole supply chain, including

- The researching institutions: Pfizer and BioNTech develop the vaccine and try to keep the formula secure, and make sure that they don't release too much information in order to prevent imitations.
- Food and Drug Administration (FDA), European Medicines Agency (EMA): The FDA and EMA check the safety of the vaccine and enforce the regulation.
- Government: The governments define the regulations.
- Doctors and Hospitals: People from the medical institutes distribute vaccine and collect research data.
- Patients: The patients participate in the research.

Stage 2: Overview Components and Processes

In stage 2, the analyst should specify the critical processes that can cause the insecurity of the formula. The ones having the access to the formula, or the relevant data might be the ones that need to secure the privacy. For instance, the researching institutions, and the FDA, EMA. In addition, the researching institutions should comply with multiple governmental regulations as listed:

- Health Insurance Portability and Accountability Act (HIPAA, 1996)
- Federal Information Security Modernization Act (FISMA, 2014)
- General Data Protection Act (GDPR, 2018)

Stage 3: Modeling Risks, Impacts, and Prevention Measures

In stage 3, the analyst should define the categories of threats by some other frameworks. And then, the analyst can further define the mapping dependencies (MD). In the report, five likely threats are listed as:

- Industrial espionage by foreign government
- Intellectual property theft by competitors
- Installing spy as insider
- Malicious insider stealing data
- Employee faking results

With the risks categorised and listed, the analyst can define the MD accordingly. Suppose the MD is calculated by the analyst and it shows:

$$MD = \begin{bmatrix} 1.0 & 0.5 & 0.8 & 0.6 & 0.05 \\ 0.5 & 1.0 & 0.6 & 0.3 & 0.2 \\ 0.8 & 0.6 & 1.0 & 0.7 & 0.6 \\ 0.6 & 0.3 & 0.7 & 1.0 & 0.4 \\ 0.05 & 0.2 & 0.6 & 0.4 & 1.0 \end{bmatrix} \quad (9.9)$$

The MD shows the probabilities of two threats happening together. For instance, the probabilities of the formula stolen by the malicious insider, while the formula is faked is 0.4.

Stage 4: Modeling Costs and Attributes

In stage 4, in order to quantify the risks and calculate RoSI, the analyst should calculate the PMC, RMC, and the threat vector (T_{cost}).

For PMC, the analyst should evaluate the total proactive costs and insurance costs for each threats. Suppose the analyst does the evaluation, and comes up the PMC term as:

$$PMC = [3.1 \ 2.9 \ 3.8 \ 2.5 \ 4.2] \quad (9.10)$$

Next, the analyst should come up with the threat vector, T_{cost} , which is the direct costs of each threats. For instance, the financial loss, the service downtime, business disruption and so on. And the threat vector is:

$$T_{cost} = [4.3 \ 5.2 \ 4.1 \ 3.9 \ 5.0] \quad (9.11)$$

And what's left is the MTC, the cost of each threats with time dependencies. The analyst should evaluate the costs of the threats evolving in time. Suppose the analyst comes up with the MTC as:

$$MTC = \begin{bmatrix} 7.0 & 7.5 & 8.0 & 7.5 & 9.0 \\ 4.5 & 5.0 & 5.0 & 5.5 & 5.0 \\ 3.0 & 4.0 & 3.5 & 4.0 & 3.0 \\ 0.3 & 0.4 & 0.5 & 0.6 & 1.0 \\ 0.4 & 0.5 & 0.5 & 0.5 & 0.9 \end{bmatrix} \quad (9.12)$$

The first column of the MTC represents the costs evolving in time once threat "Industrial espionage by foreign government" occurs. Initially, the stock market volatility might cost, and after a while, other companies might come up with useful vaccine as well. In this case, Pfizer and BioNTech may lose their profit. And this, as the time moves on, the analyst assume the costs should decrease for a short period, but rise drastically afterward. The last column of MTC denotes the costs evolving in time once threat "Employee faking results" occurs. Since this is a serious reputational damage to both companies, not only will they lose the profits of the vaccine, the reputational damage should cause the demand of both companies to decline drastically. And thus, as the time moves on, the cost will rise significantly than all the other columns.

Now, suppose the analyst wants to know the RoSI in the first year period, with the estimation of threat "Employee faking results", he can simply extract the desired values from the matrices.

$$\begin{aligned} ROSI &= \Delta T \times \sum_{i=1}^{N_{system}} \frac{(T_{cost} \times RMC) - PMC}{PMC} \\ &= \frac{(5.0 \times 1.0 \times 0.9) - 4.2}{4.2} \\ &= 0.07 \end{aligned} \quad (9.13)$$

Since the RoSI is pretty close to zero, the analyst probably think the companies should reconsider their investments and reactions toward the cybersecurity of the formula.

9.3.4 Gordon Loeb model

For organizations operating in a digital environment, the protection of their digital assets is a central business concern. The Gordon Loeb model [24] is a mathematical economic model that helps organizations to determine the right amount to invest in cybersecurity related activities. The following three parameters must be assessed for each information set which is supposed to be protected.

1. How much is the data worth? (*Potential loss*)
2. How much is the data at risk? (*Threat probability*)
3. How high is a success rate targeting this data? (*Vulnerability*)

The parameter λ represents the expected fixed or dynamic monetary loss to the organization caused by an information breach. The threat probability $t \in [0, 1]$ signals the probability of someone attempting to breach a given information set. The vulnerability parameter v is used to denote the rate of the data set being breached once a threat is realized. A company can influence the vulnerability of its information systems, but the threat probability t is considered to be fixed. $L = t\lambda$ denotes the arithmetic product of threat probability and incident loss. The product of these three parameters ($vt\lambda$) results in the median money loss without security investment by the company. The money being invested in security, with the purpose of reducing vulnerability to an information breach, is denoted by $z > 0$. The security breach function $S(z, v)$ denotes the probability that an information set with vulnerability v gets breached, given that the organization spend the amount z to protect it. The model states that a finite investment in information security can not make a vulnerable ($v > 0$) information set perfectly secure.

Expected Benefit of Investment (EBIS)

(A.1) $EBIS(z) = [v - S(z, v)]L$. It quantifies the reduction of a companies loss attributable to the extra security gained from security investment z . The EBIS value is shown in figure 9.3 as the upper curve.

Expected Net Benefit of Investment (ENBIS)

(A.2) $ENBIS(z) = [v - S(z, v)]L - z$ (A.2). It represents the net amount saved through the security investment z as the result of EBIS less the cost z . The EBNIS value is depicted in figure 9.3 as the vertical distance between the EBIS and the cost of investment.

We denote the optimal investment z with respect to the given vulnerability as $z^*(v)$. The optimal security investment into an completely invulnerable information set is always zero ($z_*(0) = 0$). ENBIS is strictly concave in z and an interior maximum $z^* > 0$ is classified by the first order condition $-S_z(z^*, v)L = 1$ (A.3). This is the product of the marginal benefits with the marginal costs of the investment. The optimal level of cybersecurity investment for an organization is at the point where the expected marginal investment costs equal the expected marginal benefits derived from the investment. The optimal investment z^* maximizes the ENBIS and it is the point where the marginal benefits are equal to the marginal cost ($EBIS(z^*) = 1$). The optimal investment z^* is much lower than the associated loss, which is expected in the absence of any security investment. Figure 9.3 shows that from a certain investment level z' , the EBIS and investment cost balance each other out, so that an investment beyond this point does not make sense [25].

Imagine a scenario where an information set with a value of 1,000,000\$ is at risk. Let the attack probability be 15%, and the chance that an attack would be successful is 80%. In this case, the expected loss is given by the product $1,000,000\$ \times 0.15 \times 0.8 = 120,000\$$. According to Gordon and Loeb, the company's investment in security should not exceed $\mathbb{E}120,000 \times 0.37 = \mathbb{E}44,000$.

Figure 9.6 shows the first type of security breach probability functions $S^I(z, v)$, depicting the expected loss for different vulnerability levels v_j and fixed investment z_i . The optimal level of investment is zero until $v = 1/\alpha\beta L$ and then increases with the level of vulnerability. This means that it might be better for a company to concentrate on high

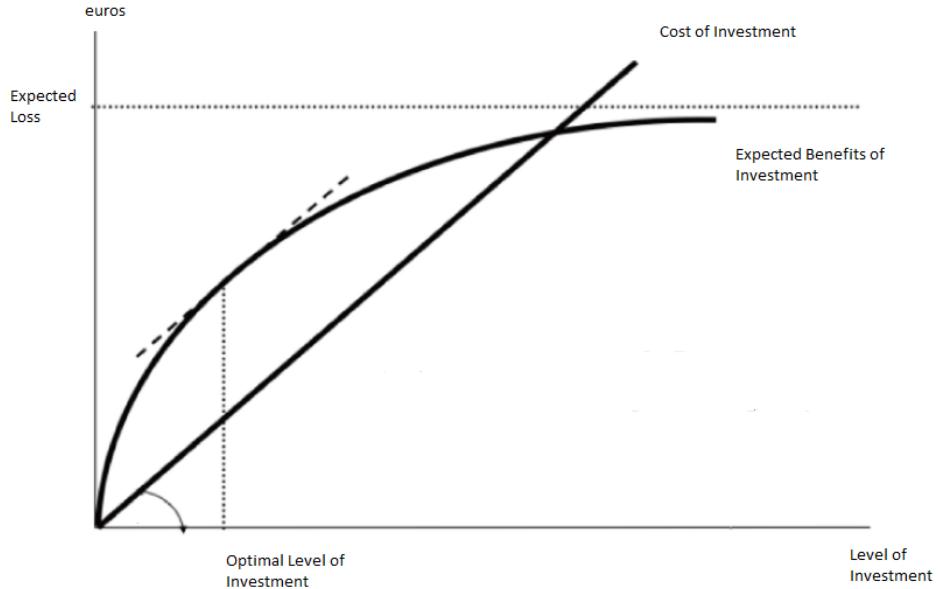


Figure 9.3: The benefits and cost of investment in information security [25].

vulnerability information sets if the security breach probability function belongs to the first type.

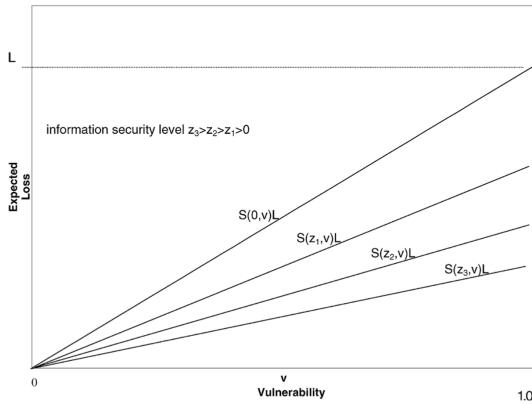


Figure 9.4: $S^I(z, v) = \frac{v}{(\alpha z + 1)^\beta}$ [25]

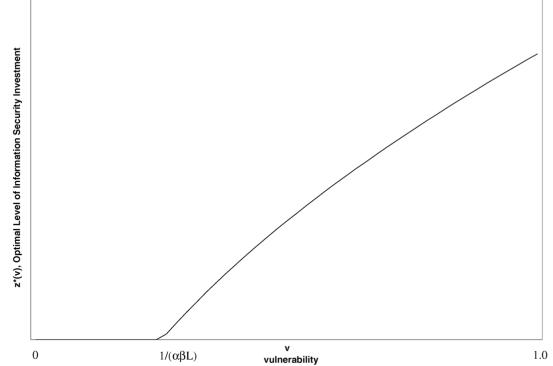


Figure 9.5: $z^{I*}(v) = \frac{(v\beta\alpha L)^{1/(\beta+1)} - 1}{\alpha}$ [25]

Figure 9.6: Type one security breach probability function and optimal level of investment.

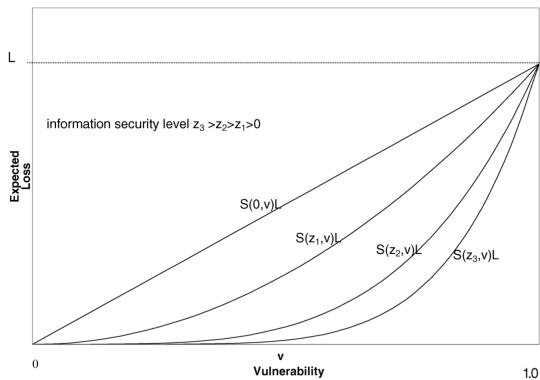
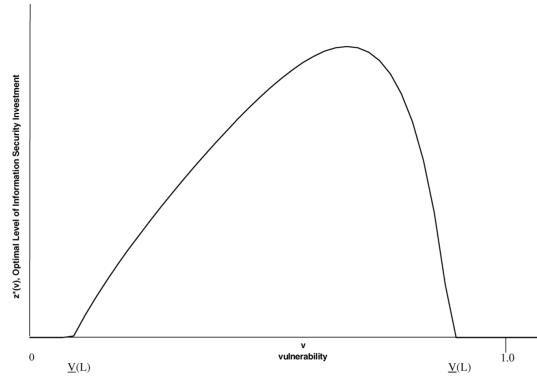
Figure 9.7: $S^{II}(z, v) = v^{\alpha z+1}$ [25]Figure 9.8: $z^{II*}(v) = \frac{\ln(1/-\alpha v L(\ln(v)))}{\alpha \ln(v)}$ [25]

Figure 9.9: Type two security breach probability function and optimal level of investment

The second family of security breach probability functions is of form $S^{II}(z, v) = v^{\alpha z+1}$. This class of security breach probability functions has the property that the cost of protecting highly vulnerable information sets becomes extremely expensive as the vulnerability increases. Therefore, it exists a lower limit $V(L)$, and an upper limit $\bar{V}(L)$ outside of which the optimal investment is zero. For this type of information set, little or no information security is economically justified for extremely high, as well as extremely low, levels of vulnerability. The optimal investment function takes on a maximum at $v = 1/e \approx 0.3679$. The Gordon-Loeb model emphasizes that a company needs to be careful in deciding where to concentrate on their information security resources. Since the benefit of security investment low vulnerability information sets is marginal for both classes of security breach functions, an investment is not justified, considering that security is already good. The model shows that generally the amount spend for the protection of a certain information set should usually be less or equal to 37% of the predicted loss [25].

9.3.5 Conclusion

Information security has become as important to modern organizations as the protection of their tangible physical assets. The models shows that, while some investment in information security is good, more security is not always worth the cost. Although it might be possible to determine the approximate value of the information set, it is usually hard to determine the threat probabilities and the vulnerability associated with it.

Thus, the second class of security breach probability functions also shows that managers allocating an information security budget should normally focus on information that falls into the mid-range of vulnerability to security breaches. Hence, a meaningful endeavor for managers may be to partition information sets into low, middle, and high levels of security breach vulnerability. Since it is a single period model, it does not consider how potential attackers of an information system could change strategies in reaction to an additional security investments. That is, our analysis does not consider the game theoretic aspects of information security.

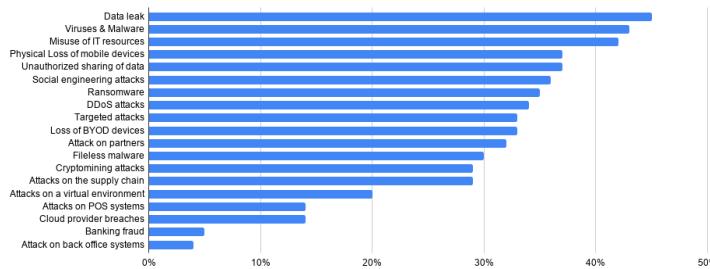
9.4 Case Studies: Economical Impacts of Cyberattacks

Being the victim of a cyberattack can become very expensive for companies. If Amazon's IT infrastructure was unavailable due to an incident, like a DDoS attack, it would lose around \$17 million in sales every hour [26]. Code Spaces, a source code hosting service, even went out of business after a hacker attack gained access to their AWS management panel and wiped out all off their data [27]. It is less expensive to prevent cyberattacks

than it is to clean up the damage afterwards [28]. However, organizations still try to conceal their losses and some are not certain about the degree of damage [29]. Since, we are dealing with an uncertain and manifold environment, it is difficult to precisely determine the right level of security investment and the areas where the funds are to be used.

9.4.1 Security in Practice

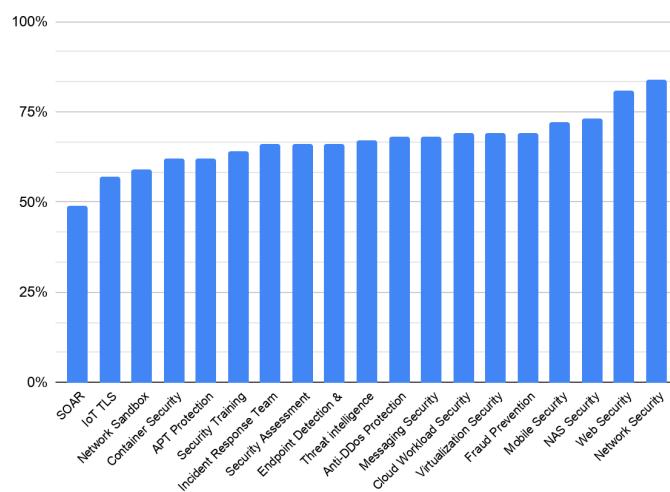
Therefore it makes sense to look at other institutions spendings. According to recent studies by Accenture and Deloitte, companies spend between 6% and 14% (10.9% on average) of their IT budgets on cyber security [30] This corresponds to approximately 0.2% to 0.9% of company revenue or \$1,300 to \$3,000 spent per full time employee [31]. In 2019, 4.58 billion euros were invested in cybersecurity in Germany alone [32]. On the criminal side, there is a marketplace for every possible type of information such as credit card numbers, health records. At first the criminals can buy tools for their attacks there and then sell stolen data for profit. On the company side, IT departments have to compete for this budget with various other departments like HR, marketing, and sales [33]. But how do companies determine the right budget to spend? There are two dominant ways to determine the companies IT security budget. One option is to rely on previous experience and follow prior protocols. The other is to analyze the internal and external environment and use tools and frameworks to derive the right amount to invest from the analysis [34].

Figure 9.10: Threats faced by companies

This amount is determined by many different factors. For example, the sort of business the company is in, the types of personal or sensitive data or intellectual property it handles, the regulatory requirements it faces, the complexity of its IT infrastructure, the likelihood of it being a target for attacks, and other elements [35].

According to data gathered by Kaspersky (Figure 9.10), the number one concern of companies is the protection of their data and the prevention of data breaches (45%). Previous incidents showed the high financial costs arising from companies loosing customers data. This made companies become more attentive. Viruses and malware (43%) make up the second biggest category, since a lot of trouble was caused by it in recent years. Ensuring compliance with security policies and industry regulations, and the cost of securing increasingly complex technology environments (42%) are the third biggest concern for decision-makers [36].

Figure 9.11 shows measures taken by organizations in order to mitigate endogenous and exogenous threats. Tactics like network (84%) and web security (81%) are implemented by a majority. These can range from very general measures, like firewalls, to highly sophisticated intrusion detection systems. Many companies also implement important personnel-related precautions such as setting up an incident response team (66%) and training employees with regard to it security (64%).

Figure 9.11: Security measures taken by companies

Three representative cyber-incidents from the last few years were selected to illustrate the negative economic effects that cyberattacks can potentially have.

9.4.2 WannaCry

Ransomware like WannaCry, Petya and BadRabbit threatens companies worldwide. The attacker gains access to a computer system, mingles with it and requests a ransom in exchange for restoration. The WannaCry attack started on 12th of May 2017 and quickly started to spread to more than 200.000 computers in over 150 countries [37]. It targets computers using Microsoft Windows and uses the EternalBlue exploit to infiltrate the system. It encrypts the victims data and demands a ransom payment to regain the unencrypted data. WannaCry was a very undirected attack because it infected computers indiscriminately. Hospitals, school districts, state and local governments, law enforcement agencies, small and large businesses were all affected [38]. Although, WannaCry was asking for a ransom payment in Bitcoin (equivalent to 300 - 600 USD) there was no intention nor a mechanism to give back the data to the victim. Ransomware is currently considered to be the main moneymaking scheme for cybercriminals and the key threat targeting almost all Internet users [39]. However, in case of WannaCry North Korea was accused of having used the NSA's EternalBlue exploit in their favour. The first wave of WannaCry could be stopped thanks to a security researcher that found the kill switch for the software.

WannaCry showed how crucial it is to keep computer systems up to date and protected. The EternalBlue exploit, found and developed by the NSA, became publicly known one month before the WannaCry attack and fixes by Microsoft were already available prior to that. Therefore, only legacy systems were vulnerable to it. Although, many organizations were affected by WannaCry the health care sector was probably the most affected one. Ransomware often targets these organisations since health data is highly valuable and the systems are often not appropriately protected. A system failure in consequence of an intrusion results in economic loss, potentially bankruptcy, and in some cases, loss of human life.

The negative effects of wanna cry were extensive. The British National Health Service (NHS) was one of the most severely affected institutions. Over 600 member organizations including 34 hospitals were infected and not able to properly use their IT systems anymore. This led to significant disruption across the NHS for patients and healthcare staff in the following weeks. This included reverting to manual processes, like reporting blood results on paper and the cancellation of outpatient appointments. The worst case of reported deaths due to wrong or missing patient data (e.g medication data) was not identified. But in consequence over 13.500 elective appointments were cancelled. The financial loss of income during this one week among infected hospitals totaled £5.9m, excluding recovery cost. If the kill switch had not been found on the same day the attack happened, the damage would have been much bigger. Patients were treated by the still functioning parts of the NHS. The healthcare sector is one of the most vulnerable to cyberattacks since it uses and relies on systems running legacy software [40]. The overall economic damage caused by WannaCry is estimated to be around 4 billion USD [41].

9.4.3 Stuxnet

Most industrial systems and facilities consider only partial security, relying on the premise of “isolated” networks, and controlled access environments. In such an environment it is closely monitored who can enter the facility and only these authorized people can access the internal network’s IT resources. However, this gives a false sense of security like shown by the Stuxnet incident in 2010 [42]. This highly sophisticated attack used multiple windows zero day exploits in order to replicate and infect new computers. Overall, it infected around 200.000 computers but remained dormant on most machines. Once the worm infected a machine it checked if a specific software for controlling industrial appliances (Siemens Step7) was installed.

The main goal of the worm was to sabotage the Iranian uranium enrichment program in Natanz. It targeted the computers controlling the centrifuges used for uranium enrichment (Figure 9.12). By quickly changing the rotation speeds, while signaling acceptable values to the operators, it destroyed them. Such a high risk target is difficult to protect. The Iranian defenders had to concentrate their efforts and money accordingly to physical and cybersecurity. Since the facility was not connected to the Internet the threat of a cyberattack was estimated as very low. On the other hand, a lot of money was spent on the physical security. For example, anti aircraft guns were installed. However, the virus was still able to find its way inside on a USB stick. The Stuxnet incident cleared up several popular but wrong assumptions. First of all it was considered highly unlikely that a cyberattack would target a highly specialized software application. Usually exploits of mass market software were preferred due to their prevalence. Additionally it was considered that a “safe” environment (implying disconnected from the Internet and with limited personnel access) was good enough protection.



Figure 9.12: Uranium Enrichment Centrifuges



Figure 9.13: Air Defense Measurements

Figure 9.14: Installments at Natanz Nuclear Facility

The Gordon Loeb model from section 9.3.4 is suitable to estimate the optimal level of security investment. The value at risk is estimated to be $\lambda \approx 100,000,000$ USD since the cost of such a nuclear is in the billions. The probability that someone will try to attack such a high risk target is fixed at $t \approx 0.5$ and the vulnerability $v \approx 0.25$. The model yields an expected Loss $d = \lambda tv \approx 12,500,000$ resulting the prediction of a security investment $I \approx 12,500,000 * 0.37 = 4,625,000$

9.4.4 Sony Hack

The hack of Sony Pictures Entertainment was unique in nature and in the way, it was orchestrated. On Monday, November 24 the movie making and entertainment unit of Sony Corporation realised that it became a victim to a broad cyberattack and its IT network has been infected by malware. However, the attack might have started one year before it was discovered.

The malware erased the data from the hard drives including the Master Boot Record (MBR). The employees were greeted by the image of a skull on the computer screens together with a threatening message. In consequence of the breach, a lot of internal data was stolen. On 27th of November the attacker group called "Guardians of Peace" started to release chunks of stolen information like unreleased movies and internal emails. Over the next three weeks nine batches of confidential files were released onto public file-sharing sites. Everything from unfinished movie scripts and internal emails to salary lists and more than 47,000 social security numbers was included. Four unreleased Sony movies were leaked to piracy websites for free viewing. The FBI stated that the attack was highly sophisticated and undetectable by industry standard antivirus software. However, the motivation was apparently ideological. The attack was linked to North Korea as retaliation for the planned release of "The Interview". The movie was about a fictional CIA plot to assassinate Kim Jong Un and North Korea did not want it to be released.

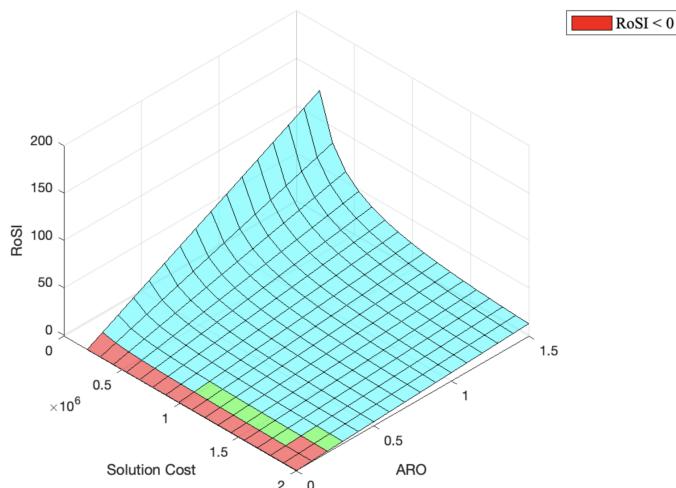


Figure 9.15: RoSI Calculation for varying Solution Cost and ARO

The Return on Security Investment (RoSI) framework can be applied to the Sony Hack in order to estimate an appropriate information security investment. According to estimates, the economic damage amounts to 35 million USD. This is set as the single loss expectancy (SLE). Since, such a big breach will on average only occur to the company once in five years, the annual rate of occurrence (ARO) is 20%. A viable solution costing 2,000,000 USD could reduce the success rate of such an incident by 50%. Therefore the mitigation factor (MF) would be 0.5. According to the formula this yields a RoSI of $\frac{(SLE * ARO * MF - Solution\ Cost)}{Solution\ Cost} = \frac{35,000,000 * 20\% * 0.5 - 2,000,000}{2,000,000} = 0.75$. If the annual rate of occurrence would be once in three years the RoSI would drastically increase to 1.625

9.5 Conclusion

Coinciding with Ross Anderson, it can be confirmed that information security is indeed difficult. The numerous challenges causing this were shown. Once security is compromised, the monetary loss can be significant and it usually requires big effort to recover

from it. Secondly, maintaining cybersecurity requires not only a lot of monetary investments, but also large organisation wide efforts. The weakest link in the chain determines the strength. Both technical and economical issues influence cybersecurity. Technical maintenance is a challenge, as well as planning the information protection investments. Although there exist practically proven models, it is obvious that the estimation of case specific model parameters can be quite difficult. They do not only depend on the used framework, but also on the analyst and the preceding assessment. Hidden biases could significantly change the outcome of the estimation. Since zero-day exploits are unpredictable and hard to defend against, it is hard to reckon with them.

In practice, security issues are not carefully enough handled by companies nor individuals. This is due to an overconfidence in perceived security, a lacking acknowledgement of the importance of information security, and the shortage of investments in such. The situation makes achieving an adequate level of cybersecurity even more difficult. However, a thorough analysis can help to make informed decisions and reduce existing bias.

Bibliography

- [1] Credit Suisse, "Mark Cuban: "data is the new gold"," <https://www.youtube.com/watch?v=8PdnOZI7H5E>, 06 2017.
- [2] White House, "The cost of malicious cyber activity to the u.s. economy," <https://www.whitehouse.gov/wp-content/uploads/2018/03/The-Cost-of-Malicious-Cyber-Activity-to-the-U.S.-Economy.pdf>, 02 2018.
- [3] A. Wirth, "The economics of cybersecurity," <https://meridian.allenpress.com/bit/article/51/s6/52/142650/The-Economics-of-Cybersecurity>, 2017.
- [4] J. Mickens, "Security economics," <https://www.youtube.com/watch?v=8PdnOZI7H5E>, 03 2017.
- [5] B. Rodrigues and M. Franco and G. Parangi and B. Stiller, "Seconomy: a framework for the economic assessment of cybersecurity," https://files_ifi_uzh_ch_CSG_staff_rodrigues_extern_publications_GECON-SEconomy.pdf, accessed: Nov.2020.
- [6] P. Maas, "Does cybercrime really cost \$1 trillion?" <https://www.propublica.org/article/does-cybercrime-really-cost-1-trillion>, 08 2012.
- [7] Accenture, "The cost of cybercrime," https://www.accenture.com/_acnmedia/PDF-96/Accenture-2019-Cost-of-Cybercrime-Study-Final.pdf, accessed: Nov.2020.
- [8] J. Bayuk, "The history of cyber security with jennifer bayuk," <https://www.youtube.com/watch?v=lvxFE-HO7oc>, 11 2019.
- [9] R. Anderson, "Economics of cybersecurity: A brief history," <https://www.youtube.com/watch?v=wAmHPJQoWEc&list=PL-YBwfqnPmvH0YyIHSFhpla7ln2Vb-l>, 02 2016.
- [10] L. Boettger, "The morris worm: How it affected computer security and lessons learned," <https://www.giac.org/paper/gsec/405/morris-worm-affected-computer-security-lessons-learned/100954>, 12 2000.
- [11] UK Government, "Computer misuse act 1990," <https://www.giac.org/paper/gsec/405/morris-worm-affected-computer-security-lessons-learned/100954>, accessed: Nov.2020.
- [12] M. T. Review, "The first ddos attack was 20 years ago. this is what we've learned since." <https://www.giac.org/paper/gsec/405/morris-worm-affected-computer-security-lessons-learned/100954>, 04 2019.
- [13] T. Terpandjian, "Operation aurora: it's perpetrators incentives and administrative responses to alleviate dangers," https://www.academia.edu/10247992/Advanced_Persistent_Threats_Operation_Aurora, accessed: Nov.2020.

- [14] Time, "What we know so far about russia's 2016 meddling," <https://time.com/5565991/russia-influence-2016-election/>, 04 2019.
- [15] A. Odlyzko, "Smart and stupid networks: Why the internet is like microsoft," <http://www.dtc.umn.edu/~odlyzko/doc/stupid.networks.pdf>, accessed: Nov.2020.
- [16] A. Hern, "Wannacry, petya, notpetya: how ransomware hit the big time in 2017," <https://time.com/5565991/russia-influence-2016-election/>, 12 2017.
- [17] T. Moore, "Introducing the economics of cybersecurity: Principles and policy options," <http://static.cs.brown.edu/courses/csci1800/sources/lec27/Moore.pdf>, accessed: Nov.2020.
- [18] J. Hendler and J. Golbeck, "Metcalfe's law, web 2.0, and the semantic web," <https://doi.org/10.1016/j.websem.2007.11.008>, NLD, p. 14–20, 02 2008.
- [19] R. Anderson, "Why information security is hard," <https://www.acsac.org/2001/papers/110.pdf>, accessed: Nov.2020.
- [20] R. Martin, "It is easier to defend than attack," <https://exploitingchange.com/2010/11/29/it-is-easier-to-defend-than-attack/>, 11 2010.
- [21] M. Hill, "Defense now far harder than attack," <https://www.infosecurity-magazine.com/news/defense-harder-than-attack/>, 10 2019.
- [22] B. Rodrigues, M. Franco, G. Parangi and B. Stiller, "SEconomy: a framework for the economic assessment of cybersecurity," <https://files.ifi.uzh.ch/CSG/staff/rodrigues/extern/publications/GECON-SEconomy.pdf>, accessed: Nov.2020.
- [23] BBC, "Boeing 737 max: Brazilian airline resumes passenger flights," <https://www.bbc.com/news/world-latin-america-55243961>, 12 2020.
- [24] L. Gordon, M. Loeb, L. Zhou, "Investing in cybersecurity: Insights from the gordon-loeb model," <https://m.scirp.org/papers/64892>, 03 2016.
- [25] L. Gordon and M. Loeb, "The economics of information security investment," pp. 438–457, 11 2002.
- [26] M. Milijic, "19+ amazon revenue statistics every user should know in 2020," <https://spendmenot.com/blog/amazon-revenue-statistics/>, 06 2020.
- [27] B. Chauhan, "4 times companies were forced to shut down due to hackers," <https://www.getastral.com/blog/911/4-times-companies-were-forced-to-shut-down-due-to-hackers/>, 09 2020.
- [28] K. Crawley, "Cybersecurity budgets explained: how much do companies spend on cybersecurity?" <https://cybersecurity.att.com/blogs/security-essentials/how-to-justify-your-cybersecurity-budget>, 05 2020.
- [29] McAfee, "The economic impact of cybercrime," <https://www.mcafee.com/enterprise/en-us/solutions/lp/economics-cybercrime.html>, accessed: Nov.2020.
- [30] Accenture, "The state of cyber resilience," https://www.accenture.com/_acnmedia/PDF-116/Accenture-Cybersecurity-Report-2020.pdf, accessed: Nov.2020.
- [31] J. Bernard, D. Golden, M. Nicholson, "Reshaping the cybersecurity landscape," <https://www2.deloitte.com/us/en/insights/industry/financial-services/cybersecurity-maturity-financial-institutions-cyber-risk.html>, 07 2020.

