University of Zurich[UZH]

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

Communication Systems XIV

TECHNICAL REPORT – No. IFI-2021.02

June 2021

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

ifi

B. Stiller et al. (Eds.): Communication Systems XIV
Technical Report No. IFI-2021.02, 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 spring term FS 2021 a new instance of the Communication Systems seminar has been prepared and students as well as supervisors worked on this topic.
The areas of communication systems include among others wired and wireless network technologies, various network protocols, network management, Quality-of-Service (QoS) provisioning, mobility, security aspects, peer-to-peer systems, multimedia communication, and manifold applications, determining important parts of future networks. Therefore, this year's seminar addressed such areas in more depth. The understanding and clear identification of problems in technical and organizational terms have been prepared and challenges as well as weaknesses of existing approaches have been addressed. All talks in this seminar provide a systematic approach to judge dedicated pieces of systems or proposals and their suitability.

Content

This new edition of the seminar entitled "Communication Systems XIV" discusses a number of selected topics in the area of computer networks and communication systems.
The first talk, Talk 1, highlights the technical foundations of Electronic Identity as well as the risks and opportunities involved. Further, selected implementations of electronic identity systems are discussed, namely the cases of Estonia, Switzerland, and the European Union. Talk 3 gives an overview of the use cases and key technologies of 5G, bringing light on various cybersecurity aspects. Talk 4 provides an overview of 6G networks, highlighting technologies that enable their implementation, such as machine learning, holographic communications, and ubiquitous connectivity. Talk 7 is a survey on quantum communication networks, introducing quantum computing and quantum Internet concepts, and providing an overview of the state-of-the-art technology of quantum communication. Talk 8 presents the InterPlanetary File System (IPFS), focusing on its underlying technologies. Furthermore, this talk shows valuable use cases, outlines advantages and disadvantages, and identifies factors of key importance for the future success or failure of the IPFS. Talk 9 provides an overview of Machine Learning (ML) for network management, describing commonly used ML methods and presenting how ML methods can be used with novel network architectures. Talk 10 is a survey on anonymization techniques for Blockchain transactions. This talk presents privacy-focused projects, analyzing technical factors, algorithms, and economic factors. Talk 11 provides a literature

review on Software-Defined Networking (SDN) and cybersecurity, analyzing specific attack schemes and vulnerabilities on SDN and highlighting mitigation strategies. Finally, Talk 12 describes the state-of-the-art research on detection and analysis for Fake News, explaining the core principles of fake news detection and illustrating their limitations and open challenges.

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, technology architectures and functionality, 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, Rafael Ribeiro, Alberto Huertas, and Burkhard Stiller. In particular, many thanks are addressed to 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

Chapter 1

Electronic Identity: Risk or Opportunity for Digital Authentication?

Niels Kübler

Nowadays, many countries provide E-Government systems for their residents. These systems often provide access to sensitive data such as criminal history records. The sensitivity of the data involved necessitates the use of trusted authentication. Electronic identity serves precisely this purpose as it provides authentication using government-guaranteed identities. The paper highlights the technical foundations of Electronic Identity as well as the risks and opportunities involved. This includes models for identity management but also current research topics such as self-sovereign identity. Furthermore, selected implementations of electronic identity systems are discussed, namely the cases of Estonia, Switzerland, and the European Union.

Contents

1.1 Introduction

Numerous countries use electronic identities to provide E-Government services to their residents. The fundamental problem is: how to prove that an Internet user is whom she claims to be? While this question can be answered by using public Identity Providers (IdP) (*e.g.,* Google, Microsoft, or Facebook), it misses an important point: E-Government applications provide access to highly confidential data, which require a trusted identity, in this case, identities that the government guarantees. In the context of this paper, Electronic Identity (EID) is referred to as a government-trusted identity for use on the Internet.
This paper is structures as follows: *Basics of Identity Management* provides an overview over electronic identities in the form of Identity Management (IdM) models, which is the basis for the discussion of *Protocols and Implementations*. The section *Electronic Identity Systems* shows how the protocols discussed in the approaches are used in the real world, namely in Estonia, Switzerland, and the European Union. Furthermore, section *Evaluation and Discussion* shows opportunities, risks and political debates that come with EIDs. Finally, a summary of the contents and conclusions are provided.

1.2 Basics of Identity Management

This section covers the conceptual foundations of IdP systems. Basic definitions, and an overview of relevant IdM models are presented, including introducing the concept of self-sovereign identity, which is the subject of current research.

1.2.1 Definitions

The upcoming sections of this paper require a common understanding of the terminology in order to avoid misunderstandings. Therefore, the following definitions are important:

- [1] defines **authentication** as "[...] the process of verifying an entity's identity, given its credentials. The entity could be in the form of a person, a computer, a device, a group of network computers, etc."
- **Authorization** refers to the process of obtaining information about what a user is allowed to do [5].
- **Single sign-on** (SSO) is an approach that "[...] permits users to log in once and to access multiple services by delegating user authentication from service providers to authentication services" [2].

1.2.2 Identity Management Models

In order to discuss EID implementations, it is essential to understand the basic models of IdM. These models define different architectures and properties, which are explained in the following paragraphs.
Isolated Identity is a widely used model for IdM, its basic idea being that a service provider is also the IdP at the same time [4]. The consequence is that a user has to use and manage a different set of credentials for each service she wants to use [3]. From a technical standpoint, this makes sense because the service providers can only accept identities issued by themselves. For the users, this leads to a trade-off between security and usability [3]: If a user uses different credentials (*e.g.,* passwords) for each service, security aspects are fulfilled, but it is very inconvenient to use. If a user uses the same credentials for different services, it might be convenient to use, but security would suffer.

The problem with using the same credentials is that if your password is exposed (*e.g.,* due to a provider leak or cyber-attack), all of the other identities using the same credentials are also endangered. Due to the isolation of identities, it is not possible to provide a single sign-on experience with this model [3; 4].
With the **Centralized Identity** model, services and identities are not provided by the same entity. Instead, IdM is provided by a central IdP, which all service providers use. With this model, it is possible to provide an SSO experience for all service providers attached to the same central IdP. This is already an improvement with regards to credential management compared to the isolated model. Centralized identities are often used inside organizations, where a central IdP can be easily established [3; 4].
The idea of **Federated Identity** is to have multiple IdPs that recognize each other's issued identities. As in the case of centralized IdM, IdPs are usually separated from the service providers, but it is not a strict requirement. Federated IdM allows to provide SSO functionality: As a result, users can use one set of credentials to access a variety of services [3; 4]. Federated identities allow use-cases in a broader range, as they can be used both inside and beyond the borders of an organization [3].
Self-Sovereign Identity (SSI) is an identity concept which aims to give the user complete control over her data [6]. It differs from conventional IdM models in that it is a decentralized model, which means there is no single point in the system that could expose all the personal data at once [7]. The decentralization is achieved by implementing self-sovereign identity based on blockchains or distributed ledgers [6; 7], which means that it relies on decentralized public-key infrastructure and public-key cryptography [5]. An example of an SSI implementation is discussed in section 1.3.2.

1.2.3 Discussion

In the context of EID, the aforementioned IdM models are not equally feasible for implementing an EID system. This section briefly discusses the suitability of each model for implementing an EID system. Furthermore, exemplary scenarios a provided.
Isolated identity is not suitable for implementing an EID system, as it is neither capable of providing SSO, nor different identity and service providers know and accept each other's identities. These properties would render such a system useless, as for each service, there would have to be a separate identity issued [3; 4]. This leads to problems regarding user acceptance: An EID system using isolated identities would most likely not be very well perceived by the users, as they would have to manage yet another set of credentials.
Centralized identity fits the needs of EID systems better, but introduces a single point of failure to the system [3]. An example for a centralized IdM would be the scenario where a government is the single IdP in the system, which is actually implemented in some countries. The problem with centralized IdM is that one single entity stores large amounts of user identities, which makes them an attractive target for cybercriminals [5]. Clearly, data breaches concerning identifying data have to be avoided, mainly if this data can be used to access E-Government and E-Commerce services.
Federated identity is able to provide SSO, as it is also the case for the centralized approach, but without the need for one central IdP [3]. Instead, multiple IdPs exist and they accept each other's identities. An example for a federated EID system could be a government delegating identity provision to third-party IdPs, which would implement and operate the EID system. Federated identity still has some downsides [5]: IdPs are still an attractive target for cyber criminals, because they store large amounts of user identities, but this is already an improvement over centralized identity systems. Also, privacy concerns have been issued, as IdPs are able to track the logins of their users.
Self-sovereign identity is able to address some of the problems of conventional (centralized) IdM models. It does so by abandoning the concept of accounts and using peer-to-peer

relationships based on blockchains and distributed ledgers instead [5]. Consequently, SSI can provide better data protection due to its decentralized nature, where the user has control over her own data. While this model does not require IdPs, it is still necessary that a trusted entity asserts identities, as this is a requirement for other peers trusting an identity [5]. An example for an EID system based on SSI could be a system as described by [9], where public attributes are stored on a public ledger, while private attributes are stored on a private ledger. Section 1.3.2.1 discusses this example in more detail.

1.3 Protocols and Implementations

IdM models on their own are not enough to build an EID system. The models discussed earlier act as a conceptual basis for the protocols in this section. Two kinds of protocols are discussed, namely conventional protocols, which rely on centralized protocols and decentralized protocols.

1.3.1 Conventional Implementations

This section presents two protocols commonly used on the Web in general and in EID systems. The term conventional is used in this case to refer to the centralized nature of the presented protocols.

1.3.1.1 OAuth

OAuth is an authorization protocol that lets web services communicate with each other on behalf of the user [11]. The OAuth protocol has the following actors, who are shortly explained [12]:

- **Client**: Application that wants to access resources on the resource server on behalf of the user.
- **Resource Owner**: Usually the end-user owning the resources on the resource server.
- **Authorization Server**: Authenticates the end-user in order to issue an access token for accessing the desired resource on the resource server (authorization).
- **Resource Server**: Server holding resources, which can be requested by clients with an access token.

The message flow of the OAuth protocol (depicted in Figure 1.1) works as follows [12]:

1. The client application sends an authorization request to the resource owner.
2. The resource owner then grants authorization and returns the response back to the client.
3. With the authorization grant, the client application then requests an access token from the authorization server.
4. If the authorization grant provided by the client is valid, the authorization server responds with an access token.
5. The client application can then use this access token to request a protected resource from the resource server.
6. If the access token is valid, the resource server returns the requested resource back to the client application.

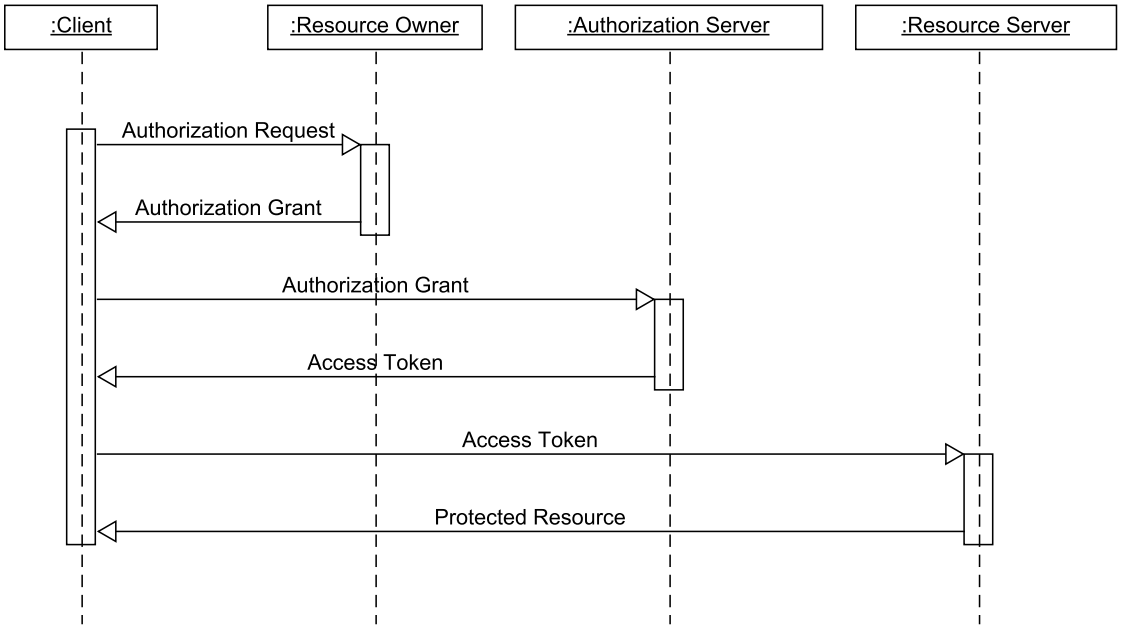

Figure 1.1: The OAuth flow based on [12]

1.3.1.2 OpenID Connect

OpenID Connect is a protocol that is based on OAuth, with the purpose of providing both authentication and authorization [13]. Authentication is provided by the OpenID protocol. It consists of the following actors [13]:

- **Client**: Also called Relying Party, is the application that wants to identify the user and access protected resources.
- **End-User**: Person using the application.
- **OpenID Provider**: OAuth server, with the task of authenticating and authorizing end-users.

The message flow of the OpenID Connect protocol (depicted in Figure 1.2) is similar to OAuth, but it additionally includes authentication (identifying) data [13]:

1. The client application sends an authentication request to the OpenID provider.
2. The OpenID provider then asks the End-User to authenticate and provide authorization.
3. If both authentication and authorization are successful, a response is sent back to the OpenID provider.
4. Now, the OpenID provider sends an authorization response back to the client application, containing an identity token as well as an access token.
5. With the access token, the client application can request user information (claims) from the OpenID provider.
6. The OpenID provider validates the access token and returns the requested user information to the client.

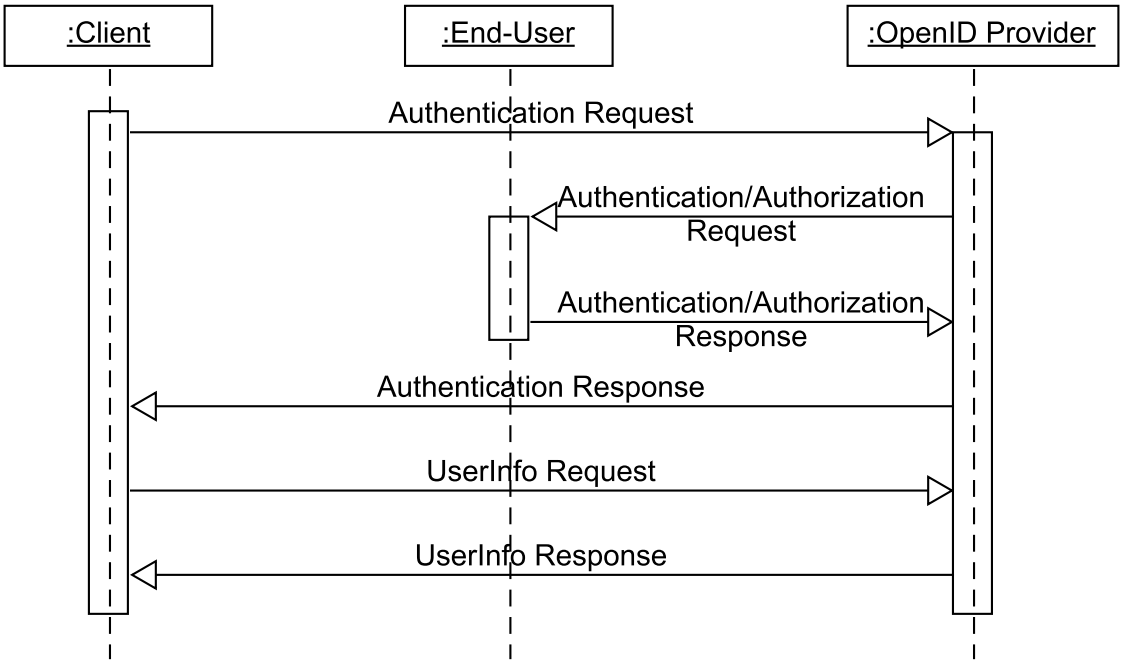

Figure 1.2: The OpenID Connect flow based on [13]

1.3.2 Decentralized Implementations

Decentralized identity systems promise to solve many of the issues of centralized identity systems. This section discusses the case of Sovrin, which is one of the most famous SSI implementations.

1.3.2.1 Sovrin

Sovrin is a decentralized identity implementation that is based on Hyperledger Indy, which is a public permissioned ledger [6], providing SSI [8]. Figure 1.3 shows an overview over the Sovrin architecture, which works as follows:
As there are no IdPs in this system, Sovrin relies on the concept of relationships between users and service providers (or between users for that matter). Each relation uses separate identifiers to prevent service providers from correlating user data. This is an improvement over centralized identity systems, where correlation cannot be prevented by users [9].
The ledger itself stores the public keys of all the peers, such that they can verify the signatures of the incoming messages from the peer they are communicating with. The messages are signed with the private key of the respective peer, which is not stored on the ledger as it is a secret that must be stored in a secure manner with the user [9].
Another important element of Sovrin are claims. Claims are identity properties of a peer, and they can be either self-asserted, or asserted by someone else. Service providers who want to verify a claim will usually ask for assertion by a trusted entity, *e.g.,* a government that guarantees for the claims to be true. Self-asserted claims can be problematic, because they are only asserted by the identity subject itself, which could claim anything it wants about itself [9]. In a way, self-asserted claims are comparable to self-signed certificates, which are not acceptable in situations where trust is crucial. In order for peers to verify the claims of an identity, Sovrin provides disclosure proofs [9]. With disclosure proofs, it is possible for identity subjects only to reveal claims that are really needed for the service to be consumed. This is called selective disclosure. An even more privacy-preserving

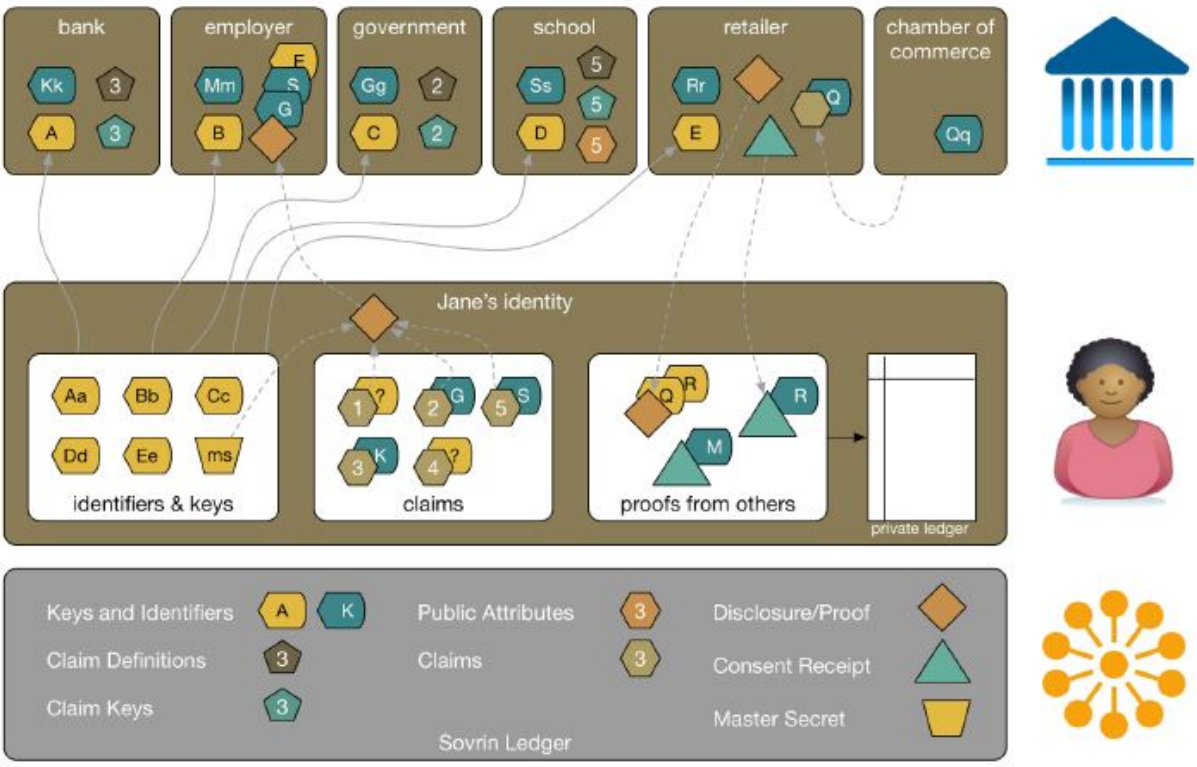

Figure 1.3: Sovrin overview [9]

method of selective disclosure is the usage of zero-knowledge proofs, which allows a user to proof, that her claim is valid without revealing its value [5].
From a technical standpoint, the Sovrin ledger has a set of actors and roles [10]: It consists of trustees, which are responsible for technical governance, which includes the approval of stewards. Stewards are responsible for operating network nodes, which can be validator nodes or observer nodes. The validator nodes run the consensus algorithm (Redundant Byzantine Fault Tolerance) in order to write transactions to the ledger, while the observer nodes are only capable of reading the ledger.

1.4 Electronic Identity Systems

EIDs are widely used across the world. This section focuses on the implementations in three regions: Estonia, Switzerland, and the European Union.

1.4.1 Estonia

Estonia was one of the first countries in the world to introduce a government-issued EID. After the fall of the Soviet Union, Estonia became independent, which meant that the existing identity documents were suddenly invalid. This meant that new identity documents (*i.e.,* passports, ID cards) had to be designed. During this design phase, they started to work on the concept of an EID in 1997, which lead up to the first EID being issued in 2002. The EID is integrated into the physical identity card [14].

1.4.1.1 Capabilities

The Estonian EID (eID) is a mandatory document for permanent residents of Estonia. It is based on the Personal Identification Code (PIC) from the Estonian Population register, which is a number that uniquely identifies a person. As it is based on the physical ID

card, the ID card contains two certificates: One for authentication and another for digital signatures. Both certificates require a separate PIN for usage. The eID is intended for both online, and real-world usage [15].
This hybrid between physical and digital identity enables many use cases [14; 15]: The primary area of application is in E-Government systems. Estonia provides a variety of services to their residents, such as their citizen portal, tax declarations but also access to their e-Health system and even remote Internet voting [14]. Another area of application are digital signatures: eID owners can user use their ID card for signing digital documents [14; 15]. On top of these online use cases, the eID can also be used in the physical world. This includes the classical use case of physical identification, but also involves using the ID card to travel within the European Union [15]. Furthermore, the eID can be used as a ticket for public transportation, as well as a proof of owning a driver's license [14].

1.4.1.2 Architecture

In the Estonian electronic identity management system (eIDMS) [14], the IdP is a government agency, namely the Ministry of Economic Affairs and Communications [15]. Certificate management and ID card manufacture are not done by the government, but by private companies, which are discussed in the upcoming paragraphs.
Issuance: Residents who need an eID contact the Police and Border Guard Board (PBGB) (formerly: Citizenship and Migration Bureau [16]) to fill in an application. The PBGB then requests Trüb AG Baltics to create a new ID card for the applicant, for which Trüb AG contacts the certificate authority AS Sertifitseerimiskeskus to order the pair of certificates for authentication and digital signature. When receiving the certificates, Trüb AG embeds the certificates in the ID card and sends it back to the PBGB. Note that the PIN codes are sent separately from the ID card. The final step for the PBGB is to hand over the ID card and the PINs to the applicant [14]. This process is shown in figure 1.4.
Usage: The usage of the EID involves two important processes, which provide alternatives for authentication [17]: This first process is to check whether the certificates used are on the certificate revocation list (CRL) (shown in figure 1.5). If they were on this list, it would mean that the certificates are invalid. The second process to check the validity of the certificate is by using the online certificate status protocol (OCSP) (shown in figure 1.6). With OCSP, the certificates can be validated in real-time, whereas the CRL approach works with a copy of the certificate revocation list periodically updated.

1.4.1.3 Stakeholders

The Estonian eIDMS consists of both governmental, and private actors.

- **Trüb Baltics AG** is responsible for manufacturing the ID cards. [14; 15]
- The **Ministry of Economic Affairs and Communications** develops the software and is also responsible for coordination the governmental IT system landscape. [14; 15]
- **AS Sertifitseerimiskeskus (SK)** is the certification authority of the eIDMS, which means that SK is responsible for managing and validating certificates. SK has also developed a software for applying and validating digital signatures [14; 15].
- The **Police and Border Guard Board (PBGB)** is the successor of the Citizenship and Migration Bureau (CMB) [16]. It is responsible for IdM, which also involves handling resident's applications for IDs [14; 15]

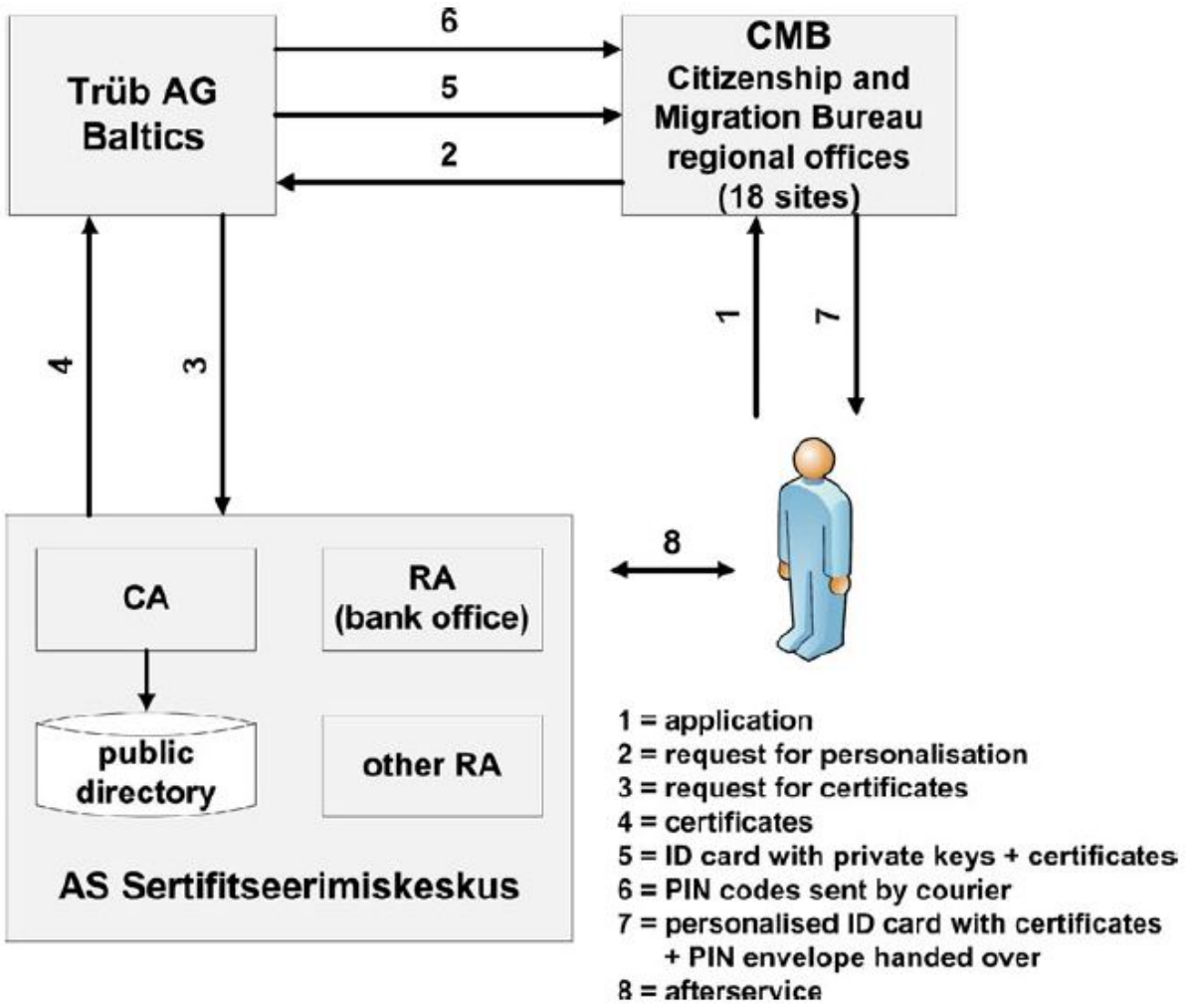

Figure 1.4: EID issuance in Estonia [14]

1.4.2 Switzerland

Switzerland introduced a law on electronic identities (BGEID) in 2019 [19]. The goal was to provide a legal basis for the development of an EID system, covering important aspects such as the responsibilities of the stakeholders and data protection. In this case, the government is not the IdP. Instead, this role is delegated to private companies and institutions [18].

1.4.2.1 Capabilities

In Switzerland, the EID serves the purpose of providing a government-guaranteed EID for use on the Internet, especially for E-Government and E-Commerce applications [28]. This means that it does not serve as a travel document (*e.g.,* passport, id card) in the real world [24]. The Swiss EID architecture is not tied to a specific technology, instead, the law explicitly states technology neutrality [19]. This is an important property, as it allows for future innovations and reduces the risk of being stuck with old technology.
EIDs are issued in three different security levels, which are *low*, *substantial*, and *high* (*cf.* Table 1.1). With increasing level of security, an EID contains incrementally more personal data and validation is required more often. Applications that require the security level *low* must also offer another method of authentication for users who do not want to have an EID, which is important to keep the EID on a voluntary basis (nobody is forced to have an EID) [19].

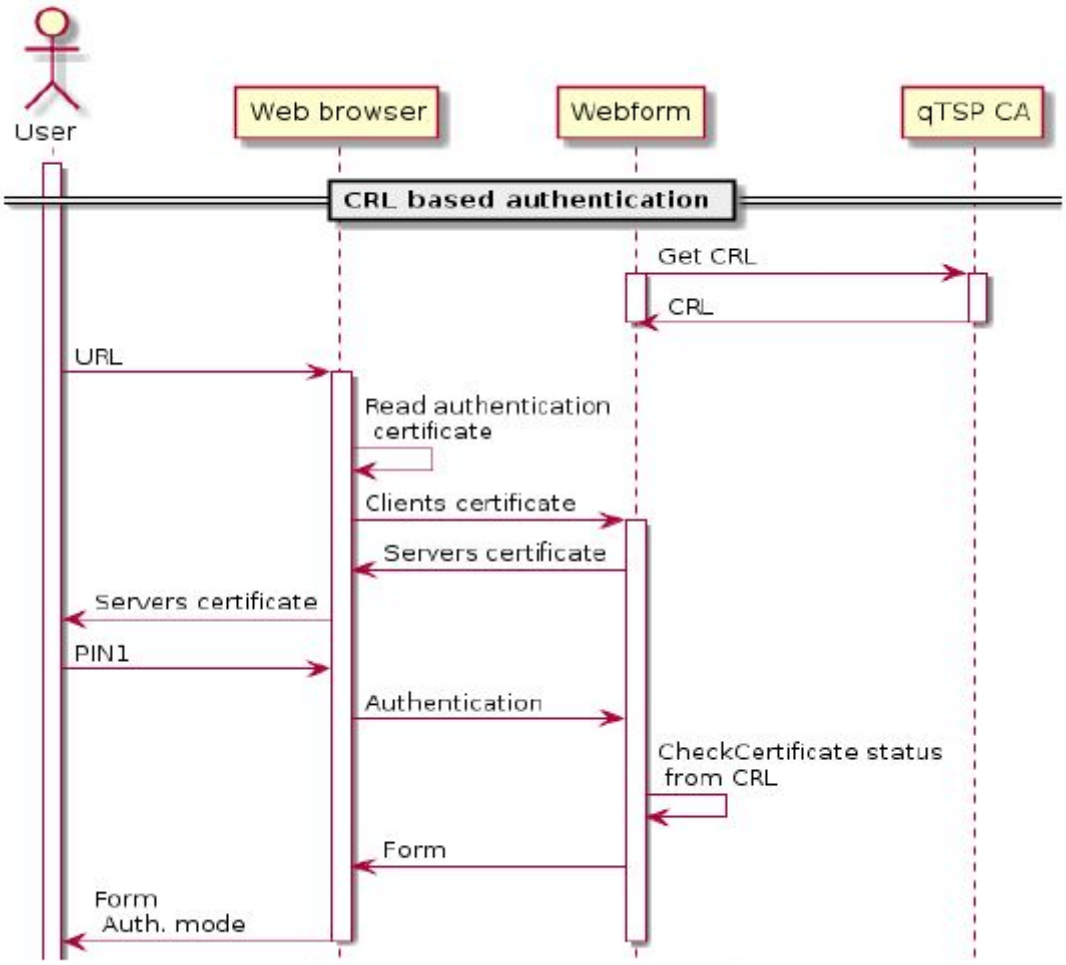

Figure 1.5: EID usage with CRL in Estonia [17]

On a technical level, IdPs are obliged to accept identities issued by other IdPs: This requirement is stated in the law, which says that interoperability between IdPs must be achieved [19].

1.4.2.2 Architecture

The basic idea of the Swiss EID architecture is that third parties (companies, cantons, or municipalities) act as IdPs. The government argues that identity provision could be much more efficiently done by private actors instead of the government itself, leading to a faster time-to-market while being audited and supervised by governmental instances [24]. This section dives into the details of issuing and using electronic identities within the Swiss concept.

Issuance: In order to apply for an EID, residents contact an IdP. They then have to show a physical identity document (*e.g.,* passport or id card) to the IdP to identify themselves. If the identification is successful, the IdP then contacts the information system of the Federal Office of Police (fedpol) and requests the transmission of identity attributes of the person applying for the EID. The fedpol information system sends back the requested identity attributes, with which the IdP can finally issue the EID to the resident [28]. This process is depicted in Figure 1.7.

Usage: When an EID owner wants to use a specific online service (*e.g.,* E-Government application) which is offered by a service provider, the user contacts the service provider by using its software. For the service provider to be able to grant access to the application, it asks the IdP for user identification. The IdP then asks the user to identify using the EID

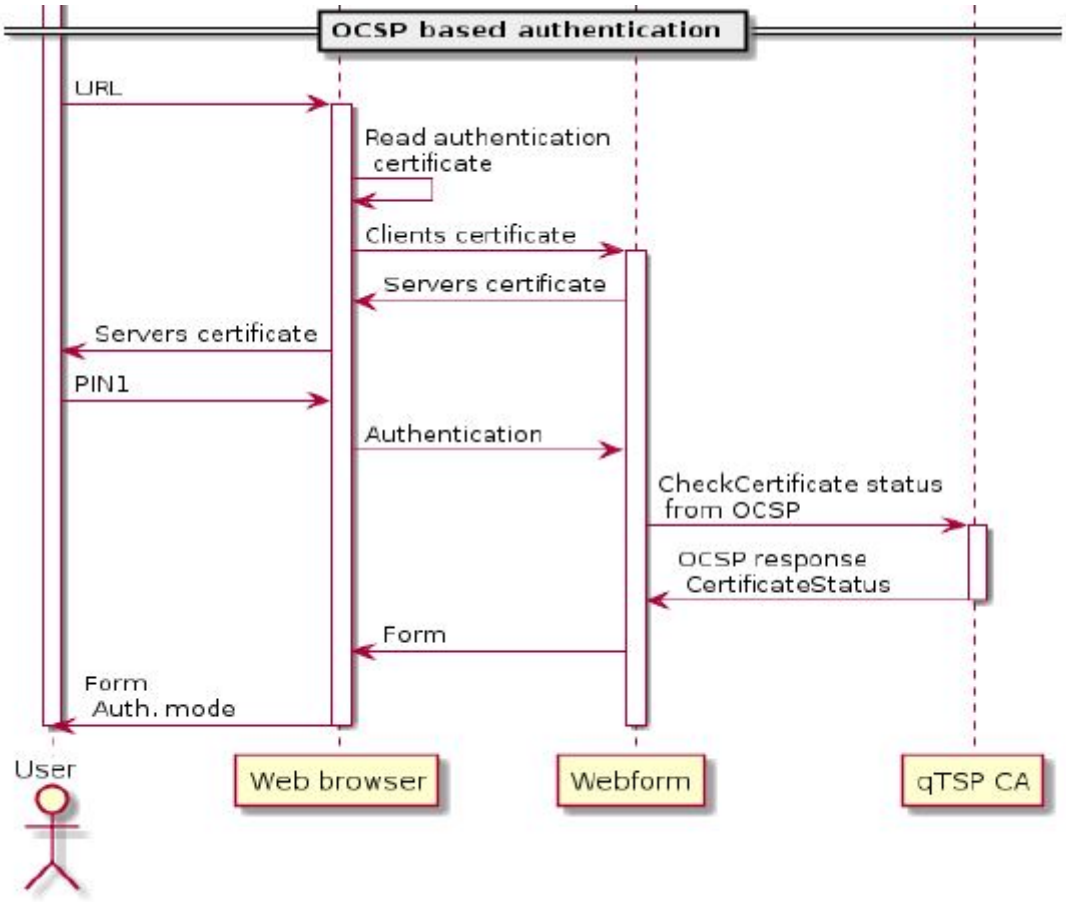

Figure 1.6: EID usage with OCSP in Estonia [17]

and to give consent to transferring identity attributes from the IdP to the service provider. The IdP confirms the identity of the user to the service provider and also includes the requested identity attributes. After this step, the service provider grants access to the user to use its service [28]. This process is depicted in Figure 1.8, which also shows one of the governments responsibilities: The IdPs are supervised by the government, which is discussed in more detail in section 1.4.2.3.

1.4.2.3 Stakeholders

The law on electronic identities (BGEID) [19] defines the rights and responsibilities of the stakeholders involved in the EID system.

- **Swiss Residents**: Residents of Switzerland, which are the users of an EID system. They are the identity subjects, meaning that they can own an EID.
- **Eidgenössiche E-ID Kommission (EIDCOM)**: The EIDCOM is responsible for recognizing and supervising IdPs. This means that the EIDCOM decides on who becomes an IdP, and it can also revoke this permission later, in case an IdP is not complying with the law [19].
- **Bundesamt für Polizei (fedpol)**: Fedpol operates an information system aggregating personal identification data from other federal information systems. This information system acts as a data source for the IdPs, including validation of EIDs. In addition to the personal identification data, fedpol also stores protocols of EID usages and user consents [19].

Table 1.1: Security levels of Swiss EIDs [19]. Use cases annotated with (?) are assumptions.

Level	Validation Interval	Required Personal Data	Use Cases
Low	Annually	EID registration number Last name First name Date of birth	Age Verification [24] Address Validation [25]
Substantial	Quarterly	Low plus: Gender Place of birth Citizenship	Criminal history record [24] File tax return (?)
High	Weekly	Substantial plus: Photograph (face)	Entry in the land register (?) Accept an inheritance (?)

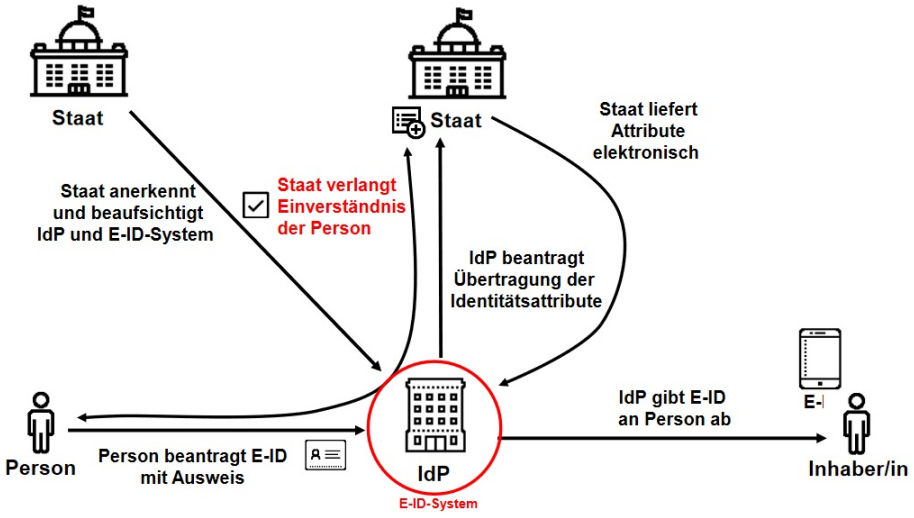

Figure 1.7: EID issuance in Switzerland [28]

- **Identity Providers**: In the Swiss EID concept, companies and organizations (*i.e.,* companies, cantons, municipalities) act as IdPs. IdPs are responsible for implementing and operating an EID system [19]. As of 2021, there are two well-known IdPs in Switzerland:
  - **SwissSign Group** and their product SwissID have been founded by major Swiss companies in 2017, with the intention to unify their digital identities. In 2020, EID functionality was implemented [20; 23]. Technology-wise, SwissID uses the OpenID Connect protocol to provide authentication and authorization [25].
  - **Procivis** is another well-known IdP for the Swiss EID. With its product eID+, Procivis has implemented an EID system for the canton of Schaffhausen and the city of Zug, which are two well-known examples [21]. With Procivis eID+, it is not explicitly known what the underlying authentication technology is, but one can assume that it is probably OpenID Connect, as they mention open authentication protocols in their blog [22].

1.4.3 European Union

The European Union does not provide its own EID system as such. Instead, they established the eIDAS regulation, which aims to provide interoperability between the national

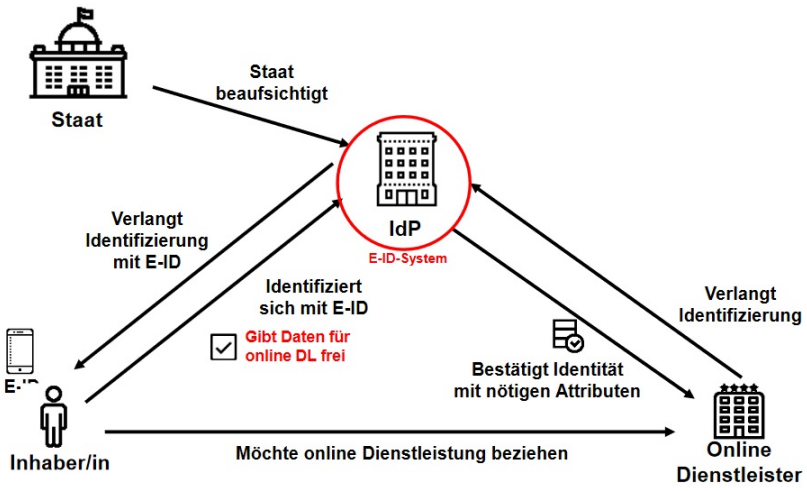

Figure 1.8: EID usage in Switzerland [28]

EID systems of the EU member states [26]. From an IdM standpoint, it is an identity federation (using the SAML protocol) between the IdPs of the EU member states [3].

1.4.3.1 Capabilities

The capability of this federation is to make it possible for the owners of any EU EID to use it for services in other EU member states. For that reason, the system is based on the idea that every country provides an eIDAS node. The eIDAS node is the element that enables identity federation between the member states, as all the other member states participating have to register with the eIDAS node of a country [3].

1.4.3.2 Architecture

Issuance: Because the eIDAS forms an identity federation, it does not issue any electronic identities. Instead, this is the responsibility of the EU member's national EID systems [3].
Usage: Figure 1.9 shows an example of an authentication process in the eIDAS system. If a user wants to use the services of a service provider that is not located in the country where her EID has been issued, the user can still log in with her EID. In this case, the user is asked to authenticate, which is done by redirecting the user to her own IdP. This is done by passing the authentication request to the eIDAS connector of the service providers country, which in turns forwards the request to the proxy service of the eIDAS node of the country where the EID has been issued. Optionally, this is also the moment where the user can consent on the identity attributes that are to be transferred back to the service provider. After authentication with the IdP, the response is returned to the eIDAS node proxy service which forwards the result to the first eIDAS node connector. The authentication result is then returned to the service provider, which completes the authentication process [3; 27].

1.4.3.3 Stakeholders

As the eIDAS system is integrating the EID systems of its member states, there is a board variety of stakeholders. This section does not focus on country specifics, but instead, stakeholders are looked at from a bird's-eye view [3]:

- **Residents of EU member states** are the users of the eIDAS system.

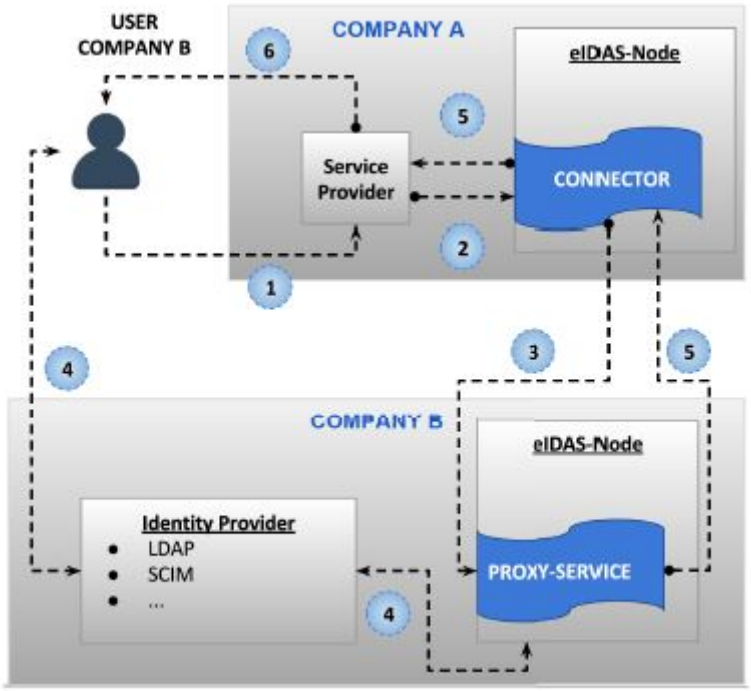

Figure 1.9: Identity federation in the European Union [3]

- **National EID IdPs** are responsible for the IdM of their citizens and residents. Their identity systems are connected to the country's eIDAS node.
- **Service providers** provide E-Governance and E-Commerce applications to EU residents. Authentication requests for foreign EID systems are forwarded to the service provider country's eIDAS node.

1.5 Evaluations and Discussion

EID systems provide both opportunities and risks. The balance between these two aspects is subject to intense political debates. Opportunities and risks are used by both proponents and opponents to form their arguments. This section discusses these aspects.

1.5.1 Opportunities

The use of electronic IdM offers a variety of potential benefits. These benefits especially affect the areas of E-Government and E-Commerce, which have been studied extensively in the studies by [3; 14; 28] as well as in the literature by [5].
The use of electronic IdM enables new and innovative use cases in E-Government and E-Commerce applications. One notable example for innovative use cases is shown by [14]. The EID system enabled the introduction of remote Internet voting in Estonia, which is a use case that requires strong authentication mechanisms due to its sensitive nature. In the area of E-Commerce applications, age and address verification can also benefit from using an EID [28]. In this example, the EID would both simplify the verification process (a user would not need to enter her data for every service provider), as well as improving its significance due to government-guaranteed identity attributes.
From a security standpoint, electronic IdM has the potential to improve the security of E-Commerce and E-Government applications [28]. According to [5], EID systems based on SSI have the advantage over centralized systems that they do not expose all the identity

data at once in case of a cyber attack. This is due to the decentralized nature of EID systems that are built according to the SSI paradigm.
Another important opportunity is increased usability: By having electronic IdM in place, SSO services can be provided to the residents of a country. This means, that they can (at least in theory) use one single set of credentials to access all the services they are interested in. Thus, it would solve the problem of cumbersome credential management [3].

1.5.2 Risks

Potential risks in the usage of electronic identities can be classified into three main categories: interoperability, data protection, and security. These risks have been covered in the context of political debates by [19; 30] as well as in the literature by [5].

1.5.2.1 Interoperability

Interoperability is primarily a concern for EID systems which consist of several IdPs, such as in the case of Switzerland [19]. But it can also be a concern when looking at multinational EID system, such as the eIDAS system [3]. The problem of lacking interoperability is that it might not suffice for users to have a single EID in order to access all the services they need. This defeats the purpose of EID systems and decreases their usability. In other words, lack of interoperability increases complexity in IdM and causes security issues, which are comparable to the isolated IdM models.
In the case of Switzerland, interoperability is accounted for in the BGEID [19], which prescribes interoperability between the EIDs of different IdPs. This means that Swiss IdPs have to access each other's issued identities, so that it does not matter for the user from which IdP an identity is issued.
With the European eIDAS federation, interoperability between the national EID systems is provided, while the implementation of national EID systems is left to the EU member countries [3].

1.5.2.2 Data Protection

Data protection is obviously a very important topic when talking about EID systems. In Switzerland, data protection concerns have been raised because of third party IdPs storing sensitive personal data [30]. People tend to distrust private organizations of storing their personal data in a safe manner, without using the data for their own purposes. This is a problem that could be solved by using a decentralized IdM system, following the SSI principle. By doing so, the user has full control over her own data [6], which eliminates the third parties storing the data. This is called privacy by design, as the system considers privacy aspects in its design from the beginning [5].

1.5.2.3 Security

Depending on the specific architecture of an EID system, security issues can arise. For centralized approaches, there is the problem of the single point of exposure (and failure). Having central databases with large amounts of personal user data makes an IdP a very attractive target for cyber attacks. This is another problem that could be solved by using decentralized IdM systems, such as Sovrin. In the case of Sovrin, the personal data would be stored by the user, which can selectively grant access to her data. No single entity stores all the personal data of the users, so there is no single point to attack from a hackers perspective [5].

1.5.3 Political Debate

This section focuses on the political debates in the aforementioned countries. Interestingly, the intensity of discussions varies largely between these regions.

1.5.3.1 Estonia

In Estonia, no debates about data privacy were held. There was one exception, concerning an incident where an LDAP directory was publicly available. The purpose of this LDAP was to hold the certificates of the eID system, which allowed everybody to find out personal details about other people [14].

1.5.3.2 Switzerland

In the run-up to the vote on the EID law on March 7th 2021 [18], both opponents and proponents prepared their arguments. These arguments are presented in this section, including the outcome of the vote.
Proponents argued that the EID law would provide a foundation for a secure and simple-to-use EID system. Simplicity would be reached because of SSO, which would help to reduce the problem of complex credential management. In their opinion, everybody would profit from the EID law, as it provides new opportunities for E-Government and E-Commerce applications. They also considered the acceptance of the new law necessary, because it would be a required basis for Switzerland to stay innovative and competitive. On top of this they state that it is not mandatory for anyone to have an EID if they do not want to [29].
Opponents expressed their concerns with regards to trust in private companies as IdPs. In their opinion, electronic identities should be issued by the government instead of private organizations. They also argued that data protection is a problem, because private companies would have access to very sensitive data. Furthermore, they mentioned that 8 out of 26 Swiss cantons were against the new EID law. Last but not least, the opponents were concerned that elderly people would de facto be forced into having an EID, because of lacking alternatives, which could be overwhelming for them [30].
The vote on March 7th 2021 turned out to provide a very clear result: 64.4% of the Swiss voting population and all the cantons voted against the EID law, preventing its introduction [31].

1.5.3.3 European Union

Similar to the case of Estonia, no indications of a public debate or controversy around the eIDAS regulation could be found. It is possible that this is due to eIDAS being a federation of the national EID systems and not a standalone EID system in itself.

1.5.4 Discussion

Comparing the progress of the Swiss EID system with the European Union, and Estonia in particular, it becomes clear that Switzerland is far behind in this regard. From a time perspective, Estonia had already reached 300'000 EID users by 2009 [14], which is 6 years before Switzerland even started working on a concept [23]. The outcome of the vote in Switzerland does not help to close the gap between Switzerland and its European neighbours, and it will be very important for Switzerland to come up with a new concept for their EID system as soon as possible.
Looking at the features of the aforementioned EID systems, it becomes clear that the Estonian solution has a wider range of use cases compared to the Swiss EID. While Swiss

EID is limited for online use [28], the Estonian EID can be used both online and in the real world [14].

1.6 Summary and Conclusion

Electronic identities are an intensely debated subject. While some countries such as Estonia already introduced their first EID systems in the early 2000s, other countries such as Switzerland are still in the process of defining their EID architecture. Opportunities arising from electronic identities include the digitalization of government services as well as increased security for E-Commerce applications. In the case of E-Government services, both the residents as well as the governments of the respective countries benefit from using systems based on electronic identities. The risks include lack of interoperability, data protection as well as security concerns. The balance between opportunities and risks heavily depends on the architecture and implementation chosen for an EID system. As a consequence, it is not possible to find a single generic answer to the question: "Electronic identity: risk or opportunity for digital authentication?"
EID will definitely gain importance over the following years, leading to a variety of innovative EID systems. The area of decentralized IdM based on blockchains and distributed ledgers offers many interesting topics for future research.

Bibliography

[1] Ebru Celikel Cankaya, "Authentication", in *Encyclopedia of Cryptography and Security*, Henk. C. A. van Tilborg and Sushil Jajodia, Eds. Boston, MA: Springer US, 2011, pp. 61-62.

[2] Victoria Beltran, "Characterization of web single sign-on protocols", *IEEE Communications Magazine*, Volume 54, No. 7, pp. 24-30, July 2016, doi: 10.1109/MCOM.2016.7514160.

[3] Jesus Carretero, Guillermo Izquierdo-Moreno, Mario Vasile-Cabezas, Javier Garcia-Blas, "Federated Identity Architecture of the European eID System", *IEEE Access*, Volume 6, pp. 75302-75326, November 2018, doi: 10.1109/ACCESS.2018.2882870.

[4] Audun Jøsang, Simon Pope: "User centric identity management", in *AusCERT Asia Pacific information technology security conference*, May 2005, p. 77, doi: 10.1.1.60.1563.

[5] Alex Preukschat, Drummond Reed, *SELF-SOVEREIGN IDENTITY: decentralized digital identity and verifiable credentials.* S.l.: O'REILLY MEDIA, 2021.

[6] Galia Kondova, Jörn Erbguth: "Self-Sovereign Identity on Public Blockchains and the GDPR", in *Proceedings of the 35th Annual ACM Symposium on Applied Computing*, New York, NY, USA, March 2020, pp. 342-345, doi: 10.1145/3341105.3374066.

[7] Asem Othman, John Callahan: "The Horcrux Protocol: A Method for Decentralized Biometric-based Self-sovereign Identity", in *2018 International Joint Conference on Neural Networks (IJCNN)*, July 2018, pp. 1-7, doi: 10.1109/IJCNN.2018.8489316.

[8] Andrew Tobin, Drummond Reed, "The inevitable rise of self-sovereign identity", *The Sovrin Foundation*, Volume 29, No. 2016, 2016.

[9] Phillip J. Windley, "How Sovrin Works - A Technical Guide from the Sovrin Foundation", 2016, Accessed: March 15, 2021. [Online]. Available: `https://sovrin.org/wp-content/uploads/2018/03/How-Sovrin-Works.pdf`.

[10] Drummond Reed, Jason Law, Daniel Hardman: *The technical foundations of sovrin*, Technical report, Sovrin, 2016.

[11] Barry Leiba, "OAuth Web Authorization Protocol", *IEEE Internet Comput.*, Volume 16, No. 1, pp. 74-77, January 2012, doi: 10.1109/MIC.2012.11.

[12] Dick Hardt, *The OAuth 2.0 Authorization Framework*, October 2012. `https://tools.ietf.org/html/rfc6749` (accessed March 13, 2021).

[13] Nat Sakimura, John Bradley, Michael B. Jones, Breno De Medeiros, and Chuck Mortimore, *Final: OpenID Connect Core 1.0*, November 2014. `https://openid.net/specs/openid-connect-core-1_0-final.html` (accessed February 27, 2021).

[14] Tarvi Martens, "Electronic identity management in Estonia between market and state governance", *Identity in the Information Society*, Volume 3, No. 1, pp. 213-233, July 2010, doi: 10.1007/s12394-010-0044-0.

[15] Ana Milena Aguilar Rivera, Kristjan Vassil, "Estonia: A Successfully Integrated Population-Registration and Identity Management System." World Bank, 2015, Accessed: March 22, 2021. [Online]. Available: https://openknowledge.worldbank.org/bitstream/handle/10986/28077/115147-WP-EstoniaIDPopregistryIDcasestudyNovweb-PUBLIC.pdf

[16] *Police and Border Guard Board - in cooperation we create security* [Online]. Available: https://www2.politsei.ee/en/organisatsioon/ (accessed March 22, 2021).

[17] *Estonian eID scheme: ID card - Technical specifications and procedures for assurance level high for electronic identification*, February 27, 2018, Accessed: March 22, 2021. [Online]. Available: https://ec.europa.eu/cefdigital/wiki/download/attachments/62885749/EE%20eID%20LoA%20mapping%20-%20ID%20card.pdf.

[18] *Elektronische Identität: das E-ID-Gesetz* [Online]. Available: https://www.ejpd.admin.ch/ejpd/de/home/themen/abstimmungen/bgeid.html (accessed February 27, 2021).

[19] *BBl 2019 6567 - Bundesgesetz über elektronische Identifizierungsdienste (E-ID-Gesetz, BGEID).* [Online]. Available: https://www.fedlex.admin.ch/eli/fga/2019/2311/de (accessed February 26, 2021).

[20] *Home | SwissID.* [Online]. Available: https://www.swissid.ch/ (accessed February 27, 2021).

[21] *Procivis eID+ | Die Smart-Government-Lösung.* [Online]. Available: https://www.procivis.ch/eid (accessed February 26, 2021).

[22] Adithya Pradeep, *eID - What is the role of the state? Lessons from around the world*, March 11, 2020. [Online]. Available: https://www.procivis.ch/post/eid-what-is-the-role-of-the-state-lessons-from-around-the-world (accessed March 14, 2021).

[23] PWC: *Digital identity Your key to unlock the digital transformation.* Accessed: February 27, 2021. [Online]. Available: https://www.pwc.ch/en/publications/2019/digital-identity-whitepaper-web.pdf.

[24] *Sechs Fragen und sechs Antworten* [Online]. Available: https://www.bj.admin.ch/bj/de/home/staat/gesetzgebung/e-id/faq.html (accessed February 27, 2021).

[25] Florian Duss, Jonas Widmer, Ronald Maurhofer, Julian Zihlmann, *Die digitale Identität der Schweiz - zentral oder dezentral gespeichert? - Applied Data Science* [Online]. Available: https://sites.hslu.ch/applied-data-science/die-digitale-identitaet-der-schweiz-zentral-oder-dezentral-gespeichert/ (accessed February 26, 2021).

[26] Diana Berbecaru, Antonio Lioy: "On integration of academic attributes in the eIDAS infrastructure to support cross-border services", in *2018 22nd International Conference on System Theory, Control and Computing (ICSTCC)*, Sinaia, October 2018, pp. 691-696, doi: 10.1109/ICSTCC.2018.8540674.

[27] *eID Documentation - How does it work? The eIDAS solution*, March 23, 2021. [Online]. Available: https://ec.europa.eu/cefdigital/wiki/pages/viewpage.action?pageId=82773030.

[28] *E-ID: Eine staatlich anerkannte elektronische Identität für die Schweiz* Accessed: February 26, 2021. [Online]. Available: https://www.egovernment.ch/index.php/download_file/772/3380/.

[29] *e-ID Schweiz - sicher und einfach im Netz.* [Online]. Available: https://www.e-id.info/de/ (accessed February 27, 2021.).

[30] *NEIN zum E-ID-Gesetz* [Online]. Available: https://www.e-id-referendum.ch/ (accessed March 1, 2021).

[31] SRF Schweizer Radio und Fernsehen, *Eidgenössische Abstimmung - Das E-ID-Gesetz wird deutlich abgelehnt* [Online]. Available: https://www.srf.ch/news/abstimmungen/elektronische-identitaet/eidgenoessische-abstimmung-das-e-id-gesetz-wird-deutlich-abgelehnt (accessed March 20, 2021).

Chapter 3

Cybersecurity in Beyond 5G: Use Cases, Current Approaches, Trends, and Challenges

Tim Brunner

With the current deployment of the 5G network, it is now the proper time to look at its cybersecurity and most importantly, think about future trends for 6G and its cybersecurity. If we look at the existing and future security issues now and try to solve them until the implementation of 6G, the chance for a secure cellular network will be much higher than if we fail to discuss its security. With this in mind, this report aims to give an overview of the use cases and key technologies of 5G and the possible use cases and key technologies for 6G. This is done by introducing the previous generations of cellular network and briefly discussing the security aspects for all of them. After examining the first four generations of cellular network, this report will focus on the fifth and sixth generation and bring light on various cybersecurity aspects.

Contents

3.1 Introduction

In the current age of information exchange, the users' mobility becomes more important than ever. This lead to the rise of new wireless communication technology in a very short time. While mobile phones could only be used for calls 30 years ago, today's devices come with a vast amount of computational power and storage capacity and thus need to be addressed like a computer. This shift from fixed network structures to flexible and heterogeneous networks brings not only advantages. The interconnections of all the new devices pose a new threat to the security and privacy of modern society. As academia and industry are advancing to new technologies in the wireless communication area, it still has to face these threats from the older generations. On top of that, private and business users will have to deal with new sorts of cybersecurity attacks and privacy threats.
Cellular networks are currently in its fifth generation (also called 5G). The current trend is that approximately every ten years, a new generation gets introduced to the customer. If this trend still holds for the next generation, 6G, then researchers have enough time to think about its requirements, and more important, its security and privacy risks. If scientists and experts can lead a constructive discussion now about privacy and security in 6G, the possibility of success are higher, as the technology advances will only aggravate our problems if we do not act on them.
As of today, there are no papers that give an overview over both 5G and 6G networks use case and key technologies regarding their cybersecurity. Such an overview is needed to discuss current as well as future issues.
To achieve the goal of having a secure future of cellular networks, this work aims to give an overview of cybersecurity aspects of the older generations of cellular network (1G to 4G), introduces the reader to key aspects of the current generation (5G) and discusses different possibilities and predictions for the future (6G).
The remainder of this document is organized as follows. First, Section 3.2 introduces the previous generations of cellular network, from 1G to 4G. In each generation, a short overview of requirements and cybersecurity issues are given. The previous generations are followed by the current generation 5G. This generation is analysed in Section 3.3 from the use cases and key technologies perspectives. Both perspectives are first introduced and then discussed according to their cybersecurity issues. After the current generation, Section 3.4 presents the next generation of cellular networks by introducing potential use cases and key technologies for 6G. The cybersecurity issues are discussed after the introduction, as in the 5G case. Finally, some conclusions are provided in Section 3.5.

3.2 Historic background for modern mobile communication

As our current generation of communication systems build up on the previous generations, this section reviews the first generation of wireless networks and then continues with the generations two, three and four before discussing the current and future generations. Since the introduction of 1G in 1979, much work has been done in the area of mobile communication. Each Generation has new functionalities and better performance in regard of the previous generation.
To be able to discuss about security threats and privacy risks along the different generations, this section will briefly look at the key technologies, use cases and vulnerabilities of each generation.

3.2.1 1G

The very first generation of mobile communication was based on analog technology. It primarily supported voice phone calls from more or less mobile devices [7]. The supported phone calls provided good voice quality but were not only unreliable, they were also expensive and insecure. The maximum speed of 1G was 2.4kbit/s [12].
In 1G, only the phone had to authenticate itself to the base station. The authentication of the base station to the phone did not happen. This was a major security flaw, as an attacker could pose as a legitimate base station and let the phone connect with the false believe of connecting to a base station of the service provider. Additionally, the phone call itself was not encrypted, it was directly modulated on the carrier signal.
Because a phone call was not encrypted, an attacker could easily listen to the conversation by tapping into the signal. This act is also called eavesdropping. If an attacker had access to an antenna, he/she could use it to build a fake base station. This base station could be used to attract victims to connect them to this antenna, which could then be used to again listen to the conversations or even manipulating the data that was being sent [7]. Such devices were also called Stingrays (which was first the name of a specific device, but later became the general name), and were used by the military and law enforcement agencies of different countries. This attack is called man-in-the-middle attack.

3.2.2 2G - GSM

The second generation of mobile communication (2G) was a major upgrade from 1G. It was introduced 1991 in Finland. 2G used the "Global System for Mobile communication" (GSM) standard. GSM is digital [12]. To use GSM, a user had to operate a mobile phone with SIM-card, to activate the phone in a given network. This allowed the phone to be operated in different countries, which was not possible in 1G. Over the course of nearly a decade, the 2G standard improved from GSM to "Global Packet Radio Service" (GPRS), EDGE and finally EDGE+. Those improvements were not formally declared into a new generation, but are informally named 2.5G(GPRS) and 2.75G(EDGE). The maximum speed of 2G was 40kbit/s with GPRS, 384kbit/s with EDGE and 1.3Mbit/s with EDGE+ [17].
With the 2G network, users were not only able to have voice calls, but it was also possible to send an SMS or an MMS. Phone calls became less expensive in 2G, which is one of the reasons why it was a worldwide success. Today, approximately 80% of the world population use 2G for wireless phone calls.
The 2G network addressed one big flaw of the first generation: It used encryption between the base station and the end device. This encryption was symmetric and the key used was known by the provider and the end device by using the SIM-card module. There were several algorithms at hand that could be used to perform this encryption. Unfortunately, the encryption algorithms were shown to be breakable and the encryption was only between the end device and the base station. Inside the network, the traffic was sent without using any form of encryption. While this flaw has been addressed, the possibility for a fake base station still existed [7].
Even though the encryption made it more difficult for attacker to eavesdrop, it was still possible as the encryption algorithm was shown to be breakable. Again, connections were still vulnerable to devices which intercept the communication between the user and the service provider. Such a device mimics the service provider/end user to the end user/service provider and can listen to or even manipulate the data transmission [7]. With the introduction of SMS and MMS, there were services that could effectively be manipulated.

3.2.3 3G - UMTS

For the third generation of mobile communication (3G), further improvements were made on the basis of the GPRS and the EDGE standards of 2G. The 3G cellular network allowed for data rates of at least 144kbit/s. By further improving 3G, the network was able to reach 14Mbit/s under High Speed Packet Access (HSPA).
As 2G improves on the security of 1G, 3G improves on 2G. 3G introduced new encryption algorithms, which still had some vulnerabilities, but where overall more secure than the algorithms used in 2G. Additionally, 3G allowed the end device to authenticate the network that it is connecting to [14].
The newly introduced two-way authentication procedure made the use of Stingrays harder but not impossible. A Stingray could demand the connection to be held on the old 2G standard. This is also called a downgrade attack. By using this attack, a Stingray could use all vulnerabilities of the old standard and thus operate as before.

3.2.4 4G - LTE

[18] The fourth generation of mobile communication (4G) is also called (LTE). Even though the first release of the LTE standard was not formally 4G, as it did not meet the technical requirements, later releases were compliant and could be called 4G. By enhancing the already existing technology, 4G could improve the download speed over the network significantly in comparison with 3G. While 3G had a top download speed of 300Mbit/s, 4G could handle even up to 1Gbit/s (if all conditions were optimal) [14].
The download speed allowed users to use their mobile device for various applications. User could watch HD-television over the mobile network, or they could have a face-to-face call with their loved one at nearly real time. Most current mobile devices support from 2G to 4G.
The security of 4G improved on the encryption of 3G network. To the best of our knowledge, there are no known vulnerabilities of the algorithm used for the encryption of the communication in 4G, thus it can be regarded as safe to use. The use of Stingrays is still possible in the 4G network. As the end devices of the 4G network will communicate closer with the internet, more traditional attacks like malware or ransomware also are to be considered [7].

3.3 Current generation of mobile communication: 5G

With the introduction of 5G nearly complete in western and Asian countries, its implementation is an interesting topic to look at right now. This generation of cellular networks is being tested since 2018 and will bring data rates of up to multiple Gbits/s, with latency times around 10ms and higher device density than before. In the first few years, 5G will be implemented alongside with 4G (also called non-standalone mode) as the standalone mode is not yet fully developed [19]. 5G comes with exiting new use cases, modern key technology and security aspects of them both. This section presents the use cases and the key technologies, to later discuss their security aspects.

3.3.1 Use cases for 5G

The fifth generation of cellular networks will provide users with a download rate of up to 10Gbit/s and with a latency as low as a millisecond. This allows for new use cases, that were not possible in 4G. This section presents some use cases of 5G.
Internet of Things, IoT:

The IoT is the term for using smart sensors in many different applications. These sensors are connected to a controller via Ethernet, cellular networks, Bluetooth or other communication standards. It is going to profit of 5G, as it will have the possibility to communicate with a variety of devices at the same time with close to no latency time. Additionally, these devices can communicate without needing a cable or a Wi-Fi connection to connect to the network. This makes IoT much more flexible than before.
These IoT solutions can be applied in various environments: the industry can use IoT devices to boost their efficiency, smart cities will improve the quality of life of its citizens, IoT in agriculture can help to decrease pesticide use while increasing crop rate and in healthcare, IoT can help the doctors and nurses to react more quickly and efficiently [20].
Autonomous driving & drones:
With the advancement of artificial intelligence (AI) and machine learning (ML), it becomes possible to have machines which can perform certain tasks on their own. Today's applications are self-driving cars, fully automated factory machines and many more.
Autonomous driving and drones can be used more efficiently, by providing them with the possibility to communicate with nearby devices, without needing a huge overhead [21]. Also, they can communicate with their controller over 5G instead of using a separate frequency, which will boost their operational range significantly up to the limitation of the devices battery. This could allow for delivery drones, which can deliver smaller packets directly to the user.
Cloud gaming & streaming:
With the recent trend for infrastructure, platform and software as a Service (IaaS / PaaS / SaaS), cloud gaming and streaming will become more prominent in the future. By offering these services as SaaS, companies have the possibility to keep the software up to date and the users profit by only paying for the time they spent, and not the whole fee. Cloud gaming will also profit from the introduction of 5G, as it allows devices with lower computational power to stream a game directly from a gaming server, while the user sends controls over the device to the server [29]. Streaming will have a much higher quality then before.

3.3.2 Key technologies of 5G

As each generation has done before, 5G introduces new technology to is mobile communication system. These technologies guarantee a better quality of service (QoS), better data rates and try to improve the security of the previous generation. The fifth generation cellular network has multiple such new technologies. While 5G also inherits many key technologies from the previous generations, this section focuses on the new technologies. This section looks at the new radio frequencies, massive multiple input multiple output, ultra dense networking, software defined networks and network function virtualization. These are not all new technologies used in 5G, but as this work aims to give an overview, it would be not in the scope to address them all.
New radio frequencies
The new 5G standard has defined two main frequency bands for the 5G network. The first frequency spectrum is is from 410 MHz to 7.125 MHz. This spectrum includes frequencies that are already in use by previous generations of cellular networks. The new spectrum includes also new, unused frequencies.
Additionally, the new standard also includes a totally new spectrum of frequencies. This range goes from 24,25 GHz to 52,6 GHz. These new ranges are also called mmWaves, because the wavelength is in the millimeter area. The 5G network can achieve higher available bandwidth than its predecessors, but is also more limited as the range of these frequencies is shorter than with traditional frequencies [24].
Massive multiple input multiple output, massive MIMO

By using multiple antennas at both ends of mobile communication, a channel can be further multiplexed by using the spatial dimensions. With this technique, antennas can increase sector throughput, energy efficiency and capacity. This can be done by splitting the cell of a cellular base station in sectors. Each sector can be served independently of the other sectors [22].
Massive MIMO is an improvement of the classical MIMO. While MIMO uses two to eight antennas, the new massive MIMO can use up to 1'024 antennas. Base stations equipped with massive MIMO are scalable with the number of antennas. It means that the more antennas a base station has, the more devices can be served. This allows the base station to not only reach more devices, but also to have the ability to reach the devices with more precision and signal power while consuming less energy. The higher precision allows the base station also to communicate with devices that are further away then before [24].

Ultra dense networking, UDN

With the introduction of the new radio frequencies, the range of the cells using these frequencies decline dramatically. The mmWaves have difficulties to penetrate through walls and windows, and their range is further limited if it rains or snows. This introduces the need for either more powerful cells or more smaller cells.
The use of so called ultra dense networks is not really a new technology, but rather a paradigm shift. It introduces new smaller cells, that can be used for a very limited space, but which are very widespread. These new cells will mostly be used in urban environments, where the device capacity is high [25].
By splitting the network in a larger number of smaller cells, the network can use its resources more efficiently. This can be done by splitting the physical space in smaller parts, which are mostly separated from each other. Thus, frequencies can be better reused in a given area [25]. This advantage can be mostly used in very dense areas like malls or sport stadiums.

Software defined networks, SDN

In the past few years, the setup of a network was a long and tedious work. Every device had to be configured by itself and was possible reachable remotely by its IP-address. Today, the management of a network can be done by a centralized network controller, by using predefined and standardized communication gateways between the network controller and the switch being used. This makes it simpler and less error-prone to incorporate changes on the network. Additionally, the network can automatically react to certain events in the network [28].
In SDNs, the network is split up in the controller plane and the data plane. Network switches become simple forwarding devices which are controlled by the centralized controller. The controller has overview over all the devices inside the network and can manage the data flow dynamically [28].
This allows to have a view over the whole network and enables administrators to react flexibly to new threats and requirements. With the centralized network controller, administrators can profit from a higher grade of automation, as the controller has all the information needed to perform basic tasks.

Network function virtualization, NFV

The trend of virtualizing has now reached the network devices. Hardware specifically designed for one task can be moved to a virtual machine (called VM) that is run in the network. This move to VMs can be done for many devices: Carrier Grade NATs, DNS, Firewalls and more. This is called the network function virtualization (NFV). This allows to place network functions at various locations in the network at run time. As the NFV are virtualized, they need a host system they can run on, this system is called hypervisor [23]. A hypervisor can host many different guest operating systems (OS).

The NFV is done by implementing various virtualized network functions (VNF). These functions are running on the VMs and are orchestrated by an centralized control plane, where resource allocation, dependencies and availability are managed [23].
This allows the network to be more flexible and dynamic. If the requirements of a network change in a short span of time, the needed functions can be adjusted accordingly by giving the VMs the resources they need. This enhances the reaction time to a given task or problem extremely. NFV is used together with SDN, as both technologies complete each other to a certain extent.

Network slicing

Network slicing allows network administrators to to split the physical infrastructure multiple logical networks, where each networks is defined according to its own requirements. This is done and controlled by a centralized controller called network slice manager (NSM). This slices coexist with each other without letting traffic from one slice transit to another slice inside the network [3].
This allows to have a network with multiple slices, where each slice has its own use case and requirements. For example, one slice can be used for IoT and sensors by using lower frequencies with smaller data rates, while another slice is used high-bandwidth consuming devices with high data rates. This allows the network to be more robust and reliable while maintaining the needed flexibility.

3.3.3 Cybersecurity of 5G

This sections aims to provide an overview over the known cybersecurity issues of the beforehand presented use cases and key technologies. This is done by firstly looking at the possible aspects of a use case/key technology and secondly by analysing known vulnerabilities and attacks.

3.3.3.1 Use cases

This section presents some cybersecurity issues and some protections offered by the use cases of 5G. Table 3.1 provides an overview of the discussed issues and protections.

Internet of Things:

A big problem for the security of IoT devices is their sheer number of devices there will be connected to the network. Most of these devices will be small and have limited computational power at hand. Additionally, with regards to energy consumption, computational power is further throttled. With the remaining computational power, it is difficult to implement secure algorithms for data transmission and other security mechanisms. This vulnerability is only made worse by the infrequent and bad update policies of both supplier and consumer of IoT devices. This makes IoT to easier targets for attackers. Furthermore, IoT devices are often used autonomously in unattended environments. This opens the possibility for attackers to gain physical access to a device, which could then be damaged or manipulated [20].
The large number combined with the limited security make IoT-devices to attractive targets for attackers. These attackers could use the IoT to create a botnet of small sensors and similar devices, which pose no threat on their on, and could unleash devastating DDoS-attacks under the command of the botnet. The Mirai malware is a rather famous example for such an attack. Another threatening attack is the False Data Infection (FDI) attack. In FDI attacks, legitimate data is corrupted which then can cause false reactions by the IoT controller. This attack could for instance pose a serious threat to power plants. IoT devices can also be cracked with the intent to gain access to personal data, as these devices possibly monitor and transmit private data [20].

Table 3.1: Cybersecurity of 5G use cases

Use case	Offered Protection	Cybersecurity issue
Internet of Things	-	Low computational power, update policy, DDoS Botnets
Autonomous cars & drones	-	Hijacking
Cloud gaming & streaming	Better protection against illegal copies	Virtualization issues

Autonomous cars & drones:
Autonomous cars & drones have the same problem as IoT devices. They have limited space and in case of flying drones also limited weight. This leads again to lowered computational power and thus open up vulnerabilities for malicious actors. The communication of autonomous cars between themselves increases the potential for attackers to manipulate such cars. Additionally, one could gain confidential data from this machine to machine (M2M) communication. Furthermore, an attacker could jam the communication between the car and other critical infrastructure and can thus endanger the car and its passengers [21].
These actors could use these devices for DDoS-attacks, that are orchestrated by botnet controllers. An attacker could also hijack the device and endanger the passenger in the case of autonomous cars or use them to crash into other objects. In modern autonomous cars, if one could gain root access to the system, the possibilities are nearly endless. As for autonomous drones, the two most observed attackers are GPS jamming and spoofing. This could hinder the proper controlling and managing of such systems [21].
Cloud gaming & streaming:
With the introduction of SaaS in the gaming and streaming community, the the producer has the control of system, on which the software runs. This enables the protection of the content the producer wants to sell. Attackers will have a much harder time to illegally copy the content and redistribute it [29].
As the content has to be hosted somewhere, which is typically done on VMs, cloud gaming and streaming will inherit the cybersecurity issues that come along with the virtualization technologies. These issues include hypervisor and guest OS vulnerabilities [26].

3.3.3.2 Key technologies

This section presents some cybersecurity issues and some protections offered by the key technologies of 5G. Table 3.2 provides an overview of the discussed issues and protections.
Massive multiple input multiple output:
By using this the further reach and the higher precision of massive MIMO, the signal strength can be enhanced. This leads indirectly to less possibilities for both intentional and unintentional jamming and the use of stingrays [10].
In the beginning of the communication between the base station and the end device, the end device sends so called pilot data in order to estimate the channel for subsequent transmissions. This pilot data can be resent by an attacker and reconfigure the base station to send the data to the attacker instead of the user. This is called a pilot contamination attack [22].
Overall, the massive MIMO technology will not only increase the QoS of a network, but also it's security by reducing the possibilities of jamming and eavesdropping. But still has some vulnerabilities that can be exploited.

Ultra dense networks:
Ultra dense networks help to improve the network security similarly to the massive MIMO technology. By increasing the number of cells, the signal strength of the antennas at the end devices also increases. This comes again with a higher QoS and better protection against intentional/unintentional jamming and eavesdropping.
On the other hand, the network needs more devices to handle all these connections. One of these devices is the relay. It connects the users with close to no signal with network. The relay is an intermediary device which reroutes the traffic to the base station. This makes the relay to an attractive target for attackers. If an attacker could gain access on such a relay, the attacker could change the MAC-address of the relay and thus make it unavailable to the network or install malicious software on the device [25].
Overall, the UDN provides multiple improvements for network security but also bring some new vulnerabilities that can be linked with the smaller less expensive devices. There are several solutions to the new threats, but many of them are of theoretical nature and need to be tested in a real world application [25].
Software defined networks:
On the one hand, the centralized controller is a weak spot of the network. If the controller is not available, the whole network goes down [10]. Such an unavailability of the controller can be reached by using a Dos or a DDoS attack. Worse, if the controller is compromised, the attacker would have access to the whole network [27]. The controller is a so called single point of failure (SPoF). On the other hand, SDN uses third party applications to implement some functions needed in the network. These applications can also contain malicious code if one does not check the application accordingly. As such an application has access to the network, such an attack could lead to the network not being available. Such an attack is also called a trojan attack, as the application poses as legitimate application and needs to be installed by an administrator [28].
While SDN is convenient for many network administrators, it is not completely safe. SDN allows administrators to react more dynamically to incoming attacks and makes the network overall more adaptable. But the controller of the network could be the new main target of incoming attacks, which could possibly affect the whole network.
Network function virtualization:
With the virtualization of network functions, these functions can be controlled in a central place. This allows for the central orchestration of various network functions, that have acted on their own beforehand. This makes the network more dynamic and flexible. The functions used can be changed on a regular basis (or randomly), which makes it more difficult for an attacker to understand the network and act accordingly. This is also called moving target defense (MTD) [23].
NFVs have the property that they use a centralized control plane, from which everything is controlled. This makes this control plane to an attractive target for possible attackers. This shows the need for secure passwords and efficient user permission control. Additionally, the hypervisor on which the VNFs are running on can also be targeted if not properly secured. If an attacker could gain access to the hypervisor, the VNFs can not guarantee a proper execution of the promised functionality [26].
If an attacker could gain access to the control plane or hypervisor of the VNFs, he/she could open the network for further attacks.
Network slicing:
The isolation of different slices influences the security positively by limiting attacks on a service to a given slice. This increases the networks robustness. Through the virtualization of various slices, network slicing allows for more detailed and individual network security for each slice. The needed isolation can only be guaranteed, if all the requirements for each slice can be met by the underlying infrastructure [3].

Table 3.2: Cybersecurity of 5G key technologies

Key Technology	Offered Protection	Cybersecurity issue
Massive MIMO	Better reliability against jamming and interferences	Pilot contamination attack
UDN	Better reliability against jamming and interferences	Relay attack
SDN	-	Controller is an attractive target and SPoF
NFV	MTD	Controller is an attractive target, virtualization issues
Network slicing	Increasing network robustness	Eavesdropping, data injection

By monitoring (eavesdropping) the traffic of the northbound/southbound interfaces (communication interface for higher/lower level services), it is possible to understand how the network slices are configured. An attacker could create a snapshot of the systems status, from which the system can be searched for vulnerabilities. It is also possible for an attacker to catch sensitive data, which would allow him/her to impersonate a service or a user and manipulate the NSM. Also, if an attacker would be able to inject data in those interfaces, it would be possible to manipulate the NSM [3].
Network slicing is a promising technology, that can allows the network to become more dynamic and accurate. Nevertheless, network slicing still has some problems that need to be solved in order to have a real impact on the networks security. These vulnerabilities are more often that not architectural problems, which means that they could be solved in the future. As network slicing builds on SDN and NFV, it also inherits their cybersecurity issues.

3.4 Next generation of mobile communication: 6G

While the introduction of the 5G cellular network is nearly finished, researchers around the globe are looking into the next generation of cellular networks [13]. 6G is estimated to be introduced around the year 2030 based on the 10 year life cycle of the previous generations. Even though we are still at the beginning of 5G, it is definitely worthwhile looking into 6G already to try to make it as reliable and safe as possible.

3.4.1 Use cases for 6G

While 5G will solve many problems like IoT-communication in the next few year, there will still be some things that are not possible with it. 6G will require even higher speeds and lower latency times than before. In 6G, we will probably talk about data rates of approximately 1 Tbit/s and latency times lower than a millisecond [12].
Multi-sensory extended reality (XR) applications
6G will bring us exiting new use cases like multi-sensory XR applications. While virtual reality takes place in the completely digital space and augmented reality is the real world combined with an information overlay, extended reality combines them both. XR lets us experience the real world with new digital things and both worlds can be manipulated. This is done by adding new and better human experiences like haptic feedback and generally including more human senses [14].

5G will already bring extended reality examples, but they still have some problems. The bandwidth of 5G will not be enough to incorporate lossless compression of picture and video. Furthermore, current latency times are not optimal for truly immersive experiences [31]. Both of these problems are likely to be solved by the future 6G network.

Connected robotics and autonomous systems

Connected Robotics and autonomous systems will also experience a boost. 6G will allow them to make even smarter decision in lesser time, by allowing them to communicate in real time with various actors. This will allow for completely driver-less cars or drones, or even a fully automated factory [2].

These automated systems are already in use in various applications, but with the higher data rates, reliability and lower latency times, new applications will come. These applications need these requirements to communicate in real time and download high resolution maps or plans. Especially the data rates and the low latency times cannot be provided by 5G [12]. Most likely, we will see such use cases with 6G.

Wireless brain-computer interaction

Wireless brain-computer interaction will allow a user to communicate with a computer, simply by thinking about it. This could improve the live of millions of people by enabling artificial limbs, that can be controlled like real limbs. Communication would be held over the cellular network, this could enable standardized communication instead of vendor specific standards. This communication requires nearly real-time latency with close to no jitter, so that the users experience is of high quality. 5G cannot provide this kind of service yet, so brain-computer interaction will have to wait for 6G [14].

Internet of Everything, IoE

In 5G, we had the IoT, now in 6G, the next step of evolution is the IoE. As the name suggests, even more products will be communicating with you and the internet than before. The IoE will bring new use cases that were not possible beforehand, as the data rates and reliability can not be delivered yet. An example for that is a fully automated factory, where the maintenance, diagnostics and operation is done by robots, based on the data of IoE-devices [13]. 6G will also give all the use cases named in 5G a boost and will be an additional step in the direction of a fully digitalized society.

The IoE will rely more on automated data exchange than the IoT, because the IoE will be heavily used in the industry 4.0. These machines and devices will need extreme high data rates, reliability and as little jitter as possible. Another requirement of the IoE is the network density. As this cannot be provided as of now, the IoE will truly flourish with the introduction of 6G [13].

3.4.2 Key technologies of 6G

The use cases mentioned beforehand come with various requirements that need to be covered so that 6G can successfully be launched around 2030. These requirements will need new technology as it is not possible today. 6G is expected to have the following requirements: 1 Tbit/s download rates, latency times lower than a millisecond and even higher connection density than we are seeing and will see in 5G. Another requirement is the broad coverage of earths surface with 6G connectivity. There are multiple technologies that will possibly being used in 6G, this section presents some of the most promising technologies.

Advanced AI & machine learning

Advanced AI and machine learning will be very important for 6G. In the 6G network, there will be an immense amount of data generated each day, that could be used to gain new insight to the network if it can be analyzed in time. Both technologies allow the network to react more dynamically to a diverse set of situations, based on the incoming

data. This data can be analyzed more quickly, accurately and with the help of AI and ML, more complex data patterns can be found [2].
These capabilities make AI and ML extremely valuable in all possible situations for the next generation network. They could be used to optimize the security policies of the network, or to optimize the QoS for for all users [6]. Unfortunately, attackers could possibly also profit from these capabilities.

Terahertz, visible light communication & beamforming

With the current use of the available frequency spectrum we will need new frequencies that can be used for wireless transmission. Also, the requirement of 1 Tbit/s download rate is only possible if we use higher frequencies. Higher data rates are reachable by having more spectrum which can then be multiplexed into more channels [4]. Furthermore, the higher frequencies reduce the size of the required antennas, which makes compact radio systems with many antennas possible [36]. This leads us to the THz and visible light communication technology.
As the name suggest, THz communication uses frequencies in the range 0,1 - 10THz for wireless data transmission, while visible light communication uses waves in the range of 400 - 800THz. Both technologies have the drawback, that the range is extremely limited even in optimal conditions. The THz waves have no chance to penetrate walls of buildings and can even be absorbed by a raindrop. Visible light is also limited to the line of sight between sender and receiver, but could already be implemented with common LEDs [11]. This would make THz waves only applicable indoors, where it can provide the high data rates. If THz waves are used together with highly directed beams, the range can be enhanced in line of sight to the antenna.

Quantum communication & computing

In the last few years, there were significant improvements made in the field of quantum computing. Even though today's quantum computers are huge and susceptible to environmental changes, this could drastically change in the next few years. Quantum computing will thus be used in the backbone of the network. It also helps to increase the channel capacity by enabling new multiple access technologies, that would have very high power demand on normal computers [14].
Quantum computing allows for new ways of encrypting data before sending it over a network. One of this techniques that is already in use is the quantum key distribution (QKD). It allows two parties to communicate with shared random key that is used to en-/decrypt messages. One advantage of this technique is that a manipulation of the message could be detected effortlessly [14].

Integration of communication methods

To fulfill the requirement to reach 6G coverage of nearly the whole surface of earth, a new way to communicate with the network will be needed. This can be provided by the integration of various technologies that we already use. This includes satellite communication and airborne drones that could provide 6G [11].
These options could be used in places where no antenna could be deployed or where it would make no sense economically. This means that a device could have 6G signal in the desert provided by a satellite. The price for the connectivity in such remote places are the higher latency times. Communicating with a satellite can have a delay of up to 0.5s.

Edge cloud computing

One of the most important part of the latency time is the time it takes the information to travel from the device to the responsible server. This time can be optimised if the needed computation is placed as close to the end device (to the edge) as possible. While this trend will also be seen in 5G, its relevance will be far more important in 6G. That is exactly the idea of the paradigm of edge cloud computing. Move the responsible servers as close to the edge as possible to reduce latency times, network traffic and storage needed on the central servers and end devices. This allows the network to be more efficient while

improving the QoS for many devices [15]. There are many different applications for edge cloud computing. An edge server could cache information or to offload computations of the end device to the server, and thus increasing their battery life [8].