- [32] S. Brehme, “Sicherheitsangst treibt ausgaben,” <https://www.computerwoche.de/a/sicherheitsangst-treibt-ausgaben,3548498,02> 2020.
- [33] G. Maria, “How to calculate your small business it security budget,” <https://www.getapp.com/resources/small-business-it-security-budget-template/>, 11 2018.
- [34] A. Moiseev, “How to decide on your company’s it security budget,” <https://www.securitymagazine.com/articles/91559-how-to-decide-on-your-companys-it-security-budget>, 01 2020.
- [35] B. Violino, “How much should you spend on security?” <https://www.csoonline.com/article/3432138/how-much-should-you-spend-on-security.html>, 08 2019.
- [36] Kaspersky, “It security economics 2020: Executive summary,” <https://www.kaspersky.com/blog/it-security-economics-2020-part-2/>, accessed: Nov.2020.
- [37] Cyware, “What is smb vulnerability and how it was exploited to launch the wannacry ransomware attack?” <https://cyware.com/news/what-is-smb-vulnerability-and-how-it-was-exploited-to-launch-the-wannacry-ransomware-attack-c5a97c48>, 06 2019.
- [38] L. Trautman and P. Ormerod, “Wannacry, ransomware, and the emerging threat to corporations,” 01 2018.
- [39] The Guardian, “Don’t pay wannacry demands, cybersecurity experts say,” <https://www.theguardian.com/technology/2017/may/15/dont-pay-ransomware-demands-cybersecurity-experts-say-wannacry>, accessed: Nov.2020.
- [40] Søren Rud Kristensen, “A retrospective impact analysis of the wannacry cyberattack on the nhs,” <https://www.nature.com/articles/s41746-019-0161-6>, 10 2019.
- [41] Kaspersky, “Ransomware wannacry,” <https://www.kaspersky.com/resource-center/threats/ransomware-wannacry>, accessed: Nov.2020.
- [42] Stamatis Karnouskos, “Stuxnet worm impact on industrial cyber-physical system security,” 11 2011.

Chapter 10

The Business Landscape of IoT Security

Andy Aidoo, Jara Fuhrer, Jonathan Stahl, Michael Ziörjen

Within the last two decades, there has been a steady rise in the production and deployment of sensing-and-connectivity-enabled electronic devices, which are replacing 'normal' physical things. The resulting Internet of Things (IoT) will soon be indispensable for many different application domains. The smart objects are continuously getting integrated within factories, cities, buildings, health institutions and private homes. Now, roughly 30 years after the birth of IoT, we are confronted with major challenges regarding its security. Due to the interconnectivity and ubiquitous use of IoT devices, cyberattacks have widespread impacts on multiple stakeholders. Past events show that the IoT domain holds various vulnerabilities, which can be exploited to generate physical, economical and health damage. Despite the knowledge of these threats, manufacturers struggle to properly secure IoT devices. This paper investigates the challenges in securing such devices by examining their particular characteristics. Furthermore, a threat taxonomy is introduced which outlines potential security gaps prevalent in current IoT systems. Lastly, countermeasures are proposed by presenting certain IoT security products from the current market.

Contents

10.1 Introduction	289
10.2 The Internet of Things	291
10.2.1 Terminology	291
10.2.2 Network Topology	292
10.2.3 IoT Architecture	293
10.3 Challenges in IoT Security	294
10.3.1 IoT Security Considerations	294
10.3.2 Characteristics of IoT devices	295
10.4 Threat Taxonomy	297
10.4.1 Security Objectives	298
10.4.2 Security Threats	299
10.5 The Business Landscape	302
10.5.1 Current Products	302
10.5.2 GDPR and its Effect on IoT	305
10.5.3 Ongoing Projects and Guidelines	307
10.5.4 Estimated Market Value & Outlook	308
10.6 Discussion	309
10.7 Conclusion	311

10.1 Introduction

The Fourth Industrial Revolution, also commonly referred to as Industry 4.0, is expected to fundamentally alter almost every business sector with unprecedented velocity. Industry 4.0 is characterized by the blurring lines between physical and virtual reality. One cornerstone of this technological revolution is the Internet of Things (IoT) [1]. The IoT is defined as an intelligent system with comprehensive awareness, reliable transmission and intelligent processing of data [2]. The first primitive device in this category was a remotely controllable toaster introduced in 1990 as a proof-of-concept. The first large-scale smart device application was an item identification system based on radio frequency identification (RFID), ten years later. Cisco, IBM and Ericson were on the forefront of educating and commercialising IoT for consumers [3]. Some IoT devices have already become the industry standard; mainly thermostats autonomously adjusting temperatures and production line sensors keeping track of machine conditions have been widely adopted [4]. It is estimated that by the year of 2023, half of all Internet-capable devices and Internet traffic will be created from machine-to-machine connections; the Internet of Things [5]. Hundreds of new devices are connected to the Internet every single second [4]. The growth rate of used IoT devices is exponential. It is estimated that around 31 billion IoT devices are currently in use and by 2021, another four billion devices will be added to this list totalling 35 billion IoT devices. By 2025 this number will have more than doubled resulting in 75 billion connected IoT devices. The size of the IoT market is even more astonishing; within seven years, the end-user spending is expected to tenfold from USD 100 billion to USD one trillion [6], as seen in Figure 10.1.

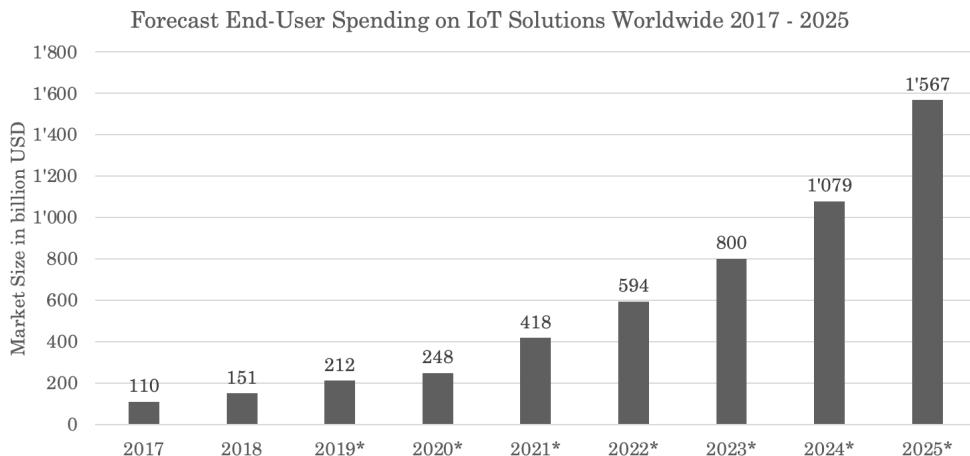


Figure 10.1: IoT end-user spending from 2017 to 2025 based on data from [7] (* = forecasts)

Applications of IoT devices are limitless; hence, the rapid growth of the aforementioned market. There are four major application categories. The *Industrial Internet of Things (IIoT)* finds various applications in production lines. Machines of the production lines communicate with each other. They can monitor each other and distribute the workload evenly between them, detect wear and tear to prevent failure and guarantee constant production, as well as provide real-time production data. The *Internet of Medical Things (IoMT)* is a further application area. IoMT's primary responsibility is the continued availability of information. A patient's heart monitor would send information to a health care providers for monitoring, analysis and possibly even remote configuration. Even a larger clientele uses fitness trackers and smart watches [8]. These IoMT devices are capable of tracking sleeping patterns, vital data and physical activity. According to Mukherjee et al. [9] and Tsai et al. [10], both physical activity and sleeping patterns play a fundamental role in preventing chronic diseases and conditions; by leveraging data from wearable IoT devices health insurance companies can offer risk-based premiums. Depending on

the future industry standard, health insurance companies not leveraging IoT devices to mitigate risk may have an insurmountable disadvantage and it could prevent them from competing in the market [11]. Another application area, maybe even the one which will be most influential for the average Joe is called *Smart Cities*. Caused by the rapid growth of urban population worldwide, there is potential for economic growth in cities. The IoT can be leveraged to manage this rapid urbanization. IoT devices in smart cities regulate traffic efficiently by recognizing traffic flow and derive therefrom optimal traffic light switches. Additionally, garbage disposal can be optimized by equipping garbage bins with sensors able to communicate with the departments of solid waste management. Instead of inefficiently driving along the streets and checking every dumpster, only filled baskets would be considered when creating a road map [12]. Cities collect, by design, a plethora of data ranging from tax payments to water consumption data, building permits and many more. Making such data accessible has plenty of benefits; the government would be more transparent, provide means for innovation and help unlock trillions in economic value, annually [13]. The final application presented are so called *smart homes*. The aforementioned, widespread thermostat and trailblazer of IoT, the Internet-capable toaster, belong to the smart home category. Other appliances are smart TVs, connected light bulbs, door locks and more [8].

Due to the more widespread application of IoT, concerns about its security arose. The traditional goals of IT-security consisted mainly of guaranteeing confidentiality, integrity and accountability of messages. However, these traditional measures are limited when applied to IoT devices (e.g., due to their computing power typically being insufficient). Furthermore, scalability issues emerged due to IoT devices' vast amount of interconnections. IoT security is important for several reasons. Without valid security models suitable for IoT applications, full user acceptance cannot be gained and trust must be established first [14]. One of the main security issues are weak default credentials of IoT devices. This vulnerability was the entry point for prominent cyberattacks like the Mirai botnet [15]. Once inside the network, more devices are infected, which patiently await instructions to commence a distributed-denial-of-service (DDoS) attack. One of the premier web host provider Dyn which hosts some of the biggest websites, including Twitter, Reddit, GitHub and Netflix became a victim of a Mirai attack resulting in the unavailability of aforementioned websites for several hours. Contrary to laptop and desktop computers, many IoT devices operate 24/7 and are therefore always available for a botnet attack. Furthermore, many IoT producers wanted to benefit from the first-mover advantage and thus released a user-friendly product which lacks security. Malware in IoT devices mostly remains unnoticed because of the minimal necessity of interactions with their user interfaces. These are the main reasons why IoT devices are particularly suitable for creating bot nets [15]. Unfortunately, the consequences of IoT security negligence does not stay in the digital realm. Hereinafter, two cyberattacks with aftereffects in the real world will be described. In 2008 a comprehensively monitored pipeline which transports crude oil from the Caspian Sea to the Mediterranean exploded without triggering a single distress signal. Immediately after the blast, the Kurdish militant and political organization Partiya Karkerên Kurdistanê (PKK) claimed credit, while official sources blame malfunctions. According to Robertson and Riley [16], however, similarly to the Mirai botnet [15] the hacker's entry points were the cameras. Thereinafter, the pipe pressure was probably increased while simultaneously manipulating the data displayed in the operational control room until the explosion occurred [16]. Two years later, the computer worm Stuxnet was discovered; this cyberattack is suspected to be a collaborative effort between American and Israeli intelligence with the aim of preventing Iran from producing weapon grad uranium. Particularly astounding in the case of Stuxnet was the precision of the cyberattack; the computer worm exploited zero day vulnerabilities to cause physical degrade in machines which were connected to each other in a completely isolated network [17]. These cases serve as power-

ful precedent from a time when the global number of IoT devices was a fraction of today's. Should we as a collective not bundle our efforts to increase security, one can only imagine what cyberattack will next give the world the creeps.

This paper is structured as followed. First, the Internet of Things with its terminology and architecture are explained in Section 10.2. Second, the characteristics of IoT devices and the resulting challenges are explored in Section 10.3. This is followed by a more in-depth look at IoT security objectives and the introduction of a new threat taxonomy in Section 10.4. Section 10.5 takes a look at the current market and the security solutions available followed by a discussion in Section 10.6 and the conclusion of our results in Section 10.7

10.2 The Internet of Things

The following section serves as preface of this paper; hereinafter terms, which are prevalently used in the network and IT security realms, are described. Furthermore, important network topologies and a reference architecture for IoT are described.

10.2.1 Terminology

Universe of Security

The umbrella-term *IT-Security* is not a static term and evolves over time with new technology. In the 1980s, IT-Security encompassed the goals of ensuring information confidentiality, availability and integrity [18]. These definitions can be ambiguous; on closer inspection, the term confidentiality can be applied to confidentiality of contents. Confidentiality can also be breached, when a third party either gains knowledge of the existence of a communication, its origin or destination. For the scope of this paper, we define the aforementioned security goals as the following: A message is confidential, if only the sender and receiver know of its existence. It has integrity if a message's content is identical for the sender and receiver, furthermore both parties are able to verify said criteria. The final goal of availability specifies that the message is readable by the sender and recipient at a moment's notice. These goals, however, are not all embracing and two decades later the accountability dimension was added. A recipient shall be able to demonstrate the origin of the message and vice-versa. Furthermore, the sender of the message cannot send it in the name of someone else [19]. According to Kenneth [20] the focus of IT-Security has shifted from mainly focusing on availability in the early days of computers to now guaranteeing confidentiality, integrity and accountability. More security objectives will be covered in subsection 10.4.1.

Universe of Security Problems

In the context of this paper, a *hacker* is defined as an entity aiming to illegitimately access a system's resource. A *malware* is a software to accomplish aforementioned goal. Generally speaking, a *risk* in IT occurs when a threat and a vulnerability are paired. Provided that a perfect hardware is coupled with a perfect software, neither a hacker nor malware would pose a risk to an IT-system. However, since they both do not exist, IT-security measures have to be taken in order to minimize harm [20]. A *threat* is any event having the potential to breach security and cause damage. Its existence requires capability of execution or favorable circumstances [21]. For instance, the installation of malware by a capable hacker always poses a threat to a system. Nevertheless it does not automatically become a risk. The second ingredient of a risk, a *vulnerability*, occurs when there is a flaw in at least either the system's design, its implementation or operation and management [21]. To continue the previous example, to pose a risk, the hacker could exploit the most error prone parameter of a system, the human. By sending a malicious e-mail, the hacker could introduce the malware into the system. When all these pre-conditions have been

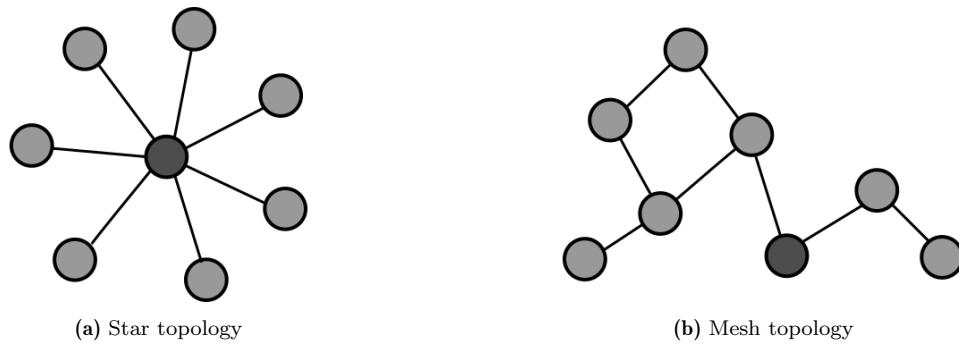


Figure 10.2: Most common network topologies in IoT, own diagram

met and actions are executed, the risk becomes an *attack* which can be further divided into intent and origin. When focusing on intent, an attack can either aim to alter system resources or gather information from the system. Former is classified as an active attack; the latter as a passive. The second dimension, origin, distinguishes between inside and outside attacks. An inside attacker is authorized to access system resources, however, does it in an unapproved way. Outside attackers do not have any authorizations to be in a system [21].

10.2.2 Network Topology

A network of computers is characterized by an interconnection of at least two autonomous endpoints, hereinafter synonymous to *node*, which exchange information. The transmission medium, which is called a link; it can be wired or wireless. The communication between nodes is governed by a network protocol. Aforementioned network aims to reliably and flexibly enable resource sharing as well as communication between network users [22].

To describe the arrangement of diversified telecommunication networks, a *network topology*, which can be divided into physical and logical topologies, is used [23]. Former illustrates how a network of computers, including switches, routers and so forth are connected. This type of network topology used to be the main source of research, when computers still were in the fledgling stages. Currently, as long as a certain robustness and scalability is provided, the emphasis lays on the design of logical topologies [24]. Latter topology depicts how information flows from one device to another and is independent of its physical aspect.

Variations of Network Topologies

The simplest topology is when dedicated links connect two nodes. A child's tin can telephone is one example of a point-to-point network, where the tin cans act as endpoints and the wire as the dedicated link. When using a telephone, dynamic *point-to-point* network are in place. In order to function neither one of the endpoints nor the dedicated links may fail. Should endpoints be connected in a series next to one another, the resulting network topology is classified as a *daisy chain*; the series connection of fairy lights around a Christmas tree is a primitive example. Information in a daisy chain travels on a so-called electrical bus, which transports signals from one node to another. Such a linear networks string endpoints in a point-to-point fashion together, while one endpoints acts as a monitor which is responsible for performing operations. It is typically used in smaller networks because of its cheap installment and expansion. However, troubleshooting is one of its weaknesses. Furthermore, the addition or removal of endpoints can disturb the whole topology. Alternatively, when all nodes are connected to a central gateway the resulting topology is called *star topology* as depicted in Figure 10.2a.

The main advantages of star topologies are their cost effectiveness, easiness of deployment and reliability. The failure or vulnerability of a node, does not compromise the whole

network; an easy to understand analogy are shunt circuits, where a defect light bulb does not prevent other light bulbs from operating. A *Mesh topology* is characterized by at least three distinct nodes; each of the endpoints aim to connect with all the other endpoints. An example of such a topology is shown in Figure 10.2b. Those links are generated dynamically and non-hierarchically [22]. According to Cilfone et al. [25] due to its scalability and reliability, the mesh topology is very attractive for the deployment of IoT-oriented networks. For instance Google Nest, Google Wi-Fi and Google OnHub all support Wi-Fi mesh networking. Historically however, the maintenance cost of said topology is high and the installation and configuration is more difficult.

Data Transmission

To transport one piece of information from one node to another, a transmission medium, also called link, is required. Electric cables, optical fiber and radio waves belong in aforementioned category. Various factors determine the capabilities of transmission medias. The bandwidth refers to the the data transmission scope, while noise absorption attributes to how external electrical noise can cause deformation of a signal. Furthermore, said capabilities are influenced by how much energy is lost as a signal is transmitted from its source.

Nodes

Gateways are a type of nodes which allow data transmission from one network to another. In the realm of IoT, gateways provide bridges between the devices. To connect two or more network segments, hubs are used. Furthermore, they amplify signals to extend the range of a transmission medium. Gateways are devices capable of operating by themselves. In a network, it regulates the information flow.

Protocols

A network of networks is commonly referred to as the *Internet*, where people and machines communicate with each other physically apart. To provide means of inter-connectivity between nodes and networks, the transfer of messages follows standard protocols. The *Transmission Control Protocol* (TCP), which is the most widely used protocol, breaks up a message into small packets and sends it via Internet to its destination where the message is recomposed using the TCP [26].

10.2.3 IoT Architecture

Due to the heterogeneous nature of IoT devices, there is no standard construction for IoT networks that fits all use cases. However, there are a number of architectures that are commonly presented in the literature. One such representation divides the IoT architecture into three layers, depending on their characteristics [2; 27; 28; 29]. Other approaches divide the architecture into more fine-grained layers (e.g., four layer architectures [30; 31] or the seven layer IoT World Forum Reference Model [32]). Another way of describing and building IoT networks is the *Fog computing paradigm* which also makes use of three layers but uses different concepts to classify the devices (i.e. edge, fog and cloud computing) [33]. For the remainder of this paper the traditional and most commonly used three layer architecture is chosen due to its simple structure and intuitive nature. It is depicted in Figure 10.3 and explained below.

- ***Application Layer*** The top layer consists of applications and middleware. Depending on the use case it can include elements of cloud computing, integrations to other applications, resolution services or web services. In general, the layer is responsible for delivering application specific services to the user [2; 27; 28; 29].
- ***Network Layer*** The middle layer consists of the network that is required for the transmission of data between the IoT devices, other network devices **or** servers. Depending on the use case there might be different network types such as mobile

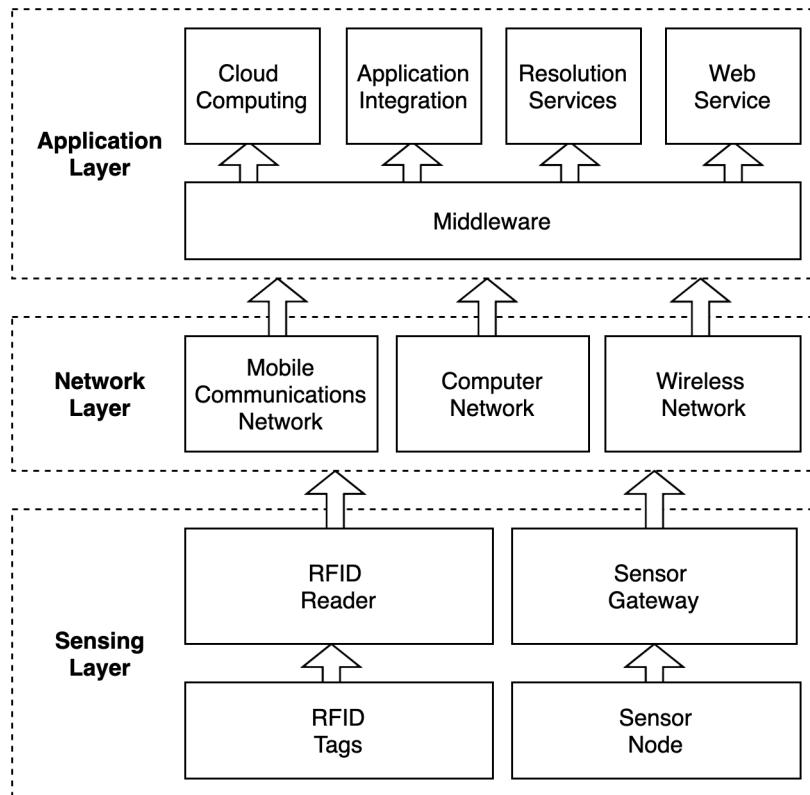


Figure 10.3: Three layer architecture of IoT, own diagram based on [27]

communication networks, computer networks or wireless networks that make use of different protocols (e.g., Constrained Application Protocol (CoAp) or ZigBee) [2; 27; 28; 29]. Since this layer is responsible for the communication between the different devices and services involved, this layer is also referred to as *Communication layer*.

- **Sensing (or Perception) Layer** The bottom layer is the physical layer which consists of the IoT devices itself (e.g., sensors, RFID readers / tags, and gateways). In many cases it involves sensors and actuators that are embedded in the environment [2; 27; 28; 29]. Since this layer consists mostly of hardware, some authors also refer to it as the *Hardware* or the *Physical* layer.

10.3 Challenges in IoT Security

The following section explores why especially IoT security ought no to be neglected and why securing IoT devices remains such a challenge.

10.3.1 IoT Security Considerations

The security aspect seems to be one of the greatest concerns within the IoT world, see for example [34; 35; 36; 37]. While security is a key requirement for any type of technology, it is even more critical in the domain of IoT. Garcia-Morchon et al. [38] mention three reasons as to why IoT security is especially important: First of all, IoT systems which are out of control cannot only jeopardize the users' privacy, but also they can cause physical harm when their sensors, actuators or other connected devices are used maliciously. As the second reason, they point out the risk for manufacturers; as attackers could get access to sensible information or proprietary assets through the IoT system, manufacturers not

only loose valuable information but also damage their own reputation. Thirdly, due to the high inter-connectivity of IoT systems, the impact of an attack goes beyond a specific device or network [38]. In this sense, the saying "A chain is as strong as its weakest link" is perfectly applicable; the IoT network is as secure as its weakest device.

The major objectives of IoT security are to ensure privacy, confidentiality, integrity and availability of the offered services [37; 38; 39], to mention only the most important ones (see section 10.4.1 for more details). Within the last years, another aspect has been added to the list: Money. Bastos et al. [40] reference a Ponemon Institute study (a research center devoted to privacy and data protection [41]) from June 2017. The institute estimated that a data breach costs on average \$141 per data record - or around \$3.6 million per incident. This huge amount of money is needed to sort out the harm caused by data breaches: for fixing the actual breach, but also for insurance protection, compliance or reputation recovery [42]. Hence, IoT security is not only applied to meet the aforementioned security objectives. It is furthermore an increasingly important part of a company's business strategy in order to prevent such high data breach costs and the reputation damage which inevitably follows.

Despite the economical benefits [40] and the particular importance of IoT security [38], the IoT industry is far away from being secure. A study led by Hewlett-Packard indicated that around 70% of the generic IoT technologies contain some kind of security vulnerability such as unencrypted data transmission or very basic passwords [35; 43]. Kolić et al. [35] expect security to be one of the fundamental design decisions of IoT systems, yet see insecure devices brought to market by vendors who do not put enough effort in securing their technologies. The authors explain this situation with the particularity of the IoT sector and devices: Customers want user-friendly, battery-efficient, and good looking products - and all this as quickly as possible. From these requirements, several features can be deducted which are particular for IoT devices. These features differentiate an IoT device from a traditional IT device like a laptop. Several researches including [35; 37; 38; 39; 44; 45] attribute the reasons of why it is particularly hard to secure IoT devices to these characteristics. They identified characteristics that make it sometimes impossible to apply traditional security measures to IoT devices. In the following section, the most mentioned ones are pointed out.

10.3.2 Characteristics of IoT devices

In contrast to standard IT systems which are often considered as "monolithic apparatuses", the novel IoT network consists of many "connected microcomputers" [35, p.85]. Regrettably, most IoT manufacturers have until now failed to treat it as such. The unique security requirements and characteristics should be incorporated within the design process of IoT devices. However, they have until now often been neglected. The consequences are IoT networks, which are vulnerable to many kinds of cyber attacks.

As IT pioneer Peter Neuman claimed, "You can't add security into something that isn't designed to be secure." [35, pp.85]. Based on contemporary literature, this seems also to be true for the IoT technology. Now it follows a look into the characteristics of IoT, trying to find out why it is so difficult to protect IoT devices and why this technology needs special attention and cannot simply be equated to 'standard' IT.

A) Usability

Customers want user-friendly products [35]. The initial setup has to be quick, the onboarding of new devices straightforward. The device should be easy to use and appealing to the eye. These requirements stand in contradiction to many security measures. Passwords, for example, that protect the device and the network it is connected to from unauthorized access, are often perceived as annoying and interrupt

the smooth usage. Manufacturers hence try to keep such enforced interactions to a minimum and usability high [46].

In case password authentication is still needed, users are asked to define one during the initial setup. Not too seldom, a simple, already used password is then chosen - or even worse: the default password is kept. Any more or less sophisticated algorithm is then able to guess these passwords. Additionally, there are manufacturers which use hardcoded passwords to keep the disruption low and to minimize the effort needed to setup a device. These insecure default settings make the IoT device vulnerable by design [43; 31].

The research and advisory company Gartner [47] published a report saying that in 2017, more than 60% of the IoT applications have been used by the consumer segment, while the business sector accounted for the rest. While businesses might employ IT specialists to evaluate and finally bootstrap the IoT devices and networks, private consumers might lack expertise and often are not as tech-savvy. The products hence must be easy to use and setup for the broad mass, which results in the renunciation of complex security measures [46].

B) Resource Limitation

Security measures are typically based on expensive schemes like encryption/decryption or signature/verification, without considering resource consumption [48]. IoT devices, however, are resource constrained, making it a challenge to secure them properly [44; 37; 39].

IoT devices are usually small and light and therefore not able to store a large battery. Additionally, they should operate autonomously without needing a person to replace the battery frequently. Hence they have limited power resources and should run energy-efficiently [48; 45; 49]. Furthermore, IoT devices need to operate resource-saving also regarding memory capacity. If there is no space for a long-lasting battery, there is often neither any for a large memory. The same is true for processors: Due to the limited space and weight, CPUs and sensor complexity are limited. This results in low data rates and IoT deployments conducted in lossy and low-bandwidth communication channels [38; 49; 45; 50].

Such resource limitations make it hard to embed traditional security measures into IoT devices and ask for specific countermeasures. Without them, IoT devices are particularly susceptible to attacks such as DDoS, which is a resource exhaustion attack that aims to limit a device's availability and which benefits strongly from the device's resource limitations [38].

C) Short Time-to-Market & Affordability

The IoT world is rapidly changing. Vendors have to deliver their products fast, if they want to keep up with the competition. They do not take time to develop sophisticated security measures and rather apply quick security fixes if necessary [35]. In doing so, they produce commercial off-the-shelf products which are not well developed and secured [50; 35], but easy to set up and affordable. Instead of being a 'fundamental design focus' [35, p.84], IoT security still relies on technologies and protocols developed for the Internet itself. The Internet however cannot be equated to the IoT domain, which is an integrated systems, interacting very closely with people (and can cause serious harm as we have seen in Section 10.3.1).

Furthermore, a well-secured but therefore expensive product would likely not be successful. Customers command respectable bargaining power, the required time-to-market is short (i.e. minimize the time that passes until an idea becomes a product ready for the market) and the additional security costs are disproportional to the cost of the device itself [44; 35].

D) Availability & Ubiquity

IoT devices are constantly connected to the Internet. They are not like laptops or phones which are shut down completely from time to time [15]. Hence, they constitute a reliable target for attackers, always available and always connected to other devices. And as if this was not enough, IoT devices seem to be everywhere, they create huge networks with access to several areas: Households, companies, transportation and factories are becoming accessible simultaneously, due to the IoT application [15; 39].

With the increasing ubiquity of IoT devices, the number of devices to be used in potential attacks increases respectively [39; 51]. Currently, around 31 billion things are connected and it is estimated that by 2025, this number will rise to 75 billion [6; 51]. Most of the devices used by private consumers are Smart Home devices like TVs, set-top boxes [47], entertainment systems, speakers or lighting, and heating sensors [52]. These apparatuses, can theoretically monitor people without drawing attention from the victims. Consumers expect some monitoring activities from the devices such that their gadgets can provide their intended functionalities. A smart light system for example is expected to listen to voice commands, however not to private conversations. For a user, it is hard or even impossible to control that in this case, only commands are being processed and private conversations maybe listened to, but definitely not processed or stored.

E) Inter-Connectivity & Heterogeneity

Inter-connectivity and heterogeneity go hand in hand. The IoT network consists of a high number of integrated and heterogeneous devices which actively share data amongst each other [50; 34]. All these open connections create multiple access points for attackers to exploit existing vulnerabilities [39]. The participating entities are heterogeneous in terms of their communication patterns, policies, protocols, features, manufacturers and, of course, their security standards. Also, they often are geographically dispersed, which means that regulations from different countries might apply [38; 39].

The mentioned characteristics of IoT networks do not favour comprehensive security standards and countermeasures to cyber attacks. Nevertheless, various working groups begin to formulate guidelines and best practices which manufacturers can use as a point of reference. They create new protocols adapted to the IoT particularities with the goal of making them more secure. Section 10.5 summarizes the most promising ones. Subsequently, there will be an elaboration on the security objectives and threats in order to better understand the underlying problem IoT is facing. Furthermore, a threat and attack taxonomy of the IoT domain concisely categorizes the potential threats to gain a better overview.

10.4 Threat Taxonomy

The following section gives an overview of threats in IoT based on the existing literature. The first section introduces security objectives that need to be achieved along with their

definition in the context of IoT. It is followed by a threat taxonomy that aggregates the findings of multiple authors to form a more holistic perspective on threats in IoT.

10.4.1 Security Objectives

The characteristics of IoT devices introduced in the previous section expose them to numerous threats. Before examining the most common threats and attack scenarios, the general security requirements for IoT devices and their network are outlined. Table 10.1 displays security objectives relevant in the domain according to the existing literature. For the remainder of this paper, the focus lies on a subset of objectives that are mentioned most often in the literature. This list of objectives lays the foundation for a more holistic understanding of security in IoT. It is presented with a short definition for each objective below [31; 38; 53; 54; 55; 56; 57].

A) Identification

The entities in the IoT system need to be able to identify other participants (i.e., they need to be aware of other entities in the network). Furthermore, the entities need to be able to distinguish friendly from potentially malicious entities. In most cases, IoT devices will reside in a certain context (e.g., belong to a group, located in a particular building, owned by a certain entity) this also needs to be taken into account. Identification refers to the process of claiming such an identity [58].

B) Authentication

Before access to a restricted resource is allowed (e.g., sensitive information) the sensing devices, users and gateway nodes must be authenticated (i.e. their identity must be verified). It must be ensured that they are who they claim to be.

C) Authorization

After the identity has been verified it must be ensured that the entity under consideration is allowed to access the data, resources or applications within the system. In the domain of IoT access to a given resource might depend on other factors such as the identity of the owner of the device (i.e., providing more information to people with certain roles) or the location (i.e., checking whether an user is accessing the device locally or remotely).