3.4.3 Cybersecurity of 6G

3.4.3.1 Use cases

This section presents some of the most well-known cybersecurity issues and some protections offered by the use cases of 6G. The Table 3.3 provides an overview of the discussed issues and protections.

Multi-sensory extended reality (XR) applications:

The use of multi-sensory extended reality allows a system to collect a huge amount of data about a person that is using this reality. This collected data could include information about the emotional state of the user, this data can be extremely valuable for some companies [14]. This makes those XR systems to attractive targets for attackers.

With the fusion of the real world with the virtual world, XR could record other highly sensitive data. This could be private bank information that are laying on the kitchen table. One solution to this problem is the blocking of such information in real time. This is called input protection [30].

Attackers could endanger the user of a XR system by manipulating the output the user gets from the system. This can be done by changing the video output so that it no longer corresponds with the real world. This is called Output protection [30].

Connected robotics and autonomous systems:

Truly autonomous system will have to heavily rely on environmental data. A possible scenario would be a self driving car, that communicates with the red light at the next intersection. Most of these devices are low in computational power to optimize battery life. That makes these devices more vulnerable to attacks. An attacker could use cracked devices to provide the cars with wrong information [14]. Additionally, such autonomous system are also very attractive to eavesdrop on or to hijack. Furthermore, robotics today were shown to have a variety of vulnerabilities. These vulnerabilities include authorization and authentication vulnerabilities [21].

Attackers could use these vulnerabilities of these devices to their advantage and cause a significant amount of damage. For example, if an attacker has root access to a self driving car, he/she could lock the doors of the car if a person is inside and demand a ”fee” to let them free. Another scenario is the hijacking of an airborne system, that could be abused as kamikaze drone [32]. One such drone is highly dangerous, but a swarm of such drones would be devastating.

Wireless brain-computer interaction:

The current trend of coming up with new ideas for wireless brain-computer interaction made it even more important to have a secure connection. There are ideas that use the emotional state of a user to complement already existing passwords. To enter a system, the user has to know the correct password and must match EEG waves based on various emotional states like relaxation or concentration. To assist the user, the scientists used the neurofeedback technique, in which the user received feedback on the current state [33].

It is possible for attacker to gain private and sensitive information of a brain computer interface (BCI) user by presenting him/her intentionally crafted stimuli. This is called a misleading stimuli attack. A misleading stimuli attack could also cause damage, if the new mental state of the user lead to incorrect actions of the BCI device. On the contrary to data extraction, it is also possible to influence the emotional state of the user by using malicious stimulation parameters. Such an malicious stimulation could reverse

Table 3.3: Cybersecurity of 6G use cases

Use case	Offered Protection	Cybersecurity issue
Extended Reality	-	Input/Output protection
Autonomous systems	-	Hijacking
Wireless Brain-Computer interaction	Emotional state password	Misleading stimuli attack, Malicious stimulation parameter
Internet of Everything	-	DDoS Botnets

the functionalities of the device, in case of a sickness that is treated with BCI, prevent the user from speaking or moving or even endanger the users life [34].

Internet of Everything, IoE:

In 6G, there will be a similar problem with IoE devices like in 5G with IoT devices. The IoE devices will probably have more computational power at hand than their predecessors, but the attackers will also have more computational power at hand. In comparison with 5G, there will be even more IoE devices in 6G, which makes them even more attractive to attackers. At the moment, updates are infrequent for IoT devices, this could significantly change over the next few years. The small devices are also a valuable target because of all the information that could be gained from them. If these devices become more widespread, attackers could gain access to highly personal or sensitive data.

Even though the update situation could improve in the future, IoE devices will never be completely safe. Attackers could use existing vulnerabilities to gain control over a large amount of small devices. Like in 5G, these devices could be used to create a botnet, which could launch far reaching DDoS attacks. The attacks could be more dangerous than today, as the number of devices will probably be much higher than now in ten years. There are also reported attacks based on industrial devices, which prove that such attacks can in fact be dangerous. Examples are the attack on Valtia in Finland, where the central heating and hot water systems were malfunctioning and the stuxnet attack, which damaged nuclear power plants in Iran [35].

3.4.3.2 Key technologies

This section introduces the most relevant cybersecurity issues and some protections offered by the key technologies of 6G. The Table 3.4 provides an overview of the discussed issues and protections.

Advanced AI & machine learning:

AI is envisioned to play an important role as defense mechanism in future cellular networks. Many interfaces today are based on REST APIs, which are a primary target due to their importance. The wide variety of APIs and the huge amount of traffic makes the identification and mitigation of threats a complex task. AI-driven API security has the capability to uncover patterns in multidimensional data, which allows better monitoring and mitigation of API attacks. AI techniques have also been considered highly relevant to implement MTD mechanisms [9].

AI and ML can unfortunately also be used to attack a given network. Attackers will use more intelligent attacks in the future, that are able to scout for various vulnerabilities in network on their own. The attackers could use these vulnerabilities to launch further attacks on the network [9].

Due to the importance AI and ML system will have in 6G, they are also attractive targets for attackers. It is possible to attack the network while it is still under construction, by poisoning the training data of the AI that will be responsible for the network. By doing this, they can change the expected reaction of the AI to given situations and possibly open up a loophole for a planned attack. If attackers have no access to training data, they have the possibility to feed the system malicious crafted samples to alter the data distribution in their favor. This is called data injection. These attacks are also called adversarial attacks [9].

Terahertz, visible light communication & beamforming:

Due to the improvement of the accuracy, THz communication helps to further reduce interference and could help against jamming by making it more costly to launch an attack [14].

While the use of these frequencies and beamforming can help against interference and jamming, it gives no guarantee that it can not happen. An attacker could still jam the network, he/she would just a need a device that could send a signal with higher power in comparison to current cells. This holds also true for other types of interference [36]. Overall, terahertz, visible light communication and beamforming help to make the network more robust and provide more uptime, but can not deliver complete secure connection.

Quantum communication & computing:

On the one hand, quantum communication gives us a possibility of encrypting data in such a way, so that the receiving party can determine if the data was manipulated. This is done by using the already mentioned QKD technology. This enables the parties involved in the interaction to have a certainty that there was no eavesdropper involved [14]. If the connection reaches a certain threshold of difference in the qubit state, the message will be discarded and sent again until the connection is safe.

On the other hand, quantum computing has been proven to be able to crack today's encryption mechanisms. The quantum computer that are used today are not advanced enough to crack the encryption, but this will change in the future. There are already new encryption mechanisms that are secure even against quantum computing, but these mechanisms are not in use yet [1].

With the further advancement of quantum communication and computing, the future 6G network will likely have to incorporate new encryption mechanisms. Whether the QKD is ready for such a widespread use or not until the introduction of 6G is not clear yet. Quantum communication will probably not be used for the communication between end device and the network, but rather inside the core network.

Integration of communication methods:

By using multiple system to provide 6G signal to the whole surface of earth, the network robustness will increase and is thus more secure against physical attacks on antennas. Additionally, the network has a backup for possible natural disasters again by being more independent of existing antennas on the ground [7].

The problem of the integration of all the various technologies are the security concerns the technologies themselves have. One problem is the use of autonomous drones for the delivery of 6G signal to various places on the earth. These drones could be hijacked, which would lead to the same problem as discussed in connected robotics and autonomous drones section [21].

Edge cloud computing:

Edge cloud computing could possibly help user with the protection of their private data, by allowing to make some first computation directly at the user site. This computations could be stripped of all unneeded private data and then be sent to a more centralized server [15]. In 6G, this could be used together with IoE devices to reduce the amount of private data, a user has to reveal to use a service.

Table 3.4: Cybersecurity of 6G key technologies

Key Technology	Offered Protection	Cybersecurity issue
AI & ML	Rest API security MTD	Zero-Day vulnerability scanning adversarial attacks
THz communication	Better reliability against jamming and interference	-
Quantum Computing	QKD	Encryption cracking
Integration of communication methods	Better robustness through more backup	Inherits problems of integrated technologies
Edge Computing	Better privacy control	Attractive target

The problem with having a dedicated server for such computations is that it makes for an attractive target for attackers. The edge cloud servers would have access to a huge amount of private data, which in turn would give an attacker more incentive to attack it [15]. As no system is completely safe, an attacker could gain access to a edge cloud server and could copy private data or even manipulate the data for his/her own interest.
Edge cloud computing will probably happen at the place of a small cell in the 6G network. Such a cell could be responsible for a huge amount of IoE devices. While an edge cloud server can provide a user with better protection for his/her private data, it also gives an attacker more incentive target such a server.

3.5 Conclusion

The first generation of cellular networks did not provide any security mechanisms for the user. It changed in 2G, where digital modulation and data encryption were introduced. The third generation further improved on the encryption of 2G, by enabling the end device to authenticate the base stations of the network. In 4G, the encryption was further improved. The fifth generation of cellular network bring many new technologies. Massive MIMO helps to reduce interference and jamming in the network, SDN, NFV and network slicing make the network more manageable and flexible but also introduce new threats. In 6G, AI and ML will play an important role as defender and attacker of the network, while also being a target itself. The introduction of new transmission technologies like terahertz frequency communication and architectural paradigms will help to increase the networks robustness against interference, jamming and natural disasters. But these advancements also come with new security problems that will need to be solved in the next few years. The security of autonomous drones and IoT/IoE devices needs to be addressed mainly because of the huge amount of devices that are affected. Also, the control planes for all the virtualization technologies will need to be as secure as possible, or else they would rather damage the network's security than help it.

Bibliography

[1] Ylianttila M., Kantola, R., Gurtov, A., Mucchi, L., & Oppermann I., *6G White paper: Researches challenges for trust, security and Privacy [White paper]*, (6G Research Visions, No.9), University on Oulu, 2020, http://urn.fi/urn:isbn:9789526226804.

[2] Viswanathan H., Mogensen P. E., *Communications in the 6G era*, in IEEE Access, vol. 8, pp. 57063-57074, 2020. https://doi.org/10.1109/ACCESS.2020.2981745

[3] Cunha VA, Silva E, Carvalho MB, et al., *Network slicing security: Challenges and directions*, Internet Technology Letters. 2019;2:e125. https://doi.org/10.1002/itl2.125

[4] David K., Berndt H., *6G Vision and Requirements: Is There Any Need for Beyond 5G?*, IEEE Vehicular Technology Magazine, vol. 13, no. 3, pp. 72-80, Sept. 2018, doi: 10.1109/MVT.2018.2848498.

[5] Katz, M, Pirinen, P. & Posti, H., *Towards 6G: Getting Ready for the next Decade*, International Symposium on Wireless Communication Systems (ISWCS), Oulu Finland, pp. 714-718, August 2018. https://ieeexplore.ieee.org/abstract/document/8877155

[6] Kato, N., Mao, B., Tang, F., Kawamoto, Y., & Liu, J., *Ten Challenges in Advancing Machine Learning Technologies Toward 6G*, IEEE Wireless Communications, vol. 27, no. 3, pp. 96-103, April 2020.

[7] Ahmad, I., Shahabuddin, S., Kumar, T., Okwuibe, J., Gurtov, A., & Ylianttila, M., *Security for 5G and Beyond*, IEEE Communications Surveys & Tutorials, vol. 21, no. 4, pp. 3682-3722, May 2019

[8] Ji, B., Wang, Y., Song, K., Li, C., Wen, H., Menon, V. G., Mumtaz, S., *A survey of Computational Intelligence for 6G, Key technologies, Applications and Trends*, IEEE transactions on industrial informatics (2021): 1-1. https://doi.org/10.1109/TII.2021.3052531.

[9] Benzaid C., Taleb T., *AI for Beyond 5G Networks: A Cyber-Security Defense or Offense Enabler?*, IEEE Network, vol. 34, no. 6, pp. 140-147, November/December 2020, https://doi.org/10.1109/MNET.011.2000088

[10] Fonyi S., *Overview of 5G Security and Vulnerabilities*. The Cyber Defense Review, vol. 5, no. 1, 2020, pp. 117-134. JSTOR, www.jstor.org/stable/26902666

[11] Chen S., Liang Y. C., Sun S., Kang S., Cheng W., Peng M., *Vision, Requirements, and Technology Trend of 6G: How to Tackle the Challenges of System Coverage, Capacity, User Data-Rate and Movement Speed*, IEEE Wireless Communications, vol. 27, no. 2, pp. 218-228, April 2020, https://doi.org/10.1109/MWC.001.1900333

[12] Ahmed, R., Matin, M. A., *Towards 6G wireless networks-challenges and potential technologies*, Journal of electrical engineering, vol. 71, no. 4, pp. 290-297, 2020

[13] Akhtar M.W., Hassan S.A., Ghaffar R. et al. *The shift to 6G communications: vision and requirements*, Hum. Cent. Comput. Inf. Sci. 10, 53 (2020). https://doi.org/10.1186/s13673-020-00258-2

[14] Wang M., Zhu T., Zhang T., Zhang J., Yu S., Zhou W., *Security and privacy in 6G networks: New areas and new challenges, Digital Communications and Networks*, vol. 6, no. 3, pp.281-291, 2020 https://doi.org/10.1016/j.dcan.2020.07.003

[15] Bhat S. A., Sofi I. B., Chi C. -Y., *Edge Computing and Its Convergence With Blockchain in 5G and Beyond: Security, Challenges, and Opportunities*, IEEE Access, vol. 8, pp. 205340-205373, 2020, https://doi.org/10.1109/ACCESS.2020.3037108

[16] Basciftci Y. O., Koksal C. E., Ashikhmin A., *Physical-Layer Security in TDD Massive MIMO*, IEEE Transactions on Information Theory, vol. 64, no. 11, pp. 7359-7380, Nov. 2018, doi: 10.1109/TIT.2018.2855058. https://doi.org/10.1109/TIT.2018.2855058

[17] 2G - Wikipedia; https://en.wikipedia.org/wiki/2G, May, 2021

[18] 4G - Wikipedia; https://en.wikipedia.org/wiki/4G, May, 2021

[19] 5G - Wikipedia; https://en.wikipedia.org/wiki/5G, May, 2021

[20] Neshenko N., Bou-Harb E., Crichigno J., Kaddoum G., Ghani N., *Demystifying IoT Security: An Exhaustive Survey on IoT Vulnerabilities and a First Empirical Look on Internet-Scale IoT Exploitations*, IEEE Communications Surveys & Tutorials, vol. 21, no. 3, pp. 2702-2733, thirdquarter 2019, doi: 10.1109/COMST.2019.2910750. https://doi.org/10.1109/COMST.2019.2910750

[21] Jahan F., Sun W., Niyaz Q., Alam M., 2019. *Security Modeling of Autonomous Systems: A Survey*, ACM Comput. Surv. 52, 5, Article 91 (October 2019), 34 pages. https://doi-org/10.1145/3337791

[22] Wu Y., Khisti A., Xiao C., Caire G., Wong K., Gao X., *A Survey of Physical Layer Security Techniques for 5G Wireless Networks and Challenges Ahead*, IEEE Journal on Selected Areas in Communications, vol. 36, no. 4, pp. 679-695, April 2018, doi: 10.1109/JSAC.2018.2825560. https://doi-org/10.1109/JSAC.2018.2825560

[23] Hawilo H., Shami A., Mirahmadi M., Asal, R. *NFV: state of the art, challenges, and implementation in next generation mobile networks (vEPC)*, IEEE Network, vol. 28, no. 6, pp. 18-26, 2014, doi: 10.1109/MNET.2014.6963800, https://doi.org/10.1109/MNET.2014.6963800

[24] Andrews J. G., et al., *What Will 5G Be?*, IEEE Journal on Selected Areas in Communications, vol. 32, no. 6, pp. 1065-1082, June 2014, doi: 10.1109/JSAC.2014.2328098, https://doi.org/10.1109/JSAC.2014.2328098

[25] Chopra G., Kumar Jha R., Jain S., *A survey on ultra-dense network and emerging technologies: Security challenges and possible solutions*, Journal of Network and Computer Applications, vol. 95, pp. 54-78, 2017, doi: 10.1016/j.jnca.2017.07.007, https://doi.org/10.1016/j.jnca.2017.07.007

[26] Vaughan-Nichols S. J., *Virtualization Sparks Security Concerns*, Computer, vol. 41, no. 8, pp. 13-15, Aug. 2008, doi: 10.1109/MC.2008.276, https://doi.org/10.1109/MC.2008.276

[27] Shin S., Gu G., *Attacking software-defined networks: A first feasibility study*, Proc. 2nd ACMSIGCOMMWorkshop Hot Topics Softw. Defined Netw. (HotSDN), Hong Kong, pp. 165-166, 2013, doi: 10.1145/2491185.2491220 http://doi.acm.org/10.1145/2491185.2491220

[28] Kreutz D., Ramos F. M. V., Verissimo P., *Towards secure and dependable software-defined networks*, Proc. 2nd ACM SIGCOMM Workshop Hot Topics Softw. Defined Netw. (HotSDN), Hong Kong, pp. 55-60, 2013, doi: 10.1145/2491185.2491199, http://doi.acm.org/10.1145/2491185.2491199

[29] Ojala A., Tyrvainen P., *Developing Cloud Business Models: A Case Study on Cloud Gaming*, IEEE Software, vol. 28, no. 4, pp. 42-47, July-Aug. 2011, doi: 10.1109/MS.2011.51. http://doi.org/10.1109/MS.2011.51.

[30] De Guzman J., Thilakarathna K., Seneviratne A. *Security and Privacy Approaches in Mixed Reality: A Literature Survey*, ACM Computing Surveys, vol. 52, no. 6 pp. 1-37, 2020, doi: 10.1145/3359626 https://doi.org/10.1145/3359626

[31] Gui G., Liu M., Tang F., Kato N., Adachi F., *6G: Opening New Horizons for Integration of Comfort, Security, and Intelligence*, IEEE Wireless Communications, vol. 27, no. 5, pp. 126-132, October 2020, doi: 10.1109/MWC.001.1900516, https://doi.org/10.1109/MWC.001.1900516

[32] Fotouhi A., et al., *Survey on UAV Cellular Communications: Practical Aspects, Standardization Advancements, Regulation, and Security Challenges*, IEEE Communications Surveys & Tutorials, vol. 21, no. 4, pp. 3417-3442, Fourthquarter 2019, doi: 10.1109/COMST.2019.2906228, https://doi.org/10.1109/COMST.2019.2906228

[33] Svogor I., Kisasondi T., *Two factor authentication using EEG augmented passwords*, Proceedings of the ITI 2012 34th International Conference on Information Technology Interfaces, 2012, pp. 373-378, doi: 10.2498/iti.2012.0441, https://doi.org/10.2498/iti.2012.0441

[34] Lopez Bernal S., Huertas Celdran A., Martinez Perez G., Taynnan Barros M., Balasubramaniam S., *Security in Brain-Computer Interfaces: State-of-the-Art, Opportunities, and Future Challenges*, ACM Comput. Surv. vol. 54, no. 1, pp. 1-35, 2021. doi: 10.1145/3427376, https://doi.org/10.1145/3427376

[35] Crelier A., *The Challenges of Scaling the Internet of Things*, Center for Security Studies (CSS), ETH Zürich; Zurich, 2019, doi: 10.3929/ethz-b-000366196, https://doi.org/10.3929/ethz-b-000366196

[36] Schulz R., et al, *Semantic Security for Indoor THz-Wireless Communication*, Cornell University, 2021.

Chapter 4

Current Vision of 6G Networks: Exploring Machine Learning, Holographic Communications, and Ubiquitous Connectivity

Konstantin Moser

Compared to its predecessor the fifth generation of wireless networks is revolutionary on many different levels. Nonetheless, it is only a matter of time until new services will demand network speeds beyond 5G capabilities. Research on holographic communication or extended reality definitely has the potential to do so. In a first step, the crucial technologies of 5G will be elaborated on. This is because 6G will possibly also rely on already implemented technologies like Massive-MIMO and Network Slicing. As a result of intense research in the area of wireless communication, there are different approaches for 6G. The idea to use visible light or terahertz waves as transmission medium is a widespread concept. Both of them have high bandwidths and could enable immense data rates. Yet, there are many hurdles to be overcome before it is possible to fully harness the potential of these spectrums. There is also the idea to use quantum communication for future wireless networks. However, this being a complex research sector the implementation is dependent on the speed of development regarding quantum technology. Moreover, it is agreed on that artificial intelligence will play a key role to optimize network functions. Also services like autonomous driving, virtual and augmented reality are flourishing at the moment and will go hand in hand with the development of future networks. In the end the development of 6G strongly depends on how 5G evolves and will presumably be a combination of many modern technologies together.

Contents

4.1 Introduction

Why are people already talking about the sixth-generation network (6G) while the fifth-generation network (5G) is only being rolled out now? This is a common question. And the answer is straight forward. 5G, the fifth generation of wireless networks, has its limits. Global mobile traffic is growing by 55% annually. In the year 2030, humanity is expected to generate over five thousand exabyte of data per month! [1]. Technology is evolving at fast pace and new services which demand extremely high data rates are at the doorstep. New forms of remote interaction that allow true immersion into distant environments are arising [2]. It could enable holographic telepresence which requires technology that allows full-motion, three-dimensional video conferencing [3]. This involves multiple-view cameras and could demand data rates on the scale of terabits per second [2]. Data rates of this magnitude cannot be supported by 5G. Furthermore, industrial processes are getting more and more automated. The key objective of industry 4.0 is reduction of human intervention. This will be achieved by futuristic automatic control systems and novel communication technologies which are the basis of high precision manufacturing. A prerequisite thereof is highest network reliability. In parallel, latencies and delay jitter must be minimal. This asks for a next generation network. For improvement of the quality of life, paired with increased sustainability, the internet of things is expected to play a key role. Healthcare technologies will be improved and the development of smart cities will continue. Additionally, services like autonomous driving and extended reality are believed to flourish. To this end a fast and reliable data network is paramount. Future successful technology must be sustainable, i.e. energy efficient. This is a major focus of 6G technology [2]. It is important to gather ideas already now of how 6G may function. In the following report the most widespread concepts and technologies will be elaborated on.

4.2 5G Technology and beyond

Before talking about 6G let us shed an eye on 5G and its three main objectives. They are, a) Enhanced Mobile Broadband (eMBB), b) Ultra-Reliable Low Latency Communications (URLLC) and c) Massive Machine Type Communication (mMTC) [2] [4]. They are the blueprint for future intelligent communication to significantly improve peoples lives [5]. Due to its unfavorable topography, there are around 20,000 mobile communication antennas in Switzerland. A good 5,300 of these are 5G antennas. Regarding the dangers of 5G radiation is controversial. Nevertheless, there is great resistance among the Swiss population against the commissioning of further 5G antennas. Until recently, the same limits applied to 5G as to 4G, although 5G works with adaptive antennas. Today, a limit of 6 volts/meter averaged per 6 minutes applies to 5G. With mobile radio diffusion using light beams, this whole issue would become superfluous. Another important factor for their success is thus certain [47].

4.2.1 Enhanced mobile broadband (eMBB)

The enhanced mobile broadband incorporates a range of services. Mainly it is about higher data rates. The target is to reach around 20 Gbit/s download and 10 Gbit/s upload [2]. Another goal of 5G is to enable access to the network everywhere while still allowing high broadband connection even in dense areas. This can be achieved through the combination of ultra-dense networking, massive-MIMO and higher frequencies.

4.2.1.1 Ultra-dense networking

With the rise of the internet of things which results in extremely high capacity requirements, ultra-dense networks have emerged as a prominent solution [6]. Ultra-dense networking describes networks with an enormous density of access points. Cells might even outnumber the amount of active users [7]. It shall ensure seamless coverage of the network even for edge users by using micro base stations. They are usually deployed in dense locations like libraries or shopping malls. This way, the server is kept close to mobile users and request latencies are minimized [6].

4.2.1.2 Higher Frequency Communication

Conventional wireless networks use electromagnetic waves as transmission medium. The higher the wave frequency, the higher is the amount of data which can be transmitted [8][9]. Each new network generation surpassed the speed of the previous network [4].
Until now wireless networks only harnessed electromagnetic waves within the radio and micro spectrum [10]. To deal with the massive growth of transmitted data, even higher frequencies need be used in the future. Millimetre wave technology describes systems which use frequency ranges between 30 and 300 Gigahertz [11]. Millimetre waves are electromagnetic waves that lie between the micro wave spectrum and the terahertz band. They enable higher broadband and lower latencies at the same time. However, if one compares the propagation characteristics of millimetre waves with microwave frequency bands, there are significant differences [12]. Especially in terms of path loss, blockage, and diffraction. Also there's higher rain attenuation and atmospheric absorption [13].
To sum up, the connectivity over distance with microwave systems is significantly more reliable than with millimetre wave systems. The signal is easily distorted by any type of interference, especially from environmental intrusions [14]. That is why the goal of enhanced mobile broadband can only be achieved through the use of higher frequencies combined with higher cell density.

4.2.1.3 Massive-MIMO

Multiple input multiple output (MIMO) is the generic term for processes that improve radio connections with several antennas used in parallel [15]. More antennas provide a better signal and can increase the overall data throughput, which is one of the main goals for eMBB. MIMO is based on the radio technology to send and receive a data stream simultaneously via several antennas through the simultaneous use of frequency, time, space and intelligent signal processing. The bandwidth of a radio channel can be increased linearly with the number of transmitting antennas [15]. This way, several terminal devices can be supplied simultaneously via the same frequency band and more data can be transmitted in total. This spatial separation of the data streams is called spatial multiplexing [16].
Massive MIMO can be seen as an extension of MIMO where far more antennas are being used [17]. There are lots of advantages to gain from using a larger number of antennas, e.g. increasing the quantity of users at the base station. This is particularly useful in dense areas [18].

4.2.1.4 Beamforming

A large amount of antennas is also a requirement for a technology called beamforming [16]. "Beamforming is a technique that focuses a wireless signal towards a specific receiving device, rather than having the signal spread in all directions from a broadcast antenna, as it normally would[19]." By using more antennas it is more effective to bundle

the transmission power and direct it to the individual user [16]. Beamforming is already implemented in modern WiFi systems nowadays. It is possible that Beamforming will slowly migrate towards 5G networks. The migration of technologies from WiFi to cellular networks is not a new concept. It has happened in the past for technologies like OFDMA. By comparing a conventional antenna with an active antenna which is used for beamforming, the performance remains more or less the same. However, previous antennas transmit with the same power for 24 hours a day. With beamforming, it is managed to only call up this performance when it is needed in this cell [20]. Beamforming is able to guarantee higher speeds and network coverage while using less energy.
Static antennas have a certain spreading field. When standing at the edge of such a network, users might experience poor data rates. The active antenna, on the other hand, virtually "sees" when a user is at the edge of the network - and directs a beam specifically towards them. As a result the user is suddenly no longer at the edge of the network but within a beam. [20]. Beamforming technologies definitely have great potential for future networks in my opinion. I personally believe Beamforming to be especially useful in rural areas. Much energy can be saved and the few users can be targeted efficiently in these less dense areas.

4.2.2 Ultra-Reliable Low Latency Communications (uRLLC)

The next target service "Ultra-Reliable Low Latency Communications" describes communication systems with error rates below 10-5 and end to end latencies around 1ms. The 5G network infrastructure is achieving this with Network Slicing based on Network Function Virtualization and software-defined networking. In conventional networks the same architecture has to adopt to serve different services. This leads to inflexibility and could potentially not satisfy users [5]. This problem is eliminated in future networks thanks to Network Slicing.

4.2.2.1 Network Slicing (NS)

Network Slicing enables multiple logical networks to operate on a shared physical infrastructure. Each of the so called "Network-Slices" create service specific networks. This allows the 5G network to be highly flexible. The goal is to logically separate sets of network functions in the physical infrastructure by building a dedicated and customized logical network [5]. The internet of things has the effect, that different users and applications have individual requirements which sometimes strongly differ from each other. That is why the flexibility of a network is so important.
Network slices include three different levels. The first one is the service instance layer. It refers to end-user services which are implemented trough or by a network slice. The next one is the network slice instance layer. This level describes groups of network functions and resources. They run these functions and form a logical network to meet network characteristics. The third level is the resource layer which describes physical & logical resources and network functions [5].

4.2.2.2 Network Function Virtualization (NFV)

On the one hand, Network Slicing is realized through Network Function Virtualization. With Network Function Virtualization it is possible to run software on virtual machines to perform network functions. Network services get decoupled from hardware. The aim is to minimize the use of proprietary hardware in networks. Devices like switches, routers or firewalls are hosted on virtual machines thanks to NFV. Individual functions of the proprietary hardware are integrated on a uniform, virtualized platform which are controlled by software. The result is enormous flexibility. This allows new services to be

implemented quickly and with little effort. NFV is particularly interesting for service providers who want to adapt their services to the requirements of the market within a very short time [21]. NFV obviously has many advantages. First of all hardware is pretty expensive. And instead of buying proprietary hardware, using cheap virtual machines to perform the network functions can save a lot of money. Furthermore, the maintenance cost gets reduced and network upgrades are made easier. Even the space needed for the network hardware is lowered and longer life cycles can be expected [22].
5G is no longer "the network". It can be seen as virtual networks operated in parallel on the basis of a shared physical infrastructure. NFV is closely related to Software-defined networking and develops its full flexibility in cooperation.

4.2.2.3 Software-defined networking (SDN)

The goal of software-defined networking is to improve network control by enabling enterprises and service providers to respond quickly to changing business needs. A network engineer or administrator is able to control traffic from a central control console without touching individual switches if a network is software-defined. In traditional network architecture devices make traffic decisions based on their configured routing tables. The central SDN controller however, instructs the switches to provide network services wherever they are needed, regardless of the specific connections between a server and the devices [23].
Just like NS also SDN consists of three layers. The application layer, the control layer and the infrastructure layer. The application layer contains all the typical functions or network applications that the companies use. Traditional networks use a specialized appliance for specific network applications like firewall or load balancer. A software-defined network replaces the appliance with an application which uses the SDN controller to manage its behaviour. The control layer acts as the brain of the software-defined network. It represents the central SDN controller software. This SDN controller is located on a server and manages the policies and traffic flow in the entire network. And the infrastructure layer consists of the physical switches in the network. Using the respective application programming interface (API) the tree layers are able to communicate [23].
Network Slicing by making use of NFV and SDN has many advantages and is already in use for current 5G networks. And I personally think that there is no reason to believe that Network Slicing should not play a role for 6G. The possibility that NFV and SDN will be used in the next generation network as well, to guarantee flexibility is certainly high.

4.2.3 Massive Machine Type Communication (mMTC)

The third core target for 5G networks is Massive Machine Type Communication. mMTC enables high amounts of interconnected devices and connects components from Machine-to-Machine Communication and the internet of things. An important requirement for mMTC is energy efficiency and high connection coverage. By reason of devices like sensors or smart counters which have to run large periods of time by battery power while being connected to the network [24].

4.3 Visions of 6G Technologies

5G definitely is a breakthrough when it comes to the design of communication networks. It represents a single platform which enables a variety of different services. "Nevertheless, looking at the increasing requests for new services and predicting the development of new technologies within a decade from now, it is already possible to envision the need to move beyond 5G and design a new architecture incorporating new technologies to satisfy new

needs at both individual and societal level.[25]" In terms of capacity, 6G should be able to connect upper trillion-level objects rather than the current billion-level like 5G does [11]. In a next step the most famous visions on 6G and its architecture will be elaborated.

4.3.1 Visible Light Communication (VLC)

Visible Light Communication describes systems which use electromagnetic waves in the spectrum of 400 to 790 THz [26]. These are even higher frequencies than microwaves and infrared. VLC could give solutions for 5G limitations as the optical spectrum bandwidth is extremely high. Communication through subterahertz and visible light links might support data rates up to terabits per second [2]. An advantage of communication trough light is that there is no electromagnetic interference with other frequencies. Also, the power consumption is lower compared to the production of radio or microwaves. Furthermore, it is a free abundant unlicensed spectrum and has a very high frequency reuse [2]. Because of this, VLC methods are believed to be powerful enablers and have huge potential for improvement in the future. Visible light cannot penetrate the skin which means, that there are no health hazards connected with it. However, for VLC communication to work, a direct line of sight connection is required. This is not the case for conventional communication methods.
Current wireless mobile networks are still far away from integrating Visible Light Communication. Regarding VLC there is also research going on in other sectors. Scientists are also trying to use visible light for Vehicle to Vehicle Communication, Underwater Communication, Visible light ID systems and Wireless local area networks (WLANs). The progress made in the different sectors for VLC, will go hand in hand with each other. But before mobile networks use visible light as a communication medium, WLANs will definitely make the step first. WLANs can be set up using full duplex LED based visible light communication. LED can be used when both communication and illiumination have to be performed simultaneously using the same device [43].

4.3.2 Terahertz Communication

Terahertz Waves fall into the spectrum of 95GHz to 3THz. It lies between the mmWave and far-infrared bands. For the longest time it stayed one of the least investigated spectrum. This has definitely changed in recent years. For beyond-5G or 6G networks higher frequencies over the terahertz band will be ubiquitous [27]. For ultrafast speeds and minimal response time, advanced versions of 5G already use millimetre wave bands. But connections only work over short distances. Submillimetre waves like terahertz waves have a even lower range while supporting an ample spectrum above hundred Gigabit per-second [27]. The THz band is thought to be the future of wireless communication. This is because it offers higher transmission bandwidths compared to the mmWave spectrum, while offering more favourable propagation settings when compared to the IR band.
With novel networking approaches, the goal is now to harness waves in the terahertz spectrum [10]. To overcome the limitations of the high propagation loss, reconfigurable antenna arrays, high-gain directional antenna systems, metasurfaces for smart radio environments are being developed [28]. Furthermore, scientist focus on understanding the coexistence of mmWaves, terahertz Waves and optical wireless communications, as this is not yet fully understood [27]. The integration of terahertz systems with optical infrastructure will gain great importance in the future. Some even describe using a ultra-broadband plasmonic modulator for THz-to-optical conversion in wireless communications [28]. Wireless communication over terahertz waves shall also provide massive connectivity, denser networks and highly secure transmissions.

At the moment, terahertz communication cannot be used on a commercial basis. However, there is progress with lasers on the lower spectrum of the terahertz band. Researchers at Osaka university were able configure a two-channel terahertz transmitter by modulating the output of a laser pair. The frequency difference was set in the 300-GHz band. This way researchers were able to transmit 8K videos which require really high data rates, with low latency and low power consumption. When using microwaves or millimeter waves it is necessary to compress the data before transmitting it. This leads to a higher power consumption and delays in transmission. With the use of lasers that are able to transmit information with terahertz waves it is not necessary anymore to compress the signal beforehand. This way researchers successfully transmitted uncompressed 8K videos which is equivalent to 48 Gbit per second [41]. Possible technologies which might be used to harness even higher datarates are the photomixer and the quantum cascade laser. The photomixer is a photonic device which converts optical radiation to terahertz waves by coinciding two laser beams with eachother. A frequency with the difference of the two base frequencies it the result. When coupled with an antenna, photomixers can be used to produce custom terahertz frequencies. Additionally, the quantum cascade laser is also a photonic technology. "Unlike a typical laser where the electron falls from a higher energy level to a lower energy level, the electrons in this laser "cascade" through a sequence of energy levels before reaching the ground state. [42]" Typical slices of quantum cascade lasers are aluminum indium arsenide and gallium indium arsenide stacked periodically on an indium phosphide substrate. The product of this periodicity are the cascading energy states. Photonics are one of the driving technologies for terahertz communications as they are able to initiate the highest data rates so far [42]. That is why I think there is a high possibility of photonic technologies beeing in use for the next generation of networks.

4.3.3 Artificial Intelligence

Research which resolves around how to exploit and explore the full potential of artificial intelligence technologies is a really hot topic around the world. Especially when it comes to using AI for future wireless networks like beyond 5G and 6G [10]. There are two sides that come with the employment of AI in wireless communications. On the one hand, it empowers intelligent resource management trough automatic adaption and powerful learning capabilities. On the other hand, current network architecture and system models do not support the use of AI in wireless communication resource management. Embracing Artificial Intelligence in wireless communication would need a restructuring of the network architecture [5]. The advances made in artificial intelligence are rather incomprehensible for most people. But basically there's two main concepts in AI. Machine Learning and Deep Learning.

4.3.3.1 Machine Learning

Machine Learning describes algorithms that analyse data. These algorithms can learn from their analysis and make decisions based on what they learned. Machine Learning are designed to work like virtual personal assistants. It affects many different industries as Machine Learning is able to perform a multitude of automated tasks that are useful [29].
For wireless communications, machine learning could get really useful for many different applications. Future networks might be able to recognize what kind of service it has to deliver to different users based on past interactions with the network. This is called service classification. Machine learning could also solve the access congestion problem by smart allocation of resources. This should preserve the network from getting overwhelmed. Another application where machine learning could get really useful is the optimization of

radio interference. The network has to know which frequencies interfere with each other and the algorithm would make the network adjust itself accordingly [30]. These algorithms require complex math and coding to eventually get the feature which is expected. Another requirement is lots of data for the algorithm to analyse. In practice there's often not enough data for analysis for the algorithm to work correctly. In that case artificial intelligence has to become smarter than just a machine learning algorithm. Then we are talking about Deep Learning [29].

4.3.3.2 Deep Learning

Deep learning can be seen as a subset of machine learning and is also called in-depth learning. The main difference between machine and deep learning is that the deep learning algorithm gets better with every calculation. The algorithm tries to forecast if its calculation will be correct or not. And learns from the outcome. Deep Learning algorithms keep trying to analyse the data with certain logical structures and thus draw conclusions similar to a human. These algorithms use a layered structure of algorithms which is called a neural network. The basic design of a neural network can be compared to a human brain. Just as we use our brains to identify patterns and classify different types of information, deep learning algorithms can be taught to perform the same tasks for machines. Whenever the algorithm receives new information, it tries to compare it with known objects. Just like a human would. This makes in-depth learning much more powerful than machine learning. This obviously sounds really promising, however the hard part is to keep the algorithms from drawing wrong conclusions, and this requires precise coding[29].
When it comes to making use of deep learning algorithms for wireless communication networks there are many potential applications. Deep learning has received significant research interests for safety critical services like vehicular communications. The goal is an intelligent algorithm which optimizes the communication between vehicles. This is an important step to make autonomous driving safer. Deep learning algorithms could also be used to optimize beamforming. Adaptive beamforming describes the prediction of how much data a user will require. In a next step the active antenna reallocates its resources to the beams which are directed to users with higher potential data usage [20]. Another application where a deep learning algorithm might be involved is smart routing for unmanned aerial vehicles.

4.3.4 Unmanned Aerial Vehicles (UAV)

Unmanned aerial vehicles are aircrafts which are piloted by remote control or on-board computers [31]. In terms of network coverage, 6G will no longer be limited to the ground. It shall achieve seamless connection of ground together with satelite and airborn networks [32]. Harnessing UAV's as flying base stations could help to meet coverage and capacity requirements foreseen in future wireless networks [33]. They could be of great use in scenarios like disaster response or temporary hotspots. There's extensive research going on for the integration of UAV's as base stations to build dynamic networks and improve the conventional dynamic structures [11]. However the step to transform the current two dimensional network to three dimensions is still an open challenge [33]. If there's a breakthrough in the field of UAV's it will be in close combination with deep learning algorithms as resource allocation. Also route optimization will play an important role for dynamic networks which also rely on artificial intelligence.
Unmanned aerial vehicles are already in use in different sectors. During the COVID-19 pandemic for example some states used UAV's to possibly spot illegal gatherings of people. Unmanned Aerial Vehicles are also in use for aerial surveillance and many more applica-

tions [45]. If and when UAV's will be used as base stations for wireless communication depends on how fast progress is made in other sectors first.

4.3.5 Quantum Communication

Computational capability of systems has increased drastically over the last years because of increasing demands. To meet the need of even higher data rates and calculation speed the concept of quantum computing could definitely have the potential to outperform conventional computing systems. "This immense power of QC comes from the fundamental concepts of quantum superposition, quantum entanglement, or the no-cloning theorem" [1]. These systems are purely based on quantum mechanics concepts which are the interaction of molecules, atoms, photons and electrons. Quantum bits are subatomic particles like electrons & photons and are also called qubits. With the help of lasers consisting of microwave beams scientists manipulate these so called qubits.
Superposition describes the fact that groups of qubits are able to represent multiple combinations of zeros and ones at the same time. While conventional group of bits obviously only embody one combination at a time. This means that a group of quantum bits have a higher processing power than the same number of binary bits.
Quantum entanglement describes the phenomenon that qubits can be entangled with each other. Even if you take them apart, the change of one qubit changes the other qubit as well. This has born the idea to make use of this entanglement and use it for quantum communication. However, there are many difficulties to tackle. One of them is decoherence. Decoherence describes that the quantum behaviour will decay easily trough influences from the environment because the quantum state is extremely fragile. This makes quantum computing a really complex challenge as qubits require exact environmental conditions. Quantum Computing assisted communications is a relatively new research area and might be used in future 6G or beyond communications. Quantum Communication is envisioned to achieve extremely high data rates which aren't possible with conventional technology yet[1]. The fact that the quantum state is extremely fragile makes quantum communication theoretically unhackable. Because intruders are not able to read data without leaving a trace. If a hacker theoretically was able to read data it would change the quantum state and the receiver would notice.
An important research sector when it comes to quantum information technologies are quantum networks and protocols. Quantum networks could possibly get used for a global scale internet or quantum key distribution protocols. Quantum key distribution as we know it works by sending photons trough optical cables. This is also the first approach to construct quantum networks. However, the signal needs to be protected against noise and decoherence. Also complex quantum error correcting codes will have to be used. Nonetheless, to achieve high security wireless quantum key distribution might be integrated into future wireless networks. Because classical cryptography systems have limitations. Using mathematical algorithms they can be hacked and there is no safe way of distributing keys. That is why in indoor environments, wireless quantum key distribution is a viable possibility. At the moment, wireless mobile QKD only exist as prototypes on a small scale. But overcoming the challenges of wireless quantum communication would be a break trough. Quantum enabled security could be added to wireless optical communication technologies and make wireless communication a lot safer. For 6G or later wireless networks this would mean the installation of QKD receivers on the ceiling, while mobile phones work as the transmitter. This is a possible future scenario for 6G [46]. When it comes to constructing wireless networks based on quantum mechanics, entanglement also sounds promising. Quantum entanglement is independent of the underlying physical channel configuration. The entangled state could allow several interesting features like a direct entanglement between two parties. Quantum repeaters are needed to establish

long-distance entanglement and will probably also be used if quantum wireless networks will ever be based on entanglement [44].

4.4 Future Services

In this section two technologies which are strongly coupled with the development of 6G will be presented. Extended Reality and Autonomous Driving will improve and evolve over time. The advancement in these sectors will go hand in hand with the evolution of the next generation network.

4.4.1 Virtual & Augmented Reality (VR & AR)

The term mixed reality (MR) includes both terms virtual & augmented reality. Mixed Reality describes services where our natural perception is supplemented by an artificially generated perception. MR creates completely new application possibilities as environments can be created for users that did not exist before.
The term augmented reality is used when physical reality is expanded to include virtual elements. This could be things which fit into a filmed image. For example additional information or filters on apps like snapchat [34]. Virtual reality on the other hand is a fully computer-generated interactive environment. For Virtual reality, appropriate VR glasses are required which completely cover the field of vision of the user. To sum up, the main difference is that in AR the real environment remains visible to users while VR completely shuts out the real environment [35]. Extended reality is an umbrella term that includes all technologies which enhance our senses and therefore include AR, VR and MR [34].
For services like AR and VR a really strong network is required. An interactive VR application will require between ten and twenty times the storage capacity used by standard HD video files. This could generate around a terabyte of data per hour. While VR applications have incredible potential, current wireless networks are struggling to handle such a high dataflow. Furthermore, the amount of detail within VR will keep increasing over time, which results in a higher data-flow [36]. However, to enhance a VR experience not only a high bandwidth is required. To minimize collision rates almost undetectable latencies are required. And this is exactly what 6G aims for [11].

4.4.2 Autonomous Driving

Autonomous driving is already becoming a reality. There are already cars and buses which function without a driver. Autonomous driving systems have many on board sensors and capture a large amount of data. They are able to sense surroundings near-real-time [38]. Commercials in the auto industry promise an urban paradise. Autonomous taxis with safe pedestrian detection and buses which are always on time. And everything silent and electric. The potential of this technology is enormous. Integrating older or handicapped people would be easy. Rural areas could be reached easily because the price for automated taxis or buses would be really low. Autonomous driving would also be a great way to reduce the amount of car accidents. This obviously depends on the degree of automation. Nine out of then car crashes are based on human error. If the majority of cars run by software and even communicate with each other the amount of car crashed could be reduced drastically. However, until the large majority of cars are not automated, accident reduction will not work as well [39].
The fifth generation of Mobile communication is being rolled out at the moment. At the same time vehicular technologies are evolving at a fast pace. Connected vehicles, also

called "V2X" are vehicles which are able to communicate with different entities. This can be other vehicles or traffic lights for example. The goal will be autonomous vehicles which use communication technologies to communicate with different entities to optimize the automation of driving. Also the sixth generation of wireless networks is at the advent. The evolution of those two research sectors, 6G and autonomous driving, will go hand-in-hand and depend on each other strongly [40]. The network which is used to communicate in autonomous vehicles has to be extremely reliable and fast at the same time, as it is a safety critical service. This definitely are goals for the next generation wireless network.

4.5 Discussion

There is a wide array of technologies for future networks. The evolution of 5G is not entirely clear. 5G can be seen as an organic entity which evolves over time [36]. In line with the speed of evolution new technologies are being introduced. The current version of 5G is release 16. There are already plans how release 17 may look like. It will likely incorporate enhancements for IoT. It may incorporate advanced applications beyond the smartphone use.
The sixth generation of wireless networks on the other hand is not under construction yet. At the moment 6G is still in the conceptual phase. As data traffic grows constantly, network requirements increase. A wider broadband calls for higher frequencies on the electromagnetic spectrum. The shorter range of the high frequency waves of these waves will be compensated with a higher cell density. It's not clear what the final frequencies will be. Will they be in the range of terahertz waves or will they be in the visible light spectrum? It could even be a combination of the two. The combination of different wavelengths is a plausible possibility. Depending on location and data usage pattern of the user, future networks may switch between different wavelengths and technologies. Researchers even talk about the upgrade of conventional light sources to base stations. Artificial light sources are everywhere in direct line of sight. Which is a prerequisite for visible light communication. In this case LED lamps would serve as a light source and data transmitter at the same time. The network could switch to high frequencies like VLC anytime if required. An automatic switch back to lower frequency communication happens if the user is out of range for VLC. This is comparable to the alternative use of 4G and WLAN like we know today. This way, VLC could provide fast and reliable network connection in dense areas like shopping malls or libraries. At the same data traffic could be taken away from lower frequencies which may be needed in more isolated areas. Like most communication technologies, WLANs will probabbly utilize VLC before mobile wireless networks will.
It is possible, that there is a break trough with quantum communication. This could change the fundamentals of how future networks are built. However, quantum communication still has a long way to go. If future mobile wireless networks would ever rely on quantum communication it would probably be done using quantum entanglement with the use of quantum repeaters. I personally think, before that happens scientist would establish a non-wireless quantum internet. And only after that, wireless quantum network stack and protocols for reliable entanglement-based networks are viable. Network and virtual reality technologies will develop fast over time and could possibly be applied for distant education, telecommuting, and advanced three-dimensional simulation [37].

4.6 Summary

We firstly asked ourselves if it is time already to think about the next generation of wireless networks. By listing some upcoming services like holographic communication or

high precision manufacturing it is clear that a network beyond 5G is a must. Furthermore, sustainable development, smart environments and the urge to keep improving energy efficiency make 6G indispensable. There are different approaches on how to tackle rising network requirements. Technologies that rely on the electromagnetic spectrum will have to increase in frequency to match higher data rates. Millimetre waves will probably be used already for later versions of 5G. Their propagation characteristics are significantly different when compared with microwave frequency bands. Fortunately they enable larger amounts of antenna elements and with it higher data rates. The next band on the electromagnetic spectrum is the terahertz band. The range of signals without disruption decreases with increasing frequency. Terahertz waves are a hot topic when it comes to 6G and have been an important research subject in recent years. Together with the visible light spectrum they shall enable extreme data rates on the scale of terabits per second. To make this possible technologies like (massive)-MIMO and beamforming will play an important role in the future. Network Function Virtualization, Network Slicing and Software Defined Networking will be used by default for future networks.

The use of machine- and deep learning to optimize networks will further increase. Machine learning could be used to recognize patterns on how users interact with the network. Based on these patterns the network automatically adjusts itself for optimal reception and speed. With service classification, machine learning will massively contribute to solving the access congestion problem by reallocating resources beforehand. The more powerful tool "deep learning", which can be compared to learning processes of the human brain, will keep improving. Deep learning could revolutionize many different applications like adaptive beamforming, vehicle communication or smart routing of unmanned aerial vehicles. Networks that are able to use satellites or UAVs as base stations can be seen as a three dimensional network which definitely is a goal for the future.

A completely different research sector is quantum communication which is based on quantum mechanic concepts. With phenomenons like superposition and quantum entanglement, unthinkable data rates and minimal latencies might be possible. However, the research field is vastly complex. The future of wireless networks depends on how fast scientists will make progress when it comes to quantum communication. The speed of development in quantum communication will decide if this technology is appropriate for 6G.

Bibliography

[1] Nawaz, Syed Junaid, Shree Krishna Sharma, Shurjeel Wyne, Mohammad N. Patwary, und Md. Asaduzzaman: *Quantum Machine Learning for 6G Communication Networks: State-of-the-Art and Vision for the Future*, IEEE Access 7 (2019): 46317–50, 2019. https://doi.org/10.1109/ACCESS.2019.2909490.

[2] Calvanese Strinati, Emilio, Sergio Barbarossa, Jose Luis Gonzalez-Jimenez, Dimitri Ktenas, Nicolas Cassiau, Luc Maret, und Cedric Dehos: *6G: The Next Frontier: From Holographic Messaging to Artificial Intelligence Using Subterahertz and Visible Light Communication*, IEEE Vehicular Technology Magazine 14, Nr. 3, September 2019. https://doi.org/10.1109/MVT.2019.2921162.

[3] *What is holographic telepresence*; https://whatis.techtarget.com/definition/holographic-telepresence, April, 2021.

[4] *What is 5G | Everything You Need to Know About 5G | 5G FAQ | Qualcomm*; https://www.qualcomm.com/5g/what-is-5g, April, 2021.

[5] M. Lin and Y. Zhao: *Artificial intelligence-empowered resource management for future wireless communications: A survey*, in China Communications, vol. 17, no. 3, March 2020. doi:10.23919/JCC.2020.03.006.

[6] Li, Bo, Peng Hou, Hao Wu, und Fen Hou: *Optimal Edge Server Deployment and Allocation Strategy in 5G Ultra-Dense Networking Environments*, Pervasive and Mobile Computing, 2021.

[7] *All about 4G LTE Technical Training*; http://www.techtrained.com/2-mins-read-on-ultra-dense-networks/, April, 2021.

[8] *Bandwidth and Data Rates*; https://www.flukenetworks.com/blog/cabling-chronicles/bandwidth-and-data-rates, April, 2021.

[9] *5G vs 4G: Everything You Need to Know*; https://www.lifewire.com/5g-vs-4g-4156322, April, 2021.

[10] Akhtar, M.W., Hassan, S.A., Ghaffar, R. et al.: *The shift to 6G communications: vision and requirements*, Hum. Cent. Comput. Inf. Sci. 10, 53, 2020. https://doi.org/10.1186/s13673-020-00258-2.

[11] Yang, Ping, Yue Xiao, Ming Xiao, und Shaoqian Li: *6G Wireless Communications: Vision and Potential Techniques*; IEEE Network 33, Nr. 4, 2019, https://doi.org/10.1109/MNET.2019.1800418.

[12] Shafi, Mansoor, Jianhua Zhang, Harsh Tataria, Andreas F. Molisch, Shu Sun, Theodore S. Rappaport, Fredrik Tufvesson, Shangbin Wu, und Koshiro Kitao: *Microwave vs. Millimeter-Wave Propagation Channels: Key Differences and Impact on 5G Cellular Systems*; IEEE Communications Magazine 56, Nr. 12, December, 2018, https://doi.org/10.1109/MCOM.2018.1800255.

[13] *Millimeter Wave - an overview | ScienceDirect Topics*; https://www.sciencedirect.com/topics/engineering/millimeter-wave, April, 2021.

[14] Goleniewksi, Lillian: *Telecommunications Essentials: The Complete Global Source for Communications Fundamentals, Data Networking and the Internet, and Next-Generation Networks*; 2001.

[15] *MIMO - Multiple Input Multiple Output (TGnSync)*; https://www.elektronik-kompendium.de/sites/net/1004251.html, April, 2021.

[16] *Massive MIMO: Antennen im Schwarmmodus*; https://www.t-systems.com/de/blickwinkel/netze/multiple-input-multiple-output/massive-mimo-800124, April, 2021.

[17] Brown, Gabriel: *White Paper: Exploring 5G New Radio: Use Cases, Capabilities & Timeline*; 2016.

[18] Vadym Slyusar: *What Is the Difference between MIMO and Massive MIMO?*; https://www.researchgate.net/post/what_is_the_difference_between_MIMO_and_massive_MIMO, April, 2021.

[19] Fruhlinger, Josh: *Beamforming Explained: How It Makes Wireless Communication Faster*; https://www.networkworld.com/article/3445039/beamforming-explained-how-it-makes-wireless-communication-faster.html, April, 2021.

[20] *Wie funktioniert Beamforming mit 5G? | Deutsche Telekom*; https://www.telekom.com/de/blog/netz/artikel/beamforming-5g-mobilfunk-570522, April, 2021.

[21] Luber, Stefan: *Was ist NFV?*; https://www.ip-insider.de/was-ist-nfv-a-605864/, April, 2021.

[22] *What Is Network Function Virtualization (NFV)*; https://www.blueplanet.com/resources/What-is-NFV-prx.html, May, 2021.

[23] *Was ist Software-defined Networking (SDN)? - Definition von WhatIs.com*; https://www.computerweekly.com/de/definition/Software-defined-Networking-SDN, April, 2021.

[24] Luber, Stefan: *Was ist mMTC (Massive Machine Type Communications)?*; https://www.ip-insider.de/was-ist-mmtc-massive-machine-type-communications-a-828903/, April, 2021.

[25] *Emilio Calvanese Strinati – 6G Wireless Summit 2020*; http://www.6gsummit.com/speakers-2/emilio-calvanese-strinati/, April, 2021.

[26] Dr. Hasan Farahneh: *Visible Light Communication for 6G Technology : The Potential and Research Challenges*; Type, Month, Year, https://asrenorg.net/eage19/sites/default/files/files/Visible%20light%20s%20as%20a%20promising%20spectrum%20for%206G%20technology(1).pdf.

[27] *Terahertz-enabled Wireless Communications in 6G: Opportunities, Challenges and Trends | Frontiers Research Topic*; https://www.frontiersin.org/research-topics/14846/terahertz-enabled-wireless-communications-in-6g-opportunities-challenges-and-trends, May, 2021.

[28] Manzalini, Antonio: *Quantum Communications in Future Networks and Services*; 2020.

[29] *Deep Learning vs Machine Learning - Was ist der Unterschied?*; https://www.it-talents.de/blog/it-talents/deep-learning-vs-machine-learning-was-ist-der-unterschied, April, 2021.

[30] Mosiane, Olorato, Nadeem Oozeer, Arun Aniyan, und Bruce A. Bassett: *Radio Frequency Interference Detection Using Machine Learning*; IOP Conference Series: Materials Science and Engineering 198: 012012, May, 2017, https://doi.org/10.1088/1757-899X/198/1/012012.

[31] *Unbemanntes Luftfahrzeug*; https://de.wikipedia.org/w/index.php?title=Unbemanntes_Luftfahrzeug&oldid=211096564, April, 2021.

[32] *In the 6G Era, Holographic Interaction Has Become a New Application Scenario, and WIMI Will Become the First Stock of Holographic AR*; https://www.globenewswire.com/news-release/2021/03/08/2188690/0/en/In-the-6G-era-holographic-interaction-has-become-a-new-application-scenario-and-WIMI-will-become-the-first-stock-of-holographic-AR.html, April, 2021.

[33] R. Arshad, L. Lampe, H. ElSawy and M. J. Hossain: *Integrating UAVs into Existing Wireless Networks: A Stochastic Geometry Approach*; 2018 IEEE Globecom Workshops (GC Wkshops), 2018, 10.1109/GLOCOMW.2018.8644504.

[34] *What's the Difference Between AR, VR, and MR?*; https://www.fi.edu/difference-between-ar-vr-and-mr, April, 2021.

[35] *Microsoft erklärt: Was ist Mixed Reality? Definition & Funktionen*; https://news.microsoft.com/de-de/microsoft-erklaert-was-ist-mixed-reality-definition-funktionen/, April, 2021.

[36] *The Evolution of 5G*; https://www.forbes.com/sites/bobodonnell/2019/11/12/the-evolution-of-5g/, April, 2021.

[37] *In the 6G Era, Holographic Interaction Has Become a New Application Scenario, and WIMI Will Become the First Stock of Holographic AR*; https://www.globenewswire.com/news-release/2021/03/08/2188690/0/en/In-the-6G-era-holographic-interaction-has-become-a-new-application-scenario-and-WIMI-will-become-the-first-stock-of-holographic-AR.html, April, 2021.

[38] Author: Yang, Bo and Cao, Xuelin and Xiong, Kai and Yuen, Chau and Guan, Yong and Leng, Supeng and Qian, Lijun and Han, Zhu *Edge Intelligence for Autonomous Driving in 6G Wireless System: Design Challenges and Solutions*; December 2020.

[39] *Autonomes Fahren: Digital entspannt in die Zukunft*; https://www.adac.de/rund-ums-fahrzeug/ausstattung-technik-zubehoer/autonomes-fahren/technik-vernetzung/aktuelle-technik/, May, 2021.

[40] Author: J. He, K. Yang and H. -H. Chen *6G Cellular Networks and Connected Autonomous Vehicles*; November 2020, 10.1109/MNET.011.2000541.

[41] *Terahertz Accelerates beyond 5G towards 6G*; https://resou.osaka-u.ac.jp/en/research/20210201_1, May, 2021.

[42] *The light way to 6G*; https://spie.org/news/photonics-focus/janfeb-2021/light-way-to-6g, May, 2021.

[43] Author: Khan, Latif Ullah *Visible Light Communication: Applications, Architecture, Standardization and Research Challenges*; May, 2017, https://doi.org/10.1016/j.dcan.2016.07.004.

[44] Author: Pirker, A., und W. Dür *A quantum network stack and protocols for reliable entanglement-based networks*; March, 2019, https://doi.org/10.1088/1367-2630/ab05f7.

[45] *List of Unmanned Aerial Vehicle Applications*; https://en.wikipedia.org/w/index.php?title=List_of_unmanned_aerial_vehicle_applications&oldid=1025849482, April, 2021.

[46] Author: Elmabrok, Osama, und Mohsen Razavi *Wireless Quantum Key Distribution in Indoor Environments*; Februar, 2018, https://doi.org/10.1364/JOSAB.35.000197.

[47] *5G-Antennen: Umstrittene Strahlen-Messung*; https://www.srf.ch/play/tv/kassensturz/video/5g-antennen-umstrittene-strahlen-messung?urn=urn:srf:video:371c2934-9f7a-4fef-8ffb-b75bf8c93ea0, May, 2021.

Chapter 7

A Survey on Quantum Communication Networks

Norina Braun

Contents

7.1 Introduction

This seminar report serves as an introduction to a future of quantum communications. It explains the basics of quantum mechanics, such as quantum superposition, quantum entanglement and teleportation, on a highly abstracted level and focuses on the results rather than the means by which it is achieved. It further describes functionalities and protocols that would be required to build an operational quantum internet. Additionally a quantum key distribution protocol is introduced and a recent achievement in the field of study of quantum communications is presented. Finally a vision for the future is proposed where quantum networks could be ubiquitous.
Whilst the general idea of quantum computing and some algorithms have been around for several decades now, it is still a relatively new field of study and the focus of many research teams around the world. Especially the early scientist working in this field have been heavily hindered by technical limitations, such that in the beginning many concepts were little more than thought experiments. Due to technical advancements, however, this has changed and contemporary researches can conduct their studies in a much more hands-on approach.

7.2 Fundamentals of quantum computing

To appreciate the technical challenges and opportunities quantum computers hold for the future, a basic understanding of quantum mechanics is necessary. This section contains an introduction to some key concepts of quantum computing and quantum communication, thereby building the foundation to comprehend the more advanced sections of this report. This section, including subsections 7.2.1 - 7.2.3 is based on [1].

7.2.1 Keyword definitions

For understanding later sections of this report, a few of the most important keywords to understanding quantum computing are defined in this paragraph. **Decoherence** denotes the uncontrolled change of a quantum state. It happens naturally over time or due to interactions with the environment and decreases the likelihood that a given qubit will collapse to a desired state. Another important concept is **fidelity**, a quantification of the quality of a qubit measurement. The closer the measured state of a qubit is to the expected state, the higher the fidelity. The concept of fidelity is closely related to decoherence. If a qubit has a high decoherence, it has a low fidelity and vice versa. In the literature a **quantum operation** is described as a unitary operation that is applied to a qubit in a superposition. It is always reversible.

7.2.2 Quantum superposition

In a classical computer, binary digits or bits are used to represent information. Classical bits can either have a value of 0 or 1, that's why they are called binary.
The quantum equivalent of bits are quantum bits, or qubits. Qubits can also be in a state of 0 and 1, denoted with $|0\rangle$ and $|1\rangle$, but they can also be in both states at the same time, which is called a superposition of the two basis states $|0\rangle$ and $|1\rangle$. This is an inherent property of the quantum world and the superposition is destroyed once the qubit is observed. So whenever a qubit is measured, it will collapse into either of the two basis states. The final state it will end up in is not predefined, but rather random.
The state of an arbitrary qubit can be described in the following way:

$$|\phi\rangle = \alpha|0\rangle + \beta|1\rangle$$

where ϕ denotes the state of the qubit, $|0\rangle$ and $|1\rangle$ are the two base values and α and β are probability factors, where $\alpha^2 + \beta^2 = 1$.
The latter describe how likely it is for the qubit to fall back to the according states once measured. So with probability α the final state of the qubit will be $|0\rangle$ and with probability β the qubit will end up as a $|1\rangle$. Naturally the distribution would be uniform, so about half the qubits would collapse to a $|0\rangle$ and the other half to a $|1\rangle$. However, it is possible to manipulate the qubits, such that one state is much more likely to be the outcome than the other. This is how information can be encoded in a qubit.
There are several different physical entities that can be used to store quantum information. One possibility are electrons. Electrons naturally are assigned four different quantum numbers, when they belong to an atom or molecule in order to uniquely identify them. One of them is the so called spin quantum number and describes the angular momentum of the electron. This quantum number is binary, it can be $\pm\frac{1}{2}$, or worded differently, an electron can have spin up or spin down. Since electrons live in the quantum realm, their spin is not predetermined, but rather they are in a superposition of both spin up and spin down until measured. Therefore the spin property of an electron can be used to store a qubit of information.
Another possibility is to utilize photons. A photon, as a light wave, can be polarized, meaning its wave function is restricted to a single plane. Light can be polarized vertically or horizontally. Or, more broadly speaking, a photon encoding two basis states, will need two planes for the wave function, which are perpendicular to each other. In this case, a photon, as a qubit, is in a superposition of both vertically and horizontally polarized light and once it passes through a filter for measurement, it will collapse to one single plane.

7.2.3 Quantum entanglement

Two qubits can be entangled together in such a way, that their individual states are not independent of each other. This has the logical consequence, that measuring one qubit of a pair of entangled qubits immediately reveals information about the state of the other.
Electrons in an atomic orbital are an example of this phenomenon. There is only ever space for two electrons in a sub-shell of an atom and they have to have opposite spin. So their spin state is entangled. As long as they are not observed, both are in a superposition of spin up and spin down, but as soon as one of them is measured to have spin up, the superposition of the other collapses and it now has spin down.

7.2.4 Bell States and Measurement

Since a single qubit can be in a superposition of both its base states, two entangled qubits are according to [2] also in a superposition of base states. As there are two qubits, that can both be in two states at once, there are now four base states. They are also called Bell states and can be described as follows:

$$|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$
$$|\phi^-\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle)$$
$$|\psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$$
$$|\psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$$

In the ϕ-states the qubits are entangled in such a way that both of them have the same state when measured and in the ψ-states, they have the opposite state. Any measurement

of two entangled qubits will collapse their superposition into one of the four Bell states. Such a measurement is called Bell measurement and it is a reversible operation.

7.2.5 No-Go Theorems

The so-called no-go theorems [3] are inherent rules of quantum mechanics. There are several different no-go theorems, three of which are mentioned subsequently. The **no-encoding theorem** states, that it is not possible to superpose an unknown pure state and a fixed pure state. The **no-cloning theorem** on the other hand forbids perfectly cloning an unknown pure state with a universal quantum transformation. This does not imply that it is impossible to have copies of a qubit, just that the person creating the copy must know the state of the qubit. Finally, the **no-deleting theorem** says, that it is impossible to delete a copy of an unknown state while leaving another untouched

7.2.6 Quantum supremacy

Quantum supremacy is a way to prove the superiority of a quantum computing device compared to a regular super computer. It is said to be achieved, when a quantum computer can perform a task that would take a normal computer an unreasonable amount of time.[4]

7.2.6.1 Google's Sycamore processor

In 2019, Google claimed that their Sycamore processor [5] had achieved quantum supremacy. To prove it, they sampled the output of a pseudo-random quantum circuit, which outputs bit-strings. Due to quantum interference within the circuit, the distribution of the bit-strings is not uniform, and some appear more often than others.
The task was to find this probability distribution and it took their processor only 200 seconds to calculate a result. They estimated that a regular super computer would take over 10'000 years to come to the same conclusion, so quantum supremacy was achieved.
IBM later claimed, that the calculation could actually be done on a regular supercomputer within 2.5 days, thus challenging their claim to quantum supremacy. Nevertheless, the study is generally accepted by the scientific community and additionally in 2020 a Chinese research team declared quantum supremacy as well, so the milestone has most definitely been achieved by now.[6]

7.2.6.2 Shor's algorithm

Shor's algorithm [7] is another quantum algorithm that is more powerful than a classical equivalent. It can be used to solve prime factorization in polynomial time. This could potentially be a threat to internet security as one of the most used encryption algorithms, the RSA encryption, heavily relies on the fact that factorizing large numbers is a classically hard problem and takes a long time on a normal computer.
The algorithm was invented in 1994 and has been proven to work, but the largest number that has been factorized using Shor's algorithm is 21, so today's quantum computer are not yet powerful enough to pose a threat to RSA encryption.

7.2.7 Teleportation

Teleportation [8] is used to send quantum information without transmitting any qubits. It uses quantum entanglement as a means to do so. In Figure 7.1 a model of how this could work is shown. Two people, Alice and Bob, share a pair of entangled qubits, A and B. Alice wants to send the information encoded in qubit a to Bob.

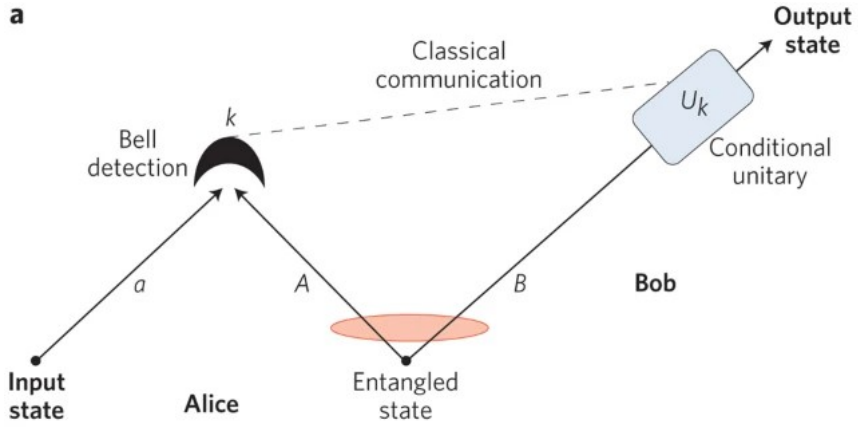

Figure 7.1: Schematic representation of quantum teleportation [8]

She entangled her two qubits a and A and performs a conjoint Bell detection measurement on them. Next, the outcome k of said measurement is communicated to Bob using a classical link. Bob can then take that outcome, perform a conditional unitary, which is the reverse operation of Alice's Bell detection, on his qubit B and transform it into the state that a was in before. The two entangled qubits they started with are consumed in the process, but Alice was able to send quantum information about the state of a qubit over a classical channel.

7.3 Quantum Internet

Once quantum computers and quantum capable devices become more common, the need for a quantum equivalent of today's internet, a so-called quantum internet, [11] will arise. It will likely consist of similar components we know and use today.

7.3.1 Quantum Networks

The quantum internet of the future will need a quantum network [9] to support its functionalities. For the foreseeable future the quantum network will exist alongside the classical one and enhance it. It will contain quantum capable end-nodes, that are connected to the network via quantum links, used to send qubits around, and quantum repeaters with quantum data planes, as intermediate stations, to provide long distance quantum links.

7.3.2 Quantum Memory

To store classical information, a bit can just be read and written onto a storing device. This does not work for qubits however, because any measurement would collapse the superposition and the quantum state information would be lost. Luckily there are ways to store quantum information, by utilizing quantum memory [10]. Considering photons as qubits, it is possible to map their quantum information to the quantum state of a material system, such as warm atomic vapours, ultra-cold atomic gases or rare-earth ion doped crystals.
In that last case, there are impurities in the crystal structure, which come from the ions of rare-earth metals. The photon can then be mapped to an internal degree of freedom of the system, in this case it would be the spin in the ions. There are also other material systems and more research is conducted to increase the lifetime of quantum storage.

7.3.3 Quantum Repeater

This section is based on [12] and [13]. Due to the no-cloning theorem any quantum signal cannot be amplified by classical means of reading and reproducing the signal. To solve this problem and ensure the possibility to send quantum information over a longer distance, two remote parties are entangled together via a process called entanglement swapping. Figure 7.2 provides a visualization of how this could be achieved. Both Alice and Bob

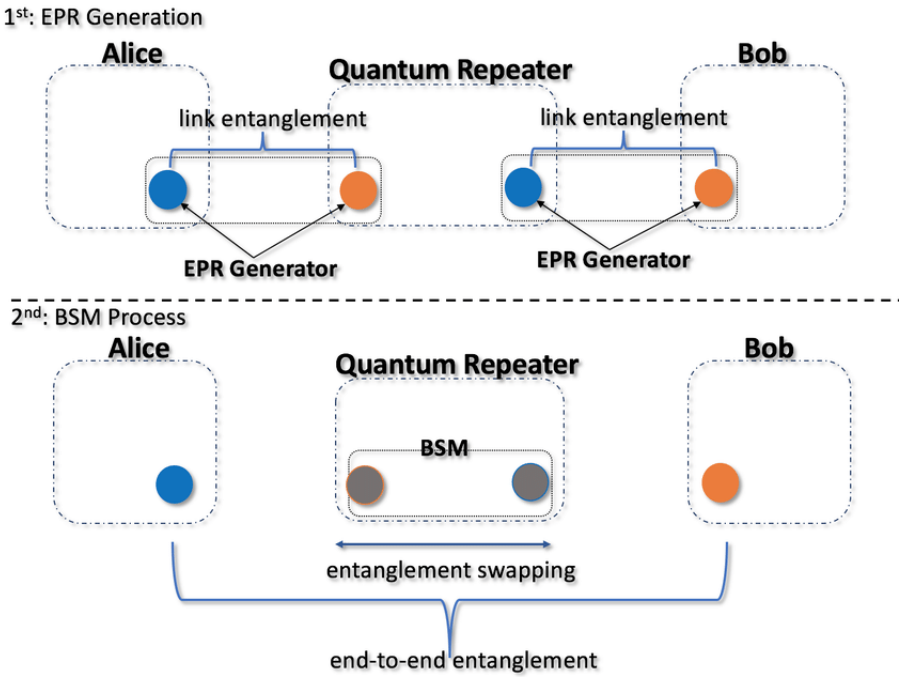

Figure 7.2: Process of entanglement swapping [14]

create an entangled pair of electrons and send one of the electrons to a common quantum repeater. The repeater is now individually entangled with Alice and Bob, but the two of them are not. In the next step, the repeater performs a Bell state measurement on the two qubits it possesses and sends the outcome of the measurement via the classical channel to Alice and Bob. With this the two qubits Alice and Bob possess respectively are now entangled, but the two qubits at the repeater are consumed in the process. Theoretically entanglement could be achieved over arbitrarily long distances using this method, however due to decoherence over time, the process has limitations.

7.3.4 Quantum Network Protocols

Just like current networks, quantum networks will require certain protocols, such that applications can use reliable services to transmit quantum information. Some proposition for these protocols are heavily inspired by today's TCP/IP stacks. [15]

7.3.4.1 Link Layer Protocol

The link layer protocol [15], or quantum entanglement creation protocol (QECP), already assumes the existence of a physical layer protocol. The latter would be responsible for actually entangling a pair of qubits, while the former ensures that the process is robust. Figure 7.3 displays a schema of both a physical and a link layer protocol. The link layer consists of four parts. The queue is there to coordinate the trigger requests to the lower layer. These requests trigger the creation of an entangled pair. The quantum memory management unit (QMMU) determines which physical qubits are to be used to generate and store the entanglement. The fidelity estimation unit (FEU) assesses the quality of

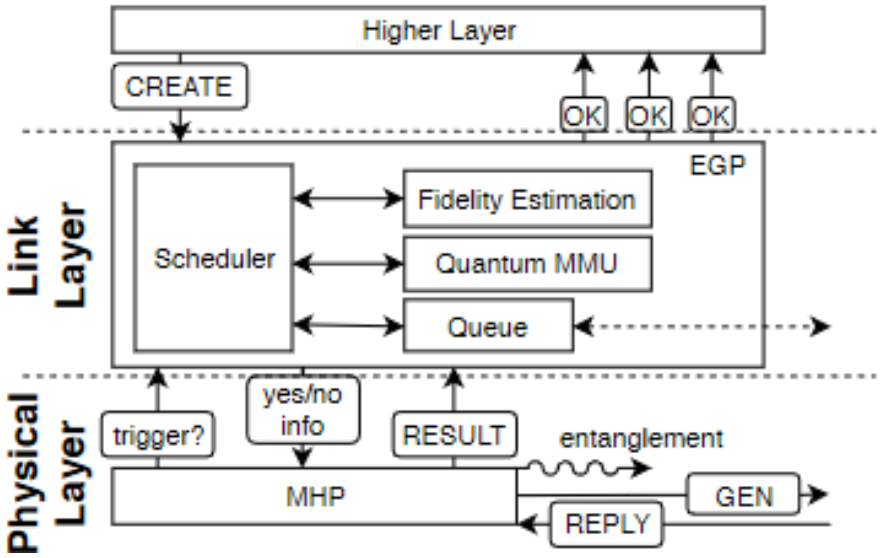

Figure 7.3: Proposed layout of a physical layer and a link layer protocol [15]

the entanglement, while the scheduler handles which requests from higher levels should be prioritized.

7.3.4.2 Network layer protocol

The network layer is responsible to provide long distance entanglement. It will likely not consist of a single protocol, but will contain several protocols [16] that work together and complement each other. The routing protocol is, similar to a classical routing protocol, there to determine an optimal path. It factors in the path length, its cost, the throughput as well as its fidelity. A signalling protocol is in place to create a virtual circuit and to manage the schedule. Finally the quantum data plane protocol is responsible to manage the link-pair generation, to perform the entanglement swapping and keep track of the swaps and to manage the quality of service.

7.3.5 Quantum Error Correction

Error correction is important to enable fault-tolerant quantum computations. There are several challenges when it comes to quantum error correction [17]. A classical approach to the problem doesn't work, as the qubits cannot be copied or measured lest they lose their quantum information.
Furthermore while classical bit flips can occur, so $|1\rangle \rightarrow |0\rangle$ or $|0\rangle \rightarrow |1\rangle$, also phase flips are possible $|0\rangle \rightarrow |0\rangle$ or $|1\rangle \rightarrow |-1\rangle$, resulting in a state change of $(\alpha|0\rangle + \beta|1\rangle) \rightarrow (\alpha|0\rangle - \beta|1\rangle)$. This together with the fact that in the quantum world errors are continuous - the angle of polarization of a photon can shift off by an arbitrary degree, not necessarily a full 90 degree flip - makes quantum error corrections more challenging than normal error correction.

7.4 State-of-the-art technology

7.4.1 Quantum Key Distribution

Quantum key distribution [18] or QKD is one of the first commercializable quantum information tasks. It ensures the creation of a secure secret key between two partners and relies on the fundamentals of quantum mechanics, such as the no-cloning theorem to

ensure that no information can be copied and the fact that any measurement alters the state, so the information cannot be looked at without changing it.
A protocol on how to achieve the distribution of the secret key was proposed by Charles Bennett and Gilles Brassard in 1984. Their protocol utilizes photons as a means to send qubits and the qubits can be measured in two different bases, either the + basis, where the light is horizontally or vertically polarized or the × basis, which is the + basis rotated by 45 degrees. For the premise, Alice and Bob want to have a shared key to encrypt their messages.
Alice begins the protocol by sending some number n photons over a quantum channel to Bob. Alice knows the basis in which she encoded it in and also knows the input value. Bob now measures the incoming photons in either of the two bases. On average they will have used the same basis for $\frac{n}{2}$ of the photons. In the next step they communicate which basis they used for which photon using a classical channel. They discard the values of all qubits where they used different bases and keep the ones where they used the same. These values now make up the raw key. Now they both reveal a part of the raw key to each other and compare values. If all values are equal, as they should be, then the protocol has worked and they have created a secret key. If the error rate is too high it could indicate that an eavesdropper has looked at the qubits and thus changed them and they will abort the protocol and start again.

7.4.2 Integrated Space-to-Ground Quantum Communication Network

This concerns a recent progress that has been achieved in China. A group of researchers has been able to create a quantum communication system that spans over 4600 kilometers that has been used for QKD [19].
Figure 7.4 shows a representation of the network. It consists of a backbone fiber link, used to send photons, four quantum metropolitan area networks, which each have quantum capable end users, two ground-satellites, one of which is directly connected to the backbone fiber link and one quantum satellite. They managed to send quantum information to a remote location using a satellite, which is a promising achievement.

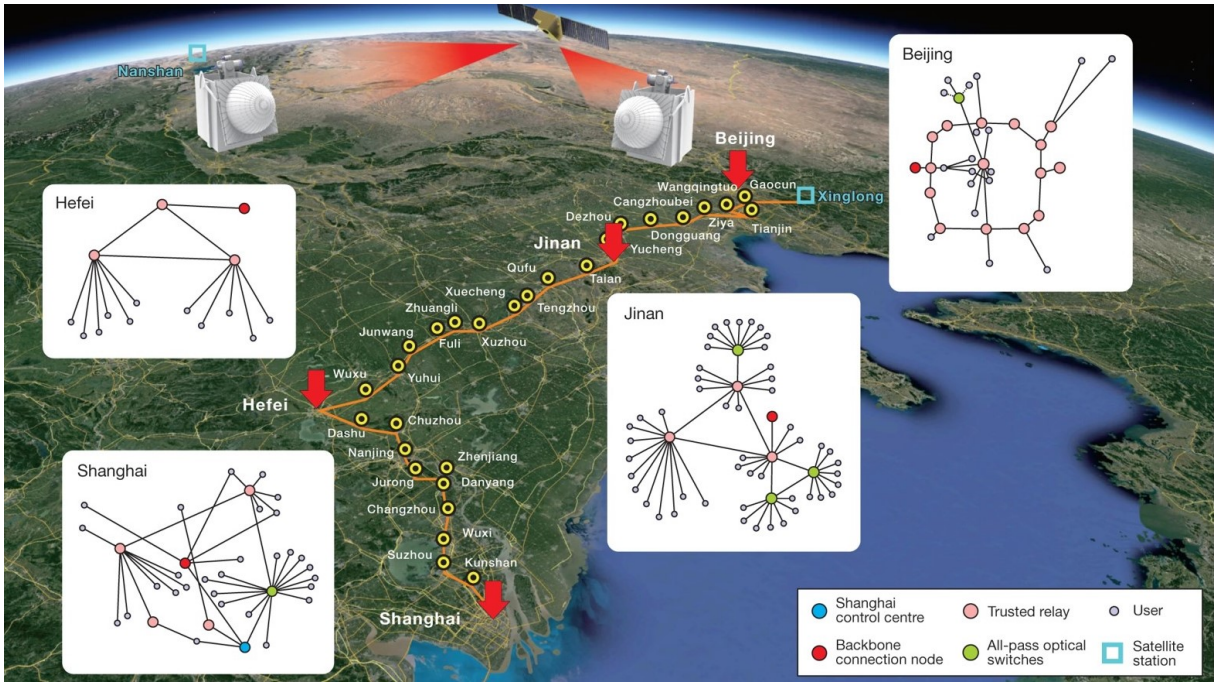

Figure 7.4: Illustration of the quantum network in China [19]

7.5 Future Scenarios

It is difficult to predict to what extent quantum systems will be used in the future, as the research is still in the beginning and many problems must be overcome before it will be in widespread use.

7.5.1 Quantum Cloud Computing

One scenario, which could be possible in the near future, would be access to a quantum computer via a cloud. This could drive innovation forward without the need for the average user to possess a quantum computer. In part, it already exists, as IBM grants access to their quantum computer Quantum Experience via a cloud service. However, this is only accessible for previously approved scientific studies. [21]
One of those studies was the simulation of a deuteron nucleus, consisting of a proton and a neutron.[20] This is a step in the direction of simulating larger atoms and molecules which could potentially revolutionize drug development and healthcare treatment. The size of molecules a classical computer can accurately model is very limited, but quantum computers can simulate much more complex systems. [22]
Especially in the pharma industry, where the relevant molecules are often proteins, consisting of thousands of atoms, the computational power of a conventional computer is just not good enough. Being able to simulate proteins, could provide an insight into whether and how they might react to various molecules, such that one that has a high change of binding to the protein can be chosen for the drug design. The possibility of modelling chemical reactions could prove useful when trying to find new ways to synthesize compounds or even how a drug is processed in the body.
With the creation of more quantum computers and achievements in the development of more powerful quantum computers, they might soon be accessible not only for a few selected research teams, but could be open to more general users. Cloud accessible quantum computers would need to provide access for users with different backgrounds and technical interests.
Quantum physicists will be more interested in the underlying technical detail and would probably like to optimize control of the qubits and improve pulse-shaping used for Fourier-transformations. Information scientists on the other hand, would concentrate on building quantum circuits, using quantum gates and provide basics for error correction, such as parity checks and conditional operations. For quantum developers, their area of interest is mostly on the application side. They will not care how it works, just that the quantum device reliably does what is expected and that it does so within a decent time period. They will create the algorithms and programs, that can then be used by the public and provide an interface for less educated users to utilize the power of quantum computers without having to understand anything about the technical details of them.[21]
For the time being these quantum computers are not accessible to the public, there are, however, other methods to drive innovation in forward in the realm of quantum programs. Some companies provide simulators for developing quantum software, where aspiring quantum developers can code for quantum applications without ever needing a real quantum computer. [23]

7.5.2 Quantum Optical Twin

The quantum optical twin [24] or QOT is a much more futuristic vision. It builds on the concept of Digital Twin Computing, a future proposed by the Nippon Telegraph Cooperation based in Tokyo as a part of their innovative optical and wireless network

IOWN proposal. Their aim is to build a twin of any physical entity and represent it digitally. It allows for communication and interaction in this created cyberspace.
QOT is the quantum equivalent of this idea. It therefore creates a digital quantum replica of a physical thing, using photons as qubits. It uses many of the previously discussed methods, such as teleportation, entanglement swapping and routing. As an example, a smart city could utilize QOT to optimize its traffic by interacting with self-driving cars.
A QOT of a human could serve as an intelligent personal assistant and could use quantum artificial intelligence to solve complex problems. Quantum optical computing could prove very useful in AI, as today's deep neural networks are energy and time consuming and quantum optical computing can compute one of their most used operation, matrix multiplication, very efficiently. Furthermore, quantum optics could help pave the way for futuristic applications, such as ambient intelligence, holographic telepresence or tactile internet.

7.6 Conclusions

While the world of quantum computing is still in its early stages, it is the focus of many scientists around the world. Progress is made consistently and some applications, such as QKD are already well developed.
Still there are a number of challenges, that need to be overcome before quantum computers can be used by the public. One concern is the lifetime of qubits, as they are still fairly short lived, and better solutions to store and memorize qubits. They are somewhat connected, as longer living qubit would decrease the need for storage, while good quantum memory could allow for shorter living qubits, which are stored more often.
Despite the drawbacks, the quantum future looks promising and with increased demand for quantum services, development and progress will be driven forward. Applications like such as large molecule simulation and quantum communication systems can look into a bright future of quantum optical computing.
In summary, this report covered an introduction to quantum communication systems. We briefly described the basics of quantum mechanics, in order to understand the opportunities and challenged that arise with quantum computing. We considered the topic of quantum communication in a quantum internet on a high abstraction level, without diving too deeply into the physical implementation of the necessary infrastructure. We looked upon the inherent properties of quantum computations, such as the security from eavesdropping due to the no-cloning theorem, as well as the different approaches to resolve the problem of quantum information storing and decoherence. An outlook on a quantum internet, which could in the foreseeable future be used alongside our current networks, was provided. Furthermore, quantum key distribution was introduced, as one application of quantum computing, which has been demonstrated in experiments. The recent achievement of a quantum communication network using satellites served as an example of the state-of-the-art technology. Finally we discussed some future applications, such as quantum optical communications and quantum cloud computing, which could be a breakthrough for researchers in physics and chemistry.

Bibliography

[1] L. Gyongyosi, S. Imre: *A Survey on quantum computing technology*; Computer Science Review, Vol. 31, February, 2019, pp. 51-71, https://doi.org/10.1016/j.cosrev.2018.11.002

[2] C. H. Bennet, G. Brassard, C. Crépeau et al: *Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels*; Phys. Rev. Lett., Vol. 70, Issue 13, March, 1993, pp. 1895-1899, https://doi.org/10.1103/PhysRevLett.70.1895

[3] MX. Luo, HR. Li,H. Lai, et al: *Unified quantum no-go theorems and transforming of quantum pure states in a restricted set*; Quantum Inf Process, Vol 16, Article No. 297, 2017, https://doi.org/10.1007/s11128-017-1754-0

[4] S. Boixo, S.V. Isakov, V.N. Smelyanskiy, et al: *Characterizing quantum supremacy in near-term devices*; Nature Phys 14, 2018, pp. 595-600, https://doi.org/10.1038/s41567-018-0124-x

[5] F. Arute, K. Arya, R. Babbush, et al: *Quantum supremacy using a programmable superconducting processor*; Nature 574, 2019, pp. 505-510, https://doi.org/10.1038/s41586-019-1666-5

[6] H. Zhong, H. Wang, Y. Deng et al: *Quantum computational advantage using photons*; Science, Vol. 370, Issue 6523, December, 2020, pp. 1460-1463, https://doi.org/10.1126/science.abe8770

[7] P.W. Shor: *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer*; SIAM Review, Vol. 41, Issue 2, 1999, pp. 303-332, https://doi.org/10.1137/S0036144598347011

[8] S. Pirandola, J. Eisert, C. Weedbrook et al: *Advances in quantum teleportation*; Nature Photonics, Vol. 9, September, 2015, pp. 641-652, https://doi.org/10.1038/nphoton.2015.154

[9] J. Haller: *Quantum networks: The next generation of secure computing*; Red Hat SysAdmin Blog, October, 2020, `https://www.redhat.com/sysadmin/quantum-networks`, last visit May 2021

[10] M.P. Hedges, L.J. Longdell, Y. Li, M.J. Sellars: *Efficient quantum memory of light*; Nature, Vol. 465, 2010, pp. 1052-1056, https://doi.org/10.1038/nature09081

[11] S. Wehner, D.Elkouss, R.Hanson: *Quantum Internet: A vision for the road ahead*; Vol 362, Issue 6412, October, 2018, DOI: 10.1126/science.aam9288

[12] N. Sangouard, C. Simon, H. de Riedmatten, N. Gisin: *Quantum repeaters based on atomic ensembles an linear optics*; Reviews of modern physics, Vol. 83, March, 2011, https://link.aps.org/doi/10.1103/RevModPhys.83.33

[13] W. Duer, H.-J. Briegel, J.I. Cirac, P. Zoller: *Quantum repeaters based on entanglement purification*; Physical Review A, Vol. 59, Issue 169, January, 1999, pp. 169-181, https://link.aps.org/doi/10.1103/PhysRevA.59.169

[14] A.S. Cacciapuoti, M. Caleffi, F. Tafuri, et al: *Quantum Internet: Networking Challenges in Distributed Quantum Computing*; IEEE Network. 2019, pp. 1-7, doi: 10.1109/MNET.001.1900092.