D) Integrity

It must be ensured that the message or the entity was not changed (i.e., modified or destroyed) during its exchange, storage, and processing. Or in other words, it must be ensured that data is not altered while it is traveling from one point to another [30].

E) Confidentiality

Resources need to be protected from access by unauthorized parties. Consequently, a failure to maintain confidentiality results in an entity gaining access to a protected resource [59].

F) Privacy

During the handling, processing, storing and deleting of data it must be ensured that the rights of individuals regarding the use of their personal information are addressed properly. This usually involves adhering to contracts or policies and applying governing regulation or law (e.g., General Data Protection Regulation (GDPR)) [60].

G) Availability

The system and its services should be available when they are required. Thus, availability refers to the probability that a system (or component) is operational at a specific point in time. As proposed by Pokorni [61] this incorporates both reliability (i.e., meeting certain performance standards in a given context) and maintainability (i.e., ability to uncouple, fix and modify components without obstructing the service) properties of a component or system.

H) Non-Repudiation

A malicious entity should not be able to hide his/her actions [53]. Thus, non-repudiation ensures that no entity can claim that a transaction did not happen when it in fact did or vice versa. It ensures that circumstances can be resolved where different parties in the system hold different views of something that happened (e.g., during a network failure) [62].

Security Objective	References
Integrity	[38] [56] [31] [53] [54] [45] [30] [39] [63] [64]
Confidentiality	[57] [56] [31] [53] [45] [30] [39] [63] [64]
Authentication	[57] [38] [55] [31] [53] [54] [30] [39] [63]
Privacy	[38] [55] [53] [54] [45] [30] [63]
Availability	[38] [56] [31] [53] [45] [39] [64]
Authorization	[38] [55] [31] [53] [54] [39]
Non-repudiation	[57] [53] [54] [39]
Identification	[57] [55] [54]
Reliability	[57] [55] [30]
Freshness	[53]
Controllability	[31]
Soundness	[64]

Table 10.1: Security objectives in IoT according to the existing literature ordered by number of mentions in the literature.

As shown in Table 10.1 there is no clear consensus in the literature about the most important security objectives in IoT. This might be due to the overlapping of certain terms and definitions. For example the term *Authentication* might also include *Identification* because the latter is a prerequisite of the former. Additionally, most authors do not provide a succinct definition of the terms used for their objectives which further complicates the comparison of their results. Furthermore, some authors introduce new security objectives such as *Freshness* which have not yet gained attraction in other literature.

10.4.2 Security Threats

There are multiple approaches to form a taxonomy of IoT security threats. Some authors classify security threats according to a layered IoT architecture [31]. Others base their

taxonomy on a single list of threats and countermeasures [63]. To aggregate the findings in the existing literature the taxonomy proposed in this paper is built on top of the three layer IoT architecture introduced in Section 10.2.3. While this architecture has been criticized for not being able to capture all the nuances in IoT systems [31; 28] it serves as a common denominator for taxonomies that involve more layers (e.g., the five layer taxonomy proposed by [57] or the four layer taxonomy proposed by [65]) and thus allows us to incorporate results from the respective authors as well.

As explained in Section 10.2.3 an IoT system usually consists of different devices with different capabilities that make use of a diverse set of communication protocols. Furthermore, various interfaces are required to enable services that make use of aggregated data collected within the system. It is therefore not sufficient to implement security measures based on traditional IT network solutions [29]. Projects such as (OWASP publish guidelines for developers and manufacturers on how to secure IoT systems. However, these guidelines focus on the most common vulnerabilities and do not provide an exhaustive list of potential threats or attack vectors in IoT systems [43]. In contrast to the broadly applicable security guidelines provided by OWASP, other authors list threats and attacks based on specific IoT use cases (e.g. smart water systems [39] or smart grids [66]). However, they do not specify if and how their approaches could be applied to other domains and use cases. An early threat taxonomy proposed by [36] incorporates different perspectives (such as identity management, storage management and physical threats) but refrains from listing more than a handful of threats for each perspective.

To address these issues, the threat taxonomy proposed in Table 10.2 incorporates threats and attack vectors listed by multiple authors and thus aims to provide a more exhaustive and holistic view. The taxonomy gives an overview of threats and categorizes them based on the traditional three layer IoT architecture. Additionally, it provides references to the existing literature where the threats are further examined and explained in the context of IoT. The list of threats is not exhaustive but contains a higher number of threats than most of the existing taxonomies. Thus, it can act as a starting point for projects that aim to secure an IoT system or for further research in this domain. The taxonomy emphasizes that IoT security needs to be addressed from multiple perspectives and it does not suffice to focus on the IoT device itself. It also highlights the heterogeneous nature of threats: Some threats involve physical access to the device (e.g., Tampering / physical damage) whereas others focus solely on the software running on the application layer (e.g., SQL injection).

The long list of threats presented in the taxonomy emphasizes the need for mature security guidelines and solutions that focus on the whole IoT system and not just individual components. The next section will outline the current IoT security market and highlight a number of solutions that are already available.

Layer	Attack	Source
Application	Data modification	[31]
	Software reverse-engineering	[38]
	Firmware attack	[38]
	Elevation of privilege	[38]
	Denial-of-Service (DoS) attack	[38] [55] [56] [53] [65]
	Many logged-in users with the same login-id attack	[53]
	Stolen-verifier attack	[53]
	Stolen/lost smart card attack	[53]
	Password guessing attack	[53]
	Password change attack	[53]
	Buffer overflow	[38] [29]
	Impersonation attack	[53]
	Memory corruption	[56]
	Code execution	[56]
	(SQL) injection	[56] [29]
Network	Cross-site scripting (XSS)	[56]
	Cross-site request forgery (CSRF)	[56]
	Collision	[57]
	Exhaustion	[57]
	Unfairness	[57] [29]
	Spoofed, altered or replayed routing information	[57] [29]
	IP spoofing	[31]
	Side channel attack	[38] [31] [29]
	Distributed denial of service (DDos) attack	[38] [31] [29]
	Selective forwarding	[57] [38] [29]
	Sinkhole attack	[57] [38] [29]
	Sybil attack	[57] [38] [53] [29]
	Wormhole attack	[57] [38] [53]
	Hello and session flooding	[57] [29] [65]
Sensing	Acknowledgement spoofing	[57]
	IP misconfiguration	[56]
	Synchronization attack	[57] [29]
	Replay attack	[31] [53] [65]
	Man-in-the-middle attack	[38] [31] [53]
	Eavesdropping	[38] [55] [29]
	Jamming	[57] [29]
Physical	Malicious substitution	[38]
	Tampering / physical damage	[57] [55] [31]
	Node capture	[38] [55] [53]
	Cloning / device replication	[38] [53]

Table 10.2: Taxonomy of threats and attacks in IoT according to the existing literature based on the three layer architecture

10.5 The Business Landscape

In this section we provide an overview of the current IoT security market and list a number of products that offer solutions to the aforementioned threats. This is followed by an illustration of the impact of current regulation on IoT. Lastly a summary of ongoing projects and guidelines in the domain is provided.

10.5.1 Current Products

Corporate / Public	Private
Software	
Hardware and Firmware	Home
Service and Cloud	

Table 10.3: IoT Security Product Categories

In this subsection a broad overview of the different types of security solutions is presented. For each type a selection of products that are currently on the market is outlined. The intention is to provide a diverse collection of products by selecting solutions that use different approaches. It should by no means be considered an exhaustive list. To structure the selection of products, product categories were defined, as can be seen in Table 10.3. The main distinction is between products in the **Corporate / Public** category and products designed for **Private** use. Within the corporate and public sector it can be distinguish between *Software*, *Hardware / Firmware* and *Service / Cloud* solutions. Since the market for IoT security solutions in the *Private* category is not as mature there are not as many different products yet. Therefore, this category only consists of a *Home* solution category.

Software

Products where security is mainly achieved through a software component fall into the category of *Software*. More precisely, security logic is added by running software on the same network as the devices to be protected. The most common approach are so-called *Intrusion Detection Systems (IDSs)*. These systems act as an additional line of defence by detecting attackers. Namely, they monitor the activities of a host or network and can trigger alerts, or launch mitigation actions when unusual behaviour is detected [50]. The following list describes two selected examples of *Software* solutions.

- I) **Armis Security** Armis Security advertises that they are the first company to offer an agentless security platform for businesses. Their product integrates into the customer's existing infrastructure. No additional hardware is required. The system typically runs in a virtual machine and can be installed on any existing server within the network to be monitored. For monitoring, they use a passive approach. This means that the network does not need to be actively scanned for devices, but the data traffic is passively analysed. By querying their device knowledge base, they can identify and classify every device on the network, whether managed or unmanaged. Their database contains profiles and properties of the devices. Based on this information, the security platform can assess the risk for each device. For example, it knows if the device is running an old OS version. In addition, the behaviour can be compared with the behavioural data stored in the knowledge base. If anomalies are detected, a warning can be issued. Not only is detection possible, but automated response actions can also be implemented. Armis integrates with network access

control products from networking companies such as Cisco. As a threat response measure, Armis could trigger a quarantine on a suspicious or malicious device. [67]

The Armis security platform can be clearly classified as an IDS. The product is only installed on a single device in the network, which is very convenient for the customer. However, it can be difficult to detect an attack if it is running in a separate part of the network [50]. To detect attacks, Armis compares network behaviour with known attack signatures/patterns and also compares a node's behaviour with expected behaviour based on historical data. Following the IDS placement strategies and threat detection methods defined by Zarpel et al. [50], Armis uses a centralised placement strategy and a hybrid threat detection method.

- II) **Bastille** Another software solution is offered by the company Bastille. It allows the monitoring of a specific area, for example an entire office, for the presence and behaviour of connected devices, which use cellular (cell phones), WiFi, Bluetooth or Bluetooth Low Energy (BLE) radio signals. It must be mentioned, however, that although the software is the focus, hardware sensors are also needed for operation. Their solution is based on the three pillars *Discover*, *Analyze* and *Act*. Bastille scans the room and *discovers* wireless transmitters. By digitally demodulating radio signals, protocols can be identified and individual devices can be plotted on a map of the room, even showing their position. The devices found are *analysed* for protocols, traffic and other devices connected to them. This can then be used to decide whether a device is under attack or performs a prohibited action. An example of a not permitted action would be, when a hearing aid establishes a connection to a device outside of the monitored area, allowing an attacker to listen what is going on inside an office. When a danger is detected, different *actions* can be taken. If a device is detected which is prohibited in that area or exhibits abnormal behaviour, it can either be physically removed or isolated by integration with network systems via software. [68]

Bastille's solution is mainly used to detect devices. It is therefore an IDS, but with a focus on the physical intrusion of devices into a monitored space. A big advantage is that devices that are not part of a specific network are also recognised. All devices that are located within the room monitored by the hardware sensors are discovered and surveyed. However, when it comes to detecting unusual behaviour, the system's capabilities are more limited than other IDSSs.

Hardware and Firmware

To make resource-limited IoT devices inherently secure without the need to install additional software, products in the category *Hardware and Firmware* can be used. The following are two selected examples.

- I) **ReFirm Labs** Every IoT device runs some kind of firmware that controls the hardware. If a firmware has vulnerabilities, such as weak passwords, backdoors, outdated components or zero-day vulnerabilities, these can be exploited by an attacker. ReFirm Labs has developed a tool that can automatically analyse firmware. It is intended to be used by manufacturers of IoT products or to check installed devices from other manufacturers for their security. Their so-called Centrifuge Platform, takes a firmware binary image as input. The output generated is a detailed security audit. Based on the available information, the manufacturers own developers can make adjustments to the firmware or, if the device is supplied by a third party, they can be informed about possible attack vectors in their products. [69]

ReFirm labs' solution addresses the problems faced by manufacturers. This is especially valuable because it hardens the security of IoT products already in the development process. That is, even before it is deployed. But on the other hand, this can also be seen as a limitation, as this solution can only be used by manufacturers or IT experts.

- II) **Zymbit** This company offers a product called Zymkey which is a plug-in hardware module for the Raspberry Pi [70]. It attempts to solve the security goals of authentication and integrity. The module can simply be plugged onto a Raspberry Pi and is controlled via an API. The following features are covered: it creates and stores a unique device ID in the hardware module, contains a strong cryptographic engine, can store public/private key pairs that cannot leave the module, and has physical tamper detection, such as an accelerometer that detects vibrations and orientation change events. [71]

The big advantage of this solution is the simplicity of integration with the Raspberry Pi. One disadvantage, however, is that it is only available for this one platform so far.

Service and Cloud

Cloud computing has been an increasingly important topic for some time. The reduction of fixed-costs (e.g., by eliminating the need to purchase hardware), almost infinite scalability and roughly 24/7 availability are arguments in favour of cloud solutions for the end customer. IoT security is no exception. And many of the big tech companies (Amazon [72], Google [73], Microsoft [74], IBM [75]) are competing for a piece of this pie. Since the solutions offered by these companies are all similar, only the solution of Amazon is covered in the following paragraph.

- I) **Amazon** Amazon has developed an entire ecosystem of IoT services in their AWS Cloud. To ensure the security of a fleet of distributed devices, they offer a dedicated "AWS IoT Device Defender" service. Two main features can be highlighted: First, the service continuously checks the configuration of all connected IoT devices. It checks whether predefined best practices are adhered to. An example would be whether all devices have valid x.509 certificates (e.g. TLS/SSL). Or whether all ports that are not required are closed. Second, it is possible to detect unusual behaviour on the devices. Expected behaviour is determined by predefined rules (e.g., restricting the entities the device is allowed to connect to or limiting how much data is received or sent). Based on these rules Alerts will be generated. [76]

A cloud service has the main advantage that the setup effort into an already existing network is minimal. One limitation of Amazon's solution is that the value of their service can only be fully realised, if also other services from AWS are consumed. For example, it is recommended to also use the services "AWS IoT Core", "AWS IoT Device Management", "AWS IoT Analytics". Of course this creates a vendor lock-in effect for Amazon products and thus may make it difficult to switch to another provider.

Home

With the increasing introduction of connected and intelligent devices in households, the danger of cyberattacks in the private sphere is also increasing. Products like Bitdefender BOX try to remedy this threat.

- I) **Bitdefender** Bitdefender advertises their BOX as an-all-in one product for a secure connected home. It can act as a standalone WiFi router, or can be connected to an existing one. By constantly scanning all the traffic, and using machine learning to process the data, the system learns the normal behaviour of the devices and can detect anomalies [77].

This product can be compared to the industrial Armis Security Solution. It can again be categorized as a centralized IDS. The detection method seems to be an anomaly-based approach, based on Zarpelão et al. [50]. It uses historical data to learn about normal behaviour and is subsequently able to detect unusual behaviour.

This section listed a number of products offering solutions for the threats introduced earlier. However, security issues are not just addressed by the market itself. In recent years new regulation has been established that also greatly impacts the domain of IoT and tries to address the security issues from a different angle. The most prominent example of such regulation and its implications for IoT are described in the next section.

10.5.2 GDPR and its Effect on IoT

With the widespread use of IoT technology and its impact on our everyday life, the need for specific regulations is growing [78]. There are billions of sensors deployed, tracking every single movement, noticing every single change. Bastos et al. indicate the massive recording of information: 'Who we are, where we are, what we do and how we do it' [40, p.1]. IoT devices collect a vast amount of data and it is thus important to not only look at how they can be secured sufficiently, but also how their handling and processing of the data can be governed. This is where regulation can have a meaningful impact.

A recent legislation with far-reaching consequences is the GDPR introduced in May 2018. GDPR's main objective is to protect and to regulate data privacy of EU citizens. The highly sensitive data collected by IoT devices must therefore be subjected to the GDPR as well. However, as Bastos et al. [40] point out, there are several hurdles in applying GDPR to the IoT environment. Subsequently, the principles of GDPR and challenges associated with it are described.

A) Consent

GDPR states that data subjects (i.e., natural persons) should be able to control which data is collected about them and that they can forbid it anytime [79]. The question arises, how these rules can be applied in the domain of IoT (e.g. What happens when a person visits a friend and the smart lock at the entrance is collecting video footage? Could guests deny the collection of their data during their visit?) It seems that current systems still do not provide such kind of control [40]. It gets even more delicate when IoT devices are placed in public areas. In this case, people are tracked without even being aware of the presence of monitoring devices. Another use case that includes third parties is when people are within the reach of sensors only for a short time. For example, when they sit in the train next to a person with IoT devices on them. Altogether, obtaining consent from third parties is even more difficult than getting consent from the active owners of IoT devices [40].

B) Data Minimisation

Purpose limitation and data minimisation are principles restricting data collection in general. Data should only be collected for a specific purpose and only as much as needed to fulfill this given purpose [79]. An example where IoT violates those principles are Smart Home systems: Sensors are constantly capturing audio to be

able to recognize user orders such as 'Turn on the light'. However, during the day it is unlikely that the light is needed for several hours, while IoT devices might still be constantly listening to voice commands.

C) **Transparent Processing**

Transparent processing refers to the user's capability to see how the data is handled, e.g., how many times a certain fact was recorded as well as where and through which channels it has been sent [40]. However, users (and passively concerned people) of the IoT technology usually are not informed about third parties nor can the manufacturers be completely sure how the data is handled. Data often jumps from device to device before arriving at the final destination where it is stored persistently. In GDPR, transparency also means the 'Right to be forgotten' (i.e., a person can at any time demand from a company to erase all their personal data) [79]. The company then has to track where all personal records are stored and remove the identified data, which might become a challenging task [40].

D) **Data Breach Reporting**

According to GDPR, companies must report a data breach within 72 hours after becoming aware of it. This rule is challenging especially within the IoT environment [40]. Finding and assessing a breach among hundreds of interconnected devices requires close cooperation of all stakeholders and vendors. Given the myriad of different devices this process might be almost impossible in reality.

E) **Privacy by Design & Data Security**

Privacy by design is accomplished when default settings do not have to be changed by the user to protect their privacy. As OWASP criticizes, this is not the case for most IoT devices [43]. Furthermore, the GDPR demands that vendors apply all necessary measures to protect the user's privacy and confidentiality. Taking the aforementioned characteristics of IoT devices into account, especially the limited resources, the deployment of such effective security mechanisms happens to be rather difficult [40].

The European Union is not the only institution which demands more stringent rules in the IoT environment. The US Federal Trade Commission (FTC) has also identified the need for action to mitigate risks concerning the lack of IoT privacy standards. Current legislative actions have a tendency towards consumer IoT devices. Only recently, this focus shifts also to government IoT, smart cities and critical infrastructures like power plants, transportation or the health system [78]. In the next Section 10.5.3, we will elaborate more about current projects and guidelines which aims to regulate the IoT market and its security.

Despite the eminent threat posed by unregulated and insecure IoT devices, many countries are reluctant in creating new regulation to not hinder innovation and economic growth. A harmonized movement around the globe could break up such motives [78]. If every single manufacturer has to comply with the same standard security and privacy rules, no country would suffer from IoT disadvantages. Furthermore, regarding the mobile nature of IoT devices, it would be reasonable to secure Swiss IoT devices in the same way as IoT devices placed a few kilometers away, on German soil.

10.5.3 Ongoing Projects and Guidelines

The products shown in Section 10.5.1 were built to protect vulnerable IoT devices. However, they do not solve all the problems that the devices have. Most often, the products are doing some sort of symptom control. But it is much more important that the products are designed to be safe from the ground up. For this reason one IoT project implementing secure access and network functionality and one IoT guideline for manufacturers, containing best practices for design, development and deployment of secure IoT services and products, are presented next. Other projects and guidelines worth to be mentioned are BITAG [80], CSA [81] and the projects by OWASP [43]. Even governmental institutions such as the American NIST [82] and Homeland Security[83] or transnational agencies like the European ENISA [84] provide guidelines.

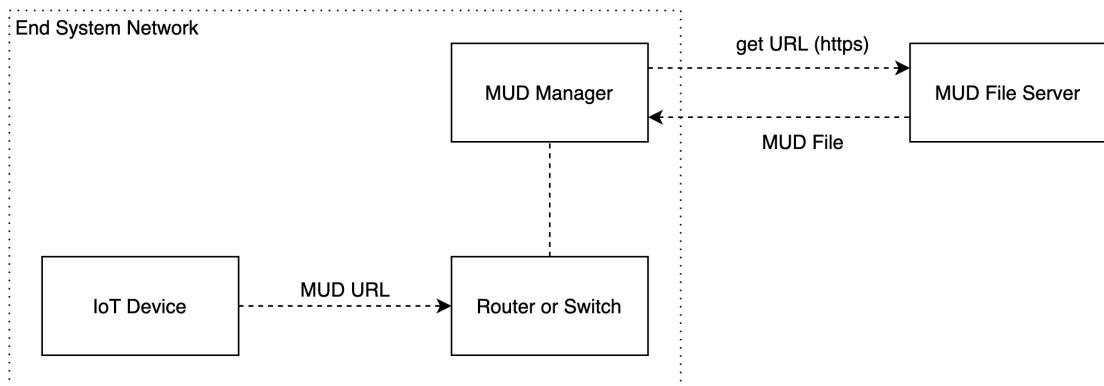


Figure 10.4: Manufacturer Usage Description architecture, own diagram based on [85]

- I) **IETF** The Internet Engineering Task Force (IETF) defines itself as "a large open international community of network designers, operators, vendors, and researchers concerned with the evolution of the Internet architecture and the smooth operation of the Internet" [86]. In order to make the attack surface of the communication of IoT devices smaller, the IETF launched a project and developed the *Manufacturer Usage Description* (MUD). It is based on the principle that each IoT device (e.g., a light bulb) has a specific purpose. And thus all other use cases are not wanted. Consequently, the MUD can be formulated from this. In the example of the light bulb, it should be able to be controlled remotely via the network and has a connection to perhaps a rendezvous service, so that it can be found by a smart phone app. In the MUD, it should then be defined that the light bulb only talks to the one rendezvous service, but not to other devices or services [85]. This ensures that the light bulb is only used for the intended purpose. The MUD and its architecture are depicted in Figure 10.4. First, the required components need to be explained. The MUD URL is a URL stored on the IoT device pointing to the MUD file server, usually provided by the manufacturer, from which the MUD can be downloaded. This is the responsibility of the MUD Manager, among others. Now a short example of how a MUD file is requested when a new IoT device is added to the network. First, the MUD URL is sent from the device (Thing) to a router or switch. This is usually embedded in the DHCP request. The router or switch passes the MUD URL to the MUD manager, which then downloads the MUD file from the manufacturer's server (MUD file server). Finally, the MUD Manager is responsible for ensuring that the specifications in the MUD file are implemented in the network.

As seen in the example above two essential components are necessary in the network, to ensure the functionality of the MUD specification. This can perhaps be seen as

the main drawback of MUD. Not only the manufacturer of the IoT device itself needs to comply with the specification also the switches and routers in the network have to implement the MUD protocol. And last but not least a MUD manager service must be run additionally. However, it seems that big networking companies such as Cisco want to adopt this protocol. Cisco announced in 2019 to provide MUD support in their enterprise network solutions [87].

- II) **GSMA** A popular guideline is published by the Global System for Mobile Communications Association (GSMA). This association "represents the interests of mobile operators worldwide, uniting more than 750 operators with almost 400 companies in the broader mobile ecosystem, including handset and device makers, software companies, equipment providers and internet companies, as well as organisations in adjacent industry sectors." [88]. Consequently, their words have quite a bit of range. Their *GSMA IoT Security Guidelines* provide detailed recommendations for the secure design, development and deployment of IoT services, networks and endpoints. Also several attack scenarios are included. It attempts to show the manufacturers how their products could be cracked, but then also how these scenarios can be prevented with common best practices. Furthermore, the guideline provides a security assessment framework, which can be used by manufacturers to test their products. If the manufacturers do not have sufficient resources or expertise, there is also the option of assessment as a service, whereby an external company goes through the assessment framework and tests their IoT solutions. [89]

10.5.4 Estimated Market Value & Outlook

The final subsection of the IoT security market analysis presents a number of statistics that display the historic growth of the market and estimate its future trajectory.

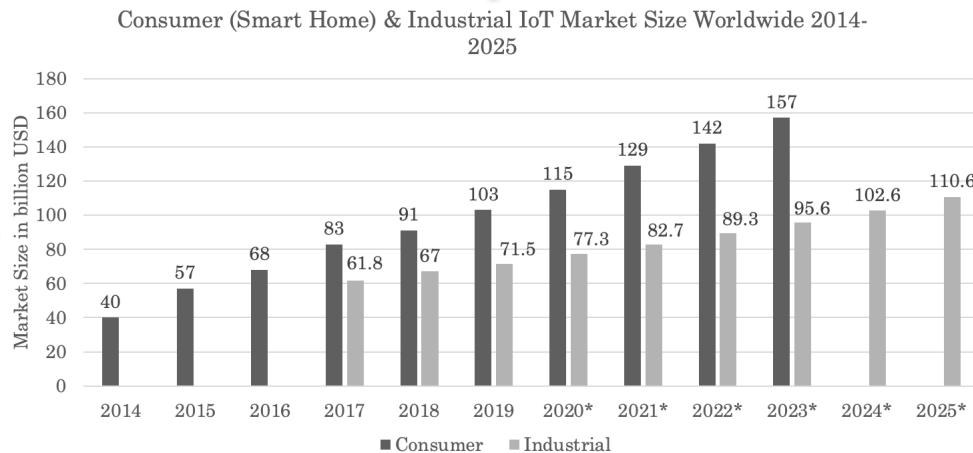


Figure 10.5: Consumer smart home and industrial IoT market size 2014 to 2025 based on data from [90] and [91] (* = forecasts)

The market for IoT devices is growing steadily. There are an estimated 31 billion IoT devices online this year (2020), and this is expected to grow to as many as 75 billion by 2025 [5]. The enormous growth can also be seen in the market sizes, seen in Figure 10.5. The consumer smart home market and the industrial market belong to the five biggest IoT sectors next to Smart Cities, Connected Health and Connected Cars. This is why those two markets were chosen as indicators for the IoT security market [92]. From 2017 to 2025, the market size in the industrial IoT segment alone is expected to rise from USD 61.8 billion to USD 110.6 billion. Which corresponds to growth by a factor of nearly two.

Similar rapid growth can also be observed in the smart home market segment. Here, USD 83 billion was spent in 2017 and is expected to grow to USD 157 billion by 2023. This again represents almost a doubling, and in comparison to the industrial market in an even shorter time span.

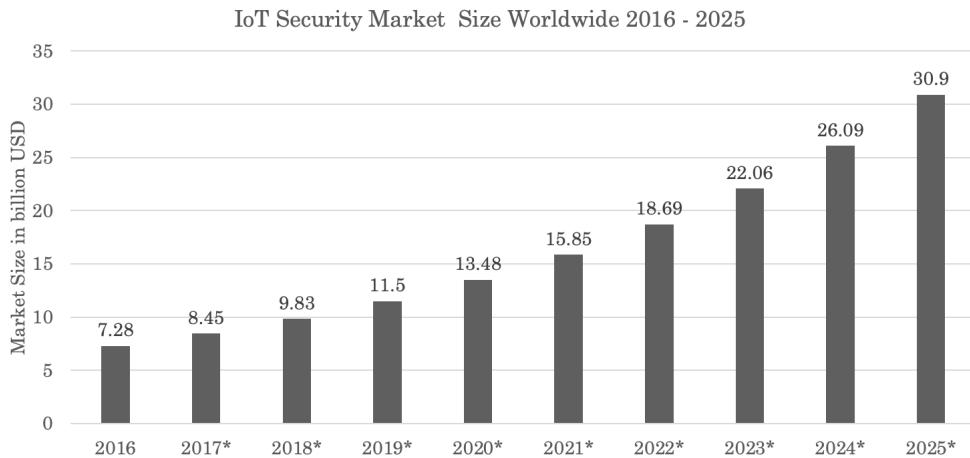


Figure 10.6: IoT security market size from 2016 to 2025 based on data from [93] (* = forecasts)

As seen in Section 10.3, many IoT systems still leave much to be desired in terms of security. Among other things, the hasty market launch of IoT products may have contributed to this fact. Business opportunities arise from such new challenges. A market study by MarketsAndMarkets [94] estimates the current market capitalisation for IoT security at USD 12.5 billion in 2020, and they expect it to rise to USD 36.6 billion by 2025. That would be more than doubling the size in 5 years. Another market estimate, from 2017, painted a similar picture as seen in Figure 10.6. It predicted a market size of USD 13.48 billion in 2020 and growth to USD 30.9 billion by 2025. This is also more than a doubling. Comparing this to how fast the industrial and smart home IoT sectors grew, it seems like a delayed response on the security side. And with a similarly large impact on the growth factor of the market size.

10.6 Discussion

Considering the vast numbers of IoT devices currently existing and the amount of devices that are estimated to exist in 2025, it is indisputable that IoT will have a major impact on our society and our economy [5]. IoT devices rapidly become part of our everyday life [45], and, sooner or later, all sectors will encounter them along their supply chain. As pointed out by the Congressional Research Service [8], some industries are already fully involved into the development of IoT applications: factories, medical institutions, homes and cities have been mentioned as the currently most interesting and developing fields. Due to this growth and entanglement into everyday life, IoT failures and attacks can be severe. Hence, IoT security is a concern of extreme significance [37]. Noor and Hassan [37] examine research projects from 2016-2018 and infer that there are several challenges in securing IoT devices and the networks. By reason of the particular characteristics of IoT devices, it is not feasible to apply traditional IT countermeasures, and dedicated security procedures need to be developed [35; 37; 38; 39; 44; 45]. Nevertheless, Noor and Hassan [37] conclude that there can be a fast progress of IoT security research identified, supported by various products emerging on the market.

Such products can be distinguished between Software, Hardware / Firmware, Service / Cloud and Home solutions, depending on where the security measures are applied to. For Software solutions, the most common approach are Intrusion Detection Systems;

systems, which monitor activities and trigger alerts when unusual behaviour is detected [50]. Hardware / Firmware solutions secure the devices without installing additional software and is especially intended for IoT manufacturers [69]. Service / Cloud products focus on securing the whole network of distributed devices by checking their configuration and monitoring their behaviour to detect unusual actions [76]. In private households, the most important objective of IoT security is to protect the users' privacy [77]. Home solutions scan, for example, the data traffic and make sure that no sensitive information is leaked.

With the emergence of IoT devices, the need for certain regulations arose. Multiple projects and working groups like IETF, GSMA, OWASP or BITAG elaborate best practices for design, development and deployment of secure IoT services and products. This shared knowledge makes it possible for smaller manufacturers without many years of expertise and with limited resources to build secure IoT ecosystems for the future. Furthermore, governmental agencies like the American Homeland Security or the European ENISA work on regulations or guidelines to protect the population. However, Bastos et al. [40] emphasize several gaps between existing regulations like GDPR and the IoT.

Limitations of Taxonomy

The threat taxonomy presented in Section 10.4.2 aims to provide a more exhaustive list of possible threat and attack vectors in IoT than existing taxonomies. However, there are also some limitations that are outlined in the following paragraph. The taxonomy is based on the three layer IoT architecture. Whereas the categorization based on layers helps to visualize where in the system the threat can occur it also limited due to its simplicity. Whenever possible the threats have been categorized based on the classification in the existing literature. However, for some threats it is not trivial where they should be positioned in the taxonomy. For example the threat of DDoS could affect the network or the application layer. Additionally, the taxonomy does not address that some threats differ in nature depending on the context: A node in the system could be on the receiving end of a DDos attack (i.e., if other nodes are sending requests to it) or on the sending end (i.e., if it was compromised and sends requests to another node). Garcia-Morchon et al. [38] highlight the importance of considering the complete lifecycle of IoT devices when addressing security issues. However, the taxonomy presented in this paper does not account for the dynamic nature of IoT systems where devices might join the network at any time and devices might belong to multiple owners during their lifecycle. It may thus give the false impression of addressing IoT security as a one-off task when in reality securing the system should be a continuous and on-going task.

Further versions of the taxonomy could address some of these issues by incorporating a more dynamic view of IoT security. Additionally they could include a ranking of the threats based on metrics such as the Common Vulnerability Scoring System (CVSS) by the National Institute of Standards and Technology (NIST) of the United States of America [95]. By incorporating data from public datasets such as the Common Vulnerabilities and Exposure (CVE) database [96] which lists known vulnerabilities an additional perspective could be gained on how common certain threats are. It might be interesting to explore how these threats can be combined to form common attack paths that a malicious user might take. Due to the limited scope and resources of this paper the taxonomy provided is not as exhaustive as it could be. Further versions could aggregate findings from more authors based on a rank threats by impact (e.g. using CVS data) allow combinations of attacks (show attack paths when threats / attacks are used in combination) include strategies / solutions to mitigate threats

10.7 Conclusion

Whereas the market of IoT in general seems to grow at hyper-speed the market for IoT security is still in its infancy. As said in the beginning of this paper, the vulnerabilities of IoT devices have been and will be exploited in cyberattacks; the Mirai Botnet or the computer worm Stuxnet will not be the last ones of their kind. However, recognizing the threats posed by the insecure IoT devices and identifying the need for more security measures is a first step into the right direction.

For the development of IoT security measures, it certainly helps to question the reasons of why it is so challenging to secure them. This has been done in this paper by analysing the particular characteristics of IoT devices. Features like usability, limited resources, ubiquity, short time-to-market and interconnectivity prove that traditional security measures cannot be applied one-to-one; dedicated models and products are needed to secure the IoT domain. The developed threat taxonomy (Table 10.2) could serve as a guideline for manufacturers to design more secure devices and to decrease the number of entry points a hacker can find to target an attack. Gratifyingly, as Section 10.5 shows, there are already several promising products on the market which aim to make the use of IoT technology more secure. Additionally to these available products, various institutions and working groups unite their forces and knowledge to formulate guidelines such that manufacturers build secure IoT devices in the first place. However, there is still a lot of room for additional security products and services as the growth trajectories of the markets show. To conclude, it can be said that the business landscape of IoT security is moving into the right direction. Nevertheless, it could enhance its speed as the numbers of IoT devices and hence the potential threats are increasing exponentially. More secure products should be developed and more affordable security measures offered. Consumers should get more responsible and security-aware, supported by regulations, guidelines and governments that pay enough attention to this market. Consumers just like manufacturers and governments have to take on their role to make the IoT world a saver place.

Bibliography

- [1] K. Schwab. (2016) The fourth industrial revolution: what it means, how to respond. Accessed: 2020-11-11. [Online]. Available: <https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/>
- [2] Q. Gou, L. Yan, Y. Liu, and Y. Li, “Construction and strategies in iot security system,” in *2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing*. IEEE, 2013, pp. 1129–1132.
- [3] P. Suresh, J. V. Daniel, V. Parthasarathy, and R. H. Aswathy, “A state of the art review on the internet of things (iot) history, technology and fields of deployment,” in *2014 International Conference on Science Engineering and Management Research (ICSEMR)*. IEEE, 2014, pp. 1–8.
- [4] C. T. Mark Patel, Jason Shangkuan. (2017) What’s new with the internet of things? Accessed: 2020-11-12. [Online]. Available: <https://www.mckinsey.com/industries/seminconductors/our-insights/whats-new-with-the-internet-of-things>
- [5] Cisco Systems Inc. (2020) Cisco annual internet report (2018–2023). Accessed: 2020-11-11. [Online]. Available: <https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.pdf>
- [6] G. D. Maayan. (2020) The iot rundown for 2020: Stats, risks, and solutions. Accessed: 2020-11-12. [Online]. Available: <https://securitytoday.com/Articles/2020/01/13/The-IoT-Rundown-for-2020.aspx?Page=1&p=1>
- [7] H. Tankovska. (2020) Forecast end-user spending on iot solutions worldwide from 2017 to 2025. Accessed: 2020-11-12. [Online]. Available: <https://www.statista.com/statistics/976313/global-iot-market-size/>
- [8] Congressional Research Service. (2020) The internet of things (iot): An overview. Accessed: 2020-11-12. [Online]. Available: <https://fas.org/sgp/crs/misc/IF11239.pdf>
- [9] S. Mukherjee, S. Patel, S. Kales, N. Ayas, K. Strohl, D. Gozal, and A. Malhotra, “An official american thoracic society statement: The importance of healthy sleep,” *American Journal Of Respiratory And Critical Care Medicine*, vol. 191, no. 12, pp. 1450–1458, 2015.
- [10] J. Tsai, E. S. Ford, C. Li, G. Zhao, and L. S. Balluz, “Physical activity and optimal self-rated health of adults with and without diabetes,” *BMC Public Health*, vol. 10, pp. 365–365, 2010.
- [11] R. Bharadwaj. (2019) The state of iot in insurance – automotive, home, and health. Accessed: 2020-11-12. [Online]. Available: <https://emerj.com/partner-content/iot-insurance-automotive-home-health/>

- [12] L. Horwitz. (2020) Can smart city infrastructure alleviate the strain of city growth? Accessed: 2020-11-12. [Online]. Available: <https://www.cisco.com/c/en/us/solutions/internet-of-things/smart-city-infrastructure.html>
- [13] J. Manyika, and others. (2013) Open data: Unlocking innovation and performance with liquid information. Accessed: 2020-11-12. [Online]. Available: <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/open-data-unlocking-innovation-and-performance-with-liquid-information>
- [14] S. Sicari, A. Rizzardi, L. A. Grieco, and A. Coen-Porisini, “Security, privacy and trust in internet of things: The road ahead,” *Computer networks*, vol. 76, pp. 146–164, 2015.
- [15] C. Kolias, G. Kambourakis, A. Stavrou, and J. Voas, “Ddos in the iot: Mirai and other botnets,” *Computer*, vol. 50, no. 7, pp. 80–84, 2017.
- [16] M. R. Jordan Robertson. (2014) Cybersecurity - mysterious '08 turkey pipeline blast opened new cyberwar. Accessed: 2020-12-09. [Online]. Available: <https://www.bloomberg.com/news/articles/2014-12-10/mysterious-08-turkey-pipeline-blast-opened-new-cyberwar>
- [17]
- [18] V. Voydock and S. Kent, “Security mechanisms in high-level network protocols,” *ACM Computing Surveys (CSUR)*, vol. 15, no. 2, pp. 135–171, 1983.
- [19] H. Federrath and A. Pfitzmann, “Gliederung und systematisierung von schutzzzielen in it-systemen,” *Datenschutz und Datensicherheit: DuD*, vol. 24, no. 12, pp. 704–710, 2000.
- [20] K. C. Laudon, *Wirtschaftsinformatik*. Pearson Deutschland.
- [21] R. Shirley. (2000) Internet security glossary. Accessed: 2020-10-22. [Online]. Available: <https://www.rfc-editor.org/rfc/pdfrfc/rfc2828.txt.pdf>
- [22] D. Evans, G. Jarboe, H. Thomases, M. Smith, and C. Treadaway, *Networking Complete 3rd Edition*, 3rd ed. New York: Wiley, 2002.
- [23] T. J. Grant, *Network Topology in Command and Control*. IGI Global, 2014.
- [24] Cisco Systems Inc., “What is network topology?” accessed: 2020-10-26. [Online]. Available: <https://www.cisco.com/c/en/us/solutions/automation/network-topology.html>
- [25] A. Cilfone, L. Davoli, L. Belli, and G. Ferrari, “Wireless mesh networking: An iot-oriented perspective survey on relevant technologies,” *Future Internet*, vol. 11, no. 4, p. 99, 2019.
- [26] J. Postel. (1981) Rfc 793 - transmission control protocol. Accessed: 2020-11-11. [Online]. Available: <http://www.ietf.org/rfc/rfc793.txt>
- [27] S. Li, “Security Architecture in the Internet of Things,” *Securing the Internet of Things*, no. Mic, pp. 27–48, 2017.
- [28] P. Sethi and S. R. Sarangi, “Internet of Things: Architectures, Protocols, and Applications,” *Journal of Electrical and Computer Engineering*, vol. 2017, 2017.

- [29] M. binti Mohamad Noor and W. H. Hassan, "Current research on Internet of Things (IoT) security: A survey," *Computer Networks*, vol. 148, pp. 283–294, 2019, accessed: 2020-12-02. [Online]. Available: <https://doi.org/10.1016/j.comnet.2018.11.025>
- [30] A. Bujari, M. Furini, F. Mandreoli, R. Martoglia, M. Montangero, and D. Ronzani, "Standards, Security and Business Models: Key Challenges for the IoT Scenario," *Mobile Networks and Applications*, vol. 23, no. 1, pp. 147–154, 2018. [Online]. Available: <https://doi.org/10.1007/s11036-017-0835-8>
- [31] J. Zhang, H. Chen, L. Gong, J. Cao, and Z. Gu, "The Current Research of IoT Security," in *2019 IEEE Fourth International Conference on Data Science in Cyberspace (DSC)*. IEEE, 06 2019, pp. 346–353. [Online]. Available: <https://ieeexplore.ieee.org/document/8923684/>
- [32] W. Stallings, "The Internet of Things: Network and Security Architecture," *The Internet Protocol Journal*, vol. 18, no. 4, pp. 2–24, 2015.
- [33] P. Y. Zhang, M. C. Zhou, and G. Fortino, "Security and trust issues in Fog computing: A survey," *Future Generation Computer Systems*, vol. 88, pp. 16–27, 2018. [Online]. Available: <https://doi.org/10.1016/j.future.2018.05.008>
- [34] S. Singh and N. Singh, "Internet of things (iot): Security challenges, business opportunities & reference architecture for e-commerce," in *2015 International Conference on Green Computing and Internet of Things (ICGCIoT)*. IEEE, 2015, pp. 1577–1581.
- [35] C. Koliас, A. Stavrou, and J. Voas, "Securely making things right," *Computer*, vol. 48, no. 9, pp. 84–88, 2015.
- [36] S. Babar, P. Mahalle, A. Stango, N. Prasad, and R. Prasad, "Proposed security model and threat taxonomy for the internet of things (iot)," in *International Conference on Network Security and Applications*. Springer, 2010, pp. 420–429.
- [37] M. b. M. Noor and W. H. Hassan, "Current research on internet of things (iot) security: A survey," *Computer networks*, vol. 148, pp. 283–294, 2019.
- [38] O. Garcia-Morchon, S. Kumar, and M. Sethi, "Internet of Things (IoT) Security: State of the Art and Challenges," *Internet Research Task Force (IRTF)*, no. 8576, pp. 1–50, 2019. [Online]. Available: <https://rfc-editor.org/rfc/rfc8576.txt>
- [39] J. Pacheco, D. Ibarra, A. Vijay, and S. Hariri, "Iot security framework for smart water system," in *2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA)*. IEEE, 2017, pp. 1285–1292.
- [40] D. Bastos, F. Giubilo, M. Shackleton, and F. El-Moussa, "Gdpr privacy implications for the internet of things," in *4th Annual IoT Security Foundation Conference, London, UK*, 2018.
- [41] Ponemon. (2020) Ponemon institute. Accessed: 2020-11-12. [Online]. Available: <https://www.ponemon.org/>
- [42] M. Drolet. (2018) What does stolen data cost [per second]. Accessed: 2020-10-28. [Online]. Available: <https://www.csoonline.com/article/3251606/what-does-stolen-data-cost-per-second.html>

- [43] “OWASP Internet of Things Project ,” 2018, accessed: 2020-12-16. [Online]. Available: https://wiki.owasp.org/index.php/OWASP{_-}Internet{_-}of{_-}Things{_-}Project{#}tab>Main
- [44] S. Notra, M. Siddiqi, H. H. Gharakheili, V. Sivaraman, and R. Boreli, “An experimental study of security and privacy risks with emerging household appliances,” in *2014 IEEE conference on communications and network security*. IEEE, 2014, pp. 79–84.
- [45] D. Minoli, K. Sohraby, and J. Kouns, “IoT security (IoTSec) considerations, requirements, and architectures,” in *2017 14th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, 2017, pp. 1006–1007.
- [46] G. Jonsdottir, D. Wood, and R. Doshi, “Iot network monitor,” in *2017 IEEE MIT Undergraduate Research Technology Conference (URTC)*. IEEE, 2017, pp. 1–5.
- [47] R. van der Meulen. (2017) Gartner says 8.4 billion connected "things" will be in use in 2017, up 31 percent from 2016. Accessed: 2020-10-17. [Online]. Available: <https://www.gartner.com/en/newsroom/press-releases/2017-02-07-gartner-says-8-billion-connected-things-will-be-in-use-in-2017-up-31-percent-from-2016>
- [48] H. Hellaoui, M. Koudil, and A. Bouabdallah, “Energy-efficient mechanisms in security of the internet of things: A survey,” *Computer Networks*, vol. 127, pp. 173–189, 2017.
- [49] C. Bormann, M. Ersue, and A. Keranen, “Terminology for constrained-node networks,” *Internet Engineering Task Force (IETF): Fremont, CA, USA*, pp. 2070–1721, 2014, accessed: 2020-12-07. [Online]. Available: <https://www.hjp.at/doc/rfc/rfc7228.html>
- [50] B. B. Zarpelão, R. S. Miani, C. T. Kawakani, and S. C. de Alvarenga, “A survey of intrusion detection in internet of things,” *Journal of Network and Computer Applications*, vol. 84, pp. 25–37, 2017.
- [51] M. Vega. (2020) Internet of things statistics, facts & predictions [2020's update]. Accessed: 2020-10-17. [Online]. Available: <https://review42.com/internet-of-things-stats/>
- [52] S. Liu. (2020) Internet of things (iot) - statistics & facts. Accessed: 2020-11-09. [Online]. Available: <https://www.statista.com/topics/2637/internet-of-things/>
- [53] A. K. Das, S. Zeadally, and D. He, “Taxonomy and analysis of security protocols for Internet of Things,” *Future Generation Computer Systems*, vol. 89, pp. 110–125, 2018. [Online]. Available: <https://doi.org/10.1016/j.future.2018.06.027>
- [54] I. Alqassem and D. Svetinovic, “A taxonomy of security and privacy requirements for the Internet of Things (IoT),” *IEEE International Conference on Industrial Engineering and Engineering Management*, vol. 2015-January, pp. 1244–1248, 2014.
- [55] R. Roman, J. Zhou, and J. Lopez, “On the features and challenges of security and privacy in distributed internet of things,” *Computer Networks*, vol. 57, no. 10, pp. 2266–2279, 2013. [Online]. Available: <http://dx.doi.org/10.1016/j.comnet.2012.12.018>
- [56] S. Tweneboah-Koduah, K. E. Skouby, and R. Tadayoni, “Cyber Security Threats to IoT Applications and Service Domains,” *Wireless Personal Communications*, vol. 95, no. 1, pp. 169–185, 2017.

- [57] M. Nawir, A. Amir, N. Yaakob, and O. B. Lynn, "Internet of Things (IoT): Taxonomy of security attacks," *2016 3rd International Conference on Electronic Design, ICED 2016*, pp. 321–326, 2017.
- [58] J. M. Stewart, "Explain the difference between identification and authentication (identity proofing). , " 2020, accessed: 2020-12-15. [Online]. Available: https://www.oreilly.com/library/view/comptia-securitytm-review/9780470404843/9780470404843_{-}explain_{-}the_{-}difference_{-}between_{-}identific.html
- [59] K. Scarfone, J. Wayne, and M. Tracy, "Confidentiality, Integrity, and Availability - Archive of obsolete content | MDN," jun 2018. [Online]. Available: https://developer.mozilla.org/en-US/docs/Archive/Security/Confidentiality,{-}Integrity,{-}and{_-}Availability
- [60] "Data Privacy vs. Data Security [definitions and comparisons]," oct 2020, accessed: 2020-10-20. [Online]. Available: <https://dataprivacymanager.net/security-vs-privacy/>
- [61] S. Pokorni, "Reliability and availability of the Internet of things," *Vojnotehnicki glasnik*, vol. 67, no. 3, pp. 588–600, 2019.
- [62] J. A. Onieva, J. Zhou, and J. Lopez, "Multiparty nonrepudiation: A survey," *ACM Computing Surveys*, vol. 41, no. 1, 2008.
- [63] J. Zhou, Z. Cao, X. Dong, and A. V. Vasilakos, "Security and privacy for cloud-based iot: Challenges," *IEEE Communications Magazine*, vol. 55, no. 1, pp. 26–33, 2017.
- [64] T. Martin, D. Geneiatakis, I. Kounelis, S. Kerckhof, and I. N. Fovino, "Towards a formal iot security model," *Symmetry*, vol. 12, no. 8, pp. 1–16, 2020.
- [65] J. Pacheco and S. Hariri, "IoT security framework for smart cyber infrastructures," *Proceedings - IEEE 1st International Workshops on Foundations and Applications of Self-Systems, FAS-W 2016*, pp. 242–247, 2016.
- [66] K. Kimani, V. Oduol, and K. Langat, "Cyber security challenges for IoT-based smart grid networks," *International Journal of Critical Infrastructure Protection*, vol. 25, pp. 36–49, 2019. [Online]. Available: <https://doi.org/10.1016/j.ijcip.2019.01.001>
- [67] I. 020 ARMIS. (2020) Agentless security forthe enterprise of things. Accessed: 2020-11-11. [Online]. Available: https://info.armis.com/rs/645-PDC-047/images/Armis_Solution_Brief.pdf
- [68] B. N. I. Security. (2020) Bastille. Accessed: 2020-12-15. [Online]. Available: <https://www.bastille.net/>
- [69] R. Labs. (2020) Centrifuge platform®. Accessed: 2020-11-12. [Online]. Available: <https://www.refirmlabs.com/centrifuge-platform/>
- [70] R. P. Foundation. (2020) Raspberry pi. Accessed: 2020-12-16. [Online]. Available: <https://www.raspberrypi.org/products/raspberry-pi-4-model-b/>
- [71] Zymbit. (2020) Zymkey. Accessed: 2020-12-15. [Online]. Available: <https://www.zymbit.com/zymkey/>
- [72] Amazon. (2020) Iot device defender. Accessed: 2020-12-16. [Online]. Available: <https://aws.amazon.com/iot-device-defender/>

- [73] Google. (2020) Google iot core. Accessed: 2020-12-16. [Online]. Available: <https://cloud.google.com/iot-core/>
- [74] Microsoft. (2020) Azure defender for iot. Accessed: 2020-12-16. [Online]. Available: <https://azure.microsoft.com/en-us/services/azure-defender-for-iot/>
- [75] IBM. (2020) Ibm iot platform. Accessed: 2020-12-16. [Online]. Available: <https://www.ibm.com/business-operations/iot-platform/>
- [76] Amazon. (2020) Aws iot device defender. Accessed: 2020-12-16. [Online]. Available: <https://aws.amazon.com/iot-device-defender/>
- [77] Bitdefender. (2020) Bitdefender box. Accessed: 2020-11-12. [Online]. Available: <https://www.bitdefender.com/box/>
- [78] M. Zervaki, "Regulating the iot: 2020 and beyond," accessed: 2020-10-20. [Online]. Available: <https://www.accesspartnership.com/cms/access-content/uploads/2020/06/Regulating-the-IoT-2020-and-beyond.pdf>
- [79] E. Union, "General data protection regulation," accessed: 2020-10-29. [Online]. Available: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679>
- [80] BITAG. (2020) Bitag. Accessed: 2020-12-16. [Online]. Available: <https://www.bitag.org/report-internet-of-things-security-privacy-recommendations.php>
- [81] CSA. (2020) Csa iot security controls framework. Accessed: 2020-12-16. [Online]. Available: <https://cloudsecurityalliance.org/artifacts/iot-security-controls-framework>
- [82] NIST. (2020) Nist iot security. Accessed: 2020-12-16. [Online]. Available: <https://www.nist.gov/internet-things-io>
- [83] H. Security. (2020) Homeland security iot. Accessed: 2020-12-16. [Online]. Available: <https://www.dhs.gov/securingtheIoT>
- [84] ENISA. (2020) Enisa security iot. Accessed: 2020-12-16. [Online]. Available: <https://www.enisa.europa.eu/publications/guidelines-for-securig-the-internet-of-things>
- [85] D. R. E. Lear, R. Droms. (2019) Manufacturer usage description specification. Accessed: 2020-12-16. [Online]. Available: <https://tools.ietf.org/pdf/rfc8520>
- [86] IETF. (2020) Who we are. Accessed: 2020-12-16. [Online]. Available: <https://www.ietf.org/about/who/>
- [87] L. Su. (2019) Mud is officially approved by ietf as an internet standard, and cisco is launching mud1.0 to protect your iot devices. Accessed: 2020-12-16. [Online]. Available: <https://blogs.cisco.com/security/mud-is-officially-approved-by-ietf-as-an-internet-standard-and-cisco-is-launching-mud1-0-to-protect-your-iot-devices>
- [88] GSMA. (2020) About us. Accessed: 2020-12-16. [Online]. Available: <https://www.gsma.com/aboutus/>
- [89] ——. (2020) Gsma iot security guidelines. Accessed: 2020-12-16. [Online]. Available: <https://www.gsma.com/iot/iot-security/iot-security-guidelines/>

- [90] H. Tankovska. (2020) Consumer spending on smart home systems worldwide from 2014 to 2023. Accessed: 2020-12-10. [Online]. Available: <https://www.statista.com/statistics/693303/smart-home-consumer-spending-worldwide/>
- [91] S. R. Department. (2020) Industrial internet of things market size worldwide from 2017 to 2025*. Accessed: 2020-12-10. [Online]. Available: <https://www.statista.com/statistics/611004/global-industrial-internet-of-things-market-size/>
- [92] ipropertymanagement.com. (2020) Global iot market distribution in 2019, by sector. Accessed: 2020-12-16. [Online]. Available: <https://www.statista.com/statistics/1095380/global-iot-market-distribution-by-sector/>
- [93] theinsightpartners.com. (2017) Size of the internet of things (iot) security market worldwide from 2016 to 2025. Accessed: 2020-12-10. [Online]. Available: <https://www.statista.com/statistics/993789/worldwide-internet-of-things-security-market-size/>
- [94] Markets and Markets. (2020) Iot security market. Accessed: 2020-11-12. [Online]. Available: <https://www.marketsandmarkets.com/Market-Reports/iot-security-market-67064836.html>
- [95] National Institute of Standards and Technology, “Vulnerability Metrics,” 2020. [Online]. Available: <https://nvd.nist.gov/vuln-metrics/cvss>
- [96] MITRE Corporation, “Vulnerabilities By Type,” 2019. [Online]. Available: <https://www.cvedetails.com/vulnerabilities-by-types.php>

Chapter 11

Randomization of MAC Addresses: Market Opportunity or Privacy Violation?

Csanad Erdei-Griff, Anna Jancso, Andris Prokofjevs

MAC address-based location tracking brings with it a variety of new use cases; we go into detail about new options for navigation, indoor positioning systems, customer and pedestrian behaviour tracking, contact tracing and the recovery of stolen devices. Furthermore, we explore how this technology can potentially be used maliciously, especially when it comes to privacy violations. We also look into how personal data is handled legally by different countries and various ways of how data collection can be used for monetary gain. Also, we explore ethical aspects of data collection and ask ourselves whether the average user should be made more aware of what kind of data they are willingly handing over to data collection companies.

Lastly, we believe that service providers should openly state in an easy-to-understand format which kind of data they collect so that users are well informed when they make the decision whether they are willing to use the product. Moreover, it should be possible to set privacy preferences on a device-level, which should then serve as the default settings for every new service the user may opt into using.

Contents

11.1 Introduction and Problem Statement	321
11.2 Background	321
11.2.1 MAC address	321
11.2.2 Fingerprinting	322
11.2.3 Tools for Wi-Fi Tracking	325
11.2.4 Spoofing	326
11.3 Use Cases	327
11.3.1 Benefits	328
11.3.2 Security Issues	333
11.4 Evaluations and Discussions	334
11.4.1 Legal Aspects	334
11.4.2 Economic Aspects	338
11.4.3 Ethical Aspects	340
11.4.4 Technical Aspects	342
11.5 Summary and Conclusions	344

11.1 Introduction and Problem Statement

The data collected from tracking people's movement, their habits and routines are a valuable resource for improvement and profit in many economical sectors [57][65][56][12][36][19]. In the past decade, a new technology emerged, which complements existing data collection methods; MAC address fingerprinting.

By using any Wi-Fi or Bluetooth capable device and having the corresponding beacon switched on, the device will periodically send out probe requests to nearby access points, being able to detect which ones it can connect with [42]. This probe request happens to contain the device's UUID (Universal Unique Identifier) in the form of the MAC address [55]. Since the MAC address is, in theory, unique to one device, this makes it possible for a wireless access point to detect whenever that specific device is within its proximity. If an organization has a whole network of such access points, they can essentially track a device's location dynamically within that area, without the owner of the device ever noticing.

One way of avoiding such tracking is a method called MAC address randomization [34]. The device will periodically generate a new MAC address on the software level, thus rendering the tracking ineffective. However, since the randomization can only take place on a software and not a hardware level, there are ways of circumventing it [18]. This report summarizes a variety of attacks which can be performed to overcome randomized MAC addresses, making device fingerprinting and tracking possible again.

Furthermore, it goes into detail about some subjectively positive and negative use cases of MAC address based device tracking, including potential benefits and dangers of the technology. Lastly, legal, economic, ethical and technical aspects of not only MAC address-based tracking, but tracking and personal data collection are discussed. This is done based on other papers on the topic.

11.2 Background

In the following sections, we introduce and explain terminology that is crucial for understanding how MAC address randomization works and what it is used for.

11.2.1 MAC address

The Media Access Control address (MAC address) is a globally unique identifier that manufacturers burn into the Network Interface Controller (NIC) of devices. The MAC address is 48 bit long and typically represented as 6 groups of 2 hexadecimal digits. The first three octets, known as the Organizationally Unique Identifier (OUI), identifies the manufacturer. The manufacturer can then assign the rest of the bits to the NICs. [44] have demonstrated that this NIC part of the address in turn can reveal information about the device and the model. An example MAC address is shown in Figure 11.1.

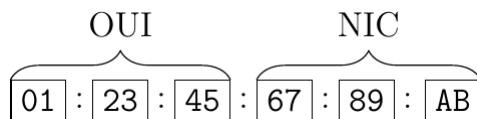


Figure 11.1: Example MAC address [43]

Apart from global, unique MAC addresses, locally assigned addresses are in use as well which are not expected to be unique across the world. They are used in different contexts such as mobile device-tethered hotspots and peer-to-peer networks but also as part of MAC address randomization (cf. Section 11.2.4) [43]. Global and local addresses may

be distinguished by the Universal/Local bit in the most significant byte. Similar to the OUI, manufacturers can buy a prefix, the so-called Company Identifier (CID), from the Institute of Electrical and Electronics Engineers (IEEE) for locally assigned addresses. Accordingly, this CID has the local bit set. However, [44] found that some companies such as Apple do not seem to adhere to these rules by utilizing prefixes that were licensed to other companies.

The MAC address occurs in a variety of technologies such as Wi-Fi, Bluetooth and Ethernet and each of their interfaces have their own separate MAC address. MAC addresses are crucial for layer 2 device communications (data link layer in the OSI model). Thus, MAC addresses are only relevant on a Local Area Network (LAN). When two devices talk to each other in the same LAN, they need to include both their (source) and the target's device (destination) MAC address in the frame's header. If the MAC address of the target device is not known beforehand, it can be found using the Address Resolution Protocol (ARP) which maps IP addresses to MAC addresses. Whenever data leaves the LAN, IP addresses are mainly used for routing and the MAC addresses changes as the packet travels from router to router.

11.2.2 Fingerprinting

In the following, we introduce what fingerprinting of the devices and identities is, using Wi-Fi technology as illustration, and how it can leverage MAC addresses for that. Later, we mention other fingerprinting techniques that work on a global scale or use different technologies such as Bluetooth.

Device fingerprinting through Wi-Fi

Many portable devices these days such as smartphones, tablets and other wearables integrate Wi-Fi for internet access. In the infrastructure mode, these devices communicate via Access Points (APs). Before a device (also called the *station*) can communicate with other devices, it needs to discover the APs within range and connect to one. There are two different discovery mechanisms [30].

- Active scanning: The station periodically broadcasts probe request frames to which APs can respond.
- Passive scanning: APs periodically emit beacons to advertise their presence to which stations can listen.

Passive scanning is less efficient than active scanning because the station has to try to listen and wait on all channels and might miss beacons when not waiting long enough. Therefore, passive scanning takes more time on average. In active scanning, each of these frames includes the MAC address of the station in plain-text since only the payload is encrypted but not the header. This is also shown in Figure 11.2.

Moreover, once the station is associated with an AP, frame transmissions and receptions also include the MAC address. This makes tracking of the movements of a device trivial as the MAC address is globally unique and therefore uniquely identifies a device. This is also referred to as device fingerprinting. Fingerprinting methods are classified into three categories [42]:

- Active fingerprinting: The fingerprinter (tracking device) can initiate a connection to the fingerprintee (tracked device).
- Passive fingerprinting: The fingerprinter only observes traffic coming from or going to a fingerprintee.

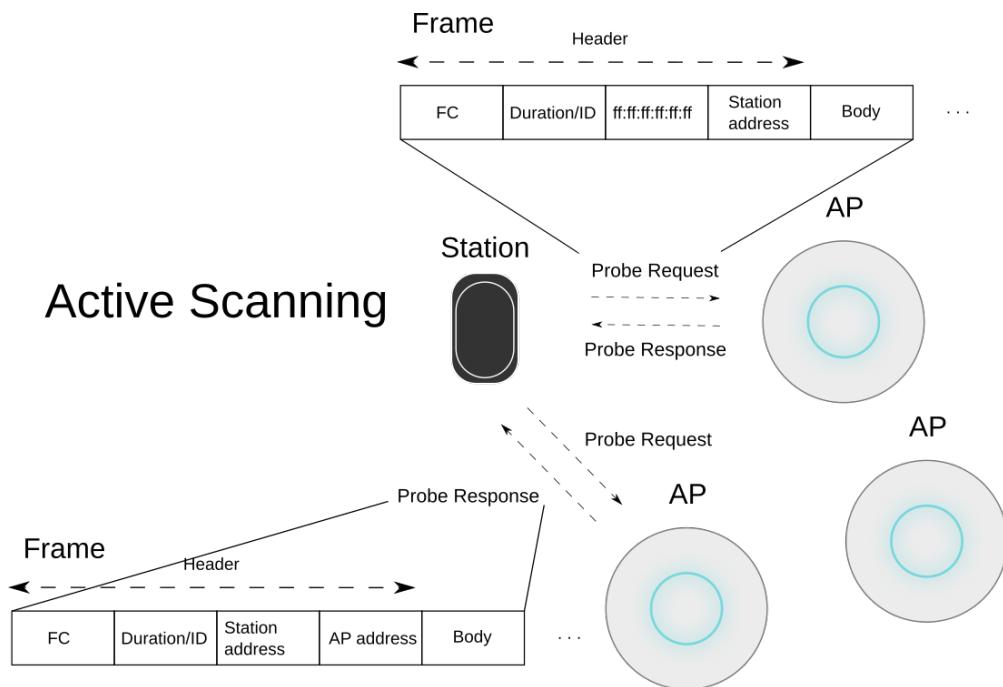


Figure 11.2: Active scanning with probe requests.

- Semi-passive fingerprinting: The fingerprintee initiates a connection to the fingerprinter which allows the fingerprinter to interact and collect data about it. A website would be an example here as the fingerprintee sends requests to it.

Passive fingerprinting cannot be detected by anyone, whereas in active and in semi-passive the tracking devices can be known.

For Wi-Fi, device fingerprinting is especially facilitated due to the increasing numbers of hotspots and by the fact that most mobile device users keep the Wi-Fi enabled most of the time.

A simple device tracking system can be described as follows: a set of monitoring nodes are deployed in an area of interest. Each node records the MAC addresses by listening into the communications and forwards them to a central server where the data is aggregated and analyzed [18]. This is the passive tracking scenario.

In the semi-passive tracking scenario, the APs routinely collect the MAC addresses and events associated with them such as association, authentication and disconnection along with a timestamp in logs [17]. In an analogous fashion, operators of APs can also push this data to a central server for analysis [10]. Institutions like universities or shopping centers already have an existing infrastructure in place that would easily allow them to track devices [60]. In both settings, this gives not just information about which devices are present, but also how long and how frequent those devices are in the area [18]. Moreover, one can infer which places a user has visited such as whether this user went to the library or took part in a political event [34].

Device Fingerprinting through Bluetooth

The same way as a device's MAC address can be read from the Wi-Fi probing signal it emits, it can be read from the any Bluetooth signal it gives off as well [2]. In practice, this means that as long as a device's Bluetooth beacon is switched on, it will become discoverable and trackable by any other device with Bluetooth capability. Thus, by using wireless headphones, a wireless keyboard or a wireless pen, a person is automatically exposed to location tracking.

Identity fingerprinting

Since many people do not share their devices with other people, it is also possible to link the address to a particular individual, once the mapping between the device and the individual has been established [34]. It is important to note, though, that the MAC address itself cannot give a direct link to the identity of the device's owner. However, [17] presented approaches for finding the link between the MAC address and the real owner's identity (here also called the *target*). To demonstrate their practicality, they formulated two properties that their methods shall uphold [17]:

- Accuracy: the MAC address that has been matched with an identity is really the MAC address of the identity's device.
- Stealthiness: the target does not notice that (s)he is being tracked.

Two factors constrain the feasibility of their methods. Firstly, devices emit probe requests at given intervals. [17] experimentally observed that for example an iPhone sends probe requests every 45 seconds whereas a Samsung device does so every 30 seconds. Secondly, devices have limited Wi-Fi transmission range which can vary according to the type of device. [17], for instance, reported 100m for the Samsung device and 30m for the iPhone. Furthermore, [17] propose the following two approaches which we will shortly describe:

- Beacon Replay Attack: This attack assumes that both the home and the work location of the target are known and that at both places the target connects to some wireless network. It works as follows: The attacker goes to the home location of the target and extracts and records the SSID and security features in the beacon frames transmitted by all APs in range. Usually, this can be done at any time of the day since most people leave their APs on 24h/365 days. This can be done without raising too much suspicion since the attacker only has to walk past the house in which the target lives, preserving the stealthiness aspect. Later, the attacker goes to the work location while the target is working there. The attacker creates a rogue AP there impersonating the home AP by replaying the beacons. This causes the target's device to make a association request to these. The attacker can then extract the MAC address from this response. Since the probability that more than one person work and live at the same places is small [28], the link between the MAC address and the target can be established with high confidence. This attack is illustrated in Figure 11.3.
- Stalker attack: The easiest way to associate a device's MAC address with a target is to be alone with the target, which implies that the distance between the stalker and the target has to be small. This, however, compromises the stealthiness property. [17] propose that the target be tracked at a greater distance. This greater distance also means that the MAC addresses of other individuals within range is recorded. To exclude those MAC addresses, the tracker needs to stalk the target by following him/her for a while at a reasonable distance but still within transmission range of his/her device. Since bystanders will move away from the target over time, ultimately, only one single MAC address will be continuously registered, which is the one belonging to the target. The authors empirically evaluated the contact length by moving around a city for a specific period of time with a tracking device and observed that the vast majority of contacts are short (less than 10 minutes). Obviously, this method only is viable if the target is alone and not in company of another person.