[15] A. Dahlberg, M. Skrzypczyk, T. Coopmans, et al: *A link layer protocol for quantum networks*; ACM Special Interest Group on Data Communication, August, 2019, pp. 159-173, https://doi.org/10.1145/3341302.3342070

[16] W. Kozlowski, A. Dahlberg, S. Wehner: *Designing a quantum network protocol*; International Conference on emerging Networking EXperiments and Technologies, November, 2020, pp. 1-16, https://doi.org/10.1145/3386367.3431293

[17] S.J. Devitt, W.J. Munro, K. Nemoto: *Quantum error correction for beginners*; Reports on Progress in Physics, Vol. 76, No. 7, June, 2013, https://iopscience.iop.org/article/10.1088/0034-4885/76/7/076001/meta

[18] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf et al: *The security of practical quantum key distribution*; Reviews of modern physics, Vol. 18, Issue 3, September, 2009, pp. 1301-1350, https://link.aps.org/doi/10.1103/RevModPhys.81.1301

[19] YA. Chen, Q. Zhang, TY. Chen, et al: *An integrated space-to-ground quantum communication network over 4,600 kilometres*; Nature 589, 2021, pp. 214-219, https://doi.org/10.1038/s41586-020-03093-8

[20] E.F. Dumitrescu, A.J. McCaskey, G. Hagen, et al: *Cloud Quantum Computing of an Atomic Nucleus*; Physical Review letters, Vol. 120, Issue 21, May, 2018, pp. 210501, https://link.aps.org/doi/10.1103/PhysRevLett.120.210501

[21] A. D. Corcoles, A. Kandala, A. Javadi-Abhari, et al: *Challenges and Opportunities of Near-Term Quantum Computing Systems*; Proceedings of the IEEE, Vol. 108, No. 8, August, 2020, pp. 1338-1352, doi: 10.1109/JPROC.2019.2954005.

[22] J. Preskill: *Quantum Computing in the NISQ era and beyond*; Quantum, Vol. 2, 2018, p. 79, https://doi.org/10.22331/q-2018-08-06-79

[23] A. Dahlberg, S. Wehner: *SimulaQron - a simulator for developing quantum internet software*; Quantum Science and Technology, Vol. 4, No. 1, 2018, p. 015001 https://iopscience.iop.org/article/10.1088/2058-9565/aad56e/pdf

[24] A. Manzali: *Quantum Communications in Future Networks and Services*; Quantum Reports, Vol. 2, No. 1, pp. 221-232, https://doi.org/10.3390/quantum2010014

Chapter 8

An Overview into the InterPlanetary File System (IPFS): Use Cases, Advantages, and Drawbacks

Christian Bieri

Negative headlines of centralized systems have accumulated in the recent past. In addition to efficiency and scalability problems, such systems are often said to no longer be able to meet the constantly growing requirements in the foreseeable future. The current noticeable trend away from centrality does not really play into the cards of centralized systems.

Therefore, the aim of this paper is to present an innovative concept, launched in 2014 by Juan Benet, a computer scientist from Stanford, that could solve the above stated problems in one stroke: The InterPlanetary File System. In addition to the background knowledge necessary for understanding the IPFS, this paper focuses exclusively on its underlying technologies, shows useful and legit use cases, outlines advantages and disadvantages, gives a subjective statement based on the authors opinion and identifies those factors that could be of key importance for the future success/failure of the IPFS. The knowledge for doing so was acquired only from scientific papers, conference deliverables, technical reports, specialist books and relevant websites. Furthermore, the author has practically dealt with the IPFS and tested a number of use cases thoroughly.

On closer inspection, it becomes clear that the IPFS, despite there exist some disadvantages, offers many advantages. It is therefore not surprising that the InterPlanetary File System is often said to have disruptive properties - it could fundamentally change the way we share content nowadays.

Contents

8.1 Introduction

8.1.1 Problem statement and motivation

In the recent past, negative headlines of centralized systems have become more and more frequent. Be it a server outage of an important provider with global relevance or the blocking of content based on questionable arguments. In addition to this, humanity is moving at full steam ahead into a new era in which data distribution becomes increasingly important. If one thinks about the global availability of cheap smartphones or the Internet of Things (IOT), it should quickly become clear what kind of scale we are talking about. Furthermore, possibly driven by the hype around cryptocurrencies, a need for decentralized/distributed systems is noticeable - people have become more suspicious and try to bypass centralized points of contact whenever possible. Last but not least, centralized systems are often associated with efficiency and scalability issues.
Although the underlying technologies of centralized systems are constantly evolving, it seems as if the numerous problems that such systems by definition entail cannot be solved easily. Critics go even further: they suspect that the above-mentioned systems may soon no longer be able to meet future requirements.
However, there is a silver lining on the horizon: Back in 2014, Juan Benet, a computer scientist from Stanford, published a scientific paper in which he presented a novel concept that could solve all of the mentioned problems in one stroke - The InterPlanetary File System (IPFS). The IPFS is an innovative concept with thoroughly disruptive characteristics. It is a content-based system with a peer-to-peer architecture which uses existing, established technologies and combines them in a completely new way.

8.1.2 Structure and procedure

Considering the aspects mentioned in chapter 8.1.1 as well as the advantages of decentralization/distribution, it becomes clear immediately, that the IPFS could take on a central role in the near future. Therefore, the aim of this article is to give a detailed insight into the IPFS with its underlying technologies, to outline useful and legit use cases and to draw a conclusion based on the knowledge elaborated throughout the seminar. In order to do so, the author has decided to structure the present paper into the following 4 main sections:
The **first section** sheds light on the most important background knowledge which is inevitable to fully understand the InterPlanetary File System such as the Open System Interconnection (OSI) reference model, cryptographic hash functions/cryptographic hashes, the most common network architectures, the Domain Name System (DNS) and the World Wide Web (WWW)/Hypertext Transfer Protocol (HTTP). **Section two** concerns the three main components - content addressing, content linking and content discovery - on which the IPFS is built upon. Additional topics which are relevant within the context of the present paper such as the concept of immutability/permanence and the role of the InterPlanetary Naming Service (IPNS) are covered as well. **Section three** is about useful and legit use cases of the IPFS and also reveals the origin of the name InterPlanetary File System. Last but not least, **section four** concludes the whole topic based on the authors opinion and made experiences in a highly subjective way. Furthermore, it gives a brief overview into how the IPFS might develop in the foreseeable future and mentions the most obvious aspects for its future success.
As can be seen, the sections are arranged in such a way that the necessary knowledge is built up step by step. The choice for such an outline was made because it not only helped the author to understand the subject in detail, but also to make the subject as understandable as possible for the target reader.

Since the IPFS is a complex subject, a great deal of information had to be gathered and processed in advance - mainly scientific papers, conference deliverables, technical reports, specialist books, and relevant websites were used. The information contained therein was condensed and written down in the form of a paper with the structure mentioned above. Furthermore, the author dealt with the IPFS in practice and tested the system thoroughly.

8.2 Background

8.2.1 The Open Systems Interconnection reference model

In 1984, the International Standardization Organization (ISO) developed the so-called Open Systems Interconnection reference model [1]. It was initially designed to enable the communication between electronic devices and consists of seven layers [1]. It must be mentioned though that the OSI reference model does not offer any possibility to transfer data but rather the framework needed for the transmission of data. The actual data transmission is done via different types of protocols [2].
The seven layers can be divided into two categories: **1) layer I - IV**: network-related (types of cables, connectors, Media Access Control addresses, Internet Protocol, routing, Transmission Control Protocol, User Datagram Protocol), and **2) layer V - VII**: application-/presentation-related (Server Message Block Protocol, standards like MFPEG, TIFF and ASCII, File Transfer Protocol, Hypertext Transfer Protocol) [3]. The on layer VII located Hypertext Transfer Protocol is especially important in the context of the present paper and will be further explained in chapter 8.2.5.

8.2.2 Cryptographic hash functions/cryptographic hashes

Since cryptographic hash functions play a crucial role in the IPFS (Content Identifiers, Distributed Acyclic Graph, PeerIDs), understanding them is necessary. Cryptographic hash functions output a value - the so-called cryptographic hash or digest - with a predetermined length. The length of the digest is independent of the size of the input value and depends on the cryptographic hash function used (see figure 8.1 for a schematic overview) [4]. However, it is important to note that the ability to randomize the hash algorithm as well as the size of the output hash is key when it comes to security of the cryptographic hash function (in the sense of avoiding collisions). A given input, typically associated with a time stamp, will always generate the same cryptographic hash (deterministic, see below). In this sense, changes in the integrity of the given input can be detected since the cryptographic hash would be completely different.
The cryptographic hash is **a) one-way**: calculating/guessing the input message from its cryptographic hash must not be feasible [4], **b) deterministic**: given an input "i1", the cryptographic hash function always delivers the same cryptographic hash for the same input [4], **c) uncorrelated**: by changing a given input "i1" slightly, the cryptographic hash function must deliver a completely different cryptographic hash [4], **d) weak collision resistant**: given an input "i1", it should be infeasible to find a different input "i2" which results in the same cryptographic hash [4] and **e) strong collision resistant**: finding two different inputs "i1" and "i2" which result in the same cryptographic hash should be infeasible [4].
The characteristics described above come in handy because by using a cryptographic hash it **a)** becomes possible to identify any piece of data via a unique cryptographic hash and exchanging them is **b)** very resource-efficient as cryptographic hash functions output a cryptographic hash with a fixed length - regardless of the size of the input [4].
From time to time though, cryptographic hash functions no longer meet the above stated characteristics and are broken. Therefore, supporting different cryptographic hash func-

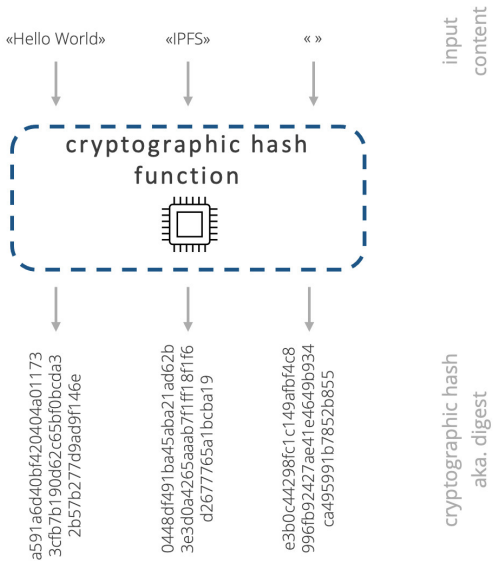

Figure 8.1: cryptographic hash function schematically

tions is key. But when supporting different functions, there must be a way to identify which algorithm was used to generate the cryptographic hash. For this purpose, the IPFS uses the multihash format [5]. According to [5] "A multihash is a self-describing hash which itself contains metadata that describes both its length and what cryptographic algorithm generated it". In the view of the author, the multihash format is essential in order to have a future proof system.

8.2.3 Network architectures

The chapters 8.2.3.1 - 8.2.3.3 give an overview into the most common network architectures and highlight advantages and disadvantages.

8.2.3.1 Centralized systems

Centralized systems use a client/server architecture. Usually, clients are directly connected to a central server. This central server coordinates/serves all the other clients in the system. Scaling a centralized system is done vertically but not horizontally [6] [7].
Such systems offer **a1) a high user experience** (easy to physically secure system, easy removal of a node and many more), are **a2) cost efficient to set up and maintain** and are **a3) easy to coordinate**. On the other hand, centralized architectures are often associated with **d1) high downtimes/availability issues** since we have a single point of failure, **d2) bottlenecks** when traffic skyrockets and **d3) security issues** because there exists a single point of attack (see figure 8.2 for a graphical visualisation). Nonetheless, most web services such as online banking accounts, YouTube or a mobile app store are centralized systems [6] [7].

8.2.3.2 Decentralized systems

Unlike centralized systems, decentralized systems distribute the coordination/serving workload across multiple central servers. Each of these central servers interacts independently with other central servers and serves as a central unit. Nodes are directly connected to a central server. Therefore, decentralized systems still rely on central servers, but more than one per network. Usually, decentralized systems can be scaled vertically but not horizontally [6] [7].

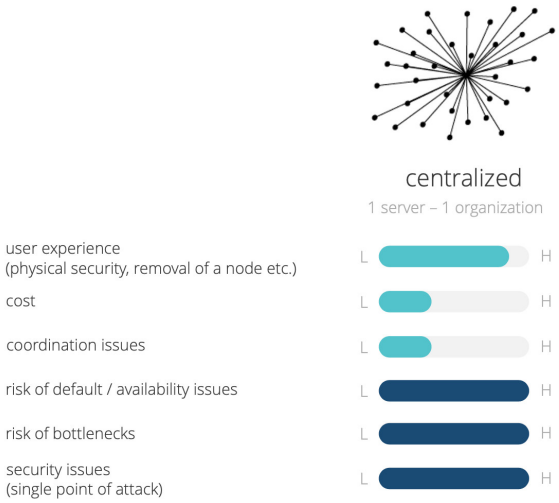

Figure 8.2: advantages/disadvantages of centralized systems

Decentralized systems are **a1) highly available** since there does not exist a single point of failure, offer **a2) faster performance and lesser bottlenecks** due to balancing the load on different central servers and have **a3) no single point of attack** which results in a more secure system. On the other hand, such systems are often associated with **d1) a lower user experience**, **d2) higher costs** due to labor-intensive maintenance and **d3) coordination issues** due to the lack of communication between nodes (see figure 8.3 for a graphical visualisation) [6] [7].

8.2.3.3 Distributed systems

Similar to decentralized systems, distributed systems avoid a single central server. Distributed systems consist of equal interconnected nodes, some of which become server nodes for the coordination of the system. Therefore, data ownership and computational resources are shared evenly across the whole system. As every node in the system makes its own decision, the final behavior of the system is the aggregate of the decisions of the individual nodes. The process by which the system votes and makes decisions is contingent on the systems consensus mechanism. Therefore, all nodes are peers of each other - they work towards a common goal. Distributed systems can either be scaled vertically or horizontally or both [6] [7].
As mentioned above, distributed and decentralized systems have many similarities when it comes to their architecture. Therefore, the advantages and disadvantages of distributed systems are quite similar to the ones of decentralized system (see chapter 8.2.3.2) but on a higher extent (see figure 8.3 for a graphical visualisation) [6] [7].

8.2.4 The Domain Name System

Usually, participants in a network not only have an address, but also a name. These names are strictly regulated and hierarchically structured. They cannot be changed but extended. As electronic devices communicate via addresses (the names are for our convenience only), there must be a system which resolves names into addresses [8]. Therefore, we use a system of distributed databases - the Domain Name System.
The DNS is hierarchically structured. On top are the root servers. These 13 servers consist of hundreds of severs which are located in different countries all around the world [9]. The root servers know the worldwide structure and refer to the Name Servers (NS) to which the Top Level Domains (TLDs) are delegated. These servers in turn delegate name spaces to other Name Servers. By further extending this structure, a tree-like structure is

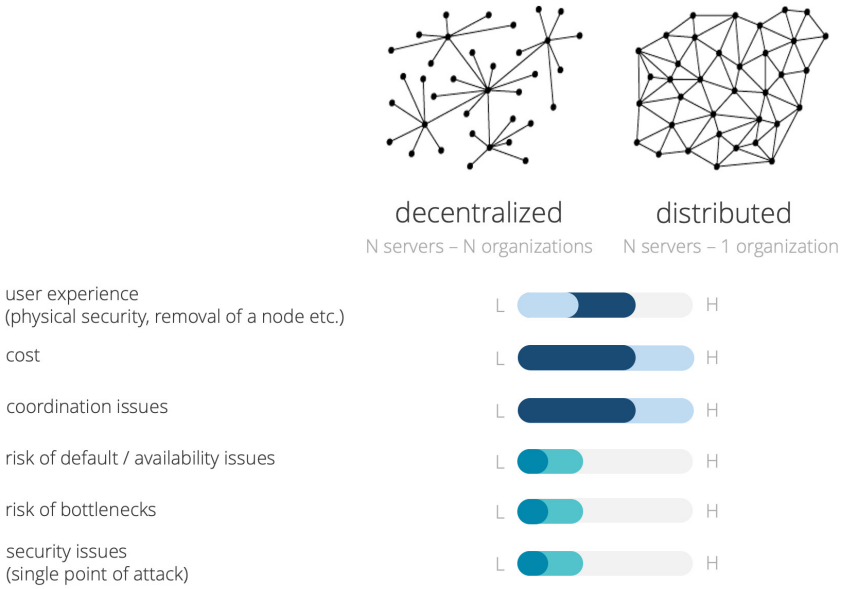

Figure 8.3: advantages/disadvantages of decentralized/distributed systems

created. It must be mentioned though, that no DNS server contains the entire database but only a subset of it [10]. As a result, names can be resolved into addresses (forward lookup) or addresses can be resolved into names (reverse lookup) [11].

8.2.5 The World Wide Web and the Hypertext Transfer Protocol

8.2.5.1 General overview

The probably best known and most often used service on the internet is the World Wide Web. To make the WWW work, the HTTP is needed. The layer VII protocol enables the communication between web browsers (clients) and web servers (servers) over the internet [12].
In order to receive content, the web browser (client) sends a request to a web server (server). Since the HTTP is a connectionless text-based protocol, a new connection must be initiated for each request [12]. As already mentioned in chapter 8.2.4, the resolution from names into addresses is done by the DNS. The WWW is an address-based, centralized system. Although such a system has advantages and a raison d'être, it does not come without flaws - these possible deficiencies will be extensively highlighted in the following chapter 8.2.5.2 [12].

8.2.5.2 Challenges of the World Wide Web and the Hypertext Transfer Protocol

The WWW can be referred to as a centralized system. As seen in chapter 8.2.3.1, centralized architectures have specific advantages and disadvantages. The following section skips the advantages and focuses mainly on the disadvantages of the WWW and takes a closer look at the most obvious and current challenges [13].
In such a system, users are not only reliant on hosts but also on administrators. It becomes clear that the majority of the WWW is controlled by ordinary people maintaining servers [13]. That is all well and good as long as everyone does a good job, but sometimes, people do not do their job properly. Additionally, people have become more distrustful towards central parties and no longer want to trust them blindly. What may happen when these people/central parties do not do what they are intended to do can be seen below:

> Google has suffered a worldwide outage, with failures reported across the company's services [...] The widespread failures caused many to highlight the risks of digital concentration, where an outage at a single company takes down a substantial proportion of online activity [14].
>
> The online encyclopaedia was blocked in Turkey in April 2017, after it refused to delete articles critical of the country's government. [...] The ban lasted 991 days [...] [15].

In the eyes of the author, such incidents are more than problematic and should be avoided at all costs. But since it is a centralized system, such outages can occur every time and censoring content is extremely easy. Furthermore, according to [16] "[...] each hop costs money to the data provider and wastes bandwidth". Last but not least, as stated by [16], "HTTP downloads a file from a single computer at a time, instead of getting pieces from multiple computers simultaneously". If one also considers that networking will become even more extensive in the coming years, the author acknowledges that a solution is needed to address the mentioned problems.
However, although unreliable, slow, manipulable and expensive, the WWW and HTTP are still widely used. According to [17], this may be the case because "[...] no general file-system has emerged that offers global, low-latency, and decentralized distribution". This might be true to some extent. In the author's eyes, however, it is not only related to technological aspects, but also to the fact that change is a long-wided process that has to be undergone in order to new technologies being applied on a large scale.

8.3 The InterPlanetary File System

8.3.1 General overview

Since we are entering a new era of data distribution and the old systems may no longer be sufficient (see chapter 8.2.5.2), the time might be right for the IPFS [17]. According to [18], "The InterPlanetary File System is a distributed system for storing and accessing files, websites, applications, and data". It was was thought up by Juan Benet, a computer scientist from Stanford University [16] and makes use of a peer-to-peer architecture [17]. Unlike the WWW which knows how to find information by its address/location, the IPFS finds information by its content. Therefore, it is also referred to as a content-based system [18].
To make such a system work, the IPFS is built upon many modular libraries which support specific parts of the system. The three main components are **a) content addressing/Content Identifiers (CIDs)**, **b) content linking via Merkle Directed Acyclic Graphs (DAGs)** and **c) content discovery via Distributed Hash Tables (DHTs)**.
Or in other words: The IPFS uses many different, straightforward and established technologies and combines them in a novel and unique approach, resulting into a file system that tries to connect all nodes with the same system of files, which has no single point of failure and where nodes do not need to trust each other [17].
The following chapters 8.3.2 - 8.3.6 not only give a detailed insight into the underlying technologies, but also outline additional knowledge which is worth being mentioned in the context of the IPFS.

8.3.2 Content addressing and content identifiers

As mentioned in chapter 8.3.1, the IPFS makes use of content addressing - it identifies content by what it is [20]. Therefore, content that uses the IPFS protocol needs a Content Identifier - a self-describing content-addressed identifier. These CIDs are generated by a

cryptographic hash function [20]. Therefore, a CID can be seen as a label which is used to point to content in the IPFS. Although it does not indicate where the content is, it can be considered as an address, based on the content itself [21].
Although most content in the IPFS is hashed by sha2-256 [22], it is mentioned that the IPFS uses the multihash format. As a result, many other cryptographic hash functions, such as Blake 2 or MD5, are supported as well [23]. Cryptographic hashes are not only used for uniquely identifying content, but also for linking it together (chapter 8.3.3) [20].
The CIDs from the IPFS can come in two different forms/version: **a) version 0 multihash CIDs** with base 58 encoding which are simple but inflexible and **b) version 1 multihash CIDs** which uses some leading identifiers which clarify the encoding used (multibase prefix), the CID version (CID version identifier) and the format of the target content (multicodec identifier) [21]. The combination of the multihash format and the leading identifiers from v1 CIDs provide future proof CIDs. As a result, the v1 CIDs from the IPFS are very well prepared for the future [5].

8.3.3 Content linking via Merkle Directed Acyclic Graphs

8.3.3.1 General overview

In chapter 8.3.2, we laid the foundation for content addressing. However, just assigning CIDs to content is not enough - content must be linked together. Therefore, a new data structure must be introduced: the so-called Merkle Directed Acyclic Graphs.
DAGs are often used in distributed systems. The IPFS uses a Merkle DAG which is optimized for representing directories and files. In a Merkle DAG, every node has a unique identifier which is a cryptographic hash produced by the nodes content (see also chapters 8.2.2 and 8.3.2). As a result, whole file systems can be represented with Merkle DAGs [24]. The Merkle DAG from the IPFS is an enormous flexible way to represent content [25]. Since every node in a Merkle DAG has a unique identifier, also known as CID, it becomes possible to identify content via the value of this CID [24].

8.3.3.2 Key characteristics

Besides the fact that every Merkle DAG has a unique identifier, there exist some further key characteristics which are worth being mentioned: Merkle DAGs are **a) structured and linked together**: every leave (nodes without children) is contained in the parent Merkle DAG; every node is the root of a (sub) Merkle DAG and contained in the parent Merkle DAG [26], **b) immutable**: changing content of a node changes its unique identifier affecting all the ascendants and resulting in a completely different Merkle DAG [26], **c) deduplicated**: nodes with the same unique identifier represent the same Merkle DAG [26] enabling storing the same content only once [25] and **d) tamper proof**: all content can be verified on its integrity with its CID, IPFS detects tampered content [25].

8.3.3.3 Building process and related topics

To build a Merkle DAG, the IPFS splits up the content into blocks of 256KB (junker) first [24]. Afterwards, the Merkle DAG can be built in a bottom-up like process. The construction process starts from the leaves. As soon as the leaves' CIDs are computed, parents can be added. Building a Merkle DAG in a top-down like process does not work - the content linking would not be possible [26]. A simple, but for understanding sufficient, Merkle DAG can be seen in figure 8.4.
Since content is split up into blocks, different parts of the content can be fetched from different sources. Additionally, content can be authenticated quickly [24]. Furthermore, similar content can share parts of the Merkle DAG. Imagine you update a website: The

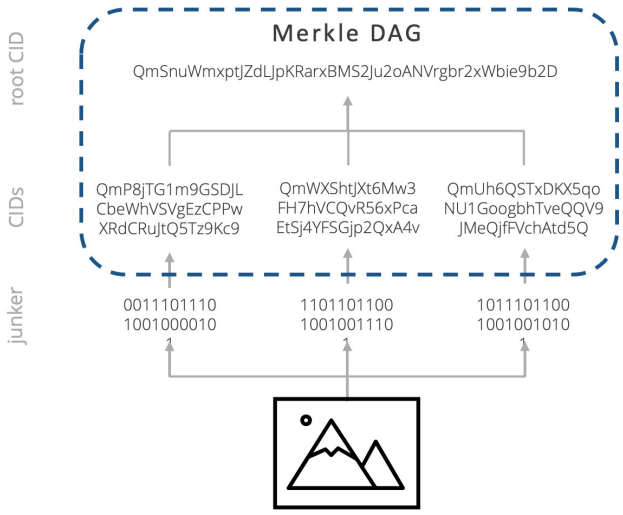

Figure 8.4: Merkle DAG schematically

updated contents - and only these - receive a new CID. Therefore, according to [24], "The old version and the new version can refer to the same blocks for everything else". In the opinion of the author this is an extremely sophisticated principle and, as a result, transmitting versions of large datasets becomes more efficient as only the updated content must be transferred [24].

8.3.4 Content discovery via Distributed Hash Tables

8.3.4.1 General overview

So far, we have seen how content addressing with CIDs works (chapter 8.3.2) and how content is linked together (chapter 8.3.3). Now, we move on to the last important piece of the IPFS: content discovery via Distributed Hash Tables which consists of three main concepts: **1) DHTs, 2) Bitswap** and **3) libp2p**.

8.3.4.2 Distributed Hash Tables

A DHT is a database of keys and values which is used for content discovery. Content discovery is used to identify which node/nodes is/are hosting the content we are looking for. It is called distributed because the table is distributed across all nodes in the network [27].
DHTs are often referred to as the coordination system of a peer-to-peer system [17]. One very popular DHT is the Kademlia DHT. It is not only equipped with an efficient lookup mechanism, but it is also used in many peer-to-peer applications and established as well. Besides the Kademlia DHT, there exist two other DHTs: **1) the Coral DSHT** which extends the Kademlia DHT mainly in terms of performance, storage and bandwidth areas and **2) the S/Kademlia DHT** which can be seen as a more secure version of the Kademlia DHT [28].
The DHT is consulted twice. **1st) for finding the peers** which are storing the block that make up the content we are after (content discovery) and **2nd) to find the location of these peers** (peer routing) [27]. Especially for peer routing (see also chapter 8.3.4.4), the Kademlia routing algorithm is of outmost importance. The node-ID-based routing algorithm uses an innovative XOR and is enormously efficient [29]. However, explaining the algorithm in more detail would go far beyond the scope of this paper.

8.3.4.3 Bitswap - a marketplace for blocks of data

As soon as we have discovered the content and have found the current location of it, we can request these blocks from other peers and send blocks to other peers. In order to do so, the IPFS uses Bitswap [27]. The message-based protocol contains either a want list or blocks and is responsible for acquiring and sending blocks from peers to peers [30]. The IPFS requests blocks from Bitswap and Bitswap fetches them from the network [31].
Therefore, Bitswap is often referred to as a marketplace for blocks of data. It is the way how the IPFS exchanges blocks of data [32] or the way how data distribution in the IPFS happens [33]. As already mentioned, content is split up into blocks of 256KB (chapter 8.3.3.3). Every block is uniquely identified by a CID (chapter 8.3.2). Understanding this concept is key when it comes to understanding Bitswap.
Fetching content is a recursive procedure. **1st)** the root CID is added to a so-called want list [32]. The want list is a list which contains the CIDs of desired blocks which a peer wants to receive [30]. When the CID of the root block is added to the want list, the following steps are carried out sequentially: **2nd a)** The desired root block is requested from all connected peers via a want broadcast. If no connected peers have the desired root block, then **2nd b)** the DHT is consulted for finding the ID/IDs of the peer/peers who have stored the desired block [30] [31]. Depending on the availability of the block, **3rd)** the block is either sent by a peer or we receive Peer ID/s where the block is available from the DHT. Peers who have the desired content are added to the session. Only these peers will be considered for further enquiries [31]. As we have the PeerID/s, we can **4th)** request the desired root block. With the root block at hand, we also know the roots child/children block/blocks in the form of the CID/CIDs. The child/children CID/CIDs are **5th)** added to the want list and the whole process starts again until we reach the leaf/leaves block/blocks of the Merkle DAG [30] [32]. A graphical visualisation of Bitswap and its procedure can be seen in figure 8.5.

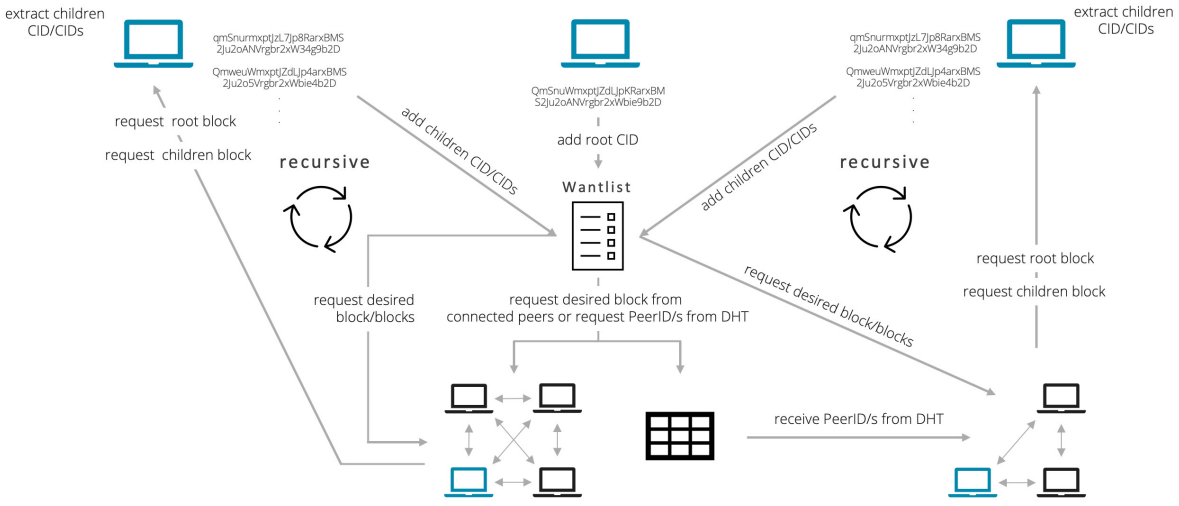

Figure 8.5: procedure Bitswap schematically

Additionally, it is important to specify which node/nodes is/are preferred to receive content from. According to [31] "peers in the session are ordered by latency and the request is split across peers". This makes, in the view of the author at least, more than sense since the IPFS tries to offer global, **low-latency** and decentralized distribution. Furthermore, since we can check the integrity of every block with its CID and its content, peers do not have to trust each other. Therefore, it does not matter which peer/peers sends/send us the desired content [30].

As we have seen, nodes come together in a marketplace for exchanging content in the form of blocks. Usually, marketplaces are thought to have a currency. A digital currency for incentivizing nodes could be implemented in the form of a Bitswap Strategy. The base case though works without any currency: According to [33] and in order to exchange blocks, "nodes have to provide direct value to each other in the form of blocks". When the nodes have what the other wants, it works fine. Otherwise, "nodes must work for their blocks". For the described procedure to work, Bitswap Credit is used which allows peers to track their balance with other nodes [33]. According to the author's understanding, such a functionality is more than desirable, as it will put a stop to freeloaders.

8.3.4.4 libp2p

libp2p was developed specifically for the IPFS as an open-source project [34]. As stated in [34], "libp2p is a modular system of protocols, specifications and libraries that enable the development of peer-to-peer network applications". Due to its modular and composable characteristics and because of the independence to the IPFS, it became very popular for many other projects as well [34].
libp2p is responsible for many tasks in decentralized systems - it addresses the following five key areas:

- **Identity/Security**: Each libp2p peer holds a private key and a public key. This key pair is not only used to set up a secure communication between peers and for authentication reasons [28] [36], but also for generating the PeerID which is the output of a cryptographic hash function of the public key encoded in the multihash format. A peer address has the following structure: /p2p/QmY[...]x5N[1] [36].

- **Addressing**: Since libp2p offers the possibility to operate across different types of networks, it must be possible to work with different addressing schemes. In order to do so, a multiaddress is used where multiple addressing information are encoded into a single structure. Using a multiaddress avoids ambiguity. When the transport multiaddress (/ip4/7.7.7.7/tcp/4242) is combined with the peer address (/p2p/QmY[...]x5N1[1]), we have a new multiaddress at hand which contains all the essential information needed [35].

- **Transport**: The so-called transports in the IPFS are responsible for transporting bits from one peer to another. As the therefore responsible protocols can be freely chosen by the developers [37], libp2p must provide an interface which supports a broad range of protocols [34]. The two core operations of transports are **a) listening**: which accepts connection requests from another/other peer/peers and **b) dialing**: which opens a connection to a listening/listening peer/peers [37].

- **Content Discovery**: Since the IPFS uses the concept of content addressing, libp2p provides an interface for content routing using a Kademlia-based DHT [34].

- **Peer Routing**: For sending a message to another peer, we need its PeerID and its location on the network. Often, we only know the PeerID. As metioned in [34], "Peer routing is the process of discovering peer addresses by leveraging the knowledge of other peers". Peer routing in libp2p uses a DHT and the Kademlia routing algorithm [34].

Besides the above-described tasks there exist some more such as **a) PubSub** which stands for publish/subscribe and is a very useful pattern for sending a message to groups of interested receivers [34], **b) Network Address Translation (NAT) Traversal** for configuring

[1] original peer address: /p2p/QmYyQSo1c1Ym7orWxLYvCrM2EmxFTANf8wXmmE7DWjhx5N

incoming connections [38], **c) Protocols** such as ping, identify/push and kad-dht with a unique ID for the protocol negotiation process [39] and **d) Stream Multiplexing** which is used to share the underlying, usually scarce, transport medium [40]. However, discussing them in more detail would go beyond the scope of this paper.

8.3.5 Immutability and the role of the InterPlanetary Naming Service

After having discussed content addressing (chapter 8.3.2) and content linking (chapter 8.3.3), understanding the concept of immutability and the role of the InterPlanetary Naming Service within this context should be straightforward.

> An immutable object is an object whose state cannot be altered or modified once created. Once a file is added to the IPFS network, the content of that file cannot be changed without altering the CID of the file [41].

Immutability is an excellent property when it comes to storing data but with content that can change or needs to be altered, immutability is problematic though. In order to enable mutability in an immutable system, we need another layer which is located on top of the CIDs [41]. The author shares the stated view and thinks that the Interplanetary Naming Service (description follows) offers an ideal solution to deal with the mentioned problem. Whenever content in the IPFS changes, we are building up a new Merkle DAG with a new CID (CID_{new}), instead of editing, updating or otherwise changing the old Merkle DAG. The new parent/root CID (CID_{new}) consists of nodes which did not change and of the new nodes which have changed [41].
As in the IPFS nothing is overwritten, the user can not only visit the old content, but also the new content by either using CID_{old} or CID_{new}. visiting content via the CID is a quite arduous approach as every time content is updated, the CID changes as well. As a consequence, we would need to distribute a new CID whenever we change content. To overcome this inconvenience, the IPFS uses a pointer which always points to the latest content instead. The system in charge of this process is called the InterPlanetary Naming Service. The IPNS is very important for a convenient use of the IPFS - especially in an always changing environment [41]. It creates an address that can be mutated [42] and therefore enables mutable naming without the need of distributing IPFS CIDs [43].
In the IPFS, every user gets a mutable namespace in the form of /ipns/<nodeId> to which a user can publish content to. The /ipns/ prefix enables the distinction between mutable and immutable paths. Instead of using the IPNS for creating mutable addresses on the IPFS, DNSLink can also be used. It is not only much faster but also uses human-readable names [42]. However, discussing DNSLink in more detail would go beyond the scope of this paper.

8.3.6 Participation, persistence, permanence and pinning

In order to have content which is permanently available, not being lost or fortuitously deleted, it is essential that peers are actively participating because the more often copies of content are stored within the IPFS, the more reliably available these copies to other users of the IPFS are [48]. This happens, at least to some extent, automatically: Whenever a peer downloads content in the IPFS, this peer shares this content with other peers for a limited period of time. In order to do so, an IPFS node needs local storage to store and obtain data. Note: all content which is available in the IPFS must be stored in some node's local storage [46].
However, as storage is a finite resource, nodes must clear out cached content from time to time. The therefore used process is called garbage collection [45] and can be seen as a form of automatic resource management which deletes content on IPFS nodes that it thinks is

no longer needed. The collector can either be triggered when **a)** a certain percentage of used storage is exceeded or **b)** after a certain period of time [47].
As the garbage collection deletes files, there must be an approach to ensure that data persists - pinning [45]. With pinning, content can be stored on one or more IPFS nodes' local storage for an indefinite period of time [46]. Pinning ensures persistence [45]. In order to pin content, many IPFS nodes use pinning service such as Axel, Eternum, Infura and many more [49].
Since there exist almost no incentive for an individual peer to pin content, Filecoin may play an integral part of the IPFS in the foreseeable future. According to [50], Filecoin can be seen as "the missing incentive layer for the IPFS". It [...] adds incentivized, persistent storage to IPFS". However, explaining Filecoin in more detail would exceed the scope of present paper by far.

8.3.7 Session of discussion about the underlying technologies

The underlying technologies of the IPFS have already been described in detail in the chapters 8.3.1-8.3.6. The following section, however, is intended to analyse those technologies in the form of a brief discussion. It also reflects the author's opinion.
The cryptographic hash functions, which are used to generate the CIDs, and the properties they provide are not only the basis for addressing content, but are also extremely well established and offer, together with the multihash format, extremely future-proof CIDs. The author is currently not aware of any technology that offers all the necessary properties in combination and therefore finds this technology suitable.
Of no less importance are the Merkle DAGs: Above all, the author regards two properties as more than remarkable: **1)** that only those blocks need to be updated that are subject to change, thus enabling efficient data exchange and **2)** that it is a simple concept, yet very effective and extremely dynamic. Before writing this paper, the author was not familiar with Merkle DAGs in detail - now he recognises their sophisticated features and is sure that they are well suited for the purpose under consideration.
Cryptographic hash functions and CIDs together with Merkle DAGs make it possible to authenticate content quickly. This is, according to the author's understanding at least, a huge advantage, as it can possibly put a stop to fraudsters. Furthermore, it could make the use of data distribution systems much safer which is especially advantageous for non-experienced users.
Although DHTs are of outmost importance, it is the Kademlia search algorithm that is, in the eyes of the author, of most interest, as it is of great importance for peer routing and is extremely efficient due to the used XOR metric. Additionally, in the view of the author, Bitswap is remarkable. This may be the case because the author is fascinated by recursion anyway. Being able to fetch blocks from different sources is an ingenious concept. The fact that peers who are consulted for requesting content are sorted by latency is the icing on the cake: However, one remark: When searching for content that is hardly available, the search process can be very time consuming, which could affect the efficiency.
Of course, libp2p was analysed within the scope of this paper. Nevertheless, to look at the technology completely in detail would have gone beyond the scope of this paper. Therefore, it would be untrustworthy to give an opinion about the underlying technologies. However, if one looks at the functionalities offered, it quickly becomes clear, at least in the eyes of the author, why libp2p is used for so many projects nowadays.
What additionally impresses the author is the fact that hurdles are not seen as problems, but that work is done to solve them with a new concept - with a strong community behind the IPFS, this is possible. For example, the concept of immutability, which is especially problematic in the context of constantly changing content, is solved with the

IPNS. Solutions are also available for the difficulty of ensuring permanently available data, which gains more significance through the garbage collector: pinning services or Filecoin. In conclusion, it can be said that the underlying technologies are simple but effective - these technologies have been combined in an ingenious way in the IPFS. The result is a perfect interplay that makes the IPFS what it is today.

8.4 Uses cases of the InterPlanetary File System

8.4.1 General overview

The IPFS offers an almost unlimited number of use cases. Covering all of them would exceed the scope of this paper by far. Nonetheless, the following section should give an overview into some useful and legitimate use cases: **a) file sharing** via desktop applications such as IPFS Desktop, Arbore/Orion or in the form of a Dead Drop [51], **b) collaboration** for the coordination of the flow of data between colleagues such as PeerPad for written documents, the InterPlanetary Version Control System or Gthr to connect event attendants [52], **c) asset storage** for storing scripts or databases of a project on the IPFS such as DTube for video hosting, Decentraland for virtual reality exploration or Qri as an open-source tool for managing large datasets [53], **d) IPFS as infrastructure** for handling key areas of networks customizable and out-of-the-box such as load balancing, deduplication, high availability and many more [54], **e) IPFS as a storage optimizer** as a result of only storing identical content once [55], **f) decentralizing data** with OrbitDB for decentralized databases, with Textile - a hosting company on the IPFS - or with Fleek to build sites and apps on the IPFS [56], **g) Building decentralized applications (dApps)** with available frameworks such as Dappkit, Fission, Fleek and many more [57], **h) blockchain** through distributing the common state on-chain and storing content on the IPFS as for example Filecoin, IOTA or MindSync [58], **i) decentralized identity** as a concept of storing personal data on a personal device through for example the personal data manager 3ID Connect or the Ceramic Protocol [59].
One use case the present paper highlights is the decentralization of the WWW itself and can be found in the following chapter 8.4.2.

8.4.2 Decentralizing the World Wide Web - advantages, disadvantages and alternatives

As already seen, the WWW is a centralized system (chapter 8.2.5) with all its advantages and disadvantages (chapter 8.2.3). If we completely decentralized the WWW, as stated in [60], content could be downloaded "from many locations that are not managed by one organization". According to the author's understanding, this is of key importance to offer a general file-system that offers global, low-latency, and decentralized distribution. As a result, we would **a)** have a highly available system with no single point of attack, could **b)** minimize bottlenecks, especially when traffic skyrockets, would **c)** no longer have to trust any central party, would **d)** be less affected by outages, could **e)** reduce costs for data providers, would **f)** no longer waste bandwidth, could **g)** circumvent efficiency problems, would **h)** make it almost impossible to censor content and it **i)** would be much more unlikely being deceived by other participants since we are able to verify requested content via the CID.
However, we would have to accept some disadvantages in return. Most probably, the user experience would **a)** deteriorate, **b)** the costs for maintaining the network would skyrocket, **c)** coordination issues could occur and there **d)** exist adaption problems because

the WWW/HTTP are so widespread in use - evolving existing protocols incrementally seems more likely than changing paradigms.
There is one more property worth being mentioned though which is especially obvious when nodes are far away or disconnected: Content can be received faster if it can be fetched from a node/nodes close by. This fact becomes particularly valuable for local networked communities without good connection to the broader internet [60].

> That last point is actually where IPFS gets its full name: The InterPlanetary File System. We're striving to build a system that works across places as disconnected or as far apart as planets. While that's an idealistic goal, it keeps us working and thinking hard, and almost everything we create in pursuit of that goal is also useful here at home [60].

The author highly acknowledges the vision stated above but believes that this concept must be established on planet Earth in a first step. However, he also admits clearly that an interplanetary system can serve as an incentive and, if established successfully on our planet, may even become reality one day.
Finally, it should be mentioned that, instead of decentralizing the WWW completely, one could also consider addressing only some parts of it. Such as for example backing up the DNS, one core component of the WWW (chapter 8.2.4), on the IPFS for improving its availability with OrbitDNS [62]. Additionally, the deduplication property of the IPNS could help archiving the WWW. With the InterPlanetary Wayback, this could be made possible [61].

8.5 Conclusion

8.5.1 Achievement of objectives and encountered difficulties

In my view, all stated objectives (see chapter 8.1.2) have clearly been achieved. There were five main factors which contributed to the success: **1) Excellent preparation beforehand** which enabled me to know since the beginning which goals I should achieve at which time. **2) An exemplary time management** which was responsible for being able to submit the deliverables ahead of time and allowed me to focus on quality instead of tinkering at the last moment. **3) Fruitful cooperation** between me and my supervisor. My supervisor always took the time to answer my questions and concerns. When problems arose, they were discussed on a bilateral basis. I have always tried to take the tips into account in the elaboration of the present paper. **4) A broad range of high-quality literature** which enabled me to understand the subject thoroughly. **5) Motivation** to get an in-depth insight into a cutting-edge and interesting topic which was completely new to me and the thought of bringing the topic closer to my classmates in an understandable form.
Throughout the execution of the work, no difficulties, which are worth being mentioned, arose. If there were any questions or ambiguities, they were solved directly with my supervisor on the spot. Due to the thorough preparation beforehand, modifications were hardly necessary during the execution. Of course, I received some tips from my supervisor - whenever possible, I incorporated them into the present paper. On the whole, however, I am more than satisfied with my approach and will try to adopt it similarly in future projects.

8.5.2 Subjective statement[2]

Not only have I been interested in cryptocurrencies for quite some time, but I am also fascinated about the underlying technology. Why do I mention cryptocurrencies within this context? These currencies have one main goal: decentralization. They try to tear the power out of the hands of a central party.
Although interested into decentralization, I have never heard/read anything about the IPFS before I participated in the seminar communication systems. I only applied for this topic because I am intrigued by planets and space. However, neither was I aware at the time of my application that the topic had little to do with this, nor did I comprehend the possible implications of the IPFS in the foreseeable future.
When I first heard/read about decentralization within the context of cryptocurrencies, I barely understood what it is about. I just guessed that a few nerds had thought up something awesome. Not at all I have understood their main goals and ideals.
Over time, however, I realized what they try to achieve and became more aware of the problems that centralization can actually cause. I am extremely pleased that I happened to come across a topic, the IPFS, that addresses this very issue as well.
In general, I am an advocate of the fact that all people should have equal opportunities - in my opinion, this is a condition which we should aspire. I guess that the IPFS could pave the way to get there. With the IPFS, it becomes almost impossible to exclude people from the system or to negatively influence someone with low bandwidths. For sure I also like the fact that the IPFS could be able to connect planets. In my opinion, however, this is of secondary importance.
Especially intrigued am I about the fact that the IPFS is an open-source project - everybody can participate and everyone can contribute valuable content. What blows my mind, however, is that the IPFS has created something great with a number of, in my view, manageable technologies that could add a lot of value in the near future.
While trying out the IPFS, I was more than surprised how easy it is to participate and how stable the whole system runs. I must honestly say that it makes absolutely no difference to me whether I use the WWW with HTTP or the IPFS - I just want to have an underlying technology that is awesome and serves its purpose reliably. It must be clearly mentioned though that I am technology affine - for technology averse people, this may be quite different.
The in my view biggest challenge is to achieve a comprehensive adaption - people do not like change and this will not be different when it comes to the IPFS. In order to achieve adaption, people should be incentivized to change. To be honest however, I reckon that there most likely exist parties which are everything but interested in people to switch to the IPFS.
In my opinion, the advantages of the IPFS outweigh the disadvantages by far. Personally, I would be quite willing to make the switch to the IPFS. However, I would like to explicitly mention that centralized systems also have advantages and that these systems may be the better solution depending on the use case.
In my view, however, such a system is more than necessary and it meets the demand of our time. In addition to the many advantages, the IPFS has a huge number of use cases which make the concept extremely interesting and even more useful.

8.5.3 Future outlook

Eventhough the IPFS comes along with many benefits and could live up to some of our ideals, it is presently difficult to say where the journey will take us. However, it is clear that

[2]Conclusion based on my own opinion and made experiences with the IPFS. It is highly subjective and may not reflect the opinion of other people.

a trend towards decentralization/distribution is currently in full swing and will certainly be strengthened by the current cryptocurrencies hype.
However, one thing must not be forgotten: The IPFS is far from being a flawless system without drawbacks. Nevertheless, it shows nicely what kind of technology we may be dealing with in the near future. Besides the more obvious advantages of the IPFS, which especially occur as a result of decentralization, and its impact on its future adoption, there are, in the view of the author at least, two main aspects which may even be more crucial to its long-term success:

- Since humanity is entering a new era in which data distribution becomes increasingly important (global availability of cheap smartphones, IOT), our current systems may no longer be able to meet the future requirements soon - The IPFS could hold the key to success.
- Fast and reliable access to global networks is more and more becoming a critical factor for success. Be it for countries, communities or individuals, without it, operating successfully in our interconnected world becomes almost impossible - The IPFS could hold the key to success.

Considering the above-mentioned aspects as well as the advantages of decentralization, it becomes clear immediately, that the IPFS may play a major role in the very near future. How and in what form, the future will show.

> At the bare minimum, it can be used as a global, mounted, versioned file system and namespace, or as the next generation file sharing system. At its best, it could push the web to new horizons, where publishing valuable information does not impose hosting it on the publisher but upon those interested, where users can trust the content the receive without trusting the peers they receive it from, and where old but important files do not go missing [63].

Bibliography

[1] R. Schreiner: Computer Netzwerke; specialist book, (Munich, Vol. 6), 2016, p. 3.

[2] R. Schreiner: Computer Netzwerke; specialist book, (Munich, Vol. 6), 2016, p. 7.

[3] R. Schreiner: Computer Netzwerke; specialist book, (Munich, Vol. 6), 2016, pp. 4-6.

[4] InterPlanetary File System: Hashing; https://docs.ipfs.io/concepts/hashing, 03, 2021.

[5] ProtoSchool: Anatomy of a CID, Lesson 2; https://proto.school/anatomy-of-a-cid/02, 03, 2021.

[6] Geeks for Geeks: Centralized, Decentralized and Distributed Systems; https://www.geeksforgeeks.org/comparison-centralized-decentralized-and-distributed-systems, 03, 2021.

[7] Gemini: Centralized, Decentralized, and Distributed Networks; https://www.gemini.com/cryptopedia/blockchain-network-decentralized-distributed-centralized, 03, 2021.

[8] R. Schreiner: Computer Netzwerke; specialist book, (Munich, Vol. 6), 2016, p. 90.

[9] Internet Assigned Numbers Authority: Root Servers; https://www.iana.org/domains/root/servers, 03, 2021.

[10] R. Schreiner: Computer Netzwerke; specialist book, (Munich, Vol. 6), 2016, pp. 91-92.

[11] R. Schreiner: Computer Netzwerke; specialist book, (Munich, Vol. 6), 2016, pp. 92-95.

[12] Stanford University: How does the internet work?; https://web.stanford.edu/class/msande91si/www-spr04/readings/week1/InternetWhitepaper.htm, 03, 2021.

[13] J. Smith: HTTP vs IPFS: is Peer-to-Peer Sharing the Future of the Web?; specialist report, 03, 2016, https://www.sitepoint.com/http-vs-ipfs-is-peer-to-peer-sharing-the-future-of-the-web.

[14] A. Hern: Google suffers global outage with Gmail, YouTube and majority of services affected; newspaper article, 12, 2020, https://www.theguardian.com/technology/2020/dec/14/google-suffers-worldwide-outage-with-gmail-youtube-and-other-services-down.

[15] BBC: Turkey's Wikipedia ban ends after almost three years; newspaper article, 01, 2020, https://www.bbc.com/news/technology-51133804.

[16] A. Case: Why The Internet Needs IPFS Before It's Too Late; specialist report, 10, 2015, https://techcrunch.com/2015/10/04/why-the-internet-needs-ipfs-before-its-too-late.

[17] J. Benet: IPFS - Content Addressed, Versioned, P2P File System; technical report, 07, 2014, p. 1.

[18] InterPlanetary File System: What is IPFS?; https://docs.ipfs.io/concepts/what-is-ipfs, 03, 2021.

[19] InterPlanetary File System: A modular paradigm; https://docs.ipfs.io/concepts/how-ipfs-works/#a-modular-paradigm, 03, 2021.

[20] InterPlanetary File System: Content addressing; https://docs.ipfs.io/concepts/how-ipfs-works/#content-addressing, 03, 2021.

[21] InterPlanetary File System: Content addressing and CIDs; https://docs.ipfs.io/concepts/content-addressing, 03, 2021.

[22] ProtoSchool: Anatomy of a CID, Lesson 1; https://proto.school/anatomy-of-a-cid/01, 03, 2021.

[23] InterPlanetary File System: Identifier formats; https://docs.ipfs.io/concepts/content-addressing/#identifier-formats, 03, 2021.

[24] InterPlanetary File System: Directed acyclic graphs (DAGs); https://docs.ipfs.io/concepts/how-ipfs-works/#directed-acyclic-graphs-dags, 03, 2021.

[25] J. Benet: IPFS - Content Addressed, Versioned, P2P File System; technical report, 07, 2014, p. 6.

[26] InterPlanetary File System: Merkle Directed Acyclic Graphs (DAGs); https://docs.ipfs.io/concepts/merkle-dag, 03, 2021.

[27] InterPlanetary File System: Distributed hash tables (DHTs); https://docs.ipfs.io/concepts/how-ipfs-works/#distributed-hash-tables-dhts, 03, 2021.

[28] J. Benet: IPFS - Content Addressed, Versioned, P2P File System; technical report, 07, 2014, p. 2.

[29] D. Mazières, P. Mamounkov: Kademlia: A Peer-to-peer Information System based on the XOR metric; technical report, p. 1.

[30] InterPlanetary File System: Bitswap; https://docs.ipfs.io/concepts/bitswap, 03, 2021.

[31] D. Vanderbist: Bitswap; technical presentation, 09, 2019, pp. 2-9.

[32] unknown: About Bitswap; poster from the IPFS developer summit, (Berlin), 07, 2018.

[33] J. Benet: IPFS - Content Addressed, Versioned, P2P File System; technical report, 07, 2014, p. 3.

[34] libp2p: What is libp2p?; https://docs.libp2p.io/introduction/what-is-libp2p, 03, 2021.

[35] libp2p: Addressing; https://docs.libp2p.io/concepts/addressing, 03, 2021.

[36] libp2p: Peed identity; https://docs.libp2p.io/concepts/peer-id, 03, 2021.

[37] libp2p: Transport; https://docs.libp2p.io/concepts/transport, 03, 2021.

[38] libp2p: NAT Traversal; https://docs.libp2p.io/concepts/nat, 03, 2021.

[39] libp2p: Protocols; https://docs.libp2p.io/concepts/protocols, 03, 2021.

[40] libp2p: Stream Multiplexing; https://docs.libp2p.io/concepts/stream-multiplexing, 03, 2021.

[41] InterPlanetary File System: Immutability; https://docs.ipfs.io/concepts/immutability, 03, 2021.

[42] InterPlanetary File System: InterPlanetary Name System (IPNS); https://docs.ipfs.io/concepts/ipns, 03, 2021.

[43] J. Benet: IPFS - Content Addressed, Versioned, P2P File System; technical report, 07, 2014, p. 9.

[44] J. Benet: IPFS - Content Addressed, Versioned, P2P File System; technical report, 07, 2014, p. 10.

[45] InterPlanetary File System: Persistence, permanance, and pinning; https://docs.ipfs.io/concepts/persistence, 04, 2021.

[46] J. Benet: IPFS - Content Addressed, Versioned, P2P File System; technical report, 07, 2014, p. 7.

[47] InterPlanetary File System: Garbage Collection; https://docs.ipfs.io/concepts/persistence/#garbage-collection, 04, 2021.

[48] InterPlanetary File System: Participation; https://docs.ipfs.io/concepts/what-is-ipfs/#participation, 04, 2021.

[49] InterPlanetary File System: Pinning services; https://docs.ipfs.io/concepts/persistence/#pinning-services, 04, 2021.

[50] Filecoin: Filecoin store; https://filecoin.io/store/, 04, 2021.

[51] InterPlanetary File System: Share files; https://docs.ipfs.io/concepts/usage-ideas-examples/#share-files, 03, 2021.

[52] InterPlanetary File System: Collaborate; https://docs.ipfs.io/concepts/usage-ideas-examples/#collaborate, 03, 2021.

[53] InterPlanetary File System: Store assets; https://docs.ipfs.io/concepts/usage-ideas-examples/#store-assets, 03, 2021.

[54] InterPlanetary File System: IPFS as infrastructure; https://docs.ipfs.io/concepts/usage-ideas-examples/#ipfs-as-infrastructure, 03, 2021.

[55] InterPlanetary File System: Lower your storage usage; https://docs.ipfs.io/concepts/usage-ideas-examples/#lower-your-storage-usage, 03, 2021.

[56] InterPlanetary File System: Decentralize your data; https://docs.ipfs.io/concepts/usage-ideas-examples/#decentralize-your-data, 03, 2021.

[57] InterPlanetary File System: Build a dApp; `https://docs.ipfs.io/concepts/usage-ideas-examples/#build-a-dapp`, 03, 2021.

[58] InterPlanetary File System: Blockchain use-cases; `https://docs.ipfs.io/concepts/usage-ideas-examples/#blockchain-use-cases`, 03, 2021.

[59] InterPlanetary File System: Decentralized Identity; `https://docs.ipfs.io/concepts/usage-ideas-examples/#decentralized-identity`, 03, 2021.

[60] InterPlanetary File System: Decentralization; `https://docs.ipfs.io/concepts/what-is-ipfs/#decentralization`, 04, 2021.

[61] InterPlanetary File System: Decentralize the Web itself; `https://docs.ipfs.io/concepts/usage-ideas-examples/#decentralize-the-web-itself`, 04, 2021.

[62] InterPlanetary File System: Decentralized DNS; `https://docs.ipfs.io/concepts/usage-ideas-examples/#decentralized-dns`, 04, 2021.

[63] J. Benet: IPFS - Content Addressed, Versioned, P2P File System; technical report, 07, 2014, p. 12.

Chapter 9

Machine Learning for Network Management: Current Status and Possible Research Directions

Ivana Mesić

Network management functions ensure that the network is working properly and there are no faults or intrusions. With the growing complexity of network topologies, larger amounts of data and more users, network management has become increasingly harder to perform. Machine learning (ML) methods have been introduced to improve or replace existing network management technologies. These methods are still not widely used in practice, but they have recently been topics of a lot of research. In this paper, an overview of the most commonly used ML methods is given and their brief explanations. Also, novel network architectures such as SDN and NFV were described and how they can be used with ML methods. Finally, the most important network management functions were presented and how the ML methods can be used to improve them in practice.

Contents

9.1 Introduction

Today, it is almost unimaginable to function without using some type of network. For example, mobile phones, Wi-fi, Internet of Things and similar technologies have become an essential part of our society, and their influence has been even more prominent during the COVID-19 pandemic. Networks provide the basis for modern-day technology, economy and overall way of life, so it is important to ensure their proper functioning. In this sense, network traffic is growing each day, with new users and devices joining and creating more and more data. Network providers are struggling to deal with such exponential growth and to keep network performances up to standard using traditional network technologies [10]. Artificial intelligence, and especially Machine Learning (ML) methods, present solutions for tackling the surging network traffic and big data.

ML has seen a fast development in the recent years, integrating itself in numerous technologies that we use every day. Now, researchers are exploring ways to improve network management tasks using ML methods. The overwhelming amount of data that is generated each day by network users is impossible to monitor and manage manually, so ML techniques are introduced to provide automation. Using ML in every part of network management is being explored, from monitoring and routing to network security and quality of standards management.

Benefits that ML methods provide are faster data analysis and management, higher quality assurance for users and higher flexibility and scalability of network management systems. Although there are numerous advantages of using ML, it is still largely unexplored topic and needs a lot of research. One of the biggest challenges in using ML in NM is also a high costs for building infrastructure to support new methods [13].

In this paper, an overview of ML methods and their use in Network Management (NM) tasks is provided. In Section 9.2 the basics of ML algorithms and how they work are explained. In Section 9.3 an overview and categorization of most common ML methods is provided. Section 9.4 provides and overview of NM tasks and some new network architectures which are best suited for integrating with ML methods. In Section 9.5 the advantages and challenges of using ML in NM are listed. In Section 9.6 it the most common uses of each ML method category are listed in SDN networks. Finally, in Section 9.7, it is explored how ML methods are used categorized by NM tasks.

9.2 Machine Learning (ML)

To define ML, first it is required to define a broader term - Artificial Intelligence (AI). AI is a field of study which seeks to build computer systems that have human-like properties. These properties can be divided into four main categories: thinking like humans, acting like humans, thinking rationally and acting rationally [1]. AI can also be used as the name for the systems that have said properties. For example, key properties that make AI human-like is the ability to collect data in its surroundings, detect patterns and acquire knowledge from the data and produce conclusions.

ML is a branch within AI. ML is defined as "a set of methods that can automatically detect patterns in data, and then use the uncovered patterns to predict future data, or to perform other kinds of decision making under uncertainty" [2]. ML is closely related to data mining and statistics, but these fields differ from ML in emphasis and terminology [2]. One of the differences between ML and statistics is that ML is applied statistics with increased emphasis on the use of computers to statistically estimate complicated functions and a decreased emphasis on proving confidence intervals around these functions [3].

ML today is used in various aspects of life, such as medical research, molecular biology, text processing, computer vision and robotics. One of its usages is also in NM, which will

be further explained in this paper. Although it has a widespread usage, ML has great potential for improvement and for finding new use cases in which it can be used.

9.2.1 Learning Process

A ML algorithm is defined as "an algorithm that is able to learn from data" [3]. Here there is a need to also define what "learning" means in the context of this definition. In [4], a definition of learning is given as: *"A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E"*. Since there are many different tasks, experiences and performance measures which can be used in a ML algorithm they will not be formally defined, but rather some of the most frequently used examples will be named.

9.2.1.1 ML Tasks

As the definition of learning states, the ML algorithm has the goal to do a certain task and it does it by learning through experience and with more experience tries to improve its performance measures. The ML task is usually the desired outcome of processing input data, called an example. An example is a set of features, usually marked with a vector $\mathbf{x}$. Tasks that the ML algorithm can be given are:

- **Classification** - Classification is the process of determining to which of the k predefined classes (or labels) the observed example $\mathbf{x}$ belongs to. When solving classification tasks, the ML algorithm usually produces a function f (f: $\mathbb{R}^n \rightarrow \{1,2,...,k\}$) which maps the values from $\mathbf{x}$ to one of the k classes, such that $y = f(\mathbf{x})$ or it can output a probability distribution over classes. Examples of classification tasks are handwriting recognition, face recognition, flower classification, and email spam filtering.

- **Regression** - Regression is a process similar to classification, but the output is a numerical continuous variable instead of categorical. Here the ML algorithm produces a function (f: $\mathbb{R}^n \rightarrow \mathbb{R}$) which maps the input example to a numerical value [3]. Some of the examples of regression tasks are predicting the tomorrow's price of stock based on today's price; predicting the exact temperature at any location in the building given weather data, time and door sensors; predicting the location in 3d space of a robot arm end effector, given control signals (torques) sent to its various motors [2].

- **Transcription** - Transcription tasks require the ML to observe an unstructured representation of some kind of data and transcribe it into discrete, textual form [3]. Two most notable examples of transcription are: text recognition - when the ML algorithm gets a picture of a text and it must produce the same text as a sequence of characters and speech recognition - when the algorithm is given an audio recording and it must produce the written transcript of the recording.

- **Machine translation** - Machine translation tasks require translating a set of characters from one language to another language. It is commonly used in natural language processing (*e.g.,* translating from German to English).

- Some other notable tasks that will not be further explained are **anomaly detection** (marking odd examples from a set), **synthesis and sampling** (generating new examples that are similar to the training data) and **imputation of missing values** (given an example which has missing values, predict the missing values based on former training data).

9.2.1.2 ML Algorithm Workflow

The following workflow for an ML algorithm can be used as a guideline to design and implement new ML algorithms [2].

1. **Data preparation and preliminary analysis**: First step in the workflow is to understand the data and prepare it for further analysis.

2. **Feature extraction**: This step includes presenting the data as a set of examples. Each example should have a set of features which can then be processed by the algorithm.

3. **Dimensionality reduction**: Dimensionality reduction is not a necessary step in the workflow. Dimensionality reduction refers to eliminating some features from the examples. This could be done either because the features are irrelevant and because of them the model is not working properly or because there are too many features and the model is becoming too complex.

4. **Model selection**: Model selection refers to the process of selecting the correct mapping function and choosing its hyperparameters. The hyperparameters of the model are one level above the usual model parameters, for example if the model is a polynomial function the polynomial degree would be a hyperparameter and the coefficients of the function would be regular model parameters.

5. **Model training**: Model training is the process of iterating the model function on the training data set.

6. **Model evaluation**: Model evaluation is a step in which the performance of the model is evaluated by testing on the validation and test data sets.

7. **Diagnostics and debugging**: In this step, modifications to the model are made if the performance is evaluated as inadequate.

9.2.1.3 Model Evaluation and Performance Measures

When preparing the data set for a ML algorithm, it is split into 3 parts: a training data set, a test data set and a validation data set. The training data set is used to train (or learn) the model. The validation data set is used to ensure that the correct model was chosen or to evaluate the model. The test data set is then used to test the performance measures of the model based on its predictions [10].
The validation data set is used to ensure that the algorithm generalizes well. The goal of the algorithm is to predict unseen data, so it needs to have an optimal level of generalization. Since big amounts of data are always hard to obtain, the most common way of producing validation sets is through k-fold cross validation. The k-fold cross validation splits the data into k separate sub-sets and then trains k-1 sub-sets and evaluates the remaining set. It repeats the procedure until all sub-sets are processed [10]. The overall performance estimation is then an average of all the iterations. The model with the best performance estimations is then chosen.
As mentioned, the evaluation process seeks to find the optimal way of generalization. This is done by measuring the validation error and the training error. The **training error** is the error on the training data set. As the algorithm keeps learning, the training error will decrease. The **validation** or **test error** is the error on yet unseen data or the test data set [2].
In Figure 9.1, it is depicted the **underfitting** and **overfitting** that can occur in ML algorithms [10]. If the generalization rate is too high, we talk about **underfitting**. This means

that the model is too simple or that it did not have enough training data or iterations to learn from. When we let the model learn the training and test error keep falling up to a certain point, as shown in the figure below. During that time, the generalization rate keeps falling and the model is improving its' complexity. The point at which the test error stops falling is the optimum of point of the model's complexity. From that point forward, the model is **overfitting**. That means it is loosing the generalization and is adapting too much to the training data set, so it performs badly on unseen data from the test data set. The goal of the evaluation and validation process is to find a balance between overfitting and underfitting [2].

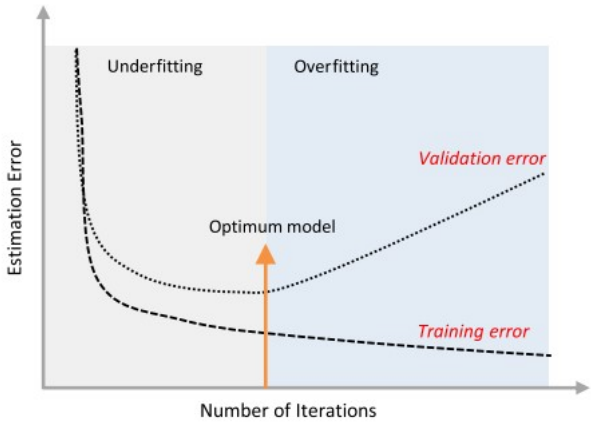

Figure 9.1: Underfitting and overfitting [10]

When the model is evaluated and the best model is chosen, we can measure the effectiveness of the model on the test data. Here, some of the most commonly used performance measures are listed. All of the measures below are used with classification tasks [10].

- Confusion matrix: An example of a confusion matrix is presented on the figure 9.2. This confusion matrix is made for a classification problem with 2 target classes. Confusion matrices are mostly used in such problems, but they can also be expanded for use with multiclass classification problems. The confusion matrix displays 4 performance indicators of a classification model. The goal is to maximize the true positive rate and the true negative rate, because they show that the classificator is working properly (classifying examples correctly). The goal is also to minimize as much as possible the false positive and false negative rates. Depending on the task of classification, the architect of the model can penalize one rate more than the other (for example, in cancer detection false negatives should be penalized more than the false positives).

- Accuracy: A = (TP + TN)/(total examples). The accuracy simply measures the rate of correctly classified examples.

- Sensitivity or True positive rate (TPR): TPR = TP / (TP + FN) The TPR gives the probability that a true outcome is actually true [10].

- Specificity (SPC) or True negative rate: SPC = TN / (TN + FP) Specificity gives the probability that a false outcome is actually false [10].

- Precision or Positive predictive value (PPV): PPV = TP/(TP+FP) The precision rate gives the proportion of actual true outcomes, given the set of examples classified as true. The precision is used when rare events are classified. In that case, there will be a lot of negative examples, so to improve efficiency and reduce the number of data to process, the focus is only on data classified as true.

- False omission rate (FOR): FOR = FN/(FN + TN) FOR is the proportion of false negatives, given all of the examples classified as false [10].

- Receiver Operating Characteristic (ROC) curve is the curve which plots the relation between FPR and TPR. The goal is to make the curve pass through point (0,1) because it would mean the TPR is 1 and FPR is 0. Since it is impossible to make such a perfect classifier, the goal is to make the curve approach that point as much as possible. If the curve is the line x=y, the classifier is worthless, since it classifies examples randomly.

- Area Under the Curve (AUC) is the area that the ROC curve closes with x and y axes. The maximum AUC is 1, and the goal is to make AUC close to 1. If AUC is 0.5, this means the curve is x=y and that signifies that the classifier is worthless.

Other performance measures that can be applied to problems that are not binary classification [10]:

- Accuracy: Accuracy, as in classification performance measures, is the rate of correct predictions made on the data set [10].

- Mean absolute error (MAE): The error is calculated as the difference between the predicted (y') and the actual (y) output $e = y' - y$. The mean absolute error is the mean of absolute values of errors for each example in the test set (mean($|e|$)).

- Mean squared error (MSE): MSE is the mean of squared values of individual errors (mean($|e^2|$)).

- R^2: $R^2 = SS_E/SS_T$ R^2 is also called the goodness of fit. It is the ratio of the expected sum-of-squares and the true sum-of-squares. The sum of squares is calculated with the following formula $SS = \sum_i (y_i - y_i')^2$

PREDICTION	ACTUAL LABELS	
	True	False
True	True Positive (TP)	False Positive (FP)
False	False Negative (FN)	True Negative (TN)

Figure 9.2: Confusion Matrix

9.2.2 Components of a ML Algorithm

In Section 9.2.1.3, it was introduced what it means when a ML algorithm learns something and improves its results. Here, it will be explained which components of the algorithm enable it to do that. Nearly all ML algorithms can be described as a combination of a specification of a data set, a cost function, an optimization procedure and a model [3]. These components can be replaced mostly independently and with those replacements we get a wide variety of ML algorithms.
The **data set** is the data that is given to the ML algorithm. The data in supervised learning will be made of input examples and their paired output labels and in unsupervised learning the data will consist only of input examples. The input examples are usually represented with a matrix $\mathbf{X}$ (N x D) and the output labels are usually represented with a vector $\mathbf{y}$. In

the ML process, the data set is usually divided into 2 or 3 parts. The two main parts are the training and the test data sets. During the first phase of the process, the algorithm learns solely on training data and then the algorithm is evaluated on the test data set. To ensure that the evaluation is correct, there cannot be overlapping between the test and train data sets.
The **model** of a ML algorithm is a set of predictor functions or hypothesis. A hypothesis is a function f (f: $X \rightarrow Y$) which predicts an output label based on the given inputs. During the optimization process this set of functions is searched through in order to find the one that best predicts the desired outputs [6]. This can also be described as searching for the optimal parameters that describe the hypothesis function. In practice, most used models are probabilistic models which define probability distributions $p(\mathrm{y}|\mathbf{x})$ or $p(\mathbf{x})$, depending on whether supervised or unsupervised learning methods are used. If a model has a fixed number of parameters, they are called *parametric models*. Parametric models are faster to use, but make stronger assumptions about the nature of data distributions [2]. In the *non-parametric models* the number of parameters grows with the size of the training data set. They are more flexible than parametric models, but have a disadvantage when dealing with large data sets because they become computationally too complex.
If we consider a function (hypothesis) f which makes a prediction y', when the truth is y, then the **cost function** is defined as the cost of predicting y', when the truth is y [2]. The most common cost function used is the negative log-likelihood (NLL), because minimizing the NLL results in maximizing the log-likelihood function and causes maximum likelihood estimation [3].
The **optimization procedure** is the use of optimization algorithms to improve the ML model. That refers to tuning weight parameters of a function or training an artificial neural network. It can also be described as searching through a set of hypothesis to find the optimal one(s). There is a general categorization of optimization algorithms into first-order and second-order optimization algorithms. Those that use only the gradient (such as the gradient descent method) are called the first-order methods, and those that also use the Hessian matrix (such as Newton's method) are called second-order methods [3]. One especially successful group of optimization methods are convex optimization algorithms, which can only be applied to convex functions. They are useful because their local minima/maxima are necessarily also global minima/maxima, but unfortunately not all problems in ML can be expressed in terms of convex optimization [2].

9.3 Machine Learning Methods

ML methods can be broadly categorized into 3 main groups: supervised learning methods, unsupervised learning methods and reinforcement learning methods. If we consider the definition of the learning process, given in Section 9.2.1, we can state that these methods have differences in the experience component of the learning process [3].

9.3.1 Supervised Learning Methods

Supervised or predictive learning is an approach in ML in which the goal is to predict the output variable(s) based on the given inputs. The algorithm gets a training set in which there are labeled input-output pairs from which the algorithm creates a mapping function. Using the mapping function, the algorithm can then predict the output variables of the test data set.
The input data set consists of N examples and each example consists of D features that are later mapped to output $\mathbf{y}$. If the output is categorical the problem is known as classification or pattern recognition and if the output is continuous the problem is known

as regression. Supervised learning methods are the most commonly used in practice and one of the reasons is because they provide a reliable way to measure accuracy. One of the problems arising when using supervised learning is that the labeled training data sets are hard and expensive to obtain.

9.3.1.1 k-Nearest Neighbors (KNN)

K-nearest neighbors is an example of a simple non-parametric classifier. This classifier determines probability $p(y|\mathbf{x}, D, K)$ that a new test example x is in class c. It calculates the probability by how many of the k closest points from the training set to the test point are members of the given class c. To determine which points are the "closest", many measures can be used, but the Euclidean distance is the most common [2]. If the method is used for regression, the predicted outcome is the average or weighted-distance average of the k-nearest neighbors [10].

9.3.1.2 Artificial Neural Networks

Artificial neural networks are ML systems which are made of layers of small units that work together to solve ML tasks. These units or neurons are interconnected and they can have various network topology. There are many types of neural networks, but the two most commonly used are the feedforward neural network (multilayer perceptrons) and the convolutional neural network (CNN).
The **feedforward neural network** or **multilayer perceptron** (MLP) is a series of units called perceptrons. The perceptron is a simple model (most commonly used is logistic regression) that takes in inputs from the former layer and sends its output to the next layer in line. The layers in between the first and the final layer are called hidden layers. There can be different number of units in each of the layers in the network. On the figure 9.3, a simple topology of an ANN with one hidden layer can be seen. The outputs from the first layer are multiplied by a set of weights to produce the inputs for the hidden layer and then the same process is repeated with a different set of outputs and weights to produce the input for the final layer. The inputs can be represented with a matrix $\mathbf{X}$, the weights with a matrix $\mathbf{W}$ and the outputs with a matrix $\mathbf{Y}$ [2]. The MLP uses the backpropagation algorithm for training the neural network. In the backpropagation algorithm, the weights are modified until the desired accuracy of the system is achieved. One can show that an MLP is a universal approximator, meaning it can model any suitably smooth function, given enough hidden units, to any desired level of accuracy [2].

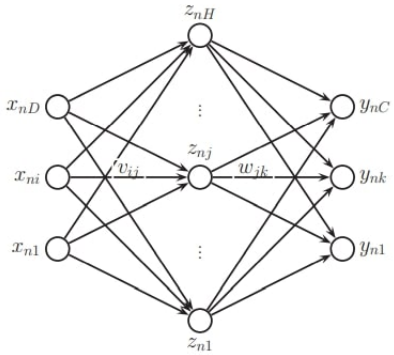

Figure 9.3: A Simple artificial neural network [2]

Convolutional neural networks are a type of MLP that are best suited for processing pictures and useful for problems where the original input features are not very individually informative [2]. For example, the individual pixels in an image are not very informative,

but when they are processed in groups, the features that they form can be detected. This process is called feature extraction, because the network learns non-linear combinations of original inputs [2]. To reduce the number of weight parameters CNNs use weight sharing. This means that during the optimization process not every individual weight of an input needs to be calculated which reduces complexity and saves time [5]. Convolutional neural networks have a property called translation invariance, which means that the network can classify a pattern no matter where on the image it is located [2].

9.3.1.3 Support Vector Machines (SVM)

Support Vector Machine (SVM) is a classification technique. This method chooses a function (hyperplane) that divides the feature space into 2 parts. The new data points are classified based on the part of the feature space they fall into [10]. The hyperplane that divides the feature space is chosen in such a way that it has a maximum margin to the two (or more) nearest points on each side. The data points that are closest to the dividing line are called support vectors. The minimum number of support vectors is 2 and the hyperplane must have an equal distance to both of those 2 vectors. On Figure 9.4 a simple SVM is shown with two support vectors.
SVM can be linear or nonlinear. In the linear case the data is linearly transformed and mapped into the feature space. If the data is nonlinear, kernel functions are used to map the input data into a feature space that can be divided by the hyperplane.

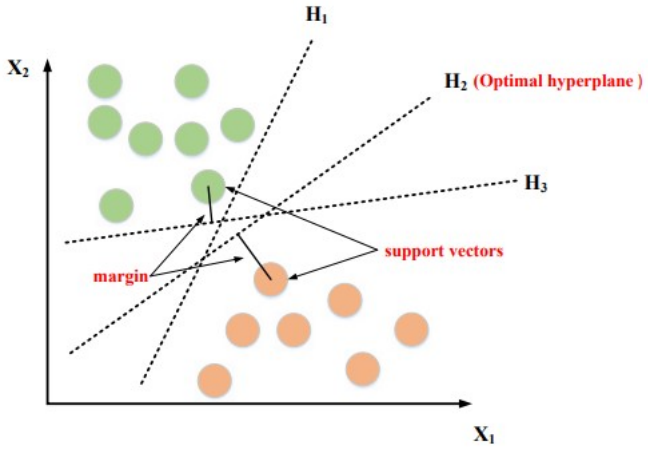

Figure 9.4: Support vector machine with 2 support vectors [14]

9.3.1.4 Decision Trees

As their name states, decision trees (DT) are ML algorithms that perform classification through learning trees. In decision trees "each node of the tree is associated with a region in the input space, and internal nodes break that region into one sub-region for each child of the node" [3]. New examples are classified into a region by comparing the example with each node of the tree [14].
If the size of the decision tree is unlimited, the method could be considered non-parametric, but since they are usually constrained in practice they are parametric methods. Decision trees can sometimes struggle when solving simple problems easily solvable with logistic regression, but they are useful when computational resources are constrained. Other advantages of DT are intuitive knowledge expression, simple implementation and high classification accuracy [14].

9.3.2 Unsupervised Learning Methods

Unsupervised or descriptive learning is also known as knowledge discovery. That is because the goal of this process is to find "interesting" data patterns from the given inputs [2]. The unsupervised ML techniques are leveraged to learn about similarities and patterns in data and generate clusters that can be used to identify classes of interest [7]. The tasks in unsupervised learning are not well defined, since we cannot tell which algorithms to look for or with which performance measures to measure the success of the algorithm.

9.3.2.1 Clustering

Clustering is "the process of grouping similar things together" [2]. The input data is divided based on some principle into subdivisions or clusters. Two categories of clustering are similarity-based clustering and feature-based clustering. In similarity-based clustering the input is a dissimilarity (or distance) matrix ($N\ x\ N$) and in feature-based clustering the input is a feature matrix ($N\ x\ D$). The advantage of similarity-based clustering is that it allows easy use of kernel functions, while feature-based clustering works well with raw data with potential noise in it.
Clustering can also be classified based on the output it produces. If the data is divided into disjoint sets, the clustering is called **flat** or **partitional**. Flat clustering is computationally simpler and faster to use.
If the output of the clustering process is a nested tree of partitions, the clustering is **hierarchical**. Hierarchical clustering can be done from the top-down, or from the bottom-up. If it is done from the top-down, the initial cluster is divided into smaller clusters in each step based on some similarity measure. If the approach is bottom-up, in the beginning each data point is a cluster itself. They are then grouped into larger clusters and the process is done when everything is joined in one large cluster. Hierarchical clustering is more complex than flat clustering, but it is often more useful.
One of the common clustering methods is the **k-mean clustering**. In k-means clustering the data is divided into k clusters. The iterative process can be described with these steps:

1. Choose k random data points - nodes. These will be the initial centroids.
2. Go through each node and pair it with the centroid that is the closest to it. To determine which nodes are the closest, a predefined distance function is used.
3. When all nodes are divided into groups formed around centroids, calculate which node is the new centroid inside the formed groups.
4. Repeat the process until a predefined convergence condition is met.

One other clustering method is called **spectral clustering**. Figure 9.5 depicts how this method is better suited for some data sets because of their shape in the feature space. The basic algorithm of spectral clustering is:

1. Construct a k-nearest neighbor graph from data. Create a weight matrix (W) of the graph.
2. Compute L = D - W. D is a diagonal matrix with d_i weighted degree of each node.
3. Find eigenvectors of the matrix
4. Construct a matrix from these eigenvectors V = $[v_2, v_3, ..., v_l + 1]$
5. Apply a clustering algorithm using V as a feature matrix.

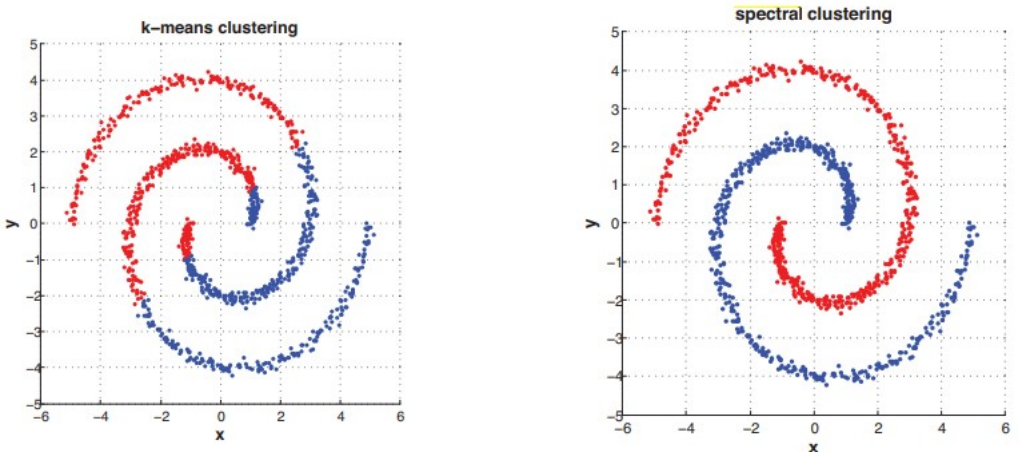

Figure 9.5: Difference between k-means and spectral clustering [2]

9.3.3 Reinforcement Learning

Reinforcement Learning (RL) methods are most rare in practical use between the 3 types of ML methods. In some cases, the output of the system is a sequence of actions and a single action is not as important as the general policy that is the sequence of actions to reach the goal. Reinforcement learning algorithms are used in cases such as these to assess the goodness of policies and learn from past good action sequences to be able to generate a good policy [5].
One further difference that the reinforcement learning algorithms have is that they do not have a fixed data set. They interact with an environment which creates a feedback loop between the learning system and its experiences [3] that can be seen on the figure 9.6. RL involves an agent, a state space S and an action space A [14]. The agent tries to maximize its long-term reward by learning from the environment, which creates the before mentioned feedback loop, also presented on the figure below. The objective of the agent is to learn the optimal behavior policy π that maps the state space S to the action space A. With this mapping it maximizes the expected long-term reward by choosing the best corresponding action, given a particular state [14].

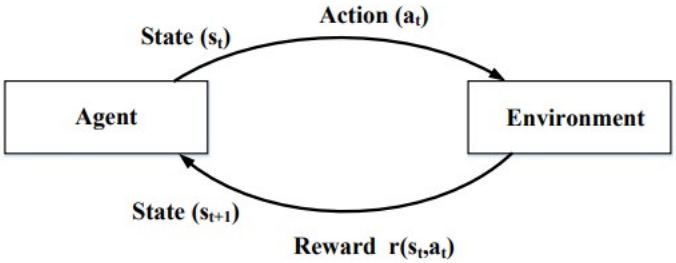

Figure 9.6: RL feedback loop [14]

Some of the use cases in which reinforcement learning algorithms are used are game playing and robot navigation through space. In each of these two examples a single action is not as important as the final goal of the system. For example, it is not as important exactly which trajectory a robot should take to reach its final destination, as long as it does it efficiently and successfully. Reinforcement learning algorithms have shown superhuman performances in complex games such as Go, but they are still largely unexplored and difficult to scale for use in large systems.

9.4 Network Management

In order to explain how various ML techniques and methods can be applied to network management, first we need to define what network management is and what does it include. Network management is a set of tools, methods and procedures whose main focus is to control, plan, allocate, deploy, coordinate, and monitor network resources [15]. Since network management is a broad term which includes many functions, it can be divided into multiple layers of management [15]:

- *Business management*: management of budgets/resources, planning, and agreements.
- *Service management*: management of delivery of services to end users (such as providing data storage or application delivery).
- *Network management*: "Management of all network devices across the entire network" [15]
- *Element management*: management of collections of similar network devices (subscriber management systems, access routers).
- *Network-element management*: management of individual network devices [15].

There are several of tasks that fall into the category of NM, but they can be broadly categorized into 3 main categories: **monitoring** (which can be further divided into monitoring for event notification and monitoring for trend analysis and planning), **configuration** and **fault management** [15] [16].
If the focus is on network-element management, or management of individual network devices, the tasks which it includes are described with the "FCAPS" model. The FCAPS stands for Fault, Configuration, Accounting, Performance, and Security management, which are also the main tasks of network-element management.

9.4.1 Network Management Tasks

In the section below it is explained what the above listed categories of management tasks include.

- *Monitoring for event notification*: an event in a network is "something that occurs that is noteworthy" [15]. Events can be various, e.g., a characteristic reaching a threshold value or a problem or alarm that occurs in the network. There are some actions that should be taken in case that an event occurs. Some of these actions could be logging the event into a log-file or notifying someone. In case there is a need for issuing an alarm (notifying someone), depending on the event, the notification could be issued to an administrator, the end user or a manager. To issue immediate notifications, the system needs to be continuously monitored and checked - a process called real-time analysis [15]. The system performs real-time analysis by gathering data from various network devices using short polling intervals to ensure the events are processed in real-time. Since the network devices produce large amounts of data, the system cannot gather all the data all the time. Thus, the number of characteristics and network devices needs to be adapted in order for the infrastructure (CPU, storage, memory) to support gathering of the data. The data gathering should also be limited in such a way that it does not impact the overall performance of the network, since large amounts of data can negatively influence network's performance.

- *Monitoring for trend analysis*: Monitoring for trend analysis also collect data from various network devices, as with monitoring for event notification. The difference is that this collecting is done in longer polling intervals, since there is no need for real-time responses to events. The gathered data is then processed and used to predict future events. This kind of monitoring is especially useful when planning for network expansion and growth.

- *Configuration management*: includes setting system parameters for turn-up; provisioning the network; configuration and system backups and restores; and developing and operating system databases [15]

- *Fault management*: includes processing events and alarms; problem identification, troubleshooting, threat mitigation; taking actions to return the network into its normal state[15]

- *Performance management*: includes implementing performance controls, network and system performance data collection and analysis; network and system performance parameters control [15].

- *Accounting management*: includes monitoring and managing subscriber service usage and service billing [15].

- *Security management*: includes implementing security protocols; mitigating threats; collecting and analyzing security data; generating security reports and logs [15]

9.4.2 Network Management Mechanisms

9.4.2.1 Monitoring Mechanisms

Monitoring is the process of collecting data, processing it and displaying the processed data. Data is collected through polling of the network devices. The polling intervals are the time intervals from one "ping" of the network device to the other. The polling intervals are adjusted based on what is being monitored during the process. After data collection, the next phase is data processing. Data processing includes extracting characteristics from the data and modifying existing data. To display the data, various techniques can be used, some of which are: logs and textual displays, graphs and charts (both static and moving), and alarms [15].