Identity fingerprinting is more difficult to realize than device fingerprinting because it involves more human intervention by stalking the person of interest while device fingerprinting is mostly automated once the monitoring nodes have been set up. This is also the

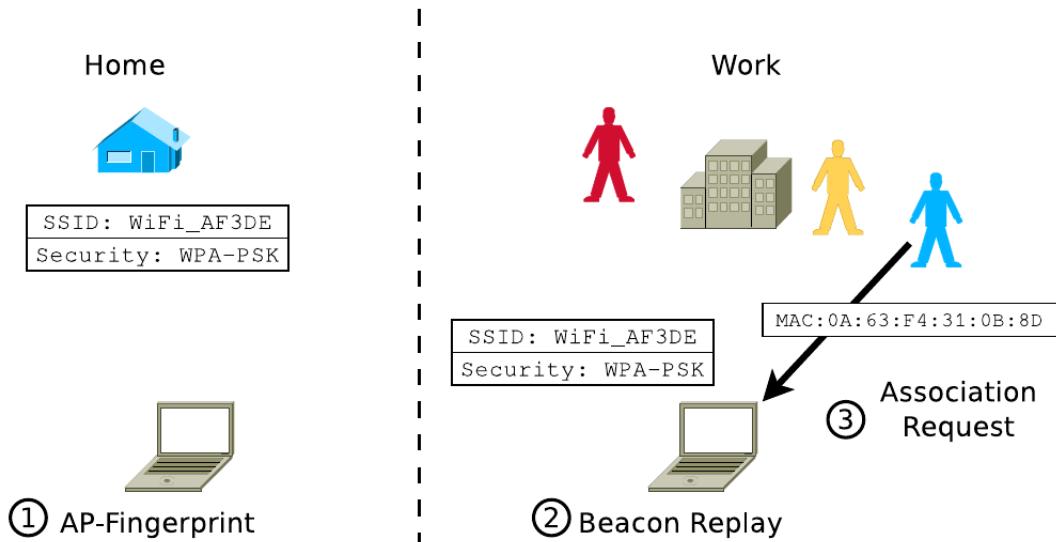


Figure 11.3: Beacon replay attack [17]

reason why, unlike device fingerprinting, it is not feasible to collect and analyze massive data sets with identity fingerprinting. Moreover, identity fingerprinting is more difficult to conceal since people have to be stalked while device fingerprinting is stationary. Especially, the stalker attack might be tricky to carry out when the transmission range is short and keeping an eye on the target is hampered by obstacles.

Other fingerprinting techniques

While MAC address tracking happens on a local level because MAC addresses are 2-layer address, global tracking can also be possible since layer-2 addresses might be encoded into the 3-Layer addresses [10]. The IPv6 Stateless Address Auto-configuration (SLAAC) [58] generates the Interface Identifier (IID) of the Layer-3 address from the layer-2 address.

11.2.3 Tools for Wi-Fi Tracking

There are different kinds of tools for Wi-Fi tracking, giving a wide array of options to choose from based on their needs and preferences.

Companies, such as the Toronto-based Aislelabs [4], provide complete out-of-the-box solutions, including all hardware, software, databases and data analysis. They deliver finished analytics to their customers, making their and other similar companies' products a viable option for non-technology centered service providers like restaurant, hotel or retail store chains.

For private persons willing to set up a smaller local tracking network, a good solution is to use a CreepyDOL (Creepy Distributed Object Locator) system [52]. It was developed in 2013 by the former US Military cyber security expert Brendan O'Connor. The system consists of multiple small sensor devices, which are each comprised of a Raspberry Pi Model A, two Wi-Fi adapters and a USB hub, which can be distributed in a desired area to create a network. The sensors will then send their collected data to a central data-processing server. However, a CreepyDOL system does not only collect MAC address data, it can also access certain applications like Dropbox and has the potential to recover sensitive personal and location information. For this reason, using the CreepyDOL system falls into a legally grey area.

A more ad-hoc method for educational purposes would be using a private access point in combination with tools such as aircrack-ng [3] and Wireshark [64][17], both of which are open-source and available free of charge. Using these with a compatible Wi-Fi interface, it

is possible to capture MAC addresses of Wi-Fi devices along with other private information such as SSIDs [17], as seen in Figure 11.4 (note that the last four characters of every MAC address are anonymized).

```

total number of devices : 32
Apple, Inc. | 40:a6:d9:ee:__:_ | -28 dB | 1 | ''
SAMSUNG ELECTRO | 20:64:32:c1:__:_ | -45 dB | 1 | ''
Murata Manufact | 00:37:6d:ea:__:_ | -88 dB | 1 | ''
Apple, Inc | 00:26:b0:7d:__:_ | -43 dB | 1 | ''
RIM | a0:6c:ec:2a:__:_ | -67 dB | 3 | 'SSID_1','SSID_2'
Apple Inc | 70:56:81:bb:__:_ | -58 dB | 1 | ''
Agere Systems | 00:02:2d:bf:__:_ | -49 dB | 1 | ''SSID_4'
Apple, Inc | f8:1e:df:d9:__:_ | -50 dB | 1 | ''SSID_5'
Murata Manufact | 00:37:6d:42:__:_ | -89 dB | 1 | ''
Intel Corporate | 00:24:d7:59:__:_ | -57 dB | 1 | ''SSID_6'
LG Electronics | 10:68:3f:4e:__:_ | -58 dB | 1 | ''
Apple, Inc. | 24:ab:81:8d:__:_ | -82 dB | 1 | ''
Apple, Inc. | 58:55:ca:f3:__:_ | -91 dB | 1 | ''
Intel Corporate | 00:21:6a:7f:__:_ | -76 dB | 1 | ''
...

```

Figure 11.4: Example Information obtained from Wi-Fi Probe Requests [17]

11.2.4 Spoofing

On the hardware level, the MAC address cannot be changed as it is hard-coded on the NIC. However, it is possible to change the MAC address on the software level. This is also called spoofing. Spoofing the MAC address is very easy. For example the following command on Linux, changes the MAC address:

```
sudo ifconfig eth0 hw ether xx:xx:xx:xx:xx:xx
```

eth0 being the network interface of your choice. Spoofing can have both legitimate and illicit purposes. Hackers might spoof the MAC address to bypass MAC filtering and gain unauthorized access to services. On the other hand, users might want to mask their identity to stop third parties from fingerprinting their devices (cf. 11.2.2). Spoofing can be detected in some situations, though. In closed environments such as in a company network, a set of addresses of authorized clients are registered. If such an authorized client uses a new, unregistered address, one can infer that (s)he spoofed the address [34].

MAC Address Randomization

One common way to enhance privacy is by randomizing MAC addresses. [34] were the first to suggest short-lived, disposable identifiers. They stated some challenges for creating such addresses:

1. The address should be unlinkable, *i.e.* the new address cannot be inferred from its former address so that an tracker cannot link two addresses to the same user.
2. The addresses have to be unique to avoid collisions.

Whether an address is randomized for privacy purposes can be determined to some extent by looking at the Universal/Local bit (cf. Section 11.2.1). If it is not set, then no randomization was performed, *i.e.* this is the global MAC address. If it is set, the address was either randomized for privacy reasons or used as part of another service. To distinguish these two uses cases, [43] had to apply elaborate techniques such as looking at IEs or specific OUIs. [43] found that 53% of addresses in their study were randomized.

Since multiple randomized addresses can be mapped to one device, one can assume that the adoption rate of randomization for devices is well below 50%.

All major operating systems by now support MAC address randomization [60; 43]:

- Apple supports randomization starting from iOS 8 [9]. Initially, it only performed randomization in unassociated and sleep mode and from iOS 9 on also in active mode [60]. In iOS 14, Apple also added randomization when connecting to networks as a default enabled setting [22]. Moreover, MAC addresses are rotated every day [22].
- Android uses randomization since version 6.0 for background scans provided that hardware and driver supports randomization [5]. Since version 9.0, it also randomizes MAC addresses in associated mode and in version 10 it became a default enabled setting [22]. However, unlike the newest iPhones, the MAC address stays static per SSID [22]. Generally, randomization is also handled differently from manufacturer to manufacturer. [43] reported that Samsung does not perform any randomization due to possible chipset incompatibilities.
- Windows added randomization support in Windows 10 [62] for probe requests as well as for authentication, association and data frames [43].
- Linux randomizes addresses since kernel version 3.18 for iwlwifi drivers and since kernel version 4.5 for brcmfmac drivers [33]. Like Windows, it performs randomization for probe requests, authentication, association and data frames [43].

The first randomization scheme was proposed by [34] who applied a forward chain of MD5 hashes starting with a random seed (i.e. the previous hash serves as the input for the next hashing). To create a valid address, they took the 3 least significant bits of the hash value as an index into the list of valid OUIs [39] and concatenated the next 24 bits to the OUI. To avoid collisions, they used reverse ARP requests with a double address switch: rather than including the current identifier into the source address (as this would reveal exactly the information that we want to conceal), they used a first random address with the actual request containing a second random address which is the address that we want to assign to the device but first need to check whether it is already occupied by another device. When the device receives a reply, it means that the address is already in use and therefore requests with other addresses have to be tried until one is found that is available. While this technique can still lead to collisions with the first random address, since this address is only used in a few transmissions, it should not cause major disruptions.

Generally, randomization schemes vary across devices. For example iOS devices randomize the complete MAC address including the OUI bits, while Android devices usually share a common prefix using either the OUI, the CID or some undefined prefix with the local bit set [43]. However, [43] also found instances where randomized global addresses were employed, namely in Motorola devices, even though this violates the rule that no two devices must share the same global MAC address.

11.3 Use Cases

People, companies and organizations have found different ways how device fingerprinting through the MAC address can be utilized. While some aim to improve their profits, others strive to improve the standard of living for many people. However, as with any technology, there are also some who wish to exploit and misuse it. In the following sections, some current use cases are discussed. The rating of either “Benefit” or “Issue” is completely subjective.

11.3.1 Benefits

The process of identifying devices based on their MAC address fingerprint brings with it certain advantages over similar alternative technologies. Examples of use cases include but are not limited to localization, gathering data about crowd behaviour and planning of public spaces.

In the following sections some practical use cases will be presented, all of which do see practical use in a real environment. The applications of this technology are constantly being improved and expanded upon, so this list is by no means exhaustive.

Better Localization and Navigation Options

In the past, maps and landmarks were the way travellers could find their desired destinations on their journeys. With the invention of satellites, GPS (Global Positioning System) localization has become the standard for navigation and marks a huge technological leap forward in the sector. In the past few years however, a new technology has emerged which enhances and sometime even completely replaces GPS localization. Wi-Fi based localization solves some of the shortcomings of GPS navigation, *e.g.*, need to be in the line of sight of a satellite.

[57] have shown that using a technique where stationary Wi-Fi access points are mapped to their geographical location, it is possible to accurately locate a Wi-Fi capable device such as a smartphone using their unique MAC address fingerprint [57]. This process works by monitoring the device's proximity to the locations of known access points in real-time and calculating its exact location based on this data.

Due to the today's ubiquity and thus density of Wi-Fi access points in public spaces, it is almost certain that a person carrying a mobile device with Wi-Fi capability such as a smartphone is in the proximity of at least one such access point at all times. As a result, it is possible to determine device location once an access point was geographically mapped. This is further facilitated by the fact that most access points tend to be located in areas with a high density of living spaces, as seen in Figure 11.5. These areas also correspond to where the demand for GPS navigation is the highest.

This new way of geographic localization is shown to be as accurate as traditional GPS localization [57], provided that there are enough known access points to ensure permanent tracking coverage. The device's location can furthermore be tracked over time with a very high accuracy and temporal resolution.

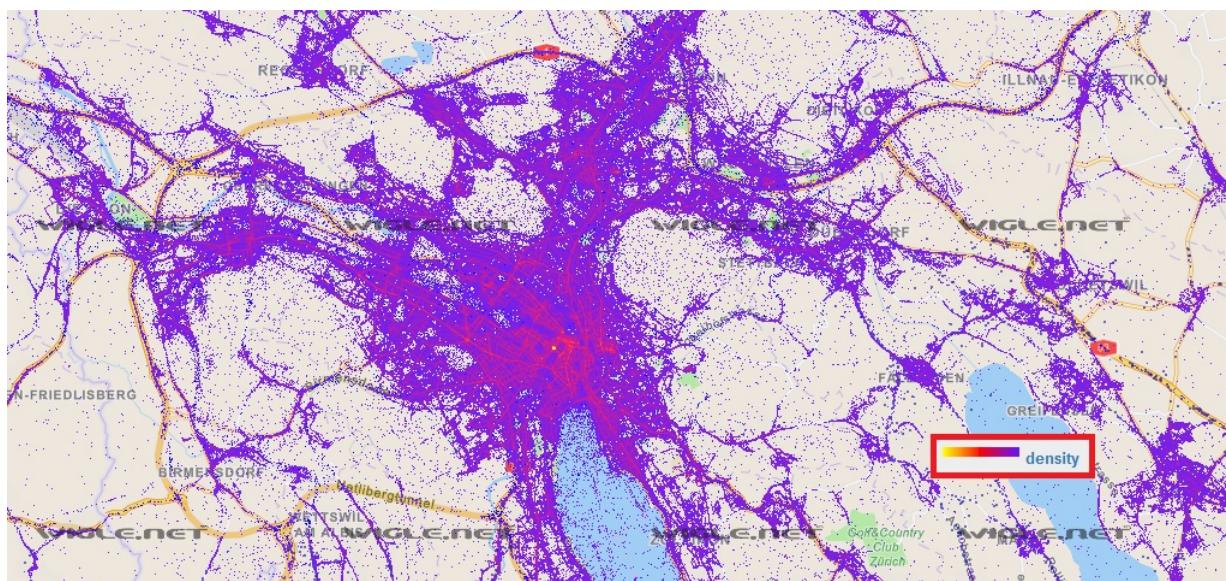


Figure 11.5: Access Points in Zuerich and Surrounding Regions [61]

A different approach to improved localization and navigation is a method called "Wardriving" [57]. This technique combines traditional GPS localization with Wi-Fi based localization to achieve a higher location accuracy. This method also alleviates the problem of indoor localization (*e.g.*, in tunnels or underground), provided that there are known access points in those areas.

Indoor Positioning System

Building further on the idea of using Wi-Fi instead of traditional GPS localization for indoor environments, there have been many use cases in the past years where the concept was tested and found to be a viable option.

One such case study was conducted in 2011 in Gangzhou (China) [65], where a Wi-Fi based indoor positioning system was used to locate resources on an underground construction site where GPS based localization was inadequate [65]. These resources included workers, machinery, building materials, and vehicles. The system was built using the Received Signal Strength Indication (RSSI) from each Access Point (AP) method.

As seen in Figure 11.6, all access points in range of a specific resource would monitor their corresponding signal strength to it. This data would then be sent to the main server where the different signals' strength would be compared using a radio map. Then, the estimated position of the resource can be determined.

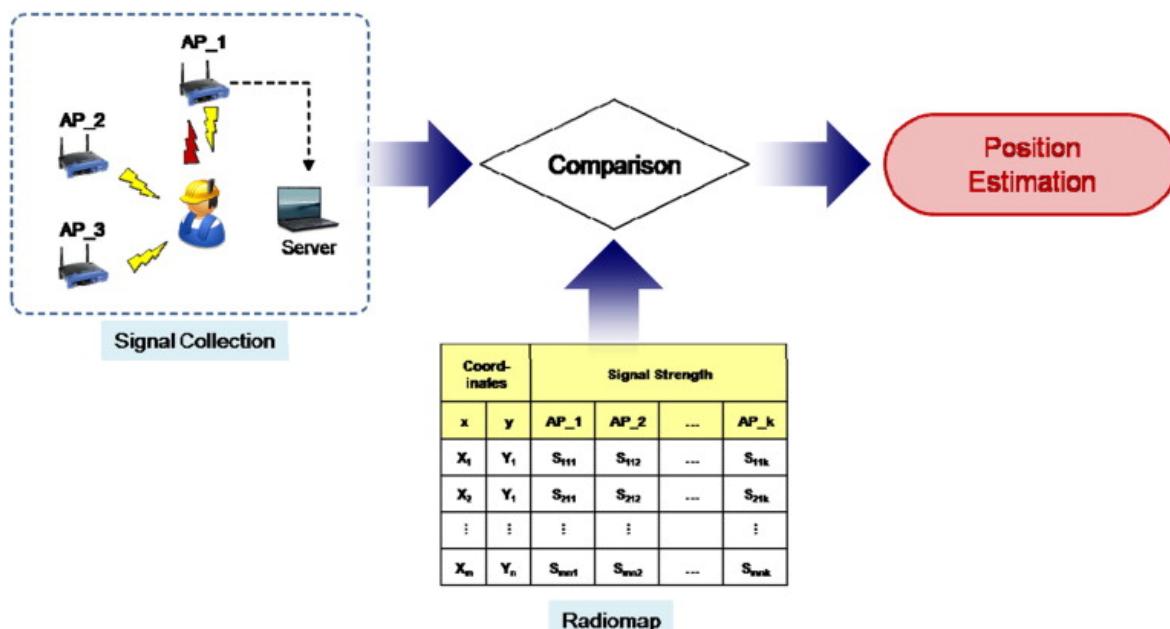


Figure 11.6: Wi-Fi based Localization System on the Gangzhou Tunnel Construction Site, China [65]

The positioning was found to be accurate within 5 meters of error and thus proved that Wi-Fi based localization is a viable and feasible complement to GPS based localization in such environments. Using the unique MAC address fingerprint of every device which is attached to a resource, it should be possible to create a timeline of locations of each one of them for later analysis. The findings could be used to improve the efficiency of future projects by planning the resource movement ahead of time.

A different use case of Wi-Fi based indoor localization could be in locations such as shopping centres, malls or adventure parks. These places often have a large amount of destinations people wish to visit, so a local positioning and navigation system might help in guiding them to where they desire.

Tracking the Shopping Behaviour of Customers

Targeted advertisement and data collection about people's shopping interests and patterns has always been a very relevant topic, with no signs of it losing importance any time soon. Traditionally, a good way of data collection for all major retailers (for example, Coop in Switzerland or Walmart in the US) have been customer loyalty programs [56]. Regular customers may request a card which they can use with every purchase to earn an arbitrary amount of bonus points which they can then use for small bonuses in the future. By doing this, customers enable retailers to connect their purchases and shopping habits to their identity.

Nowadays, a Wi-Fi based indoor positioning system can be used as a replacement or complement to the traditional bonus points program to help collect relevant data. As seen in Figure 11.7, using MAC address fingerprinting, it is possible to collect data about where people spend the most time, how they generally move in the building, where are the choke points and where visitors tend to concentrate. This data can then be used to make improvements to the infrastructure, direct people towards or away from specific key locations, reevaluate the value of individual shop locations, etc.



Figure 11.7: A Heat Map of People's Movement in a Building [54]

In fact, in as early as 2013 there were claimed to be already around 40 major MLA (Mobile Location Analysis) companies in the USA alone collecting customers' location data on behalf of various retailers [26]. This data includes information about how long customers wait at which cash register, how often customers visit a store on average, how many stores of the same chain they visit in a specific time frame, how many people visit without actually buying anything, which hot spots in the store draw customers to them, which promotions get the most attention, how customers move in the store and how many different destinations they visit in which order.

Tracking the Behavioural Patterns of the Masses

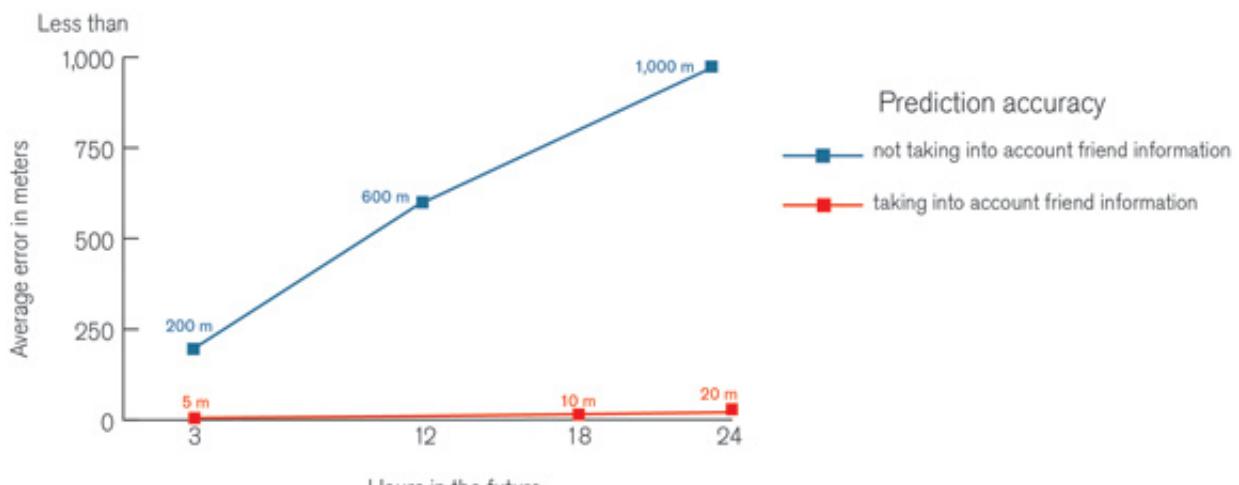
CCTV surveillance is commonplace nowadays, but nowhere is it as widespread as in mainland China. The CCP (Chinese Communist Party) under the administration of General Secretary Xi Jinping had an estimate of 200 million monitoring CCTV cameras in 2019 to keep track of their citizens, with that number growing rapidly [49]. However, traditional camera-based tracking methods have their drawbacks and there is a new emerging technology which could compensate for them: the Wi-Fi based tracking of masses.

It has been discovered that Wi-Fi tracking can be used to track pedestrian behaviour in cities with good effectiveness. Due to the ubiquity of modern smartphones and other Wi-Fi capable devices, Wi-Fi tracking has a much lower barrier of entry than other tracking techniques such as video cameras and surveys [53]. Furthermore, Wi-Fi tracking is a lot less intrusive and noticeable to pedestrians when compared to the other, traditional methods [1]. Wi-Fi based tracking is also more accurate than the former methods and has much higher coverage for the same cost. Using proximity-based MAC address tracking, spatio-temporal data can be gathered about human mass movement in public areas, which can then be used to improve city infrastructure. If needed, it is also possible to track an individual device thanks to the MAC address fingerprinting.

A study done in Switzerland by UK researchers sponsored by Nokia in 2012 concluded that given the prerequisite that smaller groups of devices which belong together (such as family or close friends) can be identified, it is possible to accurately predict the movement of individuals belonging to that group [20]. While the study was done using mobile data-based localization, it is possible to use the same method with Wi-Fi-based localization. Essentially, by knowing where people associated with you are, the system can predict your location with an average of a 20 meter accuracy for the following 24 hours [20]. As seen in Figure 11.8, this is orders of magnitude more accurate than inferring the location without the additional information of the associated devices whereabouts.

Divining Your Future Location

How using friends' mobility patterns improves prediction accuracy.



Source: Mirco Musolesi, University of Birmingham, based on an algorithm that analyzed 18 months of data from 200 smartphone users near Lausanne, Switzerland.

Figure 11.8: Localization Accuracy with and without Location Information of associated Devices [20]

However, this principle is not only limited to pedestrians. Using the same principle, it is possible to track vehicles and thus traffic patterns as well [12]. Historically, vehicle origin-

destination (O-D) data has been difficult and cost intensive to track. The process used cameras at many locations which would then send the video feed to a central computing facility to analyze the number plate of each every vehicle passing by them. This data would then be used to infer each vehicle's origin and destination, based on which cameras they have been spotted by. Using MAC address tracking, this process can be strongly simplified. Cameras can be replaced by Wi-Fi access points, which leads to being able to uniquely identify each vehicle extremely quickly. This new method is a very cost effective way of gathering vehicle O-D data in small and controlled networks. For more extensive networks, it can be used to complement the camera-based tracking instead, due to technical limitations.

A more niche, yet, still useful application of MAC address identification and device tracking can be found in the public transportation sector. Data can be collected about passenger behaviour, peak hours and the degree of utilization [36]. This helps public transportation companies analyze which routes they need to invest more into, how they can change their schedules to be more efficient and where they can afford to cut back.

Technology Aided Contact Tracing

The concept and practice of contact tracing has been around for the better part of 200 years now, with the earliest recorded usage being during the Black Death plague [19]. Infected individuals would go into quarantine and their houses would be marked for others to see and avoid. Since then, contact tracing and the platform which enables it has evolved massively. In the past few decades the technique's effectiveness has been enhanced by modern technology [19]. Nowadays, the preferred platform for contact tracing is the smartphone due to its ubiquity in modern society.

Currently, during the COVID-19 pandemic, it is also being utilized to help keep the spread of the virus under control. Many governments and organizations offer their own contact tracing smartphone applications (for example, SwissCovid in Switzerland). The technology these applications use is Bluetooth [16]. Smartphones' Bluetooth beacons broadcast a low energy signal to other nearby devices. Part of this signal is the device's universally unique identifier (UUID) [2], which is usually the device's MAC address. Every smartphone can then save all UUIDs it comes into contact with and keep the information for (in the case of COVID-19) a few weeks. If the device's holder were to be infected, they can have the application upload all the UUID it has collected and have those devices' owners notified. At the same time, the application also periodically checks the database if any of the UUIDs it has saved happen to belong to people who have reported themselves to be infected. Thus, MAC address fingerprinting is of good use when it comes to technology-aided contact tracing, as the critical mass of smartphone owners is met nowadays for it to be practically effective.

Recovering Stolen Devices

Electronic device theft is bound to become more and more commonplace with their increasing abundance. For this reason, some device and operating system manufacturers have come up with their own ways of locating stolen devices.

Apple for example has a web application called "Find My" which, as one of its main features, lets users geographically locate their Apple device [6]. Google has a similar application for locating Android devices, called "Google Find My Device" [29]. It can also geographically locate an appropriate device, provided that the owner has signed into a Google account on it at a previous point in time. Microsoft's version also offers this main functionality for its users [48].

However, there is another way of recovering stolen electronics: by using their MAC address fingerprint to determine their location [59]. Provided that the device's rightful owner keeps note of its MAC address, they can convey this information to the corresponding authorities in case the device is stolen or misplaced. Then, it can either be passively tracked by checking which access points it came across at which time, or actively sought out using a mobile access point.

11.3.2 Security Issues

Due to the completely unprotected nature of a device's MAC address, it is very easy for any potential attacker to get hold of it using a simple Wi-Fi access point. As long as the device's Wi-Fi receiver is turned on, the probing signals it emits in order to scan for available connections can be detected [18]. While this alone might not realistically pose a huge issue in isolated cases, serious problems occur if an attacker were to monitor individual MAC addresses over time or do so using multiple access points in various locations.

Privacy

It is straightforward to see how the unwanted tracking of individual MAC addresses in a spatial or temporal way may pose major privacy concerns. The easily identifiable MAC address fingerprint of any Wi-Fi capable device provides a convenient way of determining which APs it came across at which time. Any potential attacker with a dedicated infrastructure of access points or a back door into established ones may gather location data of individual or all devices they come into contact with, without the device's holder ever noticing or consenting to it. Currently, the only way of avoiding this issue other than using MAC address randomization is by turning off the device's Wi-Fi receiver every time one leaves a trusted area, which most people will not be willing to follow through with. When looking at current events, contact tracing applications (*e.g.*, SwissCovid in Switzerland) also pose a privacy concern. By using these applications, a person is directly uploading all data about their interactions directly to the application's owner's database. Even if the application's owner is trusted (*e.g.*, the government), there is no way of telling whether the database will be back-doored in the future and who's hands the data might end up in. The problem is further amplified when it comes to applications which also hold information which could identify individual users (*e.g.*, names or addresses).

Unwanted Targeted Tracking of Individuals

As it stands, it is under normal circumstances difficult but not impossible to track a device's MAC address back to its holder's identity remotely and without them ever noticing [17]. Provided that the individual in question is a high-profile person such as a politician or celebrity, it is reasonable to assume that there is enough incentive to link their identity to the MAC address of the device they carry with them at all times. This could lead to unwanted and dangerous situations for the affected persons.

As an example, it is possible to cause an event to automatically trigger as soon as a specific device and thus person enters the proximity of a specific access point. This could be exploited to trigger a bomb or alert malicious individuals to said person's arrival. This method is called "Wi-Fi booby trapping" [17]. Alternatively, by keeping track of a celebrity's location, it would be next to impossible for this individual to evade from unwanted journalists and paparazzis. Furthermore, a high profile individual might be more prone to defamation: if they were to be localized for a longer amount of time in a questionable place such as a brothel and this information were to go public, it would be bound to taint their image.

Moreover, it is not required to know a person of interest's MAC address in advance to track them. Provided that the access points keep logs and timestamps of devices they come across, it is possible to evaluate the data at a later point in time, when the person's device's MAC address is revealed. Organizations or entities with large amounts of such data thus, hold a lot of power as they can, if wanted, reverse-engineer essentially any person's movement patterns or specific locations at key points in time.

Criminal Activities

Unwanted parties gaining access to sensitive information such as location data of entire populations also poses a large public security risk. Knowing where people tend to gather in large masses at specific points in time can facilitate planning devastating terror attacks or public disruption. Conversely, by using the same data to determine when a specific location has the least amount of people around, crime organizations might be able to exploit this information and use it for their advantage in order to avoid detection or drawing attention to themselves.

On a smaller scale, MAC address fingerprinting might prove itself as a useful tool for burglars. By planting an access point near a potential target building, burglars would be able to tell how many people are inside of the building at any given time, how many of those were permanent residents and at what times which person tends to come and go. Using this information, they would be able to safely determine a time frame in which the building is empty and it is unlikely for a person to return unexpectedly. Furthermore, if a person is to return regardless, the burglars could set the system up in a way such that it automatically warns them if a specific MAC address shows up on their access point, giving them enough time to make a swift escape.

11.4 Evaluations and Discussions

11.4.1 Legal Aspects

As one would expect laws and rulings differ in different regions and countries. For the purpose of this report we look at the state of legal status of fingerprinting in three relevant regions: European Union, Switzerland and United States of America

European Union: General Data Protection Regulation (GDPR)

When talking about the legal aspects of data collection inevitably one of the most important laws to come in mind is the European GDPR. The GDPR is a privacy and data protection regulation passed in EU in 2016 and became active in 2018. Its primary goal is to grant individuals more control over their personal data. The idea of the GDPR is not to define strict rules of what can be considered as a personal data, but rather provide a legal framework by giving the main traits of personal data. This way, the law is applicable in an always changing digital environment and is not strictly bound to any technical definition or term describing personal data. According to the definition in the GDPR itself: “‘personal data’ means any information relating to an identified or identifiable natural person (‘data subject’); an identifiable natural person is one who can be identified, directly or indirectly” [23]. In a way it is not about establishing the person’s identity per se, but to be able to detect a single individual in a group. For example, detect if a person visiting a website is a unique new customer or a recurring visitor. In this notation digital

fingerprints like web browser cookies, IP addresses as well as MAC-addresses or any other comparable ways to identify a person are considered personal data¹.

Under the GDPR following simplified rules apply with regards to collecting and processing of personal data [24]:

- The data subject has given consent to the processing of his or her personal data for one or more specific purposes.
- Processing is necessary for various reasons:
 1. performance of a contract to which the data subject is party
 2. compliance with a legal obligation
 3. in order to protect the vital interests of the data subject or of another natural person
 4. performance of a task carried out in the public interest
- Processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data.

The first and last reasons are of particular interest for us in order to decide if the fingerprinting can be performed on legal grounds. First is relatively easy: in general if the user gives his/her consent for his/her data being collected it is enough to proceed with the collection. An example of this people living in Europe experience every time they visit a website, when they are prompted to agree on collection of web cookies, which on their own are considered fingerprinting. The situation with MAC address fingerprinting, however, is different, because collection can happen passively, without a direct interaction with a user, there is no possibility for a collector to ensure users consent to data collection. This can be mitigated if the consent is sent along with the MAC address in a probe request. However, that is not the case as for now, and thus direct collection of MAC addresses under the GDPR would be not possible with the consent reasoning.

The last legal reason to personal data collection under the GDPR, having a legitimate interest is a vogue term. According to Information Commissioner's Office [40], to determine if legitimate interest is present we should use a three part test:

- Purpose test - is there a legitimate interest behind the processing?
- Necessity test - is the processing necessary for that purpose?
- Balancing test - is the legitimate interest overridden by the individual's interests, rights or freedoms?