9.4.2.2 Configuration Mechanisms

Configuration includes operating and controlling network devices by setting their parameters. Configuration mechanisms include direct access to devices, remote access to devices, and downloading configuration files [15]. As different network devices require setting different control parameters, the goal is to produce a table of configuration parameters, establish the methods for configuring these parameters and understand the effects of changing each parameter. Configuration mechanisms also need to understand how network problems effect different network devices in order to be able to mitigate network problems by configuring the parameters of the affected devices.

9.4.2.3 Instrumentation Mechanisms

Instrumentation is " the set of tools and utilities needed to monitor and probe the network for management data" [15]. Instrumentation mechanisms include access to network management data via SNMP (Standard Network Management Protocol), monitoring tools, and direct access. The 3 main functions of instrumentation mechanisms are performance

monitoring, workload measurement and incident and problem management. It is crucial to build a robust and simple instrumentation mechanism, to ensure that the instrumentation system itself does not collapse when network problems occur. Building a hierarchy in data flow management and having multiple systems collecting data are good principles for building robust instrumentation systems.

9.4.3 Novel Network Architectures

Today, technology is changing and improving at a faster pace than ever. Networks are not resilient to these changes. Since there is more and more network users and services, there is also a need to continually improve the network architectures to support such growth. The three main drivers of change in network architectures can be grouped as: increase of demand, increase of supply and changing of traffic patterns.
Increase of demand is the result of more mobile traffic, increasing number of devices in Internet of Things (IoT), processing of Big Data and a shift of enterprises to cloud computing. All of these factors put pressure on existing traditional architectures of enterprise networks and the Internet [17].
As the demand on networks is rising, so is the capacity of network technologies to absorb rising loads. Introduction of 4G and 5G networks, improving Ethernet and Wi-Fi speed and increasing the performance of network devices are some of the ways the network infrastructure has improved to support the increasing demand.
Although some network elements have improved to support the rising demand, the change of traffic patterns makes traditional enterprise networks increasingly unequipped to deal with such traffic. In addition, the variety of applications impose quality of service (QoS) and quality of experience (QoE) requirements on the networks, which means that the traffic load must be handled in an increasingly sophisticated and agile fashion [17].
The deficiencies that traditional network architectures have are static complex architectures, inconsistent policies, inability to scale and vendor dependence. These are the issues that are being tackled by changing to different architecture types such as **Software-Defined Networking** (**SDN**) and **Network function virtualization** (**NFV**).
In the following sections, two new network technologies will be presented. These two technologies (SDN and NFV) have been recently introduced as a replacement of traditional network architecture to try to face with the upcoming challenges of dealing with surging demand. Their main advantages are increased automation, maintainability and on-demand scalability.

9.4.3.1 Software-Defined Networking (SDN)

In the recent times, SDN architectures have reached a tipping point and are replacing the traditional networking model. They provide an increased level of adaptability and flexibility which are needed to meet the needs of new technologies such as IoT, cloud and increased mobile networking.
In routing packets, there are 2 main elements involved: the control function and the data function. The **control function** determines the routes and relative priority of traffic while the **data function** routes the traffic based on the determined control-function policy [17].
The main difference between SDN and traditional network architectures is that SDN has a decoupled data plane and control plane, while the traditional network architectures have those planes integrated. In traditional networks, each network device (router, bridge, packet switch etc.) would perform both the control and the data function and would need to implement routing and control network protocols. In SDN all routing, naming, policy declaration and security checks are centralized in a central controller [17] while the data plane devices become simple packet-forwarding devices.

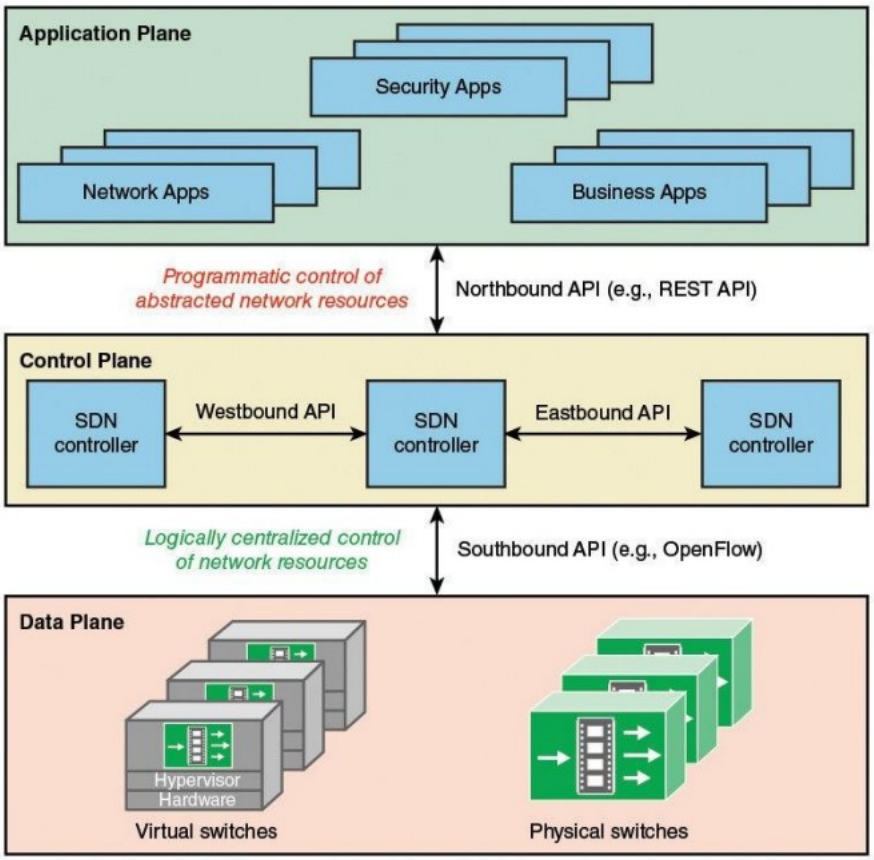

Figure 9.7: SDN Architecture [17]

On the picture 9.7, an overview of the SDN architecture can be seen. It is made out of 3 different layers or planes which are independent and connected with open interfaces. The 3 planes are:

1. **Data Plane**: The data plane is also known as the infrastructure plane and it is the lowest plane in SDN architecture. The 2 main functions that the data plane has are *control support function* and *data forwarding function* [17].

   The control support function includes interacting with the control plane through Southbound APIs (e.g. OpenFlow protocol which is the first and most common open standard Southbound interface [14]) in order to support programmability. The data forwarding function includes forwarding, dropping and modifying packets based on policies received from the control plane [14]. The key characteristic of data plane devices is that they perform these forwarding functions without embedded software to make those decisions, but rather just following the rules from the control plane.

   The data plane is comprised of forwarding devices, which include virtual and physical switches. Virtual switches are software-based switches which can run on common operating systems, while the physical switches are hardware-based switches [14].

2. **Control Plane**: The controller plane is the central plane of the SDN architecture. Its main characteristics are updating forwarding rules dynamically, and making network administration flexible and agile [14].

   Control plane is implemented as a server or a set of coordinated servers called controllers. There are a number of different controller architectures such as NOX, POX, Floodlight, Ryu, OpenDayLight and Beacon. The most prominent one is the OpenDayLight architecture which t aims to produce an extensible, open source, virtual networking platform atop such existing standards as OpenFlow [17].

A general definition of the control plane's functions is "mapping application layer service requests into specific commands and directives to data plane switches and supplying applications with information about data plane topology and activity" [17]. The specific functions of the control plane are:

- shortest path forwarding - establishing preferred and shortest traffic routes
- notification management - receiving, processing and forwarding application events
- security mechanisms - provides isolation and security enforcement between applications and services [17]
- topology management - builds and stores information about the network topology
- statistics management - collects the data on traffic
- device management - configuring switch parameters and attributes and managing flow tables

The control plane uses 3 different types of interfaces to communicate in the SDN network: southbound, northbound and east/west-bound interfaces. The southbound interfaces are used for communicating with the data plane and the northbound interfaces are used for communicating with the application plane. The east/west-bound interfaces are used in multi-controller SDN networks. A multi-controller network is a type of SDN network where there are multiple controllers, each controlling its own domain, because the scale of work would be overwhelming for just 1 controller. In order to provide a global-network view to upper-level applications, the controllers exchange information using east/west-bound interfaces.

3. **Application Plane**: The application plane (AP) is the highest plane in the SDN architecture. The scope of the application plane is not as well defined as the scope for the 2 other planes in SDN. The broad definition is that it includes a number of network applications and services that deal with network management and control.

   The applications and services which are part of the AP interact with the control plane through Northbound interfaces. The most common Northbound interface used is the REST interface. The *network services abstraction layer* (NSA layer) is a layer in SDN architecture which can be implemented as a part of the application plane, the controller plane or as a part of the Northbound interfaces. The NSA layer provides an abstract view of network resources which means that the the applications can access control plane functions without knowing the details of underlying network switches. It also provides a generalization of control plane functionality which enables the applications to work on a wide range of controller network operating systems [17].

   The applications plane's main function is to obtain the network state information through the Northbound interfaces and monitor, control and manage the network through this information. Based on the received information and business requirements, the applications can easily implement the control logic to change network behaviors and forward the modified rules to the control plane, which will further propagate the changes.

9.4.3.2 Network Function Virtualization (NFV)

Network Function Virtualization (NFV) includes switching the function of hardware devices to their virtual representations on a virtual machine running on a virtual server

infrastructure. It involves the implementation of network functions in software that can run on a range of industry standard server hardware, and that can be moved to, or instantiated in, various locations in the network as required, without the need for installation of new equipment [23]. The goal of NFV is to evolve standard IT virtualisation technology to consolidate many network equipment types onto industry standard high volume servers, switches and storage.
New network services may require additional different types of hardware devices. These devices require space and power and additional capital expenditures. Replacing them with software solutions would save the before mentioned resources. Hardware devices also have a shorter life-cycle and additional resources are wasted to design new ones and integrate them into networks. Software solutions are easier to update and provide more adaptability and independence from hardware platforms [17].
Network devices that can be virtualized with NFV are: Enterprise Access Routers, Provider Edge Routers, Enterprise Firewalls, Enterprise WAN Optimization Controllers, Deep Packet Inspection devices, Intrusion Prevention Systems and other Security appliances and Network Performance Monitoring functions [23].
Virtualization technology enables a single PC or server to simultaneously run multiple operating systems or multiple sessions of a single OS. The host operating system can support running multiple Virtual Machines (VMs) each of which has different operating system or hardware characteristics [17]. A VM Monitor (VMM) or hypervisor is the software which sits between the hardware and the VMs acting as a resource broker and ensuring that the VMs can share the host's resources effectively. The number of VMs that can exist on a host is called a consolidation ratio [17].
NFV supports *service chaining*. This means that for a given traffic flow within a given application, the service provider steers the flow through multiple virtualized network functions (VNFs) to achieve the desired network functionality. A key characteristic of NFV is also *management and orchestration (MANO)* which includes the deploying and managing the lifecycle of VNF instances. An important property is also the *distributed architecture*. A VNF may be made up of one or more VNF components (VNFC), which implement a subset of the VNF's functionality and can be deployed in one or multiple instances [17]. These VNFCs can be deployed on separate, distributed hosts which provides scalability and redundancy.

9.5 Applying ML to Network Management

As discussed in the previous section, networks are growing more and more complex with the growing demand and increasing number of users and devices. Existing infrastructure with traditional network architecture is inadequate to process and manage the increasing traffic. That is why the networks are changing their architecture to software-defined networks which enable automation and are more maintainable and adaptable. ML techniques can be applied in combination with the SDN architecture to improve the network management processes.
Network management is essential to operate and maintain any network. Operating costs dominate equipment costs for most Telecom networks [16], making automation of network management tasks an economically beneficial goal to achieve. Use of ML techniques with the combination of transferring to SDN could dramatically improve the network management systems in place and save companies money, time and resources.
Since monitoring and configuration processes require collection and processing big amounts of data, ML techniques provide the best solution for handling these big data sets. In current network architectures, a lot of these tasks are done manually and are as such susceptible to human errors. The automation of these processes would provide key benefits

such as synchronization of multiple data sets, faster processing of data and minimization of errors. Also, the programmability of SDN enables that the optimal network solutions (e.g., configuration and resource allocation) made by ML algorithms can be executed on the network in real time [16].
To summarize, the main benefits of using ML techniques in network management are:

- *Better data management*: includes systematic approaches to collecting and processing large amounts of data during monitoring and configuration processes which are best tackled with ML techniques
- *Higher quality assurance*: "Network operators are neither comfortable nor ready to deploy live traffic on untested configurations, and the concepts of self-optimization are highly suited for such problems." [10]
- *Reducing complexity*: ML methods produce reusable and automatable software, which are relatively easy to implement and have a better performance [13]
- *Scalability and cross-compatibility*: ML methods are generally more scalable and modifiable in comparison to existing methods in network management. The use of ML methods provides reusable software and tackles the problems of vendor-dependency of traditional network architectures.

9.5.1 Challenges of Using ML in Network Management

It was previously explained what benefits ML methods could bring to network management, but their use in this field does not come without its challenges.
Using ML methods for NM is a relatively new and unexplored idea. There is still no blueprint for designing and operating learning-based networks at scale [10]. This means that incorporating ML methods into a network requires a lot of research and development. Since this field of study is relatively new and not all methods are explored, a lot of time goes to selection and optimization of ML algorithms for a specific function in the network management systems. These research efforts require a "new mix of interdisciplinary skills not necessarily present in the industry so far" [13].
The other challenge that presents itself is the infrastructure needed to support the integration of ML methods. Before the network data can be effectively consumed by ML applications, a fairly substantial infrastructure must be put in place, which requires time and money [13]. Also, the "toolchain to integrate ML frameworks with network orchestration and SDN/NFV at scale is missing" [10].
Another problem is acquiring data. ML algorithms need data for training and most of the time they need labeled data to learn from. Although network management uses large amounts of data, most of the data is not labeled. Acquiring labeled data can be expensive and time-consuming, which is another substantial investment in integrating ML in networks.

9.6 ML Methods in SDN

9.6.1 Supervised Learning Methods in SDN

K-nearest neighbors is a common classifying method used for various purposes. Some of the advantages of KNN are that it is easy to implement, not sensitive to outliers and it has high accuracy [12]. Since KNN is usually implemented by linear scanning (calculating distances between the training and test data sets) it becomes computationally complex and time-consuming when the data sets are large. In [18] KNN is used for marking

unknown IP adresses as malicious based on known IP blacklists. One other use of KNN is presented in [19] where it is used as a high-accuracy detection mechanism for detecting Distributed Denial of Service (DDoS) attacks.
SVM is used mostly for binary classification tasks and if multi-classification is needed it is reduced to multiple binary cases. SVM can be embedded in the controller and used for DDoS attack detection, where it classifies benign entries from normal traffic and malicious flow entries from DDoS attacks [12]. For example, [21] presents a mechanism for DDoS attack detection and compares results of linear SVM and polynomial SVM. In [20], an approach to anomaly detection is presented that combines SVM and entropy-based detection. Some of the advantages of SVM are: it is stable, it has a lower false-alarm rate and when designed at the SDN controller level, its complexity does not impact the efficiency of the SDN [12].
Decision trees are mainly used for packet classification in networks. Decision trees are widely used for high speed classification switches, since they can handle a large number of packets [12]. The benefits of decision trees compared to KNN and SVM are that it is easily understandable and fast to implement. However, when decision trees are used for classification problems with many categories, errors can arise more quickly.

9.6.2 Unsupervised Learning Methods in SDN

The standard k-means algorithms are not commonly used in SDN, but variant k-means algorithms are being used more frequently. Hierarchical k-means algorithms, other variant K-means algorithms using heuristic methods and k-means algorithms with cooperative game theory initialization are used for controller placement problems [12].
Algorithms which compare performances of supervised and unsupervised learning methods are being developed and tested for traffic classification problems. The goal of the algorithms is to determine the advantages and disadvantages of each of the chosen methods. The findings have shown that the k-means algorithms are faster but KNN and Bayes methods are more accurate. In practice, unsupervised and supervised learning methods are combined on large data sets to extract advantages from both types of methods [12].
Principal Component Analysis (PCA) is is used for feature selection before applying classification. Clustering algorithms such as K-means and other distance-based learning algorithms are often used for anomaly detection. Self-Organizing Mapping algorithm (SOM) is an artificial neural network algorithm used to reduce the payload in network intrusion detection [22].

9.6.3 Reinforcement Learning Methods in SDN

As mentioned in the ML methods section, reinforcement learning includes an agent, an environment and a set of actions. The agent's goal is to find the optimal policy π that maps a set of states to a set of actions [14].
When using RL in SDN, the controller works as the agent and the network is the environment. Figure 9.6 illustrates the cycle of a RL algorithm. Here the controller (the agent) monitors the state s_t and chooses an action a_t at each time step. It then immediately receives a "reward" from the network (the environment), which is actually an indicator of how well suited the action a_t was as a reaction to the state s_t. It then transitions to the state $s_t + 1$ and by repeating this process the controller learns to choose the best suited actions.
Some of the advantages of using RL in SDN are that it works well without the prior knowledge of the mathematical model of the environment and that it makes fast decisions when trained. The disadvantage is that it becomes too complex when dealing with high-dimensional state space and action space [14].

9.7 Network Applications with ML Methods

9.7.1 Traffic Classification

Traffic Classification (TC) tasks match the traffic in networks with pre-defined classes of interest which can be classes of applications, applications and classes of services. Traffic classification is needed to perform some essential network management functions. These include capacity planning, security and intrusion detection, QoS and service differentiation, performance monitoring, and resource provisioning [7].
Traffic classification methods can be divided into four broad categories: port number, packet payload, host behavior or flow features classification. Port number classification has been proven to be ineffective, but some classifiers use the port number method in combination with other techniques [7].

9.7.1.1 Payload-based Traffic Classification

Payload-based TC searches through the payload for known application signatures. It has high computation and storage costs and is difficult to adapt the signatures to a growing number of applications and their dynamics. Despite those disadvantages, the payload-based classifiers achieve high accuracy and are often employed to establish ground truth [7].
Finamore et al. [24] present a payload-based classifier which uses long-lived UDP traffic to extract application signatures. Their goal is to automatically discover application-layer header format, without caring about specific values of the header fields and to let the protocol format emerge automatically. They use a χ^2 test to determine the randomness of the first N bytes of each packet and evaluate the distance between observed values and a reference uniform distribution. The results of the test are then used to compactly represent application fingerprints, which they call Chi-Square Signatures. They then build an SVM classifier which has signatures as inputs and the target classes as output and works with accuracy of 99.6%.
Another approach to payload TC is presented in Ma et al. [25]. They used unsupervised learning to classify traffic without any previous knowledge of application classes. They train their classifiers based on the label of a single instance of a protocol and a list of partially correlated protocols, where a protocol is modeled as a distribution of sessions. A session is a pair of unidirectional flow distributions, one from the source to the destination and another from the destination to the source. A hierarchical clustering analysis (HCA) is performed to group the observed protocols and distinguish between the classes of interest. They compare different protocol models, and the product distribution protocol model has the lowest misclassification rate of under 5%.

9.7.1.2 Host Behavior-based Traffic Classification

Host behavior-based TC overcomes the disadvantages of misused port numbers and encrypted packet payload. It uses the inherent behavioral characteristics of hosts on the network to predict the classes of interest. It examines traffic between hosts and examines each application based on the different traffic patterns it produces.
In Schatzmann et al. [26] a new approach to host-based TC is presented which discriminates between mail traffic and other HTTPS traffic based on 3 characteristics of mail traffic. (i) Determining server proximity - they identify webmail servers by identifying the SMTP, IMAP and POP servers using port numbers because those servers are found in proximity to wembail servers. (ii) Determining common characteristics of mail servers (e.g. usage patterns). (iii) Detecting periodic patterns caused by application timers in mail servers. Based on these characteristics, 4 features are measured and used to classify

mail and non-mail traffic: service proximity (in IP address distance), activity profiles, session duration and periodicity. With these 4 features as input parameters, an SVM is built and it classifies the traffic with accuracy of 94.2% and precision of 75.8% in the mail class.
Bermolen et al. [27] also present a solution of traffic classification using a ML method. They study the classification of P2P-TV traffic, which is an increasingly important topic. P2P-TV is a growing internet service, but its traffic can deteriorate the quality of Internet service or can pose security risks or create copyright infringement. For those reasons, it is important to be able to classify the traffic which these P2P-TV applications produce. The proposed approach is to derive application signatures from the packets and bytes exchanged between peers in small time windows. These signatures are used to train a SVM classifier to discriminate between applications belonging to the P2P-TV classes (i.e. PPLive, TVAnts, SopCast, Joost). The classifier results in a worst-case TPR of about 95%, with FPR well below 0.1

9.7.1.3 Flow Feature-based Traffic Classification

Unlike payload-based and host behavior-based TC methods, flow feature-based TC methods observe a communication session which consists of a pair of complete flows. A complete flow is an "unidirectional exchange of consecutive packets on the network between a port at an IP address and another port at a different IP address using a particular application protocol" [7]. A flow can have features such as packet length, packet inter-arrival time, flow duration, and number of packets in a flow. Flow feature-based TC uses these features to classify traffic to classes of interest.
For flow feature-based TC various ML techniques are used. In Roughan et al. [8] they use k-NN and Linear Discriminant Analysis (LDA). They observe the average packet sizes and flow duration. They notice similarities between streaming applications and bulk data transfer applications. Because of those similarities, they introduce a further feature - inter-arrival variability to distinguish between these types of applications. The results show that k-NN performs better than LDA and has a lowest error rate of 5.1 and 9.4% for four and seven class classification, respectively.
There are also multiple approaches which include building single or multiple-class SVM classifiers. A traffic classification problem with more than two classes, naïvely transforms the SVM into N one-vs-rest binary subproblems, resulting in a higher computation cost for a large number of classes [7]. Some approaches are dealing with this disadvantage by creating tournament-design SVMs where they organize target classes into pairs and effectively reducing the target classes by half. This reduces the computational cost, but can result in higher misclassification.
Shi et al. [9] compare several supervised ML methods on 2 different datasets. The first data set includes 249 features, so they introduce a FCBF algorithm for feature extraction. They conclude that all supervised learning methods have a shorter computation time when using only selected features and not the full set. On the other hand, some methods such as SVM and decision trees show a smaller accuracy when combined with the FCBF feature extraction algorithm since they themselves have feature reducing strategies which come into conflict with FCBF.
Flow feature-based TC can also be made using unsupervised learning methods. Hard-clustering means that an unknown data point must belong to a single cluster, whereas, in soft clustering, a data point can be mapped into multiple different clusters [7]. Hard-clustering using k-means algorithms has been used in flow feature-based TC. However, soft-clustering methods are more appropriate for classifying membership, since flow features from applications such as HTTP and FTP can exhibit high similarity [7]. There have been various approaches to soft-clustering using EM techniques. EM is an iterative and

probabilistic soft clustering technique, which guesses the expected cluster(s) and refines the guess using statistical characteristics, such as mean and variance [7].

9.7.2 Traffic and Network Flow Management

A network traffic flow is a sequence of data packets [12]. Traffic and network flow management includes flow prediction, throughput prediction, queue control and congestion management.

Traffic prediction is an important research issue for traffic management and also routing optimization. Traffic prediction aims to predict future traffic volumes by analyzing previous trends. The past trends are acquired through continuous monitoring of network performances and analyzing the data. Monitoring in itself is a challenge because of the large amounts of data that each network device produces. ML methods provide a reliable and automated way for sampling and analyzing these big data sets. In SDN, the controller uses traffic prediction to route the packets in such a way that it prevents congestion.

Traffic prediction has traditionally been made using time series forecasting (TSF). TSF models construct a regression model capable of drawing accurate correlation between future traffic volume and previously observed traffic volumes [7]. TSF models can be divided into statistical models and ML models. In the last decade, different types of NNs and supervised learning methods have been used for TSF in traffic prediction. The NNs receive the minimum, maximum and average number of bits per second used on that path in the last epoch and from those data predict the available bandwidth on the path. Approaches using NNs show high long-term and short-term prediction accuracy at low complexity with limited number of features and limited number of layers and neurons [7]. TSF approaches are restrictive in general because they are only possible if past observations on the prediction variable are made. Non-TSF approaches can be used when it is technically or computationally difficult to conduct required measurements that TSF needs for prediction. Non-TSF methods infer traffic volumes from flow count and packet header fields. Although they have higher prediction error rates, these rates remain relatively low not only for NNs but also for other supervised learning techniques [7]. Some researchers suggest that a more complex NN (in terms of number of layers and neurons) might be required to achieve better accuracy in a non-TSF setting.

Congestion control can be observed as an independent NM function, but since it can also affect the network flow here it is presented as a part of the network flow management function. Congestion control ensures network stability, fairness in resource utilization, and acceptable packet loss ratio [7]. It is responsible for throttling the number of packets entering the network. Research efforts have been focused on TCP congestion mechanisms and there have been improvements such as Delay-Tolerant Networks (DTN) and Named Data Networking (NDN). Still, there is room for improvements in areas of packet loss classification, queue management, Congestion Window (CWND) update, and congestion inference.

TCP classifies all packet losses as network congestion, which produces unnecessary congestion control when the packet is lost for other reasons (e.g. packet reordering, fading and shadowing in wireless networks). TCP then reduces its transmission rate each time it detects a packet loss, effectively lowering its throughput rate. ML methods provide a solution for accurate classification of the cause of packet loss, so that congestion control is utilized only when needed. The classifiers mostly use metrics readily available at end-systems, they are trained offline and tested on synthetic data on network simulators. From unsupervised learning techniques, Expectation-maximization is used for the Hidden Markov models in numerous studies and it performed well on simple network topologies. It is important to note that all tested classifiers perform only binary classification [7].

Queue management is another important function performed at intermediate network nodes by dropping packets to control queue length. The conventional method is Drop-tail which utilizes the First-In-First-Out (FIFO) principle to handle incoming packets. When the queue reaches its maximum length, the incoming packets are dropped until places are made available in the queue. Active Queue Management (AQM) is a proactive technique that removes some of the advantages that Drop-tail has. In AQM the packets are dropped before the queue becomes full, which allows the end-systems to respond to congestion before queue overflow. Significant research has been conducted to apply ML for building an effective and reliable AQM scheme. AQM schemes apply different supervised techniques for TSF (Time Series Forecasting) and reinforcement-based methods for deducing the increment in the packet drop probability. The supervised learning methods are various types of NNs, which use TSF to predict future queue length or traffic volume. From the RL methods, there have been significant research efforts in PIDNN. PIDNN is a NN that is incorporated within a Proportional Integral-Derivative (PID) controller. All these ML-based AQM schemes improve and speed up the queue stabilization over non-ML AQM schemes for varying network conditions [7]. It is important to note that all of the ML methods have not been tested in real-world scenarios but rather on network simulations, so their performance measures should be taken with caution.

9.7.3 Routing Optimization

Network routing optimization includes selecting the optimal path for the packets to reach their destination. In SDN, routing is one of the central functions of the SDN controller used to achieve network load balancing and autonomous control [12]. Inefficient routing paths can lead to overloading of network links and increasing end-to-end transmission delay.

Routing optimization has so far been handled using Shortest Path First (SPF) or heuristic algorithms. SPF algorithms do not make the best use of the network resources, while heuristic algorithms become too complex in SDN architectures. ML algorithms have advantages over SPF and heuristic algorithms. ML methods can give near-optimal solutions quickly and they do not need to know the exact mathematical model of the underlying network to produce their decisions [14].

The task of routing optimization can be centralized or decentralized. In a **centralized routing operation** the first packet of a flow is transmitted by the switch to the controller. The controller then calculates the optimal packet route and updates the forwarding tables of the switches along the path. The decision is made based on QoS requirements (minimize/maximize a certain metric) that are also used to control the weight of each metric in the reward function [7]. A **decentralized routing operation** means that each routing node is a learning agent that makes local routing decision from data inquired from the environment. These nodes can make their decisions independently or jointly using multi-agent collaboration [7].

Although some supervised and unsupervised learning methods are explored in routing optimization, reinforcement learning has dominated research in this field [7]. When using RL in routing optimization, the controller works as an agent and the network as the environment. The state space is composed of network and traffic states. Since RL works on an action-reward principle the action is the routing solution and the reward is influenced by optimization metrics (e.g. network delay).

Q-learning is a simple, model-free technique in RL that has been widely used for routing optimization. Q-learning gets its from the Q-value. Q-value is the estimate of the remaining cumulative reward (total reward accumulated by the time the packet reaches its destination) associated with each state-action pair. A Q-learning agent learns the best action-selection policy by selecting the action with highest expected Q-value. Once the

action at is executed and the corresponding reward is known, the node updates the Q-value of the state-action pair [7]. Implementation of Q-learning to routing optimization is called Q-routing. In Q-routing, the agent is the router and the Q-value is the estimated time of the packet reaching its destination. The router determines a certain policy (e.g. packet will reach its destination x through node y) and maps it to a Q-value. When the packet reaches its first stop (node y), the node sends the updated Q-value to the router and the process is continued until convergence. Q-routing outperforms the SPF routing algorithm in terms of average packet delivery time and shows better stability and robustness to topology changes under higher loads [7].

9.7.4 Resource Management and Allocation

Resource management includes "controlling the vital resources of the network, including CPU, memory, disk, switches, routers, bandwidth, AP, radio channels and its frequencies" [7]. Network providers need to predict demand for a service and based on those predictions allocate the needed resources. Underestimating or overestimating the predicted load could cause unsatisfactory service or financial losses. Therefore, the need for estimating the traffic loads and appropriately dynamically managing the resources is an essential ability for network providers. Although ML methods are widely used for load prediction and resource management in cloud data centers, various challenges still prevail for cellular networks, wireless networks and ad hoc networks [7].
SDN architecture enables decoupling of control and data planes and a centralized control through the SDN controller. This property is especially useful in resource management and allocation, since the SDN controller achieves unified control over computing, storage and network capabilities. Resource management can be broadly divided into **admission control** and **resource allocation**. In SDN architectures, resource management can also be divided based on the planes it affects, so there is data plane resource management and control plane resource management. In SDN admission control is part of the data plane management, while resource allocation is part of both the control plane and data plane management.
Admission control is a process of managing a large number of service request by filtering (accepting or rejecting) new incoming requests based on resource availability [14]. If a request is accepted it generates revenue to the network provider, but it can lower network performance because of the overload. Therefore, network providers are faced with a trade-off between generating revenue and maintaining QoS standards. Admission control does not include load prediction, which is usually hard to do accurately.
There have been various approaches using supervised learning, mostly using NNs, random NNs and multi-layer perceptron NNs. The output values of those NNs vary significantly, since this is not a simple classification task. Some models include evaluating user experience, QoS fulfillment ratios or average packet delays, while others build a decision maker with an output of accepted/rejected as an answer to the request. Reinforcement learning is an approach well suited for these tasks. Using the action-reward system, explained previously, the system can evaluate which actions are proper responses to certain types of requests. Deep reinforcement learning(DRL) is especially prominent in this field and it is shown that DRL methods can be applied to large scale systems [12].
Resource allocation is a decision problem that actively manages resources to maximize a long-term objective, such as revenue or resource utilization [7]. The challenge in resource allocation is in predicting demand variability and future resource utilization. ML-based methods can be used to learn indicators which can help in making appropriate resource management decisions. Reinforcement learning is the dominant method in this field. The advantage of using RL is that it can be deployed without any initial policies, and it can learn to adapt to the dynamic demands for a reactive resource allocation. The

disadvantage of using RL is that it can quickly have an overwhelming number of states for a moderate size network and become computationally too complex. Decomposition, decentralization, and approximation have been used to deal with the dimensionality issue of applying RL [7].
A resource allocation problem that arises in SDN is *controller placement*. The controller location has an impact on the network performance since the distance between the controller and the switches can increase traffic flow processing delay. Models using different supervised learning methods (DTs, ANNs and logistic regression) have been used to estimate the best controller location [14]. The input of the training data sets was traffic distribution, and the output was the corresponding controller placement solutions of heuristic algorithms.

9.7.5 Network Security

Network security is a crucial task in network management. It includes protecting the network against cyber-threats that may compromise the network's availability, or yield unauthorized access or misuse of network-accessible resources [7]. Security is provided in part by anti-threat applications (such as firewalls, anti-virus software and spyware detection software), but network behavior analysis is also an important part in security management. The main tasks of security management are intrusion detection, attack detection and fault diagnosis.

9.7.5.1 Network Intrusion Detection

Network intrusion detection (NID) provides real-time protection against internal attacks, external attacks, and accidental operations [12]. Intrusion detection is especially simplified when using a SDN since the programmability of SDN enables fast reactions to network attacks and the global network view of the SDN controller simplifies the collection and analysis of network traffic.
An Intrusion Detection System is a device or a software application whose purpose is to monitor the network and identify potential threats and attacks. There are two types of IDSs: misuse-based (also known as signature-based [14]) and anomaly-based [7]. Singature-based IDSs use signatures of known attacks created by humans or extracted from data and compare new traffic flows to those signatures to find possible malicious activities. Singature-based IDSs have high accuracy, but have some disadvantages. They can only detect attacks whose signatures are known and they have high time consumption since all signatures have to be compared. Anomaly-based IDSs define a "normal" behavior of the network and mark all behaviors that deviate from the "normal" as anomalies and possible attacks. ML methods are widely used in anomaly-based IDS by training a model to identify normal activities and intrusions [14]. Since intrusion detection can be modeled as a classification task, the most commonly used ML methods are supervised learning methods.
ML has been introduced in **misuse-based NID** because manual inspections of signatures have become virtually impossible because of the large volume of generated network traces. Cannady [28] is one of the earliest works that deployed ML methods for NID. This method of NID involves a creation of an ANN. Some of the advantages that ANNs provide in NID are:(i) flexibility of analyzing data from the network even if they are incomplete or distorted (ii) the speed of ANNs enables immediate intrusion detection (iii) because the output of a neural network is expressed in the form of a probability the neural network provides a predictive capability to the detection of instances of misuse [28] (iv) ANNs can "learn" the characteristics of misuse attacks and identify instances that are unlike any which have been observed before by the network. Other ML methods used for misuse-

based intrusion detection are decision trees, SVM and Naive Bayes classifiers. In [29] Naive Bayes and decision trees were compared. They have shown that Naive Bayes is faster, but decision trees show a slightly higher accuracy.
A disadvantage of misuse-based intrusion detection is that it fails to predict new attacks, since it can recognize only attacks from that have known signatures. That is why **anomaly-based NID** methods are used. A disadvantage of anomaly-based IDSs is that they can produce false alarms, since it is difficult to completely represent "normal" behavior [7]. ML methods have been very successful in anomaly-based NID since they offer autonomy and robustness in learning and adapting profiles of normality as they change over time. Anomaly-based NID can be divided in 2 categories: flow feature-based and payload-based detection. In recent years, besides those methods, deep learning and reinforcement learning have made a lot of improvements in this field and are becoming the preferred methods for anomaly-based detection [7].
Flow-feature based detection learns and models the "normal" behavior of the network based on flow features. It was previously stated that misuse-based detection methods rely mostly on supervised learning techniques, since they need to tackle classification tasks. Anomaly-based methods rely on unsupervised learning or a combined supervised-unsupervised learning techniques. In this approach the algorithm is fed with an unlabeled training set to find a structure, or a hidden pattern, in the data [7]. The hypothesis behind this approach is that normal network behavior is more common, whereas malicious behavior appears less often. In the model, the larger clusters of data will present normal behavior and smaller, more distant clusters or data points will present malicious behavior. Jiang et al. [30] show that the size of the cluster is not enough for detecting normal behavior, but that the detection is improved when adding a parameter of distance from other clusters. The majority of works in anomaly-based IDSs employ a combination of supervised-unsupervised ML methods. Some of the best performing methods are random forrests with nested dichotomies and enhanced SVMs - a combination of soft-margin SVM (supervised method) and a one-class SVM (unsupervised method).
Payload-based detection learns normal network behavior from packet payload, which means that it can detect attacks embedded in the packet payload that wouldn't be detected with the flow feature-based method. There have been multiple methods for payload-based detection using the n-gram method. This method uses a sliding window of size n to scan the payload while counting the occurrence/frequency of each n-gram and computing the mean and standard deviation. From those characteristics payload models are created for each service, port, direction of payload, and payload length range [7]. The algorithms measure the deviation between incoming packets and the payload models and mark the packets with larger deviations as potentially malicious.

9.7.5.2 Network Attacks Detection

Network Attacks Detection (NAD) can be included within the intrusion detection tasks, but some works divide it as a special network management function. Malicious attacks can significantly deteriorate network performances or completely halt network traffic. For those reasons, it is imperative to have a system which can immediately identify an attack and take appropriate measures to diminish its effects. Among the most predominant attacks on the SDN controller layer, Link Discovery Attacks and Address Resolution Protocol (ARP) Spoofing Attacks are fundamental since they are the gateways to many other SDN threats and attacks [12].
Distributed Denial of Service (DDoS) attacks are also one of the biggest security concerns and are increasing every year. DDoS attacks try to exhaust system resources by simultaneously sending a large number of fake requests using many puppet machines which results in resources being unavailable to legitimate users [14]. In SDN, the DDoS attack

can exhaust the networking, storage and computing resources in the data plane and the control plane, which will make the SDN network unavailable. To solve such problems, systems of detection are implemented in SDN controllers. Some of the models implemented in those systems use deep, recurrent and convolutional Neural Networks (NNs). These tasks are better suited for such models, rather than simple classifiers as SVM and k-NN, because NNs can perform feature extraction and attack detection in their process. Some researchers warn that using ML methods for attack detection can deteriorate the overall performance of networks. For that reason, it is imperative to optimize the ML algorithm in such a way that it does not interfere with network traffic and still do it's main function of detecting attacks.

9.7.6 Network Fault Management

Network fault management (NFM) includes "detection, isolation and correction of an abnormal state of the network" [7]. Faults can increase operating costs for network providers and lower network performances for users, so detecting them is essential. Fault detection and diagnosis is becoming increasingly difficult because networks are becoming bigger in size, complexity and dynamics. There are 2 approaches to NFM: **naive fault management** and **fault prediction**. Naive fault management is a reactive method and consists of predicting fault, localization of the root cause of fault and fault mitigation. Fault prediction is a proactive method and tries to prevent faults by constantly monitoring the network, predicting faults and initializing mitigation procedures to minimize performance degradation [7].

Fault prediction includes monitoring the network behavior and trying to predict a fault before it has occured. Networks produce large amounts of data which can be collected and processed to determine abnormalities in behavior. To include all of the data that network devices generate in this process would be redundant and would require too much computational resources. That is why feature selection and dimensionality reduction are especially important when building algorithms for fault prediction. ML methods frequently used for fault prediction are supervised learning methods such as decision trees, SVM, enhanced SVM, linear regression etc. Recently, deep NNs are also being used for this purpose and they have been shown to outperform other ML methods [7].

Fault detection is, unlike fault prediction, reactive. It includes identifying and classifying a fault after it has occured. Fault detection can also be made as a classification task, but it is imperative to be able to do it in fastest possible time. Neural networks and recurring NNs have been used for these tasks using supervised learning. Unsupervised learning methods that were used are k-means and self-organizing maps.

Localizing the root cause of fault minimizes the time to repair of networks that do not have a proactive fault prediction system. Supervised learning methods that have been used for fault localization are decision trees, sometimes enhanced with heuristics algorithms, and Bayesian networks. In unsupervised methods algorithms such as discrete state space filtering, Winner-take-all and Neural Gas algorithm were used. Unsupervised learning approaches showed ability to detect fault locations in real time, but were also prone to sounding false alarms.

9.7.7 QoS/QoE Management

User perception and satisfaction are becoming more and more important to both network operators and service providers. **Quality of Service** (QoS) parameters (e.g., loss rate, delay and throughput) are related to network Key Performance Indicators (KPIs) such as packet size, transmission rate and queue length. Previously, the user experience was not

differentiated from network QoS. Nowadays, service providers have come to the conclusion that it is more important to evaluate the service quality from the user's perspective [7].
Quality of Experience (QoE) are the metrics used to illustrate the user's perspective. Measuring QoE is complex since it involves an individual's expectations and perception and cannot be measured by measuring only network performances. There have been 2 approaches to QoE assessment: subjective testing and engagement measurement through objective quality modeling. A widely used metric, which is easy to implement and compute, is the mean opinion score (MOS). In this method, users are asked to evaluate the service and the scores are then averaged into the MOS metric. Users' opinions can be biased and they cannot be forced to leave an evaluation of the service, so objective metrics were proposed to objectively model service quality assessment. Some of these objective metrics are the video quality metric (VQM), the perceptual evaluation of speech quality (PESQ) metric and the E-model for voice and video services. Recently, data-driven analysis of QoE have produced engagement metrics such as service time and probability of return. They draw the impact of user quality perception to content providers more directly.
ML methods can be used for QoS and QoE prediction. QoS prediction can be formulated as a regression task using network KPI as input features. QoE prediction using ML methods relies mostly on establishing a correlation between QoS and QoE and learning how QoS metrics impact QoE metrics.
Supervised learning is the preffered method for **QoS prediction** since it can be formulated as a regression task. In [11] 4 different classifiers are compared for determining service performance of a network. They propose an automata with 4 possible states (Best, Good, Fair and Bad) to determine the standard of service. They use KPIs as input parameters to the models and the classifiers produce which is the most probable state of the network. They compared Linear Regression, Decision Trees, Random Forrest and Gradient Boosting. Gradient Boosting and Random Forrest achieved highest performance with accuracy and precision of 88% and 87% respectively.
QoS/QoE correlation models have been proposed in the literature for different media types (e.g. voice, video and image). QoS/QoE correlation models can also be seen as QoE prediction models, since they establish a way of inferring QoE parameters from given data. There are various ML methods used for building these correlation models: simple regression models, NNs, DT, RF, SVM and etc. The problem in QoE assessment is the lack of a standardized QoE measure that would allow to compare different correlation models. There is also a need for a better and clearer quantitative description of the impact of QoS on QoE. Some researchers correlate QoE to a single QoS parameter while others point out that the QoE is influenced by multiple QoS parameters jointly.

9.8 Further Research and Development

There are numerous ways in which ML algorithms can be improved and new purposes for which they can be used. The application of ML to NM is one of the fields where there is a lot of research potential. Networks will need to support an exponentially higher number of users and traffic in the future. They will need to achieve these capabilities without significantly raising the operational and capital expenditures or customer tariffs [7]. For that reason, it is imperative that new and more efficient technologies (such as ML methods) are implemented into NM systems.
There is a lot of time and research needed to choose the correct ML method/model for a specific NM function. This would be made easier if some of the ML methods were proven to be optimal for specific usages. To improve the decision-making process, **systematization** in evaluating and testing on networks is needed.

Most of the explored ML usages were only tested on simulation systems, or on limited parts of networks, so there is a need of **testing** the methods **on real-world systems**. One of the examples is admission control ML solutions that only consider traffic from limited range of applications. They also disregard that there is a difference in QoS requirements for different applications. For that reason, there is a need for more practical ML solutions. It is also a challenge to produce real-world training data because it is difficult and costly to obtain and usually needs preprocessing before being given to the ML algorithm. Maybe it would be reasonable to develop ML models for faster and easier extraction of network data from the operating networks. These models would then solely provide training data for research and development of new ML solutions.
Many ML methods consume a lot time and computing resources when performing traffic prediction, classification, routing and congestion control on intermediate nodes in the network. This means that ML methods can degrade network performance because of their complexity, even though they are implemented to improve them. The complexity of ML also varies based on the data it consumes. For that reason, **well-rounded evaluation metrics** that will help in **assessing the complexity** of given ML techniques are needed. These metrics would provide needed systematisation of ML methods, and would allow easier choice of methods in future implementations.
There is also a need for **standardized performance evaluation metrics**. Many newly proposed ML methods are being compared to outdated and deprecated solutions. This leads to false comparisons between ML methods and the **de facto** technologies, since these technologies are no longer being used. Standardized performance metrics would provide an unbiased comparison between various ML-based approaches to different problems in networking.
Probably the most difficult challenge to overcome in the future will be designing ML models that can **adaptively retrain themselves**. Network characteristics such as traffic volume, network topology and security attack signatures are dynamically changing as the network operates. Because these parameters change, the ML models which use this data for input parameters should be retrained with new data in order to function properly. This is time and resource-consuming. There is still no practical solution to this problem, but this topic will need to be further explored in the future, since it is essential for proper functioning of the models.