As an example, we can extend one used in the same source: An insurance company wants to process personal data to spot fraudulent claims on the basis of legitimate interests. Firstly, it considers the purpose test. It is in the company's legitimate business interests to ensure that their customers do not defraud it out of money. Also at the same time the non-fraudulent customers and the public in general have a legitimate interest in ensuring that fraud is prevented and detected. In order to enforce this interest a collection of personal data is necessary, this provides a necessity test. Important is not to collect more data than needed to proceed with the declared goal. To ensure the balance test, the

¹Important to emphasize is, that in the context of the GDPR it is not so important if the user's true identity is established, the deciding factor is, if the data collected can potentially be used to establish the identity.

insurance company has to make sure to not cross the boundaries of rights and freedoms of other individuals. This includes, but is not limited to, for example, data anonymization. The main idea of this point is to consider potential risks for the individual by the data collection and mitigate or refuse those from the data collection if the risks are unreasonably high and unbalanced in regard to the interest.

If we apply this knowledge to our MAC address consideration, taking for example a legitimate interest of tracking of the shopping behaviour of customers or behavioural patterns of the masses, we can conclude that legitimate interest of the public, business and users can be reasoned with improved services for the public and users as well as increased revenues for businesses. This will provide purpose test ground. Non-exhaustive and anonymized data collection should then fulfill necessity and balancing tests.

The GDPR also treats different entities differently by providing some exemptions [41]:

- “Personal or household activities - personal data processed in the course of a purely personal or household activity, with no connection to a professional or commercial activity, is outside the GDPR’s scope. This means that if you only use personal data for such things as writing to friends and family or taking pictures for your own enjoyment, you are not subject to the GDPR.”
- “Law enforcement - the processing of personal data by competent authorities for law enforcement purposes is outside the GDPR’s scope (for example, the Police investigating a crime).”
- “National security - personal data processed for the purposes of safeguarding national security or defence is outside the GDPR’s scope.”
- Extensive list of exemptions applied depending on the purpose of the data collection includes categories like: journalism, research and archiving, references and exams and health, social work, education and child abuse etc.

Switzerland: Bundesgesetz ueber den Datenschutz (DSG)

The laws controlling the data collection on Swiss soil are listed in the DSG, the document created in 1992. As many law documents it has been updated multiple times, and the last update occurred in 2019 and is expected to come into effect in 2021 which brings the DSG closer to the European GDPR.

For example an updated the DSG version states [14]:

- “Personal data: all information relating to a specific or identifiable person.”
- “Personal data may only be processed lawfully.”
- “Your processing must be carried out in good faith and must be proportionate.”
- “Personal data may only be processed for the purpose that was specified during the procurement, that is evident from the circumstances or that is provided for by law.”
- “The acquisition of personal data and in particular the purpose of their processing must be recognizable to the data subject.”

We expect the reader to find these examples familiar. Basically the goal of a Swiss legislator seems to be a formation of a data privacy setting comparable to the GDPR, which ensures data safety for individual in Switzerland and also ease international companies

in operating in different countries, by reducing the need for adaptation of processes for differing regulations.²

United States of America: California Consumer Privacy Act (CCPA)

USA in comparison to Europe have a different legal environment. It is important to acknowledge, that while Europe uses civil law, USA legislation is based on common or case law. This makes the comparison of two systems fairly complicated. Additional difference lies upon absence of federal-level data privacy protection law in USA. If data privacy is regulated in some places of United States, then it is controlled by single state laws. Analysing all of them would be out of the scope of this work. This is the reason why we will concentrate on the single most relevant one: CCPA.

The CCPA is a new law introduced by California State Legislature in 2018 and became effective in 2020. The CCPA similarly to the DSG is using the GDPR as a reference and trying to give users more control over their private data, but has differences. For example, contrary to the GDPR individuals who are subject to the CCPA should actively opt out of data collection in order to stop sharing their privacy data [15].

To give the reader a sense of how the CCPA is similar to the GDPR and the DSG we provide a personal information definition along with some granted rights from the CCPA [15]:

- “Personal information is information that can be used to reasonably identify, relate to, describe, be capable of being associated with, or be linked, directly or indirectly, to a particular consumer.”
- “The right to know about the personal information a business collects about them and how it is used and shared.”
- “The right to delete personal information collected from them (with some exceptions).”
- “The right to opt-out of the sale of their personal information.”
- ”The right to non-discrimination for exercising their CCPA rights.”

Implication of this law to the rest of the United States can be compared with the influence of the GDPR on the rest of the world: If businesses have to comply with some ruling in one geographical region of their business activity, there is already a cost of adjustment present. So, to save money in absence of contradicting regulation in other regions, businesses with high probability will decline on running two parallel systems and just concentrate on one data privacy approach for all their regions of operations. This dynamic was observed with the implementation of the GDPR in Microsoft [13].

Nevertheless, in the absence of federal law, news titles like these are not a rare occasion in the USA: “US: End Bulk Data Collection Program”[37], “Carriers Swore They’d Stop Selling Location Data. Will They Ever?”[7], “A Location-Sharing Disaster Shows How Exposed You Really Are”[50]. For this reason, we can not consider the USA as a privacy preserving space, where fingerprinting is strictly regulated, at least for now. Apparently operating in grey zone is worth the legal risk and brings good money for the businesses.

²We would like to remind the reader, that officially the updated definitions will come in force in 2021. For the purpose of this report however we are interested in future developments of different aspects of fingerprinting. Thus treat future legislation as the relevant reference point.

11.4.2 Economic Aspects

Money is inevitably the main driver of data collection. Big tech giants depend on their user's data a lot. Facebook made 70.7 billion US dollars in 2019, 98.5 percent of which comes from advertisements [35]. Google has a similar earnings structure, 70.2 percent of Alphabet's 171 billion dollars in expected revenues for 2019 were made in ads [25].

Why data has value

To be able to serve the ads effectively a platform has to know their customers. This means knowing people's age, sex, geographical location etc. It helps sending more relevant ads their way, which should result in more sales for the advertised product. Increased sales in turn make ads more valuable. Consequently since the engagement of an interested person brings more sales, the advertisement platform that can distinguish an interested person for a particular product is more valued. The cornerstone of the whole process is personal data.

However, not only personal data can be used for increased revenue. There are multiple strategies on capitalising on different types of data, for example, a distributor of some product can up- and down-sell data to their product buyers and providers. Let's look at a simplified example: fruit distributor having data on how many apples are ordered by different trading spots can down-sell this data to competing trading merchants. They are then able to analyse it in order to evaluate the effectiveness of their marketing and sell strategies against the ones of their competing trading merchants. Also, an apple distributor can up-sell the data to apple producers, for example, giving them info if green apples are selling better than red ones in a particular time or location, thus giving the producers a vital information on how to adjust their production strategies in order to increase their profits.

If we project these ideas on MAC addresses it becomes obvious, that the data which can be potentially collected by multiple beacons across the city, can be analysed and sold to different parties with different intentions: businesses, public service providers, researchers etc.

Data brokers

Exactly this is the business model of data brokers: entities whose business model is to collect, enrich, and resell the data. Data brokering is a multi-billion dollar industry and consists of over 4,000 data brokering companies [63]. In general we can distinguish three types of data brokers [31]:

- People search sites - by providing usually a full name of a person, users can find out some information on this individual for some payment. Often the accessible data includes: family members, physical addresses, social network profiles, email addresses and phone numbers.
- Marketing data brokers - collect data on individuals usually enriching it, by forming groups of people. For example, dog owners, young families or people with a particular hobby. This list can then be sold and resold to advertisers, who in turn use them to target their ads.
- Risk analytic brokers - they collect the data and enrich it by providing a risk index. The components of the index can include information on previous fraudulent activities associated with for example, particular email address or if the Social Security number is associated with a deceased person. Usually, these indexes are of particular interest for insurance companies and banks, who are eager to detect and prevent fraud.

The sources of data, which end up in data brokers' hands can be pretty wide. Usually, it is collected from public sources, like government resources who publish data on particular people or entities in connection to law suits, property ownership, marriages and social network profile data. Data also can include purchase histories, geographical locations and interests of people, by buying it from third party apps, social network quizzes and different service providers. Finally, data brokers, exchange and purchase information from each other.

Big Tech

We mentioned Facebook and Google as big data players whose entire business model is based on data above. However, an important difference between data brokers and big tech is that, tech companies are not interested in selling the data directly, but rather sell an opportunity to reach relevant audience for marketing purposes through these companies services and data. There are two main reasons for big companies for not willing to sell the data directly.

First of all, by doing so they lose their information advantage. By selling the data they collected on people, they essentially give away their unique opportunity to sell advertisement services to advertisers and, thus, enable competition to use their data to provide their services. The second reason is more important. Big tech relies on its users, and by selling their data directly, companies most likely will lose the trust of their users and potentially lose the user base. This is why, Google is especially interested in providing as much security and privacy from third parties to their customers as they can, while collecting the data themselves [51].

One of the instruments these companies use to achieve user privacy, while collecting the data is differential privacy. “Differential privacy is the statistical science of trying to learn as much as possible about a group while learning as little as possible about any individual in it.” [32]. The incentive to use it as described in same source is improved privacy for users, due to imperfections of traditional anonymization approaches. Traditional anonymization can be reversed by cross referencing additional non anonymized databases. To avoid the identity of users being assigned to the data differential privacy concept makes use of three transformations: “Hashing, a cryptographic function that irreversibly turns data into a unique string of random-looking characters; subsampling, or taking only a portion of the data; and noise injection, adding random data that obscures the real, sensitive personal information.” [32].

Price of data and black market

As multiple reports confirm, government agencies and law enforcement of different countries have legal basis to access citizen data, often getting this data from private service providers, like cell phone carriers. Often business charges government for data provision. “AT&T also said that it charges 100 USD to start tracking a phone and 25 USD a day to keep tracking it” [46].

Depending on the type and difficulty of obtaining of information it can be priced differently: “Basic information, such as old addresses, current addresses, and family connections can be bought for as little as 95 cents.”, “The Social Security number of someone with good credit, for instance, can sell for between 60 USD and 80 USD.” [21].

Of course if there is a possibility to make a profit, there will also be people who will try to use it illegally. Stolen identities have their prices too: “Basic stolen identity information on a US citizen, which only includes the Social Security number, full name and birth date, can range from 1 USD to 8 USD per person. But in some cases, hackers will package the

offering with the victim's stolen credit card information, and charge from 20 USD to 75 USD.” [45].

Applying all of the discussed in this section to MAC addresses, we can see the value of the data, which can be potentially extracted from MAC addresses, especially the movement data. However, large companies do not need the MAC address to establish the geolocation, since they have more efficient means to get it. And since the new laws may make these approaches not usable, due to potential user identity identification, differential privacy mitigates this possibility and provides big tech, as well as other data collectors with a way to proceed with their business without potentially harming users' privacy.

Smaller businesses may be interested in getting this data, however most likely will lack the means to collect it, since it requires extensive infrastructure. The only real possibility for this technology to be used in business seems to be a collection of data through a specialised service provider with beacons, and selling it at competitive prices to smaller businesses. If such a business model is sustainable is a bigger point of discussion. Hackers may be interested in MAC address data collection too. However, since the method needs extensive infrastructure, most likely they will keep on traditional ways of stealing the data, by hacking into databases and not collecting data themselves. Law enforcement is then the only one left entity who may be interested in the technology. In an absence of data linked to specific users by big tech, because of the deployment of differential privacy, government structures are the ones who have resources to build the needed infrastructure and interest to identify specific people.

11.4.3 Ethical Aspects

Social aspects of data collection have interesting points that shall be mentioned for the fullness of the picture for the reader. Usually speaking about ethical aspects of data collection, a debate is happening around the questions of who is allowed to collect the data, how much data should be collected and how it should be used. And even if legislators of multiple countries are now actively working to answer these questions and codify the rules, general public is concerned with safety of their private data, because of the risks that occur if this data is used with evil intentions: identity stealing, doxing and opinion manipulation are just some of the risks the public is usually concerned about.

Privacy paradox

Interesting observation can be done, however, in relation to this concern about data privacy, if one explores what an average person actively does to protect himself/herself. Turns out “While users claim to be very concerned about their privacy, they nevertheless undertake very little to protect their personal data.” [8]. This phenomenon is known as the privacy paradox.

Researchers observed in multiple studies, that people when faced with a choice of giving away their personal data in exchange for a service or not, often decide to disclose their data, even if they previously claimed their privacy to be of great importance. This behaviour is often linked to the lack of transparency and understanding from the user side, of what happens to the data and how it is being used. As [8] put it “privacy is not yet integrated into the social presentation of a smartphone and hence, will consequently lead to failed privacy awareness.”. The same authors name multiple aspects that could play a role in a formation of the privacy paradox ³:

- Under- and/or overestimation of risks and benefits
- (Immediate) gratifications

³important to mention, that most of the research is based on users behaviours on mobile platforms

- Difference between the judgments of risks and benefits
- Little to no risk assessment
- Value of desired goal outweighs risk assessment
- Privacy valuation failed
- Knowledge deficiency due to incomplete information

To a large extent these reasons can be mitigated with appropriate user interface design decisions and increased awareness campaign of data privacy topics, which would have its goal in “shifting the reference point from “not mine” to “mine” goes along with higher risk perception which leads to the development of psychological ownership. This might elicit a higher valuation of private information, resulting in risk-averse decision-making.” [8]. Regarding the system design, some of the possible solutions would be:

- Increased transparency of what data is shared and for what purposes: for example this could include user friendly privacy policies, which are easy to read. Similar to the “simple language” options in some government websites. Not rare are cases when the service itself does not collect the data on users and advertises it this way. However, in the terms and conditions it explains sharing users’ data to third party, which has no obligation on what data it collects and how it uses it.
- Ability to use services even if sharing of information is declined. This could potentially restrict some functionality of services, but the user should have a choice. At the moment a lot of services provide an all or nothing choice, which puts users in a great pressure to accept the data policy. One could argue that, in this case users should not use the service at all, but in a modern world a lot of social apps and services are almost a must in order to maintain social interaction and it should be considered a normal behaviour of business to give the control to user.
- Regular permissions check. This is a solution for the data sharing permissions that were given by users a long time ago, and he or she would like to remove it, but is not longer aware of its existence. A possible solution would be expiration dates introduced to consents for data sharing.
- Privacy disclaimers reminding people of what data will be shared as a result of a particular action [8]. For example, in case of liking a page on Facebook, it would be beneficial to remind the user from time to time, that this information will be used to target personalized ads.
- Use user interface design solutions to avoid habitual consenting to data policy, the simple example of that is to randomize “yes” and “no” button position.

Police profiling and Overton window

Another social concern is an extensive police profiling. Since the terrorist attacks in 2001, USA adopted the Patriot Act, which basically extended surveillance right of national security agencies. This was seen as a necessary measure to combat terrorism. However, as we have seen in the dynamics of new legislations like the GDPR, the end user is being protected from business interests, but law enforcement units are often exempted from these regulations.

This brings us to a dilemma greatly formulated by [47]: “On the one hand we have the necessity to advance technology and to use it in situations in which it is advantageous to

the whole community, on the other hand this same technology can impinge on the rights of individuals (if we let it), through sweeping changes to legislation.”

This phenomenon is called the Overton window. It basically describes the range of policies politically acceptable to the mainstream population at a given time. The theory [66] goes as follows: at each given time there are acceptable policies. Movement of social norms to extremes can take two directions: less freedom and more freedom. Depending on the amount of change it will be perceived by the population as a popular, sensible, acceptable, radical or unthinkable measure.

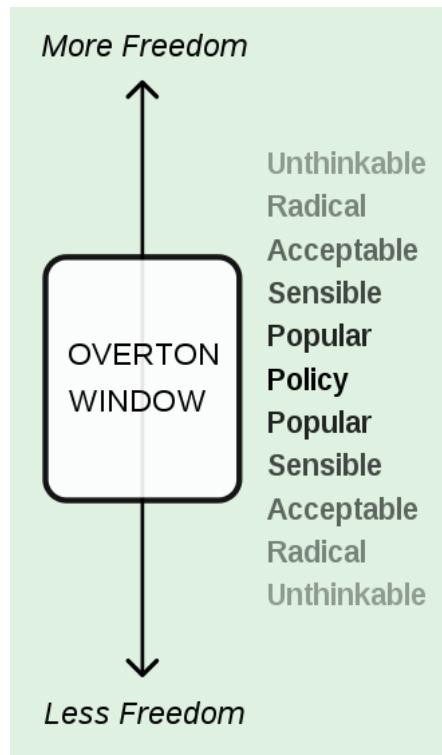


Figure 11.9: Overton window policy range [38]

The shift of public opinion can be done, by gradually shifting policies in to the areas of popular, sensible or acceptable range. After a time the new policy will become a new norm, and the Overton window will shift. These policies that were perceived unthinkable before, will now be perceived as acceptable.

This example is brought to show the importance of maintaining the balance. Ethics should be a never ending debate, because it is impossible to stress how important it is to reevaluate if the systems are doing what public wants them to do all the time. Also, the Overton window should be taken into consideration. Right now it might be acceptable to fingerprint criminals, but not the general public. However, moving from criminal to people who willingly for example, for research purposes, altruistically or for money rewards are allowing to track them is not such a big step. After it becomes a new norm, tracking anonymous data of all people becomes a norm, and then just tracking people with or without consent might become a usual happening.

11.4.4 Technical Aspects

Finally, let's discuss technical possibilities in regard to MAC addresses and privacy. As mentioned above, we could theoretically avoid sharing user devices MAC address by using passive scanning with signal emitted from AP, as discussed above this approach has its downsides. Another way to enhance users privacy in regard to MAC address collection is its randomization. However, there are a few downsides to this approach too.

11.4.4.1 MAC address randomization

Firstly, as [10] notes a simple default enable/disable policy for randomization is not sufficient because in some situations address randomization is desirable and in other cases it is not. For example, while randomization is desirable for increasing privacy, MAC addresses might also be used for restricting access to certain services and therefore randomization can there not be done without careful management. There are also use cases, where the user actually wants be tracked for example with a medical device. The authors propose a configuration for randomization based on certain context variables such as type of network (visible/internal) or location.

Secondly, and perhaps even more important, MAC address randomization does not prevent user from being identified. MAC address randomization can contribute to more privacy, but does not totally prevent device fingerprinting because it only prevents trackers from using the source field of the frame header as an identifier. [18] outlines several attacks identified in the literature that defeat MAC address randomization. They classify them into content-based and timing attacks:

Content-based Attacks Other fields in the frame header as well as in the payload also leak information that allow for fingerprinting the device.

- [34] noted that the sequence numbers need to be reset as well as these do not change with an address change and therefore make consecutive MAC addresses linkable.
- [11] found a way to defeat randomization at the physical layer: For performance reasons, the Orthogonal Frequency Division Multiplexing (OFDM) uses a scrambler which XORs the input data with a pseudo random sequence. They found that the seed used by OFDM follows certain patterns which allow to link MAC addresses. Since this issue occurs at the physical layer, it is harder to overcome because it requires modification on the hardware level.
- [60] also identified Information Elements (IEs) as carrying fingerprinting information. IEs are variable-length fields in the frame body consisting of a numerical label (type), the size of the data and the value. Each message can carry multiple IEs and each IE gets interpreted differently based on its type. They mainly serve to communicate the capabilities and features of a device such as supported rates or frequency hopping parameter set. For example, [27] lists over 20 IEs and the combination of their values usually allows to link the MAC addresses of a randomizing device. [60] name one particular IE, Wi-Fi protected setup (WPS), that is susceptible to a re-identification attack because the Universally Unique Identifier (UUID) field within it is derived from the MAC address.

Timing-based Attacks Sending patterns represent a second way of revealing identifying information about a device. In active service discovery mode, devices send probe requests in specific time intervals. For example in Wi-Fi channels, these probe requests are sent in bursts during which a device transmits frames in a specific time window of less than 500ms. This is also shown in Figure 11.10. There are two different ways to leverage timing information:

- Measuring the time between bursts (time between group of blue boxes)
- Measuring the time between probe requests in a burst (time between blue boxes)

By creating a temporal fingerprint of device, a device can still be isolated despite masking the identity. [42] exploits timestamps contained in TCP headers to estimate clock skews of a device.

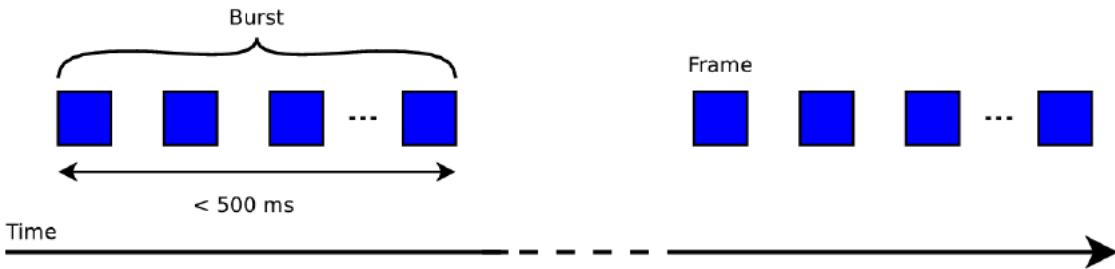


Figure 11.10: Defeating MAC address randomization by leveraging timing information of probe requests. [18]

Those various attacks show that MAC address randomization is just one, but crucial step towards enhancing privacy. In combination with passive scanning it could improve users privacy, but in general these two examples just show the impossibility to solve this problem once and for all. Similarly to ethical aspects, technical ones are a continuous arms race between anonymization and de-anonymization technologies.

11.5 Summary and Conclusions

As we saw in this report, MAC address usage and the problems associated with it have multiple facets. For one, there is the technical aspect. Devices with Wi-Fi and/or Bluetooth capability constantly send out probing signals to check for access points within their proximity, as long as the corresponding beacon is switched on [42]. By doing this, they expose their uniquely identifiable MAC address to these access points, which can then save logs of these interactions. If an organization has access to a large network of such APs within a specific area, they can track the locations of all devices which these APs detect through a method called RSSI (Received Signal Strength Indication) [65].

Through describing different kinds of use cases for MAC address based device localization and tracking, we have shown that this technology is a powerful tool, which needs to be utilized responsibly. On one hand, it can help make our lives more convenient, aid us in staying healthy during large scale virus outbreaks and provide valuable information for the optimization of public spaces. On the other hand, due to the plainly exposed nature of the MAC address, any potentially unwanted party with a network of access points they administrate can easily gain access to a device's location data within that network, which poses large privacy concerns. This is especially the case because it happens entirely without the tracked device's owner ever noticing. That being said, as it stands, it is difficult to trace a device's MAC address back to its owner's identity [17] (two ways of attempting it are the so-called Beacon Replay Attack and the Stalker Attack [17]).

It is possible to try and avoid this tracking by using specific countermeasures. One way to protect one's self against tracking is MAC address Randomization [34]. This method uses software to send out randomized, temporary MAC addresses instead of the device's hardware MAC address in probing signals, which makes it difficult if not impossible to pin it down across multiple access points.

However, there are a variety of attacks to circumvent MAC address randomization. Attacks are classified into the content-based and timing-based categories [18]. Content-based attacks exploit the fact that some other fields in the probe request header as well as the

payload also contain uniquely identifiable information. Timing-based attacks use specific patterns in the interval between individual probe requests to fingerprint devices.

Recently it has been discovered that there might be a way of completely avoiding our devices beaming out their MAC address at all times altogether. Through passive scanning, devices would never send out probe requests to begin with [30]. That being said, this technology is not widespread and device manufacturers still tend to opt for the traditional, active way of scanning.

All of these examples show the seemingly unending “arms race” between the techniques which improve the accuracy of device fingerprinting and countermeasures implemented to avoid such identification. As of now, it seems this will be a long race if the fundamental rules will not change. For example, legally pushing manufacturers to using passive scanning as the default option will most likely solve some problems, although also introducing new incentives to research ways to beat passive scanning security benefits.

The legal aspects of MAC address based fingerprinting are not any easier to navigate than the technical ones, as there are heterogeneous answers to the privacy question, depending on the regions we look at. We investigated three main legislation regions in this report: the European Union, Switzerland and the United States of America. In the EU, the European GDPR (General Data Protection Regulation) handles the legal aspects of MAC address tracking and personal information collection [40][41]. The Swiss equivalent to the GDPR is the DSG (Bundesgesetz über den Datenschutz), with the legalities being very similar between the two sets of laws [14]. The USA’s legal environment is fairly different to the aforementioned two. Here, there are no federal level privacy laws. Instead, the USA opts for single-state laws. So, we showed the CCPA (California Consumer Privacy Act) as an example [15].

As we saw in all of these regions, legislators are being pushed more and more as societies are increasingly trying to regulate and secure the data collection space in order to protect the end user. This, however, is just the beginning of the work in the few regions. Most places have little to no data protection laws. Realistically, we should expect it to take a fair while longer until it is finished. And even then, experience from other legislation suggests that there will always be legal loopholes and people willing to exploit them. Thus, while the legal framework is extremely important, it is not a final solution to all of the fingerprinting dangers.

Looking at economical aspects, we spotted an obvious truth about our present: good data has a very high economical value associated with it. Large tech companies such as Facebook and Google mainly use it for the purposes of pushing targeted advertisement [35][25]. Moreover, customer data can also be used to up-sell products to customers who are known to be inclined to pay more, or down-sell them to others who might just find the price to be slightly too high. Due to data having such a high value, a whole new multi-billion dollar industry of data collecting companies has emerged; so-called *data brokers* [63]. These companies specify solely in collecting, cleansing and selling data. However, not all data is collected legally. There exists a black market for buying and selling illegally collected data such as social security numbers, personal information or credit card details [21][45]. The low barrier of entry of MAC address tracking is only bound to facilitate such practices. The main question which needs to be addressed here, is whether there is a possibility to destroy the value of data, so that the incentive to collect it in legal or illegal manner vanishes. Our personal belief is that as long as people are willing to keep their privacy, there will be other individuals willing to break through their protection and obtain their private data. The monetary incentive this creates is the main driver for all of the problems we spoke about in this paper.

So, if the root of the problem lies in peoples’ desire to retain their data privacy, the obvious idea would be to look into the origins and structure of this desire. This is exactly what we do by looking into the ethical aspects of data collection. While most people tend to

state that they care about their privacy and personal data, this sentiment is not at all reflected by their actions in most cases [8]. In reality, this leads to most people using services which collect their personal data without a second thought or blindly agreeing to unread terms and conditions, tempted by the prospect of a “free” to use product.

This all leads to an interesting point of the discussion. Should we, as a society, push and strive for a more secure market, where a user should be warned about their own actions and should be educated more about the implications of their decisions, or should we go the way of opening up to the world as individuals, by sharing our data, thus destroying market value of this data? As we saw in recent years, people share their information on social networks willingly, so this is a viable option. However, to which extend this openness should be realised is also an important question that needs to be considered with all the responsibility. We do not attempt to answer this question due to its complexity and the boundaries of this work. Also, we do not see ourselves as entitled to do so. Obviously this work is not intended to decide on society’s desired development vector but its goal is rather to identify outlying possibilities and spark a discussion, which over time, can lead to more informed decisions on an individual and societal level about how we want to approach our private and collective data protection and privacy.

Bibliography

- [1] Naeim Abedi, Ashish Bhaskar, Edward Chung: *Tracking spatio-temporal movement of human in terms of space utilization using Media-Access-Control address data*; Applied Geography, Elsevier, Vol. 51, 2014, pp. 72-81.
- [2] Dave Addey: *iBeacons*; <https://web.archive.org/web/20131203014352/http://daveaddey.com/?p=1252>, September, 2013.
- [3] Aircrack-ng: *Official Website*; <http://aircrack-ng.org/>, November, 2020.
- [4] Aislelabs: *Official Website*; <https://www.aislelabs.com/>, November, 2020.
- [5] Android developers: *Android 6.0 changes*; <https://developer.android.com/about/versions/marshmallow/android-6.0-changes.html>, April, 2020.
- [6] Apple: *If your iPhone, iPad, or iPod touch is lost or stolen*; <https://support.apple.com/en-us/HT201472>, March, 2020.
- [7] Brian Barrett: *A Location-Sharing Disaster Shows How Exposed You Really Are*; <https://www.wired.com/story/locationsmart-securus-location-data-privacy/>, May, 2018.
- [8] Susanne Barth, Menno D.T. de Jong: *The privacy paradox - Investigating discrepancies between expressed privacy concerns and actual online behavior - A systematic literature review*; Telematics and informatics, Elsevier, Vol. 34, No. 7, 2017, pp. 1038-1058.
- [9] Mike Beasley: *More details on how iOS 8's MAC address randomization feature works (and when it doesn't)*; <https://9to5mac.com/2014/09/26/more-details-on-how-ios-8s-mac-address-randomization-feature-works-and-when-it-doesnt/>, September, 2014.
- [10] Carlos J. Bernardos, Juan Carlos Zuniga, Piers O'Hanlon: *Wi-Fi internet connectivity and privacy: hiding your tracks on the wireless internet*; 2015 IEEE Conference on Standards for Communications and Networking (CSCN), Tokyo, 2015, pp. 193-198.
- [11] Bastian Bloessl, Christoph Sommer, Falko Dressler, and David Eckhoff: *The scrambler attack: A robust physical layer attack on location privacy in vehicular networks*; 2015 International Conference on Computing, Networking and Communications (ICNC), 2015, pp. 395-400.
- [12] Miranda Blogg, Conor Semler, Manu Hingorani, Rod Troutbeck: *Travel Time and Origin-Destination Data Collection using Bluetooth MAC Address Readers*; Australasian transport research forum, Vol. 36, 2010.
- [13] Julie Brill: *Microsoft's commitment to GDPR, privacy and putting customers in control of their own data*; <https://blogs.microsoft.com/on-the-issues/2018/>

- 05/21/microsofts-commitment-to-gdpr-privacy-and-putting-customers-in-control-of-their-own-data/, May, 2018.
- [14] Bundesversammlung der Schweizerischen Eidgenossenschaft: *Bundesgesetz ueber den Datenschutz*; <https://www.admin.ch/opc/de/classified-compilation/19920153/index.html#fn-#a4-1>, June, 1988.
 - [15] *California Consumer Privacy Act*; https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180AB375, June, 2018.
 - [16] Hyunghoon Cho, Daphne Ippolito, Yun William Yu: *Contact Tracing Mobile Apps for COVID-19: Privacy Considerations and Related Trade-offs*; arXiv preprint arXiv:2003.11511, 2020.
 - [17] Mathieu Cunche: *I know your MAC address: targeted tracking of individual using Wi-Fi*; Journal of Computer Virology and Hacking Techniques, Springer, Vol. 10, No. 4, 2014, pp. 219-227.
 - [18] Mathieu Cunche, Célestin Matte: *On Wi-Fi Tracking and the Pitfalls of MAC Address Randomization*; Nouveaux défis de l'Internet des Objets: Interaction Homme-Machine et Facteurs Humains, 2016, pp. 18-18.
 - [19] Daily History: *What is the History of Contact Tracing?*; https://dailyhistory.org/What_is_the_History_of_Contact_Tracing%3F, November, 2020.
 - [20] Manlio De Domenico, Antonio Lima, Mirco Musolesi: *Interdependence and Predictability of Human Mobility and Social Interactions*; Pervasive and Mobile Computing, Elsevier, Vol. 9, No. 6, 2013, pp. 798-807.
 - [21] Max Eddy: *How Companies Turn Your Data Into Money*; <https://uk.pcmag.com/privacy/117876/how-companies-turn-your-data-into-money>, October, 2018.
 - [22] Eleven Software: *How MAC Address Randomization Can Affect the Wi-Fi Experience*; <https://blog.elevensoftware.com/how-mac-address-randomization-can-affect-the-WiFi-experience>, July, 2020.
 - [23] European Union: *General Data Protection Regulation, Article 4*; <https://gdpr-info.eu/art-4-gdpr/>, May, 2016.
 - [24] European Union: *General Data Protection Regulation, Article 6*; <https://gdpr-info.eu/art-6-gdpr/>, May, 2016.
 - [25] Forbes: *Is Google Advertising Revenue 70%, 80%, Or 90% Of Alphabet's Total Revenue?*; <https://www.forbes.com/sites/greatspeculations/2019/12/24/is-google-advertising-revenue-70-80-or-90-of-alphabets-total-revenue/?sh=6ef5ba944a01>, December, 2019.
 - [26] Brian Fung: *How stores use your phone's WiFi to track your shopping habits*; The Washington Post, Vol. 19, 2013 <https://www.washingtonpost.com/news/the-switch/wp/2013/10/19/how-stores-use-your-phones-wifi-to-track-your-shopping-habits/?arc404=true>.
 - [27] Matthew S. Gast: *802.11 Wireless Networks: The Definitive Guide*; O'Reilly Media, Inc., Chapter 4, 2005.

- [28] Philippe Golle, Kurt Partridge: *On the anonymity of home/work location pairs*; International Conference on Pervasive Computing, Berlin, Heidelberg, 2009, pp. 390–397.
- [29] Google: *Find My Device*; <https://www.google.com/android/find>, November, 2020.
- [30] Frederik Goovaerts, Gunes Acar, Rafael Galvez, Frank Piessens, Mathy Vanhoef: *Improving Privacy Through Fast Passive Wi-Fi Scanning*; Nordic Conference on Secure IT Systems, Springer, 2019, pp. 37-52.
- [31] Yael Grauer: *What Are ‘Data Brokers’ and Why Are They Scooping Up Information About You?*; <https://www.vice.com/en/article/bjpx3w/what-are-data-brokers-and-how-to-stop-my-private-data-collection>, March, 2018.
- [32] Andy Greenberg: *Apple’s ‘Differential Privacy’ Is About Collecting Your Data - But Not Your Data*; <https://www.wired.com/2016/06/apples-differential-privacy-collecting-data/>, June, 2016.
- [33] Emmanuel Grumbach: *iwlwifi: mvm: support random MAC address for scanning*; Linux commit effd05ac479b.
- [34] Marco Gruteser, Dirk Grunwald: *Enhancing location privacy in wireless LAN through disposable interface identifiers: a quantitative analysis*; Mobile Networks and Applications, Vol. 10, No. 3, 2005, pp. 315-325.
- [35] Kris Gunnars: *Facebook revenue breakdown*; <https://stockanalysis.com/how-facebook-makes-money/>, Feburary, 2020.
- [36] Arief Hidayat, Shintaro Terabe, Hideki Yaginuma: *Determine Non-Passenger Data from WiFi Scanner Data (MAC Address), A Case Study: Romango Bus, Obuse, Nagano Prefecture, Japan*; International Review for Spatial Planning and Sustainable Development, Vol. 6, No. 3, 2018, pp. 154-167.
- [37] Human Rights Watch: *US Bulk data collection*; <https://www.hrw.org/news/2020/03/05/us-end-bulk-data-collection-program>, March, 2020.
- [38] Hydrargyrum, Wikipedia: *Overton Window Illustration*; https://en.wikipedia.org/wiki/Overton_window, January, 2015.
- [39] IEEE: *OUI assignments*; <http://standards-oui.ieee.org/oui.txt>, November, 2020.
- [40] Information Commissioner’s Office: *GDPR: What is the legitimate interests’ basis?*; <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/legitimate-interests/what-is-the-legitimate-interests-basis/>, November, 2020.
- [41] Information Commissioner’s Office: *GDPR: Exemptions*; <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/exemptions/?q=dpa>, November, 2020.
- [42] Tadayoshi Kohno, Andre Broido, Kimberly C. Claffy: *Remote physical device fingerprinting*; IEEE Transactions on Dependable and Secure Computing, Vol. 2, No. 2, 2005, pp. 93-108.
- [43] Jeremy Martin, Travis Mayberry, Collin Donahue, Lucas Foppe, Lamont Brown, Chadwick Riggins, Erik C. Rye, Dane Brown: *A study of MAC address randomization in mobile devices and when it fails*; Proceedings on Privacy Enhancing Technologies, 2017, No. 4, 2017, pp. 365-383.

- [44] Jeremy Martin, Erik Rye, Robert Beverly: *Decomposition of MAC address structure for granular device inference.*; Proceedings of the 32nd Annual Conference on Computer Security Applications, 2016, pp. 78-88.
- [45] Louise Matsakis: *The WIRED Guide to Your Personal Data (and Who Is Using It)*; <https://www.wired.com/story/wired-guide-personal-data-collection/>, February, 2019.
- [46] Theodoric Meyer: *No Warrant, No Problem: How the Government Can Get Your Digital Data*; <https://www.propublica.org/article/no-warrant-no-problem-how-the-government-can-still-get-your-digital-data>, June, 2014.
- [47] Katina Michael: *The legal, social and ethical controversy of the collection and storage of fingerprint profiles and DNA samples in forensic science*; 2010 IEEE International Symposium on Technology and Society, 2010, pp. 48-60.
- [48] Microsoft: *Devices*; <https://account.microsoft.com/devices>, November, 2020.
- [49] Paul Mozur: *Inside China's Distopian Dreams: A.I., Shame and Lots of Cameras*; The New York Times, Vol. 8, 2018 <https://web.archive.org/web/20191016213347/https://www.nytimes.com/2018/07/08/business/china-surveillance-technology.html>.
- [50] Lily Hay Newman: *Carriers Swore They'd Stop Selling Location Data. Will They Ever?*; <https://www.wired.com/story/carriers-sell-location-data-third-parties-privacy/?GuidesLearnMore>, January, 2019.
- [51] Lily Hay Newman: *The Privacy Battle to Save Google From Itself*; <https://www.wired.com/story/google-privacy-data/?GuidesLearnMore>, November, 2018.
- [52] Brendan O'Connor: *CreepyDOL*; Technical Paper, Malice Afterthought, Inc., June, 2013 <https://media.blackhat.com/us-13/US-13-OConnor-CreepyDOL-Cheap-Distributed-Stalking-WP.pdf>.
- [53] Andreea-Cristina Petre, Cristian Chilipirea, Mitra Baratchi, Ciprian Dobre, Maarten van Steen: *WiFi tracking of pedestrian behavior*; Smart Sensors Networks, Elsevier, 2017, pp. 309-337.
- [54] Pinterest: *Heat Map*; <https://i.pinimg.com/originals/9d/72/3a/9d723a435d05997bdbaa5a4f212cfba28.jpg>, November, 2020.
- [55] RF Wireless World: *Home of RF and Wireless Vendors and Resources, WLAN Probe Request Frame*; <https://www.rfwireless-world.com/Terminology/WLAN-probe-request-and-response-frame.html>, 2012.
- [56] Hannu Saarijärvi, Hannu Kuusela, PK Kannan, Gauri Kulkarni, Timo Rintamäki: *Unlocking the transformative potential of customer data in retailing*; The International Review of Retail, Distribution and Consumer Research, Taylor & Francis, Vol. 26, No. 3, pp. 225-241.
- [57] Piotr Sapiezynski, Arkadiusz Stopczynski, Radu Gatej, Sune Lehmann: *Tracking Human Mobility Using WiFi Signals*; PloS one, Vol. 10, No. 7, 2015, e0130824.
- [58] Susan Thomson, Thomas Narten, Tatuya Jinmei: *IPv6 Stateless Address Autoconfiguration*; RFC 4862 (Draft Standard), Internet Engineering Task Force, December, 1998, <http://www.ietf.org/rfc/rfc4862.txt>.

- [59] Lisa Vaas: *US cop goes wardriving to sniff out stolen gadgets by MAC address. Naked Security by Sophos* <https://nakedsecurity.sophos.com/2015/09/10/us-cop-goes-wardriving-to-sniff-out-stolen-gadgets-by-mac-address/>, September, 2015.
- [60] Mathy Vanhoef, Célestin Matte, Mathieu Cunche, Leonardo S Cardoso, Frank Piessens: *Why MAC address randomization is not enough: An analysis of Wi-Fi network discovery mechanisms*; Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security, 2016, pp. 413-424.
- [61] Wigle: *Official Website*; <https://wigle.net/>, November, 2020.
- [62] Winkey Wang: *Wireless networking in Windows 10*; Windows Hardware Engineering Community conference (WinHEC), March, 2015.
- [63] WebFX: *What Are Data Brokers - And What Is Your Data Worth?*; <https://www.webfx.com/blog/internet/what-are-data-brokers-and-what-is-your-data-worth-infographic/>, March, 2020.
- [64] Wireshark: *Official Website*; <https://www.wireshark.org/>, November, 2020.
- [65] Sunkyu Woo, Seongsu Jeong, Esmond Mok, Linyuan Xia, Changsu Choi, Muwook Pyeon, Joon Heo: *Application of WiFi-based indoor positioning system for labor tracking at construction sites: A case study in Guangzhou MTR*; Automation in Construction, Elsevier, Vol. 20, No. 1, 2011, pp. 3-13.
- [66] OvertonWindow: *Website of Mackinac Center for Public Policy, whose research introduced the model of Overton window*; <https://www.mackinac.org/OvertonWindow>, November, 2020.

Chapter 12

Is it All About Money? An Analysis of Company Investments in Cybersecurity

Adrianna Marszal and Pascal Marty

Since the dawn of a new century the world has seen many cyber attacks performed on companies for financial gain. This report investigates how companies are protecting themselves against such cyber attacks. To enable the analysis, the report starts with introducing definitions of various threats and risks. Later it presents the current cybersecurity market and current security technologies. Furthermore, it moves on to show various threats in different industry sectors: healthcare, finance and telecommunications, as well as with distinctions between small- and medium-sized enterprises in comparison to multinational enterprises. Afterwards, this report analyzes cybersecurity investments which companies make in the three aforementioned sectors, also with juxtaposition of SMEs and MNEs. Lastly comes a discussion of findings and the future of cybersecurity investments which we enrich with recommendations – also, but not limited to, COVID-19 crisis.

Contents

12.1 Introduction	354
12.1.1 Methodology	354
12.2 Background	355
12.2.1 Distributed Denial-of-Service	355
12.2.2 Malware	355
12.2.3 Phishing	356
12.2.4 Insider Threat	356
12.2.5 Risks of Cyber(in)security	357
12.3 Cybersecurity Market	357
12.4 Cybersecurity Threats	359
12.4.1 Healthcare	359
12.4.2 Finance	360
12.4.3 Telecommunications	361
12.4.4 Small- to Mid-size Enterprises vs. Multinational Enterprises	363
12.5 Analysis of Cybersecurity Investments	364
12.5.1 Healthcare	365
12.5.2 Finance	365
12.5.3 Telecommunications	366
12.5.4 Small- to Mid-size Enterprises vs. Multinational Enterprises	366
12.6 Discussion	367
12.6.1 The Future of Cybersecurity Investments	367
12.7 Conclusion	369

12.1 Introduction

Since its advent in the beginning of the 1990s, the World Wide Web has enjoyed rapid and widespread market adoption, with its user-base already comprising nearly half of the world's population by 2018 [1]. The digitization of society and economy has led to an increasing dependence on the availability of functioning IT infrastructure. Many modern business models, such as on-demand video streaming or online shopping would not even be possible without today's broad diffusion of internet communication technology. Although the internet revolution has fundamentally changed many business sectors, most often in a positive way for both consumers and businesses, its success has also led to a rampant growth in cybersecurity threats and cybercrime. From the beginning of 2016 onwards, more than 4000 ransomware attacks occurred worldwide daily – an increase of 300% compared to 2015. When looking at Europe, around 80% of companies have experienced at least one cybersecurity incident in 2016. Global annual damages caused by cybercrime are estimated to grow to around EUR 4.8 trillion by 2021 [2].

With more and more connected users and devices worldwide, the surface for possible cyberattacks keeps on growing [3]. To illustrate the expanding cyber attack scene, one can look at the growth of various facets of the digitized economy. It is forecasted that more than half a billion wearable devices will be sold globally in 2021 compared to roughly 310 million devices in 2017, a growth of over 60%. The wearables category includes devices such as smartwatches, head-mounted displays, body-worn cameras, wireless headsets and fitness monitors [4]. Especially when looking at fitness monitors (and connected physiological measurement devices in general), concerns regarding data privacy and protection arise quickly. Furthermore, the world's available digital content cumulatively is expected to grow from 4 billion terabytes in 2016 to 96 billion terabytes by 2020 [5]. Much of this data is estimated to comprise of enterprise data significant for businesses, for example in the utilization of big data analytics.

This report discusses what kind of cyber threats businesses from three different sectors are confronted with today. Thus, this report presents the cybersecurity situation of the healthcare, financial and telecommunication sectors to discover whether these quite different industry sectors also face different types of threats. The report assesses what kind of investments these industries make to tackle the cybersecurity challenges and how effective they are. With this assessment, it is then possible to discuss whether it is all about money or not when it comes to cybersecurity investments. Before concluding, the report gives recommendations on possible improvements through future investments, also given the current COVID-19 pandemic.

12.1.1 Methodology

In this report we used sources such as papers from academic journals. Where we could not find information in academia, as technology often moves faster than science, we reached for materials from the industry, i.e. corporate and technical reports, articles, company posts, and blog posts from cybersecurity experts.

In terms of industry sectors presented in this report, we decided to focus on the healthcare, financial and telecommunications sectors. Our choice was motivated by the importance of these sectors' services for society as a whole. A well-working healthcare sector is fundamental to a society's well-being; most people keep their money in a bank and want to keep their resources secure; continually available telecommunication channels are something which a society as a whole increasingly depends on in the age of digitization. Furthermore, due to their importance, these industries are subject to attacks which could compromise democratic systems with high stakes at play.

12.2 Background

According to [6], hacking is an attempt to exploit a computer system or a private network inside a computer. Simply put, it is the unauthorized access to or control over computer network security systems for some illicit purpose. This definition is most commonly used in the media. However, computer science experts refer to hacking as compromising computer systems, personal accounts, computer networks, or digital devices [7]. Therefore, while the former definition is pejorative, the latter one is rather neutral. Usually hacking is performed by an intelligent and highly skilled individual for financial gain, political reasons (espionage), ideological reasons (hacktivism) or destruction (due to revenge). Hackers can be split into white hats, black hats and grey hats. White hats perform hacking for a company which they are employed by. This can be done by doing penetration testing of a system or network. However, black hats hack such systems for personal gain. In the middle of these definitions lay gray hats who discover weaknesses found in a system or network and report them. However, both the black hats and grey hats perform illegal activity. Even though hacking is thought by many to be an inherently neutral action, in this report we refer to hacking as a malicious activity performed by black hats, thus using the firstly introduced definition in this paragraph.

12.2.1 Distributed Denial-of-Service

There are various attacks that happen in the cyberspace. One of them is a Distributed Denial-of-Service attack (DDoS) which is one of the most powerful attacks online. This type of attack targets websites and aims at making them inoperable. The way attackers do this is by overwhelming the server or network with more traffic than expected. The traffic can consist of incoming messages, requests for connections, or fake packets. This way the website is “brought down by hackers”, as newspapers describe. Sometimes, victims are barely threatened with such an attack and are expected to pay a ransom. Types of websites that are targeted using this attack are multi-fold: news outlets, web shops, search engines, financial, software and media companies. Occasionally, DDoS attacks are used to divert attention from other types of cyberattacks, such as installing malware or stealing data. These days the attackers include cybercriminals who are motivated by high profits, hacktivists and also nation states.

When looking at an average annual costs caused by DDoS attacks, they are in the top third of cybercrime attack types. The Ninth Annual Cost Of Cybercrime Study has found that damages caused by DDoS attacks grew by 10% from 2017 to 2018, rising from an average annual costs of \$ 1,565,435 to \$ 1,721,285 [8]. In order to protect against this type of attack, a company has to determine its vulnerabilities. However, it is difficult to protect a company against a DDoS attack. Hence it is crucial to react quickly when it happens. As Norton writes, there are four methods to protect a company. Firstly, a company should use technology that can recognize spikes in network traffic, and notify its Internet Service Provider immediately to figure out whether the traffic can be re-routed and counter the attack. Secondly, firewalls and routers should be configured in such a way that they can keep only the real, non-malicious traffic. Also, they should be updated with the latest security patches. Last but not least, nowadays it is possible to use artificial intelligence for protection and the use of blockchain technology is currently being researched [10].

12.2.2 Malware

This type of attack is an abbreviated form of “malicious software”. A malevolent application can perform a variety of tasks: create a persistent access to a network, spy on a user in order to obtain valuable data or simply cause disruption, all this usually without the

owner's knowledge. The most common type of malware is software called ransomware. These programs are designed to encrypt a target's files and then will ask the victim to pay a ransom in order to decrypt the data. Other types of malware include spyware (stealing sensitive information), viruses and worms (malware which replicates itself), Trojan horses (software which installs a trapdoor) and adware (forced advertising).

Malware creation is on the rise due to the profits which come through organized online crime. From 2017 to 2018, Malware attacks grew by 11%. As of 2018, Malware attacks were the most costly type of cybercrime attack with an average annual damage caused by them being estimated at \$ 2,613,952 in 2018. Moreover, businesses are overlooking the danger, as according to [11] only one out of ten enterprises is concerned about malware infection, even though it is the costliest security incident for them. Also, when only looking at ransomware one can even see an increase of 21% from 2017 to 2018 [8]. The most common advice for protection against various types of malware seems common sense. It is mostly being careful about opening email attachments, staying away from suspicious websites, and keeping an updated antivirus program [12].

12.2.3 Phishing

This attack is performed by cyber criminals in order to steal sensitive information or data from victims, such as bank account details or even identity. The criminals disguise themselves as a trustworthy source and use various channels. They send fake emails and texts in mass volume, hoping that a small percentage of them will trick the users. Some messages have poor grammar, look unprofessional and include suspicious-looking URLs, but more and more messages look genuine and therefore may even trick the most aware users. Scammers can try to impersonate legitimate companies, personalize messages, replicate messages, target high-ranking executives and use pop-ups. One of the recent phishing attacks from 2018 pretended to come from Netflix and asked users to update their billing information, thus stealing their bank details. Phishing attacks, while not growing as fast as other types of cybersecurity threats, are still growing by a considerable margin, seeing an 8% growth between 2017 and 2018 [8].

In order to protect oneself against phishing, employees of companies must learn how to recognize phishing scams, for example through internal cybersecurity training programs. It is crucial to never: open suspiciously looking emails, click on a link coming from an unknown source (especially related to net banking and login credentials), send financial information via email, nor click on pop-up ads. Overall, phishing is an increasingly common cyberthreat that all kinds of employees should be made aware of [13].

12.2.4 Insider Threat

This type of attack is performed by a person connected to an organization – either a current employee, a recently-fired employee, an external consultant, or a business partner. An insider is anyone who has knowledge and/or access to the company's confidential data, IT or network resources. Depending on the source, there are different categories of insiders but in general there are two types: malicious agents and inadvertent agents. Malicious agents try to steal company's intellectual property or data, while inadvertent agents simply happen to carelessly help in that. An insider threat could be stealing documents from the company and disclosing them to a competitor; taking a picture of specialized production machinery and sharing it with others; losing a laptop on public transport; or emailing a sensitive document to the wrong person. Malicious insiders are one of the larger cybersecurity threats, causing an average annual cost of \$ 2,275,024 in 2018. Apart from causing considerable damages, insider threats are also seeing a considerable grow, increasing by 15% from 2017 to 2018 [8].

There are various ways to protect against insider threats. For example, an organization should monitor its activity on the core data sources. Furthermore, it should identify where the sensitive files are stored and review who has access to them. Additionally, a company may apply user behavior analytics to alert on suspicious behaviors. And last but not least, employees must be trained to adopt a data security mindset. As one can see, there are various types of cyberattacks and the list goes on: man-in-the-middle attack, zero-day exploit, SQL injection, DNS tunneling, business email compromise, cryptojacking, drive-by attack, cross-site scripting attack, password attack, eavesdropping attack, AI-powered attack and IoT-based attack. And these are just the most common types [14].

12.2.5 Risks of Cyber(in)security

In many firms, especially those with a long history, there is little understanding about IT. In these companies business and IT departments are often separate and lack a common language; business managers very rarely deal with IT employees. This must change because technological changes are happening rapidly around the world and touch every company. According to a former Cisco CEO [15], there are two types of companies: those that have been hacked and those who do not yet know they have been hacked. If a company is behind with software updates, firewalls and upcoming technology, it will sooner or later lose some part of its profit due to being hacked. Hence a company faces various dangers by not reacting to voices of the IT part of its business.

Thus, there are different threats facing these companies, such as a data breach which exposes data of a business' customers, ruins its reputation, causes damages on its technical infrastructure, and often has financial and legal consequences.

It is critical that companies cover cybersecurity basics such as updating, patching and integrating its systems, as well as educating all employees by providing information security training. Furthermore, a company must have a cybersecurity policy in place. A firm must have a recovery plan and constantly monitor evolving risks, too.

Data breaches have a massive negative business impact and often come from insufficiently protected data. Hence each company should have not only an IT risk management division, but also a cybersecurity risk management division. Furthermore, protection from cybersecurity threats cannot be purchased as a package but is rather an ongoing process in which a company must react to new technological developments. An organization can never be too secure; rather, it most probably is not secure enough.

12.3 Cybersecurity Market

To protect themselves against cyberattacks, companies invest in protection services and response teams for ensuring business availability and to protect crucial services as well as its infrastructure [16]. Growing demand for cybersecurity protection services has facilitated the development of a flourishing cybersecurity market. According to Cybersecurity Ventures, global spending on cybersecurity will cumulatively exceed \$ 1 trillion for the five-year period from 2017 to 2021 with the cybersecurity market growing by 12-15% year-over-year through 2021 [1]. The cybersecurity market provides companies with a variety of technologies that can be used to tackle the issue of cybercrime. In the following section we highlight some of these technologies and how they benefit their users.

This section presents different security technologies, ordered from the ones providing the largest net technology savings to those that provide the least. Net technology savings hereby refer to the amount of money a company can save by the use of technology after deducting the costs incurred due to the said technology. The rest of this section is based on the annual report of cybercrime published in 2019 by Accenture [8].

When looking at what kind of a security technology creates the highest net savings for a company, security intelligence and threat sharing technology stands at the top. This category encompasses measures facilitating the sharing of knowledge concerning security and threat intelligence between companies, so that all parties involved can learn from each other. Such measures are crucial if companies have to deal with large-scale threats like botnets where they may only see some bots of a larger botnet on their own servers and therefore would not know about the full scale of the present threat. To get a complete picture of such threats companies can utilize security intelligence and threat sharing technologies. Facebook's ThreatExchange, a platform for sharing of security threat information, is one example of such a measure [17]. Partners can utilize the provided APIs to query available threat information or to publish own information for all participating organizations or just a subset of them. In 2018, security intelligence and threat sharing benefited the companies using it with net savings of \$ 2.26 million with 67% of companies utilizing the technology.

The second most cost-saving technology was found to be automation, AI and machine learning which provides net savings of \$ 2.09 million. One promising use of machine learning in cybersecurity is in the prediction and forecasting of future cyberattacks and a sequence of actions the hackers will take when attempting to compromise a system [16]. Despite substantial net savings facilitated by this category of technologies, only 38% of surveyed companies make use of it. This represents a lost opportunity for many while also being a highly interesting field for future cybersecurity investments.

Advanced identity and access management technologies earn companies utilizing it net technology savings of \$ 1.83 million. These technologies enable administrators fine control over what data and services members of an organization have access to. This detailed management of information system access and the ability to cordon off and secure more sensitive parts of a system plays a fundamental role in cybersecurity. Out of the surveyed companies, 63% were found to utilize advanced identity and access management technologies. As security intelligence and threat sharing technology before, advanced identity and access management technology therefore represents a type of cybersecurity technology which is highly beneficial to companies using it while also being widely distributed.

Cyber and user behavior analytics technologies are at the end of front-runner technologies when looking at net technology savings incurred. This technology leads to \$ 1.72 million in net technology savings while being utilized by 32% percent of companies surveyed. Such technologies make it possible to identify users or groups of users that are prone to doing malicious operations early by calculating an individual security rating derived from their usage patterns, frequency of mistakes and duration of good online behavior [18]. As with technologies such as automation, artificial intelligence and machine learning before, it is an area of cybersecurity many companies are yet to invest in.

Moreover, cryptography technologies are utilized by 55% of surveyed companies and result in \$ 0.85 million net technology savings. These technologies are utilized to enhance the security of information systems and the data processed by them.

In regards to enterprise governance, risk and compliance technologies, they are in use by 45% of surveyed companies, while benefiting those companies with more conservative net technology savings of \$ 0.2 million. This category contains technologies that assist an organization in implementing their devised governance and risk plans and in increasing compliance with them.

Furthermore, automated policy management technologies are only used in 27% of the companies in the survey, resulting in \$ 0.09 million net technology savings. This type of technologies assists companies in managing policies throughout their life cycles.

Additionally, data loss prevention technology is utilized by 51% of companies with slim but still positive net technology savings of \$ 0.08 million. This group of technologies contains solutions which help organizations reduce the risk of sensitive data leakage. They do

this by performing both content inspection and context analysis of data sent via channels such as messengers, emails and instant messaging over the network or on managed devices. Based on the specific rules and policy of the organization, they then execute responses to address the risks posed by inadvertent or accidental leaks or exposures of sensitive data outside of authorized channels [19].

Last but not least, advanced perimeter control technologies are utilized by 58% of surveyed companies. These are solutions which help companies control and manage their so-called "perimeter", i.e. a barrier separating company's networks and systems from the "outside". This barrier is nowadays getting more and more unclear with mobile devices being in wide use[20]. It is the only type of cybersecurity technology listed in the Accenture Cost of Cybercrime Study that does not result in positive net technology savings. With \$ -0.16 million net technology savings this technology has been found to result in net technology losses for the companies using it.

To summarize, not all cybersecurity technologies are created equal. The technology bringing the highest net savings for companies is security intelligence and threat sharing technology, benefiting companies by \$ 2.26 million. Knowing this, it might not come as a surprise that it is also the most often utilized cybersecurity technology, being used by 67% of companies surveyed.

12.4 Cybersecurity Threats

This section provides details regarding the challenges which different sectors are facing regarding cybersecurity. More specifically, it discusses the most common and dangerous threats for the sectors of healthcare, finance, and telecommunications.

12.4.1 Healthcare

The healthcare sector is facing major risks when it comes to the digitization of patient care and treatments. The continuous availability of medical equipment is of utmost importance with human life being at stake. As most other industrial sectors, the healthcare sector has benefited greatly from new technology – be that through devices that monitor health or deliver medicine, telemedicine technology enabling care to be delivered remotely and/or faster, and more efficient storage of patient records in the form of Electronic Health Records (EHR) [21].

When it comes to cybersecurity, companies in the healthcare sector must defend their non-public patient data from malicious actors such as hackers. Hacked patient data could be leaked, used for clinical fraud or other unscrupulous activities in order to gain financial advantage. Compromised systems in healthcare, whether being inaccessible due to a ransomware attack or providing altered and incorrect data to clinicians, may decide between the life and death of unfortunate patients. With eHealth structures utilizing cellular data and cloud systems the healthcare sector's surface for potential cyberattacks keeps growing as well. Therefore, it becomes increasingly difficult to provide effective protection for crucial health data of customers and patients. Consequently, it is of utmost importance for the healthcare sector to not only harvest cost-cutting benefits which the new eHealth systems deliver, but also to improve the increasingly important protection of sensitive medical records [22].

While new technologies enjoy growing diffusion in the sector, legacy systems – such as the Windows XP operating system – are still being used in other areas. As Windows XP has not been supported since 2014, computer systems running this operating system are vulnerable to hackers and malware which easily avoid detection by the outdated system [21].

One example of a major incident where this vulnerability was taken advantage of can be found in the 2017 WannaCry attack. The scale of the WannaCry ransomware attack has been unprecedented, despite sufficient warnings being issued prior to the attack and security patches, which would have protected systems against the attack, being available but not installed. The ransomware infected more than 300'000 computers globally, demanding users to pay ransoms in Bitcoin. While not being specifically targeted by the WannaCry attack, the healthcare sector fell victim to it on a worrying scale with widespread damages incurred resulting from the attack. Just in the UK, fifty hospitals experienced system-wide lockouts, delays to patient care and function loss in connected devices such as MRI scanners or storage refrigerators. Unfortunately, the WannaCry attack does not seem to be a one-off attack, but rather a symbolic attack of a systematic issue concerning ransomware attacks in healthcare, some of which specifically target this sector. Reports indicate that between 2015 and 2016, 50% of UK NHS trusts fell victim to some form of ransomware attack. Concerning ransomware attacks in healthcare outside of the UK, one can take a look at the 10-day Hollywood Presbyterian Medical Centre shutdown caused by ransomware thought to have originated from a phishing email. The shutdown had to be resolved by paying a \$ 17,000 ransom [21].

12.4.2 Finance

It was expected that the global user base of internet banking would grow to 2 billion users by 2020. In fact, this milestone was already hit a few months later in 2018, two years earlier than expected [23].

Following the rapid adoption of the internet during the past three decades, the financial sector, like almost every other industry, has moved some if not most parts of its business online or created completely new business models that previously were not possible. Nowadays, financial organizations try to adapt new digital channels, automation, artificial intelligence, blockchain and other advanced technologies – while still struggling with removing legacy systems from their organizations. In this endeavor, firms are navigating a landscape with various cyber risks. Let us now explain some of the most prominent cybersecurity statistics in the financial sector.

In one of the most expensive data breaches in corporate history, Heartland Payment Systems was hit by malware in 2008. The company was a U.S.-based payment processing and technology provider. They had their debit and credit card numbers stolen which they did not know about until a year later. Overall, the firm had to pay \$ 140 million in legal fees and other costs [24].

Moreover, recently another expensive data breach happened to a firm in the financial sector. CapitalOne is an American bank holding company specializing in credit cards, auto loans, banking, and savings accounts. In 2019 a hacker gained access to 140,000 American social security numbers and 1 million Canadian social insurance numbers, together with 80,000 bank accounts as well as an undisclosed number of client names, addresses, credit scores, credit limits, and balances as well as other personal information. This hack cost the company between \$ 100 million and \$ 150 million. This includes customer notifications, credit monitoring, tech costs and legal support due to the hack [25; 26].

In 2019, over 70% of American banks ranked cybersecurity as their top concern. This should not come as a surprise because according to Clearswift, 70% of financial institutions have experienced a security incident within the previous 12 months. The leading cause of most incidents were employees' failures to follow the security procedures. 32% of the attacked were cause by using BYOD (Bring Your Own Device), while 25% by file and image downloads and 24% by unintentional data sharing by employees. Furthermore, according to Boston Consulting Group, financial institutions are 300 times more likely to be hit by a cyber attack than any other industry is. When that happens, the average

annual cost of cybercrime per company within the financial industry in 2019 was \$ 18.5 million. Therefore it seems reasonable that financial institutions dedicate 0.3% of revenue and 10% of their IT budget to cybersecurity. Moreover, according to Akamai, the vast majority of cyber attacks on financial institutions are executed using four methods: SQL injection, local file inclusion, cross-site scripting and OGNL Java injection [27].

As one can see, the financial sector is not spared from cyber attacks at all. Therefore, it is pivotal that banks and other financial institutions learn how to protect themselves against hackers. In the end, this is in order to protect their finances, brand reputation and customer loyalty.

12.4.3 Telecommunications

Nowadays, the public is heavily reliant on the omnipresence of mobile computing, including smartphones and telecommunications infrastructure. This technology plays an increasingly significant role in today's businesses and society at large. The cybersecurity of critical telecommunication infrastructure is not only about protecting the confidentiality of information, but also about the integrity and availability of data and systems which nations depend on. This is even more important when looking at newly emerging technologies such as the Internet of Things (IoT) and 5G technology. Consequently, it is unsurprising that governments around the world ought to look closely at what kind of telecommunication infrastructure, be it in the form of hardware or software, is being adopted nationwide and by whom it is developed, provided and maintained. Due to the importance which their infrastructure holds, the telecommunications sector is not only targeted by smaller independent malicious actors, but also by foreign government entities. Concerns in regard to the latter have arisen especially in light of the recent 5G technology adoption.

The threat landscape that telecommunications companies face is vast. In their 2020 Mobile Telecommunications Security Threat Landscape Report the GSMA identified the following threats the industry faces: cloud threats, device threats, Internet of Things threats, 5G-related threats, security skills shortage, signaling service threats, software threats and supply chain threats [28].

The currently ongoing rollouts of 5G mobile networks and technology around the world promise to revolutionize how enterprises and consumers interact with mobile operators. The 5G era introduces numerous new technologies, new ways of working and unprecedented increases in scale. This brings new challenges but also new opportunities for cybersecurity. According to an AT&T survey, 73% of respondents rated their concerns regarding potential impacts of 5G at medium-high to high, of which 42% believe that 5G will have a very significant impact on their network and that they may require a new security stack and/or entirely new set of processes [29].

It is crucial that the approaches taken to implement and operationalise 5G architecture, as well as its underlying technologies, are well-thought-out and, on top of that, that they are assessed thoroughly. Insufficient approaches may result in missing out on opportunities provided by the secure-by-design 5G standards. These 5G standards outline a service architecture that closes several of the gaps currently being exploited, such as fraud or security issues. Current 5G network deployments consist of non-standalone (NSA) 5G networks which do not make full use of of this standards-based security, much of which can only be utilized when a 5G core (5GC) is deployed. Therefore the significant security enhancements possible with 5G are yet to be realized. With the right approach to 5G rollouts and service launches it is possible to embed security and prevent various known threats before they impact the network [28].

Another threat identified by the GSMA originates within cloud and virtualization technology. The usage of cloud services continues to grow year-by-year, including IT and

telecommunications alike. Fortunately, virtualization and cloud threats are well understood. Protection against these kinds of threats requires a combination of traditional IT hygiene controls with the recognition of structural and supply chain changes affecting one's network. Traditional IT and hygiene threats can result from poor patching practices, virtualization-aware malware, lack of network visibility and inappropriate access controls. When looking at a network structure, threats can arise from the usage of insecure APIs and interfaces as well as misconfigured isolation controls. To provide proper resilience it is critical that the used services are as independent as possible from vendors and their geographical location. Otherwise, potential bans of partners or vendors going bankrupt could be an immense threat for a telecommunications network, as well as for the continuity and availability of its services [28].

Devices and their applications utilizing a telecommunications network pose another security threat. When users fail to update applications on their devices, the results are outdated privacy measures remaining in the telecommunications ecosystem. The existence of potentially harmful apps (PHA) or data leaking apps that are not blocked or controlled using the latest updates is almost a certainty. To prevent unauthorized use of consumer data, security patches need to be installed with further updates. The failure to update one's device with up-to-date security patches drastically increases the threat posed by the malicious applications. With over five billion unique subscribers and mobile devices accounting for 50% of internet traffic, the surface of attack is immense. While largely no direct security threat to telecommunications companies, this issue can endanger customer relations. On the one hand, consumers expect to be able to run all of their lives from their devices. On the other hand, increasing awareness concerning inadequate privacy controls and unauthorized use of data diminishes consumer trust in the entire mobile telecommunications sector [28].

Another threat which the telecommunications sector faces is a lack of supply chain resilience, a subject that has come to the forefront with the ongoing 5G rollout. Two resilience issues are of major concern. The first resilience issue is the fading trust into suppliers and geopolitical relations for critical national infrastructure (CNI) and security sensitive components. To illustrate this issue, one can look at the recent issues concerning adoption of Chinese 5G technology. The Chinese telecommunications companies Huawei and ZTE are major global players in establishing 5G networks. Huawei offers an end-to-end approach which means that the company provides hardware, software and continuous operational support to its customers. This gives the company a competitive advantage which enables them to undercut the competition in pricing. This end-to-end approach also means that identifying vulnerabilities, providing updates, installing patches, designing and distributing hardware and software upgrades all lie within Huawei's hand. For this to work, customers of Huawei's 5G infrastructure solutions have to put a lot of trust in the company. Said trust put in Chinese companies is faltering because of China's high level of cyber espionage for commercial and state purposes. While many nations such as the United Kingdom, Australia and Canada initially adopted or planned to adopt telecommunications infrastructure provided by Huawei or ZTE, many have since stepped back from such arrangements due to rising concerns regarding Chinese government's involvement and potential backdoors in the provided telecommunications infrastructure [30].

What makes the Chinese telecommunication infrastructure providers especially risky regarding the cybersecurity of nations looking to adopt their products is Article 7 of China's 2017 Intelligence Law. This law obliges Chinese organizations and citizens to support, assist and cooperate with intelligence work. It is therefore implied that Chinese companies such as Huawei and ZTE are required by law to keep open some backdoors for when the Chinese intelligence service requests their assistance in intelligence work. Despite the Chinese government, Huawei and ZTE refuting claims of compromised infrastructure, nations worldwide seem to move away from telecommunication infrastructure provided by

Chinese companies [30]. For example, British mobile providers are being banned from buying new Huawei equipment after December 31, 2020, and must remove all of Huawei's already existing kit from their networks by 2027. This decision was taken by the UK government despite delaying the nation's 5G rollout as a consequence by a year [31].

A robust supply chain has access to equipment offered by a diverse set of suppliers. This availability of various suppliers is vital for market economies and decisive in preventing vendor lock-in. Diversifying a supply chain also minimizes the impacts of an individual supplier becoming unavailable, be that by going out of business, having insufficient capacity or due to government sanctions. The second resilience issue supply chains in the telecommunications sectors face is a lack of robustness [28]. As the report showed before, popular 5G network suppliers such as Huawei offer end-to-end solutions, resulting in a strong lock-in of customers. We have also seen the impacts that the removal of a supplier such as Huawei can have on its customers, with the UK telecommunications providers having to remove and replace all Huawei-supplied telecommunications solutions.

With all the attention given to the 5G technology at the moment, it is easy to forget that 2G and 3G networks are still deployed and used globally. It is unlikely that these legacy networks will disappear from the ecosystem anytime soon, which means that legacy threats will continue to require compensating technologies and controls to protect consumers using these dated technologies. Signaling threats that can be found in these networks are data, call, email and text message interception, digital identity theft, denial of service attacks, financial fraud and theft as well as (unauthorized) location tracking [28].

12.4.4 Small- to Mid-size Enterprises vs. Multinational Enterprises

This section explores the differences between Small- to Mid-size Enterprises (SMEs) and Multinational Enterprises (MNEs) in regards to digital innovation and cybersecurity technology adoption. SME hereby refers to companies of up to 250 employees [2] or even up to 1000 employees [32], depending on the source. On the other hand, MNEs are defined as companies larger than SMEs. While both SMEs and MNEs are facing a variety of challenges when it comes to cybersecurity investments, both SMEs and MNEs face similar issues. The difference is the scale at which the issues are encountered. Here, we look more into areas where SMEs are lagging behind MNEs, as MNEs often have the resources to tackle the cybersecurity threats which SMEs do not.

In regards to differences between SMEs and MNEs, it must be said that breaches are more than twice as common in the larger companies than in the small ones. However, there are more similarities than differences. The proportions of threat actors are distributed rather similarly in both MNEs and SMEs: external actors comprise above 70%, internal actors – above 20% and partners include only 1%. Furthermore, actors' motives are also distributed uniformly: hackers perform cyber crimes due to financial reasons (79-83%), for espionage purposes (8-14%), for fun (2-3%) and because of grudge (2-3%). When it comes to compromised data, it is also a comparable situation in both SMEs and MNEs: they are mostly credentials (above 50%), personal data (about 20-30%) and then internal data (12-14%) [32]. Moreover, the two most frequent incidents targeting both SMEs and MNEs are inappropriate use of IT resources by employees and malware infection of company-owned devices [11].

Regarding SMEs, they form the backbone of the European economy, comprising 99% of European companies while accounting for two-thirds of its total employment. Despite their vital part in the economy, SMEs keep lagging behind MNEs when it comes to digitization. For example, in 2019 only 12% of SMEs were using some type of big data source compared to 33% of large enterprises, a staggering technology gap that would require at least five million SMEs adopting said technology to be closed. The rest of this section is based on

a publication from the European Union about cybersecurity, IoT and big data for SMEs [2].

There are several barriers which SMEs face in the context of adopting new technologies. As a big strength of SMEs lies in the ability to easily serve niche or specialized markets, major parts of SMEs' human resources are comprised of domain specialists. This specialization leads to a poor personal coverage in more generic functions which are beneficial in spotting new business opportunities and trends outside of the respective domain which SMEs operate in. Compared to SMEs, MNEs have a sufficient volume of employees to easily cover more generic functions without sacrificing specialised personnel.

Moreover, both SMEs and MNEs are affected by a growing shortage of qualified IT specialists on the labor market. This skills shortage is especially serious for SMEs as they, hard as it already is, have difficulties competing with large enterprises for scarce digital talent. As a result of this shortage, it is estimated that in Germany alone companies will suffer economic damage of around EUR 10 billion in revenue. It is therefore crucial for companies to invest in strategies such as up- or re-skilling implemented by training their employees. This is not a small hurdle to overcome given that improving more basic digital skills is already challenging enough. Up- or re-skilling strategies are currently dramatically underutilized in SMEs, with less than 10% of SMEs providing training to ICT specialists and less than 20% of SMEs offering training to other employees. It follows that over 90% of Europe's SMEs see themselves lagging behind in terms of digital innovation.

As seen before, SMEs generally lag behind MNEs when it comes to digital innovation and acquisition of IT talent. As one would expect, SMEs are also lagging behind MNEs when it comes to cybersecurity. Shockingly, only 32% of European SMEs have been found to have a formally defined ICT security policy in place, meaning that about 17 million SMEs in Europe alone did not have the required skills and talent in their organization and will have to acquire cybersecurity skills either by investing in training or by hiring talent externally.

Contrary to common belief, not only large enterprises are targeted by cyber crime. More and more SMEs fall victim to cyber threats with many owners of SMEs still underestimating the likelihood of their company becoming a target of a cyber attack. At the same time, SMEs become increasingly dependent on their information systems and networks in order to provide services and products to their customers. A majority of SMEs rely on some form of an information system. Many of them have an online presence as electronic communication networks, interconnected information systems and digital services become an increasingly essential part of their business models.

12.5 Analysis of Cybersecurity Investments

This section discusses the current cybersecurity situation of the investments in cybersecurity of each of the sectors introduced before, i.e. healthcare, finance and telecommunications. Furthermore, potential solutions to address open challenges and problems are proposed.

Over the previous sections we have seen that cybersecurity encompasses a vast variety of facets concerning software, hardware, people, organizations and their operational environment. Through a better understanding of cybersecurity threats and the impact associated with them, organizations can better leverage their limited capital and determine the right amount of investments into the right kind of cybersecurity. Cybercrime is not going away any time soon, with the number of cyber attacks and related costs rising year-after-year. The average number of security breaches faced by organizations questioned in light of the ninth annual cost of cybercrime study rose from 130 in 2017 to 145 in 2018 – that is an increase of 11% in only one year – while there has been an increase of 67% when looking

at the last 5 years. Meanwhile, the average cost of cybercrime rose from \$ 11.7 million in 2017 to \$ 13 million in 2018 marking an increase of 12% while the 5-year growth in cost of cybercrime is at 72% [8]. It follows that business will have to increasingly prioritize cybersecurity investments to face these growing threats.

In order to determine the best cybersecurity investment options for a certain industry, one has to consider different factors [9] which include: the average cost of different types of cyberattacks, how effective certain investments are at counteracting these different types of cyberattacks and finally which types of cyberattacks are most relevant to the industry in question.

12.5.1 Healthcare

Although the healthcare sector has been quick to adopt new technology, the same thing cannot be said when talking about taking up the responsibility of protecting the newly acquired technologies against cybersecurity threats. Despite the importance of cybersecurity in the healthcare sector, findings about the current state of the sector are alarming. While under-staffing or even non-existence of cybersecurity-related positions is in no way isolated to the healthcare sector only, the scale of this issue seems to be especially large in proportion to the involved risks. Three out of four hospitals do not have a designated employee for addressing cybersecurity troubles [22]. While healthcare organizations dedicate sufficient funding to become more integrated, there is not enough spending on keeping software updated and systems secure. This issue is further exaggerated by the general lack of cybersecurity expertise in the industry as well as the consequently following substantial expense incurred by the scarce and desired cybersecurity personnel that there is [21].

In 2015, healthcare has fallen victim to more cyberattacks than any other industry [22] with reports highlighting the growth of attacks and the rise of medical identity theft facilitated by millions of medical records being stolen globally.

Cybersecurity breaches in healthcare arise mainly from hacking, malware and insider threats. As malware is the most costly type of cybercrime in general, the sector should focus on this type of cyber threat when considering cybersecurity investments [21].

12.5.2 Finance

A report by Deloitte [33] from this year shows that financial organizations reported an increase in cybersecurity spending. It went up from 0.34% of revenue in 2019 to 0.48% of revenue in 2020. This also means that total IT spending rose from 10.1% of revenue in 2019 to 10.9% of revenue in 2020. Out of all financial institutions, financial utility companies spend the biggest share of their revenue on cybersecurity, i.e. 0.8%. The rest of this section is based on said report.

Moreover, for the last three years firms identified rapid IT changes and rising complexities as their main cybersecurity challenge, and a lack of skilled cyber professionals as the second one. Third most important challenge in 2020 is business growth and expansion. The next one is difficulty prioritizing options for securing the enterprise tied with inadequate functionality and interoperability of security solutions.

Furthermore, for the past three years the top two business issues with security implications for MNEs have been embedding security into new products and services and introducing new channels. In 2019 and 2020 the third business issue with security implications for MNE have been cost reduction, probably highly affected by COVID-19 crisis which started in the first quarter of 2020.

Cybersecurity is often included within IT functions at companies and CISOs usually report to the CIO or CTO, according to the surveyed MNEs. This shows that cybersecurity and IT need to go hand in hand in corporations. However, financial organizations should keep

some level of independence for cybersecurity in order for IT constraints to not overshadow risk management decisions. This can be done by: (a) maintaining autonomy in risk management decisions, (b) establishing linkages between cybersecurity and business and (c) prioritizing cybersecurity at board level.

Additionally, companies mentioned their top cybersecurity investment priorities in the years 2018-2020 as cloud and data analytics. In 2020 the next options are artificial intelligence and/or cognitive computing, robotic process automation and mobile solutions. The reasons behind this are: access control, protective technology and data security. The latter has been becoming more and more significant over the past three years [33].

12.5.3 Telecommunications

Out of the three selected industry sectors, the telecommunications sector sees the lowest annual cost of cybersecurity and cybercrime with an average annual cost of \$ 9.21 million faced by the communications and media sector in 2018 – a 22% increase compared to \$ 7.55 million in 2017 [8]. This relatively low cost of cybercrime can be partially explained by the adeptness of telecom operators at protecting their networks. It is also important to note that most cyber adversaries utilize telecommunications infrastructure as their primary transport method when carrying out their attacks and therefore rely upon a robust network. Consequently, the amount and types of adversaries who seek to attack the telecommunications industry are slim – typically limited to anti-establishment hackers or nation states seeking to use advanced persistent threats (APTs) [34].

Due to the vast cybersecurity threat landscape which the telecommunications sector faces, investments into cybersecurity measures alone will not suffice to ensure the security and availability of telecommunications infrastructure. This is especially visible when one looks at the telecommunication sector's supply chains, their lack of resilience and the resulting issues. To mitigate the threats arising from vulnerable supply chains, telecommunications companies need to understand who they do business with. They need to prioritize and risk-assess each supplier and put a specific focus on redundancy, flexibility and the technical and procedural ability to switch out a supplier if necessary. Furthermore, they need to map and assess the criticality of each individual component and/or service offering within the supply chain and manage operational security accordingly. The strengthening of robustness of telecommunications supply chains can be facilitated by building business continuity plans that consider the removal of critical vendors. Such plans shed light on the possible impact of such a supplier being removed. Telecommunications also ought to work with local legislators and regulators so that they can better understand how potential decisions with regard to supplier bans could affect them; the world has seen the importance of such a case with the recent Huawei bans. They should also support and engage with international standards such as LTE, which was the first global standard for mobile networks. A move away from such standards for 5G would impact its deployment and long-term security [28]. There are no quick fixes for addressing these supply chain resilience issues and specific investment numbers for the measures mentioned cannot be determined without considerable effort.

12.5.4 Small- to Mid-size Enterprises vs. Multinational Enterprises

According to Kaspersky [35], MNEs cut their cybersecurity spending from \$ 18.9 million in 2019 to \$ 14 million in 2020. This is due to COVID-19-related expenses. However, the proportion of budget spent on IT security has grown. On the other hand, SMEs increased their cybersecurity budgets from \$ 267,000 in 2019 to \$ 275,000 this year.

Moreover, in 2020, a data breach cost an MNE on average \$ 1.09 million, compared to \$ 1.41 million in 2019 while an SME had to pay on average \$ 101,000, compared to \$ 108,000 in 2019 [36].

At the same time, the predictions are that the total cybersecurity investments will increase over the next three years in 71% of companies. Regardless of company size, respondents said that the reasons behind this increase are rising complexity of IT infrastructure and the need to increase employees' capabilities. Nevertheless, about 12% of companies are considering budget cuts due to overall optimization or due to a belief that previous investments already helped resolve their issues [35]. Moreover, Kaspersky created a useful tool for companies which can be used no matter the sector or size in order to get a recommendation for cybersecurity investments [37].

12.6 Discussion

We asked ourselves whether companies these days invest enough in cybersecurity. Our answer is that they do not because attacks still take place on a large scale. It is pivotal that companies in the healthcare, finance and telecom industries rethink their cybersecurity strategies in order to bulletproof their business against hackers.

The main question of this report is whether cybersecurity strategy is all about money. We think that this is not the case. More money invested in cybersecurity does not equal better cybersecurity. This is because the same amount of money can be invested in various ways. Rather, it is critical for organizations to reach capabilities which will allow them to protect the business against cyber attacks. Once cybersecurity specialists are in place and given enough freedom by the top management, they should know how to invest company resources in the most robust available technology. However, the question stays whether a larger share of the budget spent on cybersecurity makes it more secure. A little goes a long way, but it is not true that more money means more security. Rather, if more money is available from the budget, it should be spent wisely.

Another question is about what is more important for cybersecurity, i.e. what should the resources be spent on. There are various possibilities. For example, companies can educate employees in cybersecurity through internal training programs. They can also invest in newer hardware by replacing all employees' laptops. Another option is to implement a robust antivirus for everyone and update the software which the company uses. In the end, this is all about assessing risk as some firms are more prone to online break-in than others due to varying degrees of financial profits for criminals.

12.6.1 The Future of Cybersecurity Investments

When it comes to security, a company is only as strong as its weakest link. Hackers will always find a way to breach a firm's network security via the route of least resistance. Moreover, cyber attacks are increasing year-after-year in sophistication and magnitude of impact across all industries. In order to protect themselves from cyberattacks, companies should invest wisely to build a robust security strategy. Let us first introduce general recommendations including suggestions for each industry. Afterwards we mention how the world might have to act due to the current COVID-19 pandemic.

12.6.1.1 General Recommendations

Our first advice is to revisit and check the security level and relevance of the company's network infrastructures, processes, compliance of connected mobile and PC devices, IoT etc. Furthermore, it is crucial to increase cybersecurity staff [38]. Thirdly, a company must

educate all employees about social engineering tricks. Lastly, the technology managers must ensure that the network security is tightened.

According to another source, a business can reduce the cost of data breaches in various ways depending on company size. A quick detection can lower the loss by 32% in MNEs and by 17% in SMEs. Moreover, a proactive disclosure to customers and other stakeholders that a data breach has happened can lower the financial damage by 28% in MNEs and by 40% in SMEs. Another factor are timely updates – both of software and hardware – whose lack, according to the report, increases the cost by 47% in MNEs and by even 54% in SMEs. Last but not least, if a company collects customer data, it might be forced to pay 62% more than companies who do not collect customer data. This is the case for MNEs, while for SMEs this number is at 37% [36].

Moreover, regarding specific industries recommendations from experts vary. For the healthcare industry three options are suggested. Firstly, a healthcare organization should implement a security awareness and training program in which users are educated both on malicious attacks and on accidental breaches. Secondly, it is endorsed to sufficiently utilize boundary defense measures such as firewalls, network monitoring, proxies and multi-factor authentication. Thirdly, an organization should protect its data by controlling access to the sensitive information. This includes maintaining an inventory of sensitive information, encrypting sensitive data and limiting access to authorized cloud and email providers [32]. Furthermore, the following six key imperatives for future improvements of cybersecurity in the healthcare sector have been recently laid out by [22]:

1. Define and streamline the management, governance, and expectancy for healthcare cybersecurity.
2. Improve clinical tool and fitness information technology safety and resilience.
3. To develop the needed and necessary healthcare workers by prioritizing and ensure adoption of cybersecurity awareness and skills.
4. To improve and increase industry readiness with better cybersecurity detection and education.
5. Identify mechanisms to guard research and development efforts and intellectual assets from attacks and exposures.
6. Improve information sharing of enterprise threats, dangers and their mitigation.

For the financial industry three options are suggested too. The first two overlap with the recommendations for the healthcare industry, i.e. (a) implementing a security awareness and training program and (b) boundary defense. The last recommendation are secure configurations. This means ensuring and verifying that systems are configured with only the services and access needed to achieve their function [32].

Lastly, for the telecommunications industry recommendations are missing [32].

12.6.1.2 Advice Influenced by COVID-19 Crisis

The recommendations above are valid for the usual times where society & business "is as usual". However, during the times of crisis certain actions become more relevant than others. This can be seen with some vulnerabilities becoming more prevalent in the course of the COVID-19 pandemic. For example, the setup of working from home has opened a door for various cyberattacks. These come from the fact that employees might forget to use a VPN tunnel when connecting to a companies network or from psychological stressors which make employees bypass controls as they are overwhelmed with the new situation. Also, social engineering tricks are on the rise as employees cannot just turn

to a colleague and ask them for advice. Furthermore, cyber criminals use websites with weak security to deliver malware, for example a ransomware embedded in coronavirus heat maps. Moreover, these days public organizations are experiencing acute pressure that comes from attacks such as denial-of-service or ransomware.

In an article from March this year, Boehm, Kaplan and Sportsman from McKinsey recommend a four-fold approach to fight the cyber challenges during the current pandemic [39].

Firstly, they suggest to focus on critical operating needs such as maintaining security operations, mitigating risks of remote access to sensitive data and software development environments, and implementing multi-factor authentication to enable employees to work from home. The other needs can wait until the situation becomes more stable for businesses.

Secondly, McKinsey recommends testing plans for managing security and technology risks. This means plans for incident response, business continuity, disaster recovery etc. If these plans do not exist, an organization should make it a priority to make them and test them in order to reduce the risks.

Thirdly, the article's authors recommend monitoring for newly emerging cyberthreats from different angles: in collaboration tools, networks, employees and endpoints.

Last but not least, they propose to balance protection with business continuity. When everything seems urgent, it is crucial to keep a cool head. Even though a Chief Information Security Officer might usually decline requests for approving cybersecurity policy exceptions, it might be a good idea to grant waivers during certain times of insecurity such as the COVID-19 pandemic. Tolerating a slightly higher risk in the short run might prove beneficial in the long-run. This is not to say that the organization's risk should be weakened permanently; rather, that this is a temporary situation in which organizations need to re-learn the art of trade-offs.

12.7 Conclusion

This report investigates different types of cybersecurity threats and how much damage they cause. Malware stands at the top when it comes to the average annual cost caused, while its more specific variation of ransomware is the fastest growing type of cyber attack, seeing a 21% growth from 2017 to 2018.

Afterwards, the report presents what kind of technologies companies invest in to counteract these cybersecurity threats. Security intelligence and threat sharing technology stand out as the technologies through which companies manifest the largest net technology savings, \$ 2.26 million on average per year. At the same time, the security technology was most often used with 67% of companies utilizing it.

Then the report introduces and analyzes the cybersecurity situation of different industry sectors: the healthcare sector, the financial sector and the telecommunications sector. Thereupon, it explores the different challenges which small- to mid-size enterprises and multinational enterprises face when it comes to cybersecurity.

When looking at the healthcare sector, the gap between the life-depending devices and data handled and the insufficient cybersecurity measures protecting them provided is worrying. The healthcare sector has fallen victim to more cyberattacks than any other industry just a few years ago and only three out of four hospitals have a designated employee for addressing cybersecurity issues. Despite this lack in cybersecurity measures, the healthcare sector is rapidly adopting new technologies such as networked medical devices while still operating vulnerable legacy systems such as Windows XP.

Cybersecurity is a crucial part of today's financial security, especially in regards to the rapid digitization the industry sector has seen. With the large amounts of capital pro-

cessed in the financial sector, large damages follow after successful cybersecurity attacks with companies facing an average annual cost of cybercrime of \$ 18.5 million in 2019. Out of the three industry sectors which this report inquires about, the telecommunications sector seems to be the most successful when it comes to minimizing damages caused by cybercrime. Companies in this sector had an average annual cost of cybercrime of \$ 9.21 million in 2018. Despite relatively low costs of cybercrime, the threat landscape the sector faces is vast and it is not just limited to the cyberspace. The recent rollout of 5G technology and bans of Chinese technology companies shed light on how communications technology security is not only dependent on the system and its users, but also on the supply chain providing and maintaining it.

Furthermore, the cybersecurity issues which small- to mid-size enterprises and multinational enterprises face are quite similar, although data breaches are twice as common in larger companies as in smaller companies. The big difference lies in the ability of SMEs and MNEs to tackle these issues with a large technology gap separating the two company types. Despite technological advantages for larger firms, both MNEs and SMEs face challenges when it comes to recruiting new cybersecurity talent, with the labor market for such experts being scarce. As a result, both MNEs and SMEs ought to utilize currently underutilized up- or re-skilling strategies to fill the skills-gap. It should be also noted that SMEs are getting targeted more and more often by malicious actors whose goal is to enter a supply chain's information system through the weakest link.

Lastly, the report gives recommendations for companies in terms of cybersecurity investments. It is suggested to revisit and check the security level and relevance of the company's network infrastructures, processes, compliance and to configure the network securely. Boundary defense such as firewalls, network monitoring, proxies and multi-factor authentication should also be in place. Hardware and software should both have timely updates. Furthermore, a company must protect its data by controlling access to the sensitive information. If possible, it is recommended to avoid collecting customer data. Moreover, companies should increase cybersecurity staff and educate all employees about cybersecurity through security awareness and training programs. Also, a firm should make sure it can detect cyber attacks quickly. When they happen, it is suggested that they proactively disclose this information to customers and other stakeholders to avoid bad publicity and financial damage in the future.

As the world is undergoing a COVID-19 crisis, cyber attacks are on the rise. In these special circumstances, especially when companies are exposed to more cyber threats as many employees around the globe now suddenly have to work remotely, it is proposed that a four-fold approach is to be implemented. Firstly, focus on critical operating needs. The other needs can wait until the situation becomes more stable for businesses. Secondly, test plans for managing security and technology risks. Thirdly, monitor for new cyber threats from different angles: in collaboration tools, networks, employees and endpoints. Fourthly, balance protection with business continuity.

The world is changing fast, and hackers are most often faster at adapting to it than companies are. Therefore it is decisive for business continuity that cybersecurity divisions are given enough freedom in allocating resources to what they deem is most critical for business.

Bibliography

- [1] S. Morgan: 2019 Official Annual Cybercrime Report; Annual Report, December, 2018, pp. 1-12, <https://www.herjavecgroup.com/wp-content/uploads/2018/12/CV-HG-2019-Official-Annual-Cybercrime-Report.pdf>.
- [2] Capgemini Invent, European Digital SME Alliance, Executive Agency for Small and Medium-sized Enterprises (European Commission), Technopolis: Skills for SMEs – Cybersecurity, Internet of things and big data for small and medium-sized enterprises; public report, December, 2019, pp. 1-25, <https://op.europa.eu/en/publication-detail/-/publication/82aa7f66-67fd-11ea-b735-01aa75ed71a1/language-en>.
- [3] S. Morgan: Cybercrime To Cost The World \$10.5 Trillion Annually By 2025, 2020, <https://cybersecurityventures.com/hackerpocalypse-cybercrime-report-2016/>, last visit December, 2020.
- [4] Gartner: Gartner Says Worldwide Wearable Device Sales to Grow 17 Percent in 2017, 2017, <https://www.gartner.com/en/newsroom/press-releases/2017-08-24-gartner-says-worldwide-wearable-device-sales-to-grow-17-percent-in-2017>, last visit December, 2020.
- [5] J. Saade: Thoughts on IoT and Finance, 2017, https://www.huffpost.com/entry/thoughts-on-iot-and-finan_b_11298656, last visit December, 2020.
- [6] The Economic Times: Definition of 'Hacking', <https://economictimes.indiatimes.com/definition/hacking>, last visit December, 2020.
- [7] I. Belecic: What Is Hacking? Everything You Need to Know, 2020, <https://www.avg.com/en/signal/what-is-hacking>, last visit December, 2020.
- [8] Ponemon Institute LLC & Accenture: Ninth Annual Cost of Cybercrime Study - Unlocking the Value of Improved Cybersecurity Protection; Corporate report, 2019, pp. 1-23, https://www.accenture.com/_acnmedia/PDF-96/Accenture-2019-Cost-of-Cybercrime-Study-Final.pdf.
- [9] M. Franco, B. Rodrigues, E. Scheid, A. Jacobs, C. Killer, L. Granville, B. Stiller: SecBot: a Business-Driven Conversational Agent for Cybersecurity Planning and Management; 16th International Conference on Network and Service Management (CNSM 2020), Izmir, Turkey, 2020, pp. 1-7.
- [10] S. Weisman: What is a distributed denial of service attack (DDoS) and what can you do about them?, 2020, <https://us.norton.com/internetsecurity-emerging-threats-what-is-a-ddos-attack-30sectech-by-norton.html>, last visit December, 2020.

- [11] Kaspersky: IT security economics in 2019, 2019, https://go.kaspersky.com/rs/802-IJN-240/images/GL_Kaspersky_Report-IT-Security-Economics_report_2019.pdf, last visit December, 2020.
- [12] Norton: What is malware and how can we prevent it?, <https://us.norton.com/internetsecurity-malware.html>, last visit December, 2020.
- [13] K. Porter: What is phishing? How to recognize and avoid phishing scams, 2020, <https://us.norton.com/internetsecurity-online-scams-what-is-phishing.html>, last visit December, 2020.
- [14] B. Jefferson: The 15 Most Common Types of Cyber Attacks, 2020, <https://www.lepide.com/blog/the-15-most-common-types-of-cyber-attacks/>, last visit December, 2020.
- [15] Cisco: What Are the Most Common Cyber Attacks?, <https://www.cisco.com/c/en/us/products/security/common-cyberattacks.html>, last visit December, 2020.
- [16] M. F. Franco, B. Rodrigues, B. Stiller: MENTOR: The Design and Evaluation of a Protection Services Recommender System; 15th International Conference on Network and Service Management (CNSM 2019), Halifax, NS, Canada, 2019, pp. 1-7.
- [17] M. Hammell: ThreatExchange: Sharing for a safer internet, <https://www.facebook.com/notes/protect-the-graph/threatexchange-sharing-for-a-safer-internet/1566584370248375>, last visit December, 2020.
- [18] M. Husak, J. Komarkova, E. Bou-Harb, P. Celeda: Survey of Attack Projection, Prediction, and Forecasting in Cyber Security; IEEE Communications Surveys & Tutorials, 21(1), 2019, pp.640-660.
- [19] McAfee: What Is DLP and How Does It Work?, 2020, <https://www.mcafee.com/enterprise/en-us/security-awareness/data-protection/how-data-loss-prevention-dlp-technology-works.html>, last visit December, 2020.
- [20] A. Buecker, P. Andreas, S. Paisley: Understanding IT Perimeter Security, 2009, <https://www.redbooks.ibm.com/redpapers/pdfs/redp4397.pdf>, last visit December, 2020.
- [21] L. Coventry, D. Branley: Cybersecurity in healthcare: A narrative review of trends, threats and ways forward; *Maturitas*, 2018, pp. 48-52.
- [22] C. Thyagarajan, S. Suresh, N. Sathish, Dr. S. Suthir: A Typical Analysis And Survey On Healthcare Cyber Security; *International Journal of Scientific & Technology Research*, March, 2020, pp. 1-5.
- [23] V. Wang, H. Nnaji, J. Jung: Internet banking in Nigeria: Cyber security breaches, practices and capability; *International Journal of Law, Crime and Justice* 62, September, 2020.
- [24] Firmex: The 10 Most Expensive Data Breaches in Corporate History, <https://www.firmex.com/resources/blog/the-10-most-expensive-data-breaches-in-corporate-history/>, last visit December, 2020.
- [25] VirtualArmour: The 8 Most Expensive Cyberattacks of 2019, 2020, <https://www.virtualarmour.com/the-8-most-expensive-cyberattacks-of-2019/>, last visit December, 2020.

- [26] R. McLean: A hacker gained access to 100 million Capital One credit card applications and accounts, 2019, <https://edition.cnn.com/2019/07/29/business/capital-one-data-breach/index.html>, last visit December, 2020.
- [27] Stealth Labs: Cybersecurity in Financial Sector: 8 Important Facts and Statistics, 2020, <https://www.stealthlabs.com/blog/cybersecurity-in-financial-sector-8-important-facts-and-statistics/>, last visit December, 2020.
- [28] GSMA: Mobile Telecommunications Security Threat Landscape; Corporate report, January, 2020, pp. 1-28, <https://www.gsma.com/security/wp-content/uploads/2020/02/2020-SECURITY-THREAT-LANDSCAPE-REPORT-FINAL.pdf>.
- [29] AT&T Business: AT&T Cybersecurity Insights Report - Security at the Speed of 5G, 2019.
- [30] M. Shoebridge: Chinese Cyber Espionage and the National Security Risks Huawei Poses to 5G Networks; Macdonald-Laurier Institute for Public Policy, November, 2018, pp. 1-13, https://macdonaldlaurier.ca/files/pdf/MLICommentary_Nov2018_Shoebridge_Fweb.pdf.
- [31] L. Kelion: Huawei 5G kit must be removed from UK by 2027, 2020, <https://www.bbc.com/news/technology-53403793>, last visit December, 2020.
- [32] Verizon: 2020 Data Breach Investigations Report, 2020, <https://enterprise.verizon.com/resources/reports/dbir/>, last visit December, 2020.
- [33] J. Bernard, D. Golden, M. Nicholson: Reshaping the cybersecurity landscape, 2020, <https://www2.deloitte.com/us/en/insights/industry/financial-services/cybersecurity-maturity-financial-institutions-cyber-risk.html>, last visit December, 2020.
- [34] M. Lobel: Communications Review - Security risks and responses in an evolving telecommunications industry; PwC, 2013, pp. 1-9, <https://www.pwc.com/gx/en/communications/publications/communications-review/assets/cyber-telecom-security.pdf>.
- [35] H. Aver: Cybersecurity economics, 2020, <https://www.kaspersky.com/blog/it-security-economics-2020-main/37205/>, last visit December, 2020.
- [36] Kaspersky: IT Security Economics 2020: Part 2, <https://www.kaspersky.com/blog/it-security-economics-2020-part-2/>, last visit December, 2020.
- [37] Kaspersky: IT Security Calculator, 2020, <https://calculator.kaspersky.com/>, last visit December, 2020.
- [38] D. Raywood: The Short-Term Impact of #COVID19 on the Cybersecurity Industry, 2020, <https://www.infosecurity-magazine.com/news-features/short-impact-covid19-industry/>, last visit December, 2020.
- [39] J. Boehm, J. Kaplan, N. Sportsman: Cybersecurity's dual mission during the coronavirus crisis, 2020, <https://www.mckinsey.com/business-functions/risk/our-insights/cybersecuritys-dual-mission-during-the-coronavirus-crisis>, last visit December, 2020.

- [40] IBM Study: Security Response Planning on the Rise, But Containing Attacks Remains an Issue; <https://newsroom.ibm.com/2020-06-30-IBM-Study-Security-Response-Planning-on-the-Rise-But-Containing-Attacks-Remains-an-Issue>, last visit December, 2020.