9.9 Conclusion

Over the past few decades, networks have experienced an exponential growth. The growing traffic and number of users is becoming harder to manage. ML methods have been successfully applied for improving or replacing existing network management mechanisms. In this paper, the most commonly used ML techniques were presented and explained. Different ML techniques have been used for all purposes in NM, but the focus of this paper was on traffic engineering, network security and performance optimization.
One one hand, implementing ML methods in networks has its challenges such as high cost of infrastructure and long process of development. On the other hand, it provides many benefits to the system, such as better data management, reducing complexity and higher scalability. Combining new network architectures such as SDN with ML methods enables centralized and automated NM functions.
In the future, there is immense potential for further research on this topic. There is need for systematization in research, which would be achieved by introducing standardized evaluation metrics for ML models' performance and complexity. Also, a large gap exists between research and industry implementations. Testing on real-world systems and using real-world data for model training would be the first steps to reduce this gap. Many of the

presented ML methods still need further research and development. Probably the biggest challenge in the future will be trying to keep up with development of other new network technologies and building solutions that will support surging traffic loads.

9.10 Summary

Networks today are growing in the number of users, traffic and data generated. Network management tasks are becoming increasingly difficult to perform and require a lot of infrastructure, resources and costs. Efficiently performing these tasks is crucial for the network providers, so that they do not lost revenue, and for network users, because they want high quality standards and network performance.
ML methods have been introduced as a solution to efficiently, systematically and quickly perform difficult network management tasks. ML methods can be divided into supervised, unsupervised and reinforcement learning methods. Supervised learning is mostly used for classification tasks, unsupervised learning for finding common features and patterns in data sets, while reinforcement learning is used for creating action policies through an action-reward system.
Even though ML methods provide many advantages for performing NM tasks, their use in this field is still largely unexplored. Some of the challenges are high cost of infrastructure to support the implementation of ML methods and the high cost of data needed for training ML models. SDN architecture provides a good basis for implementing ML methods, since it has decoupled data and control plane. The control plane with the SDN controller enables a centralized control of NM functions and easier changes to new ML methods.
Traffic classification can be presented as a classification task, and most ML methods for TC use supervised learning. Traffic and network flow management includes traffic prediction and congestion management. For traffic prediction, NNs using TSF are used. Routing optimization includes selecting the optimal path for packet to reach their destination. Reinforcement learning is used for routing, because it selects the best path using an action-reward system. Resource management allocation includes admission control and resource allocation. For admission control, mostly supervised learning is used, especially NNs. For resource allocation, reinforcement learning methods are best suited. Network security includes protecting the network from intrusions and attacks. Since intrusion detection can be formulated as a classification task, supervised learning methods such as SVMs are used. For QoS prediction, supervised learning is the preferred method since it can be formulated as a regression task. For QoE prediction and QoS/QoE correlation models different supervised learning methods are used, from NNs to SVMs and DTs.
All of the methods presented still need research and testing. In the future, efforts should be focused on systematization of performance metrics, which would simplify further research. Significant efforts should also be made to adapt ML models for real-world applications since they are now mostly being tested on simulations.

Bibliography

[1] Stuart J. Russell, Peter Norvig: *Artificial Intelligence: A Modern Approach*; book (New Jersey), 1995, 5-7.

[2] Kevin P. Murphy: *Machine Learning : a Probabilistic Perspective*; book (Cambridge, Massachussets, 1st edition); 2012.

[3] Ian Goodfellow, Yoshua Bengio, Aaron Courville: *Deep Learning*; book (The MIT Press, Cambridge, MA, USA), 2016., 98-110

[4] T. M. Mitchell: *Machine Learning*, book (McGraw-Hill, New York, 1st edition), 1997, 2

[5] Ethem Alpaydin: *Introduction to Machine Learning*; book (MIT Press, Cambridge, Ma, USA; 3rd edition); 2014

[6] Shai Shalev-Schwartz, Shai Ben-David: *Understanding Machine Learning: From Theory to Algorithms*; book (New York, USA, 1st edition); 2014, 13-20

[7] Raouf Boutaba et al.: *A comprehensive survey on machine learning for networking: evolution, applications and research opportunities*; in Journal of Internet Services and Applications, 2018, Vol 9, No. 16; June 2018 `https://jisajournal.springeropen.com/articles/10.1186/s13174-018-0087-2`

[8] M. Roughan, Y. Zhang, Z. Ge, A. Greenberg: *Network Anomography*, (Conference: Proceedings of the 5th Conference on Internet Measurement 2005, Berkely, California, USA), 2005.

[9] S. Dong, D. Zhou, and W. Ding: *The Study of Network Traffic Identification Based on Machine Learning Algorithm*; in International Conference on Computational Intelligence and Communication Networks; Mathura, Uttar Pradesh, India, November 2012, 205-208

[10] D. Rafique and L. Velasco: *Machine Learning for Network Automation: Overview, Architecture and Applications [Invited Tutorial]*; in IEEE/OSA Journal of Optical Communications and Networking, Vol. 10, No. 10; pp. D126-D143, October 2018

[11] Y. Turk, E. Zeydan and Z. Bilgin: *A Machine Learning Based Management System for Network Services*; International Conference on Wireless and Mobile Computing, Barcelona, Spain, p. 1-9, October 2019

[12] Y. Zhao, Y. Li, X. Zhang, G. Geng, W. Zhang and Y. Sun: *A Survey of Networking Applications Applying the Software Defined Networking Concept Based on Machine Learning*; in IEEE Access, Vol. 7, pp. 95 397-95 417, July 2019.

[13] D. Cote: *Using Machine Learning in Communication Networks [Invited]*; in Journal of Optical Communications and Networking, Vol. 10, No. 10, pp. D100-D110, October 2018

[14] J. Xie, F.R. Yu, T. Huang, R. Xie, J. Liu, C. Wang, and Y. Liu: *A Survey of Machine Learning Techniques Applied to Software Defined Networking (SDN): Research Issues and Challenges* in IEEE Communications Surveys and Tutorials, vol. 21, no. 1, pp. 393-430, 2019

[15] J. D. McCabe: *Network Analysis, Architecture, and Design*; book (Morgan Kauffman, Burlington, MA, USA, 3rd edition); 300-330, 2007

[16] R. Ramaswami, K. N. Sivarajan, G. H. Sasaki: *Optical Networks (Third Edition)*; book (Morgan Kaufmann), 2010, 469-510

[17] W. Stallings: *Foundations of Modern Networking: SDN, NFV, QoE, IoT, and Cloud*; book (Indianapolis, Indiana, USA); 2015

[18] H. Chang, T. Lin, T. Hsu, Y. Shen and G. Li: *Implementation of ransomware prediction system based on weighted-KNN and real-time isolation architecture on SDN Networks*, (IEEE International Conference on Consumer Electronics, (ICCE-TW), Yilan, Taiwan), 2019

[19] L. Zhu, X. Tang, M. Shen, X. Du, and M. Guizani: *Privacy-preserving DDoS attack detection using cross-domain traffic in software defined networks*, (IEEE J. Sel. Areas Commun., vol. 36, no. 3), pp. 628–643, March 2018

[20] K. M. Aung and N. M. Htaik: *Anomaly Detection in SDN's Control Plane using Combining Entropy with SVM*, (17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Phuket, Thailand, 2020) pp. 122-126, 2020

[21] A. T. Kyaw, M. Zin Oo and C. S. Khin: *Machine-Learning Based DDOS Attack Classifier in Software Defined Network*, (17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Phuket, Thailand, 2020), pp. 431-434, 2020

[22] J. Liu and Q. Xu: *Machine Learning in Software Defined Network*, (2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), Chengdu, China) 1114-1120, 2019

[23] P. Goransson, C. Black, T. Culver: *Software Defined Networks: A Comprehensive Approach*, book (Morgan Kauffman, Cambridge, MA, USA), 241-252, 2017

[24] A. Finamore, M. Mellia, M. Meo, D. Rossi: *KISS: Stochastic Packet Inspection Classifier for UDP Traffic*, IEEE/ACM Transactions on Networking, VOL. 18, NO. 5, October 2010

[25] J. Ma, K. Levchenko, C. Kreibich, S. Savage, G. M. Voelker: *Unexpected Means of Protocol Inference*, In: Proceedings of the 6th ACM SIGCOMM conference on Internet measurement, p. 313–26, 2006

[26] D. Schatzmann, W. Mühlbauer, T. Spyropoulos, X. Dimitropoulos: *Digging into HTTPS: Flow-Based Classification of Webmail Traffic*, (Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement), p. 322–27, January 2010

[27] P. Bermolen, M. Mellia, M. Meo, D. Rossi, S. Valenti: *Abacus: Accurate behavioral classification of P2P-TV traffic*, Comput. Netw. (2011), doi:10.1016/j.comnet.2010.12.004

[28] J. Cannady: *Artificial neural networks for misuse detection*, (Proceedings of the 21st National information systems security conference, vol. 26. Virginia), p. 368–81., 1998.

[29] N.B. Amor, S. Benferhat, Z. Elouedi: *Naive bayes vs decision trees in intrusion detection system* (In: Proceedings of the 2004 ACM symposium on Applied computing. ACM), p. 420–4, 2004

[30] S. Jiang, X. Song, H. Wang, J.J. Han, Q.H. Li: *A clustering-based method for unsupervised intrusion detections*, (Pattern Recog Lett.;27(7)), 802–10., 2006

Chapter 10

A Survey on Anonymization Techniques for Blockchain Transactions

Maximilian Achakri

Anonymity in payments is an underrated feature in today's world. Privacy and personal data are for a long time a subject of discussion and can be priced.[17] Cryptocurrencies get more popular day by day. The most common cryptocurrency is Bitcoin.[21] The Bitcoin brought a new way of thinking about money and transactions. But a big drawback is the lack of anonymity which cash provides.[13] A key reason for that is the public ledger used to store all the transactions.[21] There are already existing solutions. The oldest one is the Zero-Knowledge Proof System.[20] This technology was in invented in the context of cryptography and was adapted to fit the blockchain.[27] Another existing technique is ring signatures used by the cryptocurrency Monero. This technique started with the idea of not only one possible signer of an transaction but rather have multiple, a ring, of possible signers.[23] Monero is not the only new cryptocurrency which has a solution for anonymity. Another one is Zcash. Zcash offers strong anonymity guarantees. Zcash gives the user a feature called shielded pool with which the users can spend special and private coins called shielded coins.[11] One more cryptocurrency worth talking about is Dash. Dash wants to improve privacy like the cryptocurrencies mentioned above.[1] But Dash also wants to impress with a system focused on transactions the waiting time during a transaction caused by the approval process.[9] Further on in this paper we are going to discuss why anonymity is important and look more into those current projects and if they can succeed in the future. For that we will analyze technical factors, their algorithms and solutions for anonymization, but also some economical factors, like the respective market cap.

Contents

10.1 Introduction

Over the last few years cryptocurrency rose to one of the most talked about subjects. Cryptocurrencies disrupt the current payment systems in place. Cryptocurrencies are not only good and have flaws. This new industry is working daily on new solutions trying to fix current flawed or unfinished systems.[9] Currently the most known and most valuable cryptocurrency is Bitcoin. It was introduced in the year 2009 as the first cryptocurrency. Today it is known as a decentralized currency which tries to change how we think about money. It uses the Blockchain technology which erases the middleman needed in traditional transactions.[1]
The Blockchain is the key part to Bitcoins success. Multiple Blocks, all of them containing valid transactions, are chained together to the so-called Blockchain. This system eliminated an intermediary for transactions. This innovation made it possible to use a currency without a centralized institution regulating it. All transactions done with Bitcoin use the Blockchain as a public ledger. Public ledgers save all transactions made with Bitcoin. Before every transaction the public ledger is checked if this transaction is possible or not. If possible the transaction goes through immediately and is recorded in this ledger.[15]
The Blockchain is and will be adapted and changed for other cryptocurrencies or other applications in the future. Bitcoin as groundbreaking as it was 2009 has multiple issues. One of the issues, this paper will focus on is privacy and the ability to make anonymous payments. Privacy and Anonymity are a rising a issue concerning all topics of our actions online.[1] The reason for that is if you are not anonymous it makes you predictable. In microeconomic models most of the time agents are assumed without a name. This makes predicting outcomes for these agents really hard.[17]

10.2 Background

10.2.1 Anonymity

The Cambridge Dictionary defines anonymity the following way: ”a situation in which a person is not known by or spoken of by name”.[2] This definition can also be used in the technical space. For this work the term ”Anonymity” will be used like given above.

10.2.2 Confidentiality

Considering the term confidentiality there are key differences to the term Anonymity. Confidentiality is generally defined by Ueli Maurer as a way of communication where no information is leaked during the process.[24] The main difference to anonymity is that the term with confidentiality both sides know from each other. Being anonymous means that no one knows who you are.

10.2.3 Unlinkability

The main difference between unlinkability and anonymity is that unlinkability is not only used for humans and their actions. Unlinkability however is a part of anonymity. If you can string multiple actions of a human together its anonymity is broken.[25]

10.2.4 Privacy

With multiple data-leak scandals over the past years, privacy has risen to a big concern of our society when using the internet. To definitely define privacy is a hard task as the definition varies from field to field.[8] For this paper I will use my own definition, taking

information from life experiences. Privacy of data means that every person, group of people or organization can decide how their data is handled. If a person, group of people or organization does not want to share their data it is possible under this definition. However companies who provide services can ask for the data and if not given, they can deny access to their service.

10.2.5 Cryptocurrency

This definition is given by Ujan Mukhopadhyay and others: "A Cryptocurrency is a peer-to-peer digital exchange system in which cryptography is used to generate and distribute currency units" [26]

10.2.6 Online Transactions

For online transactions i will use my own definition. In this paper an online transaction means that a coin with a value from a certain currency is transferred from one entity to another online.

10.3 Anonymization techniques for Blockchain transactions

In the following section we will present about current anonymization techniques used for cryptocurrencies which use the Blockchain to encrypt their transactions. After that we will talk more about projects, which use current techniques to anonymize online transactions. All of these projects want to enlarge privacy compared to the market leader, considering value, Bitcoin. All of these alternatives are not as widespread as Bitcoin is. This is observable if you look which wallets are supporting the use of the alternatives. The market cap at the moment of writing this paper is for all currencies over 4 billion. Monero has the largest market cap which shows that at the current moment it is the biggest out of the three. Compared to the bitcoin they are still 150 to 175 times smaller.[4][5][6] An important sidenote to make here is that these facts are all very momentarily because the crypto-market is one if not the fastest changing and volatile market of this world.
Later on in this paper we will discuss advantages and disadvantages of the presented techniques and the presented projects.

10.3.1 Techniques

10.3.1.1 Zero-knowledge proof

A definition of a Zero-Knowledge proof (ZKP) system given by Wanxin Li: "In the context of cryptography, a ZKP protocol is a method by which one party, termed prover, can prove, through a cryptographic commitment scheme, to another party, termed verifier, that they know a secret x, without conveying any information apart from the fact they know the secret x"[27]. The ZKP has been adapted to more anonymous transaction systems like the Zero-Knowledge Range Proof (ZKRP). The ZKRP is close to the ZKP. The difference is that the secret that needs to be verified is a number within a range and not one specific number. The ZKRP can be used to send transactions from one Blockchain Network to another. This can be very helpful to enlarge Blockchain networks. When looking at ZKP's there are different ways the the relationship between verifier and prover. You have interactive systems and non-interactive system where the key difference is that in an interactive system the prover has to answer the sent message and in an non-interactive system the prover has not to answer.[27]

10.3.1.2 Ring Signatures

The Ring Signature Technique advanced after the issues of cryptocurrencies like Bitcoin emerged. The goal was to come away from a coin, used with Bitcoin, where there is only one possible signer to a coin where in one single transaction there is a group of possible signers. Going from one to multiple possible signers helps the anonymity of the transaction.[23] This is achieved by making it possible that one member of the group can sign in behalf of the group. This group is created spontaneously and has no leader. Other applications then cryptocurrencies could be whistle blowing or general anonymous groups.[16]

10.3.1.3 Homomorphic Encryption

Over the last years a lot of homomorphic Encryption schemes have been made.[3] Homomorphic Encryption systems can be used to hide the IP when accessing other systems. Homomorphic Encryption can be used do operations on the ciphertext without having to decrypt it first. This means that the service provider can directly handle the ciphertext without decrypting it into plain text.[3]

10.3.2 Current Projects

10.3.2.1 Monero

Monero is a cryptocurrency that uses the Ring Signature technique to allow for anonymous payments.[23] It is the most valuable cryptocurrency presented in this paper at the moment of writing.[4] Monero was created on the basis of CryptoNote. Monero uses one type of Ring Signature which is named Multi-layered Linkable Spontaneous Anonymous Group. Like other Ring Signatures the main part of the signature is that there is not only one possible signer of a transaction but rather a whole group of possible signers.[23] In addition to the Ring Signature technique Monero uses one-time public keys to better the anonymity of an transaction further. The one-time public key is generated to give a new public address for each new transaction. If you put both techniques together it gets rather impossible to follow the real addresses of an transaction.[7] Each transaction done with Monero is signed using their Ring Signature scheme. This scheme which makes a key image, that confirms that it is not an attempt to double spend. On top of that it uses the Confidential Transaction scheme, which is similar to the Homomorphic Encryption scheme, to hide the value of an transaction.[22]

10.3.2.2 Zcash

Zcash has two types of transactions. Transparent and shielded ones. Zcash is very similar to Bitcoin but has some key differences concerning the shielded transactions.[19] Zcash, in comparison to Monero, does not try to hide the identity of the sender to reach the goal of anonymous payment. It attacks the link between sender and receiver of an transaction. This is done with the shielded pool. Every transaction that is done using ZCash has to go through the shielded pool. The Shielded Pool is not only used for transactions but also said to be used as a storage for shielded addresses(z-addresses).For these z-addresses the shielded pool is used to make the payment unlinkable . The part where ZCash breaks the link between sender and receiver of a transaction the vJoinSplit, existing in the Shielded Pool. vJoinSplit contains a list for outgoing payment addresses and incoming payment addresses, two shielded outputs and two double-spending tokens. All of the features above help to provide more unlinkability for the payments. To explain it in an easier way and in

plain English. These features make a bigger ”mess” of a single transaction and try to hide it. It makes it harder to follow. The biggest contributor concerning anonymity is the ZKP which is attached to the system. The ZKP ensures that the sender of an transaction is anonymous.[11] For that they use a system called zk-SNARKS(Zero-Knowledge Succinct Non-Interactive Arguments of Knowledge). To make a payment happen the sender has to proof Zk-SNARKs, in zero-knowledge, their ability to afford the payment in question. To register to use ZCash there is a trusted setup stage. This stage is there to make sure that everyone that uses this system uses it with no harming intentions.[22]

10.3.2.3 Dash

Dash is a cryptocurrency with built in privacy features. The privacy feature is called PrivateSend. PrivateSend is build on the basis of CoinJoin and has been extended with decentralization, a chaining approach, denominations and passive ahead of time mixing.[9] A secondary network of full nodes, called the Dash Masternode Network, helps to more privacy when making transactions with dash.[19]

10.4 Discussion

In the following section i will talk about advantages and disadvantages of multiple points. We will talk about advantages and disadvantages of the current techniques and projects compared to the techniques used with Bitcoin. Further we will talk generally about advantages and disadvantages of anonymous payments and compare cryptocurrencies to other payment categories like credit or cash.

10.4.1 Advantages

One of the key advantages of all new techniques and projects shown is that compared to Bitcoin in all of the above the privacy and anonymity factor while doing online transactions is enhanced or fully provided.

10.4.1.1 Zero-knowledge Proof

The ZKP technique makes it possible to hide the identity of the sender of an message. This is a big leap forward to anonymous payments.[20] Another big advantage of ZKP is that the technique can be used to hide the IP-Address. This will not only provide anonymity during a transaction of money, but can also help to transfer data anonymous. Maybe this technique can be refined and used in even other systems which use transactions we do not need today but maybe in the future.[20]

10.4.1.2 Ring Signatures

The biggest plus point which helps the anonymity a lot is that you can hide yourself behind a group. This makes it very hard as an attacker to find out who made that one transaction. As mentioned before Ring Signatures can not only be used for anonymous payment systems. Using them for whistle-blowing could help a lot of people, who are to scared to share important information, sharing that important information.[16]

10.4.1.3 Homomorphic Encryption

Even though Homomorphic Encryption is not fully finished yet does not mean it never will. I can see that a fully working Homomorphic Encryption functionality can be extremely powerful. Hiding your IP is already possible with ZKP. But doing operations on an still

encrypted file could be a very useful and also groundbreaking feature in the future.[3] One of the current use cases which use a similar method is Monero with Confidential Transactions. It is used to hide the value of an transaction.[22]

10.4.1.4 Monero

Monero uses the Ring Signature technique which brings along all the plus points talked about above. Another big advantage for Monero are the one-time public keys. This means that the public keys are generated newly after every transaction. These different public keys make it harder to link multiple payments of an user together.[23]

10.4.1.5 ZCash

ZCash has a lot of potential to be an anonymous payment environment. Especially the shielded pool would work really good when used right. Zcash also takes all advantages talked about in the Zero-Knowledge Proof section because it uses it to make sure that the user that spends the coin does not reveal information about him.[11]

10.4.1.6 Dash

The Dash Masternode Network provides much more incentives to run than the traditional Blockchain nodes. The concept of the PrivateSend feature works better regarding anonymization then the one from Bitcoin. The focus is set to provide a system that can guarantee privacy. Regarding our definition of privacy it provides some good ideas. One feature they provide to enlarge privacy is that they improved the CoinJoin feature. CoinJoin merges different payments. The problem with a simple CoinJoin implementation was that this exposed the users and made it possible to follow the transactions. PrivateSend makes it possible that after joining the transactions the transaction can not be unmerged.[9] For that merge to happen you need three users to make a pool. This pool then distributes the coin to the newly formed addresses belonging to those three users.[18]

10.4.1.7 Anonymity in transactions

To talk about advantages and disadvantages of anonymous payments are more about politics then about the technical side. There are still important to be mentioned. A clear cut advantages of anonymous payment systems is the freedom to buy whatever you want. The additional advantage of decentralized online anonymous payment systems like cryptocurrencies are that not even state authorities can know how much money you get and how much you spend and on what you spend it on.

10.4.2 Disadvantages

10.4.2.1 Zero-knowledge Proof

There are not many disadvantages for the ZKP. As a technique it works really well. But if it is used it leads to full anonymity. With that comes one big issue. The only Person or Entity that knows the password to access the system is the user. When the user forgets or looses the password it is lost forever. [14]

10.4.2.2 Ring Signatures

There is one large issue with Ring Signatures known as Key Exposure. This happens when a secret key of signature is compromised. In this event all signatures given become worthless because you are able to forge signatures from that moment on. This is an issue

that is also known from the normal public/private key technique but is transferred over to the Ring Signature technique. But in the Ring Signature scheme it gets more serious. Not only can the signature be forged for the key holder but also for the entire group. This makes the whole scheme dangerous, because of one leakage multiple user will suffer. After an event like this happening all public keys and private keys need to be changed because of the Ring Signature scheme the it is not possible to find out what signature was made by the attacker.[16]Another issue with Ring Signature is that if a large enough sample exists the anonymity can be broken because the transactions are not fully unlinkable. [23]

10.4.2.3 Homomorphic Encryption

Homomorphic Encryption is a good concept but hard to implement in the real word. There are big concerns with the performance of these systems as well as if it really works. On top of that there are some specific softwares for it to run. This makes it rather unattainable for business owners to use. This is mostly the case for softwares that use Homomorphic Encryption. For Cryptocurrencies there is no coin which has used this technique successfully today. This can change in the near future.[12]

10.4.2.4 Monero

Looking at Monero this issue is being fought by the one-time key you will use. This key has a disadvantage as well, because it will create a lot of so called "dust transactions"[23]

10.4.2.5 ZCash

Zcash provides two types of payment. Transparent and shielded ones. Because this cryptocurrency provides both types there a transaction is not fully unlinkable.[19] One reason that adds to this is that most users do not use the privacy features given and even the ones that use them use them in a way that the transactions are still linkable. Another disadvantage regarding anonymity is the trusted setup page. This stage has surely advantages when it comes to preventing illegal activities. But this contradicts full anonymity because you have to sign-up with all your information to use the system in the first place.[22]

10.4.2.6 Dash

Dash needs at least three participants to use the PrivateSend feature. The process to find three users delays the time in which a payment can be made. The time in which a payment is completed is a big point to compare cryptocurrencies on. Cryptocurrencies which take to long to fulfill a payment loose a lot of value. On top of that it is not a true decentralized system because of the Masternode Network it uses for its transactions. Furthermore the unlinkability, which is needed for full anonymity, is not fully guaranteed.[19]

10.4.2.7 Anonymity in transactions

The main disadvantage of anonymous payments are that whatever you can buy anonymously creates a market of criminal buyers and sellers of illegal goods. If all payments are anonymous and decentralized there will be no way you are able to find out what happened.

10.4.3 Offline vs Online transaction systems

Cryptocurrencies want to solve one main problem with our current currency structure. They all try to decentralize payment.[9] There is however one point which they have not

met yet. That is the anonymity feature our current currency system provides. Cash is the most anonymous payment system known to us. Since the introduction of credit cards the banks that give us our cards are able to trace all our transactions.[10]

10.5 Summary

We learned that the Bitcoin Blockchain has its issues. But regarding the fact that the Bitcoin was created in 2009 it is still impressive. Today other cryptocurrencies tried, sometimes more sometimes less, successfully to solve the anonymity issues of Bitcoin. Monero, Zcash and Dash are all good options. Some of them work better than others, but the technical side of all cryptocurrencies are interesting. The question for all these currencies is if people need them. Zcash is a good example of that. If all users of Zcash would use the shielded pool as it is intended to be used, the anonymity would be a lot higher. But it is not used in that way and this brings up the question if the market has a demand for these special features.

Something we have seen multiple times today is that the difference and dependence between anonymity and unlinkability. Even though the anonymity of one single transaction maybe given, the possibility to break the anonymity would still be given if all options in the group have different behaviours when it comes to buying things. With an data-analysis the transactions could still be linked to each other and depending on the amount of transactions the anonymity of one of these transactions can be harmed. Another way we have seen unlinkability fail was when not all users of a system use the features provided to make anonymous payments.

The cryptocurrencies we talked about are at the moment more of a substitution for the credit cards as we know them. For all of the previously presented cryptocurrencies are some disadvantages given that make it hard to fully change to them. In comparison to the Bitcoin you can see that a lot of progress has been made over the last few years. It will be very interesting to follow new projects in this space because in the future there might be a cryptocurrency which would act as a full substitute to money and currency as we know it. Not only for the online world but maybe even for the day-to-day transactions.

10.6 Conclusion

The space of cryptocurrencies is evolving in a very fast speed. There are efforts to solve issues concerning privacy and anonymity in the online world. The solutions available today are not perfect but some of the systems in place come real close to anonymous online transactions.

Bibliography

[1] Alex Biryukov and Sergei Tikhomirov: *Security and privacy of mobile wallet users in Bitcoin, Dash, Monero, and Zcash*, Pervasive and Mobile Computing Volume 59, October 2019 https://www.sciencedirect.com/science/article/pii/S1574119218307181.

[2] Cambridge Dictionary: *Definition Cambridge Dictionary* https://dictionary.cambridge.org/dictionary/english/anonymity.

[3] Caroline Fontaine and Fabien Galand: *A Survey of Homomorphic Encryption for Nonspecialists*, October 2007 https://link.springer.com/content/pdf/10.1155/2007/13801.pdf.

[4] Coinmarketcap.com: *Monero* https://coinmarketcap.com/currencies/monero/.

[5] Coinmarketcap.com: *Zcash* https://coinmarketcap.com/currencies/zcash/.

[6] Coinmarketcap.com: *Dash* https://coinmarketcap.com/currencies/dash/.

[7] Dimaz Ankaa Wijaya1, Joseph Liu, Ron Steinfeld1, Dongxi Liu: *Privacy on the Blockchain: Unique Ring Signatures.* University College London, Dec 2018 https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8456034.

[8] Dmitry Epstein, Merill C. Roth and Eric P.S Baumer: *It's the Definition, Stupid! Framing of Online Privacy in the Internet Governance Forum Debates*, Journal of Information Policy 4 (2014): 144-172. https://www.jstor.org/stable/pdf/10.5325/jinfopoli.4.2014.0144.pdf.

[9] Evan Duffield and Daniel Diaz: *Dash: A Payments-Focused Cryptocurrency*, Whitepaper, August 2018.

[10] Foteini Baldimtsi, Melissa Chase, Georg Fuchsbauer and Markulf Kohlweiss: *Anonymous Transferable E-Cash*, 2015.

[11] George Kappos, Haaroon Yousaf, Mary Maller, and Sarah Meiklejohn: *An Empirical Analysis of Anonymity in Zcash*, 27th USENIX Security Symposium, Baltimore, MD, USA, August 15–17, 2018.

[12] Harold Byun: *The Advantages and Disadvantages of Homomorphic Encryption*, 2019 https://baffle.io/blog/the-advantages-and-disadvantages-of-homomorphic-encryption/.

[13] Ian Miers, Christina Garman, Matthew Green, Aviel D. Rubin: *Zerocoin: Anonymous Distributed E-Cash from Bitcoin* , The Johns Hopkins University Department of Computer Science, Baltimore, USA, 2013.

[14] imaginenext: *5 advantages (and 1 disadvantage) of zero knowledge authentication* https://imaginenext.ingrammicro.com/data-center/5-advantages-and-1-disadvantage-of-zero-knowledge-authentication.

[15] Jerome Kehrli: *Blockchain explained*, October 2015 https://www.niceideas.ch/blockchain_explained.pdf.

[16] Joseph K. Liu and Duncan S. Wong: *Solutions to Key Exposure Problem in Ring Signature*, Advandces in Coumputer , Nov 2005 https://eprint.iacr.org/2005/427.pdf.

[17] Malte Moeser and Rainer Boehme: *The price of anonymity: empirical evidence from a market for Bitcoin anonymization*, Journal of Cybersecurity, Volume 3, Issue 2, June 2017, Pages 127–135, August 2017 https://academic.oup.com/cybersecurity/article/3/2/127/4057584.

[18] Nida Khan Mohamed Nassar: *A Look into Privacy-Preserving Blockchains*, 2019 https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9035235.

[19] Niluka Amarasinghe, Xavier Boyen and Matthew McKague: *A Survey of Anonymity of Cryptocurrencies*, Association for Computing Machinery, 2019. https://dl.acm.org/doi/pdf/10.1145/3290688.3290693.

[20] Oded Goldreich and Yair Oren: *Definitions and Properties of Zero-Knowledge Proof Systems* , Department of Computer Science, Technion, Haifa, Israel, 1992 https://link.springer.com/content/pdf/10.1007/BF00195207.pdf.

[21] Peter D. DeVries: *An Analysis of Cryptocurrency, Bitcoin, and the Future* , University of Houston, USA, September 2016. https://ijbmcnet.com/images/Vol1No2/1.pdf

[22] Rebekah Mercer: *Monero Ring Attack: Recreating Zero Mixin Transaction Effect*, 2018 https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8456034.

[23] Shen Noether: *Ring Signature Confidential Transactions for Monero*, Cryptology ePrint Archive, Report 2015/1098, December 2015 https://eprint.iacr.org/2015/1098.pdf.

[24] Ueli Maurer, Andreas Ruedlinger, and Bjoern Tackmann: *Confidentiality and Integrity: A Constructive Perspective*, 2003 https://link.springer.com/content/pdf/10.1007/978-3-642-28914-9_12.pdf.

[25] Sandra Steinbrecher and Stefan Koepsell: *Modelling Unlinkability*, Lecture Notes in Computer Science (LCNS, volume 2760) https://link.springer.com/chapter/10.1007/978-3-540-40956-4_3.

[26] Ujan Mukhopadhyay, Anthony Skjellum, Oluwakemi Hambolu, Jon Oakley, Lu Yu and Richard Brooks: *A Brief Survey of Cryptocurrency Systems*, 2017 https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7906988.

[27] Wanxin Li, Mark Nejad: *Privacy-Preserving Traffic Management: A Blockchain and Zero-Knowledge Proof Inspired Approach* , University of Delaware, Newark, USA, September 2020 https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9210529.

[28] Wei Liang, Dafang Zhang, Xia Lei, Mingdong Tang, Kuan-Ching Li, Senior Member, IEEE, and Albert Y. Zomaya: *Circuit Copyright Blockchain: Blockchain-based Homomorphic Encryption for IP Circuit Protection* , Advandces in Coumputer , 2019 https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9091234.

Chapter 11

Software-Defined Networking (SDN) and cyber security: The Current Scenario, Opportunities, and Challenges

Julius Willems

Software Defined Network (SDN) is a new type of network design that separates the control and data plane of traditional networks. The SDN controller becomes a central entity with a global view and control of the network configuration. SDN provides major benefits over traditional networks, such as centralized configuration and standardized open interfaces. Over the last years SDN, and its applications and vulnerabilities in the field of cyber security, gained research interest. This report provides a literature review over the current security challenges and opportunities faced by SDNs. We analyze specific attack schemes on SDN and highlight both, mitigation strategies, as well as vulnerabilities in the SDN architecture. We conclude that SDNs provide effective means to mitigate cyber threats, and at the same time expose new vulnerabilities, such as the centralized SDN controller. At the end of the report we discuss possible future developments in SDN networks in general.

Contents

11.1 Introduction

Computer networks are complex to manage due to the diversity of devices that participate in it. Routers, switches, firewalls, and many other physical devices operate together in a distributed fashion, and together make up the network. Typically, these devices consist of a hardware and a software part. The software is referred to as *control plane* and defines the rules how data is forwarded. The hardware part is called *data plane* and forwards the data according to the rules of the control plane. While the communication *between* devices is based on standardized, open protocols, each device runs its own, proprietary and closed software [1]. With the absence of a standardized configuration interface, the process of maintaining and configuring a device is, thus, a vendor or even a model specific task. As a consequence, network administrators need to configure each device individually, resulting in a considerable overhead. The lack of a centralized network configuration interface limits the flexibility of a network, and increases infrastructure and management costs of the network [1].
The need for more flexible and programmable networks led to the emergence of key components present in Software Defined Networks (SDN) today. From the mid 1990s - 2000s, programmable network functions were introduced. Between 2001 and 2007, control and data plane separation led to the development of open interfaces between the two planes. And finally, from 2007 onwards, the adaption of the OpenFlow standard enabled greater scalability of SDNs [1].
The growing popularity and adoption of SDN over the recent years led to discussion how SDN can be leveraged to enhance security aspects in networks. In 2008 first works related to security through SDNs were published. At the same time research presented new vulnerabilities and attack vectors in SDN networks. Thus, SDN networks can be leveraged to enhance security, and at the same time introduce additional vulnerabilities by their characteristic architecture [4].
This report focuses on a literature review on SDN, and highlights its benefits and vulnerabilities for the field of cyber security. The rest of this report is organized as follows. Section 11.2 introduces the basics of SDN and covers an industriy use-case of an SDN. Section 11.3 covers the opportunities and challanges of SDN in cyber security. In section 11.4 the future of SDN related to cyber security is discussed. Finally, the report ends with a conclusion in section 11.5

11.2 Software-Defined Networking

SDN is a combination of attempts to address the aforementioned issues mentioned in section 11.1 and can be grouped into four key concepts [3]:

1. Logical and physical separation of the control plane and data plane
2. Centralized controller and global view of the network topology
3. Open interfaces between control and data plane
4. Programmability of the network by client applications

Figure 11.1 introduces the SDN architecture and the components of each of its layers. Each one of these layers and its main functionalities are described as follows.
The top layer of the SDN architecture is called application layer and refers to concept number 4, "Programmability of the network by client applications" from above. The application layer, or management plane, consists of network applications that define control and operation logic. The management plane ultimately implements abstract policies that

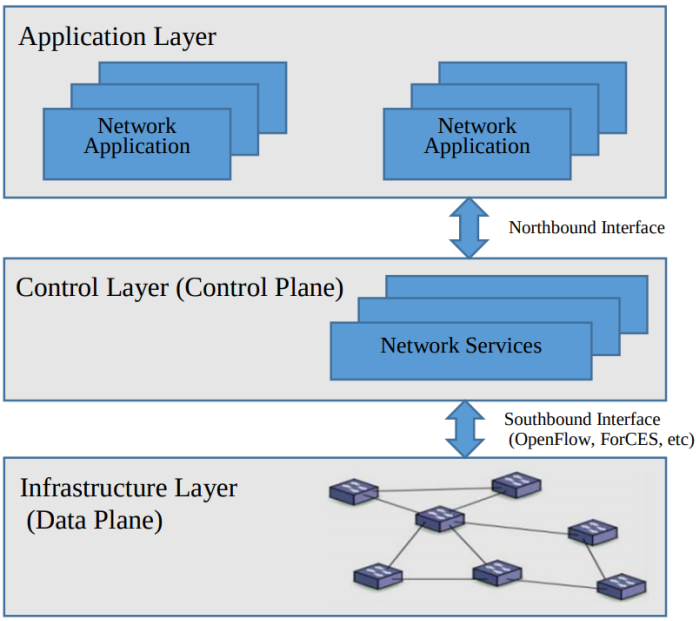

Figure 11.1: SDN Architecture [2]

are pushed down to the control plane through the Northbound Interface. This provides great flexibility, and allows more application specific network customization as network configuration is programmable and easily deployable. Examples for such applications are routing, firewalls, load balancing and network monitoring [4].
One level below the application layer is the control layer, which refers to concept number 2, "Centralized controller and global view of the network topology". It is the most intelligent layer in the SDN architecture, and is also referred to as control plane. It consists of one or multiple controllers forwarding rules and policies defined by the management plane to the control plane. In contrast to 'traditional' networking, where control and data plane communication is vendor specific, SDNs perform this task through the southbound interface and follow a vendor agnostic, open standard [3].
Below the control plane sits the data plane, also termed as infrastructure layer. The separation of control and data plane is characteristic for SDN and reflected by concept number 1, "Logical and physical separation of the control plane and data plane". The data plane consists of the actual devices (routers, switches, etc.). They receive the rules and policies form the control plane through the southbound interface [3], and are responsible for forwarding, dropping and modifying packets [4].
The last key concept, "Open interfaces between control and data plane", is enabled by OpenFlow. It is an open protocol that manages the communication between the control and data plane. OpenFlow is considered one of the first SDN standards [5]. It traces back to the work of a PhD student in 2006 at Stanford University. Nowadays, the Open Network Foundation (ONF) is tasked with the maintenance and evolution of the standard. OpenFlow defines the communication between the control plane (SDN controller) and the data plane [6]. OpenFlow enabled switches can be controlled from a single SDN controller without the need of vendor specific configuration. Forwarding information is abstracted as flows and can be configured dynamically, enabling more flexible and convenient configuration of networks. Further, analytics and statistics about flow information can be extracted and used to provide support for further applications [19].

11.2.1 Use case: Software Defined WAN

One of the early movers to exploit the benefits of SDNs was Google in 2013. In their paper [7] they present a global WAN connecting Google's data centers. One specific requirement of their data centers is the need for huge bandwidth capabilities. One major issue with their current situation was the reliability of the system. In order to achieve acceptable system reliability, WAN links needed to be over provisioned to compensate for various types of failures and errors. The over provisioning meant higher redundancy and low WAN link utilization. On average, the utilization of networking infrastructure was around 40% [7]. The low utilization of high-tech networking equipment became a considerable cost overhead. Another aspect they looked for was a way to centrally manage their network infrastructure. To address these issues, Google started to develop a Software Defined WAN tailored to their need of high bandwidth and central network configuration [7].
The successful adoption of a SDN enabled them to quickly deploy new network control services and, thus, improved flexibility and agility. One of these services, centralized traffic engineering, allocates bandwidth among competing applications based on dynamic priorities. This dynamic bandwidth allocation enabled a link utilization of nearly 100% for extended periods [7].

11.3 Opportunities and Challenges of SDN in cyber security

The goal of this section is to focus on opportunities and challenges regarding SDNs and cyber security. While SDNs are able to mitigate certain threats, new attack vectors are exposed as well. Especially the centralized controller is a high-value target. The following paragraphs cover a selection of attack types and elaborates on potential threats and mitigation strategies.

11.3.1 Denial-of-Service (DoS) Mitigation

Typically, Denial-of-Service (DoS) attacks aim at overloading a system by continuously flooding the system with requests until the overall performance degrades to the point where the system is not able to serve legitimate users anymore. In the SDN architecture, the control plane (SDN controller) is the most vulnerable target as it represents a single point of failure for the whole network.
OpenFlow defines that switches (data plane) receiving a packet look up the respective output port in their flow table. If there exists no such entry for the given packet, the switch requests a new flow rule from the SDN controller (control plane). This is done by sending a 'packet-in' message. The latter responds to the packet-in message with a new flow rule and time-to-live (TTL) until this flow rule is valid [8].
Multiple attack vectors exist on the SDN controller. Overloading the SDN controller causes packet-in messages to be stuck in the controller's queue. Thus, no new routing rules can be generated, and unknown flows arriving at the data plane are stuck as the switch device can't receive flow rules for new messages.
Another attack vector is the link bandwidth between data and control plane. Due to the separation of the control and data plane in SDN networks, communication between the two are routed over the network. Flooding the network in order to generate many packet-in messages can exhaust the bandwidth in the link between the switch and SDN controller. The result is similar to the first scenario where the switch device can't receive any new flow rules. Thus, the network is blocked and can't handle new messages.

Lastly, TCAM overflow describes an attack of the data plane itself. The attacker continuously sends new flows causing the switch device to request new flow rules from the SDN controller. Forwarding tables storing the flow rules become full, and the switch has to constantly delete old entries to make space for new rules. Repeating this process causes the delays and ultimately can block the network [8].
To address these vulnerabilities, the authors of [9] propose rate limiting, a common solution to prevent certain clients from penetrating a system. Their solution is integrated in the data plane, and monitors the packet flow of clients and simply drops messages if they arrive in a frequency that exceeds a certain threshold. Thus, a potential attack on the SDN controller is mitigated. On the other hand, legitimate traffic is affected as well.
Another, more complex approach involves three components. A flow management module manages routing paths of individual flows, and specifies dynamic timeouts of TCAM entries based on a calculated threat probability. A rule aggregation module is responsible for aggregating flow entries to reduce the amount of TCAM entries. Finally, a monitoring module collects statistics on flows and provides them to other mitigation systems. Although the approach introduces more complexity, the authors were able to achieve good results in their experiment [8]

11.3.2 Intrusion detection

Intrusion Detection Systems (IDS) are designed to continuously monitor the activity in a network and analyze it to detect malicious behaviour. An IDS that monitors the network is also referred to as NIDS. NIDS can be further categorized in signature based and anomaly based approaches. Signature based approaches analyze the collected data and aim to identify known attack patterns. They deliver high detection results but perform poorly when having an unknown attack scheme at hand [10].
On the other hand, anomaly based approaches rely on a statistical model for normal network traffic. Analogously to signature based approaches, the network traffic is continuously monitored and compared with the statistical model. Whenever a discrepancy is detected between the network traffic and the statistical model, an alarm is triggered. While the anomaly based approach performs better on unknown attacks, it delivers less accurate results and suffers from a high false positive rate [10].
Machine learning (ML) and deep learning (DL) based NIDS make up the majority of recent research in this area [11]. While these techniques promise more accurate detection results, challenges specific to ML/DL arise, too. One such typical challenge is called feature selection/engineering. It is the important task of selecting the right features from the dataset and combining them into new ones. Although this challenge belongs to the ML/DL domain rather than NIDS itself, it represents a current issue, nonetheless.
The lack of recent and large enough datasets in the field of NIDS is yet another challenge, which researchers are facing. In order to train accurate DL models, typically a vast amount of data is needed that is split in a training and testing part. Due to the limited availability of such datasets, researchers fall back to artificially generated datasets.
To implement an NIDS that monitors the network traffic in real-time or in high-speed networks, performance also becomes an issue and a potential bottleneck [11].

11.3.3 Proactive defense

SDN network security technologies can be approached in a passive or active way. Firewalls and IDS are examples for such passive systems. Here we focus on the other end and look at active systems. Figure 11.2 depicts the structure of a scan detection attack at the data layer. The network consists of trusted and untrusted switches and hosts. Untrusted switches can scan and analyze network traffic, discover the network topology

and ultimately provide support for further attacks such as DoS. In this scenario, an

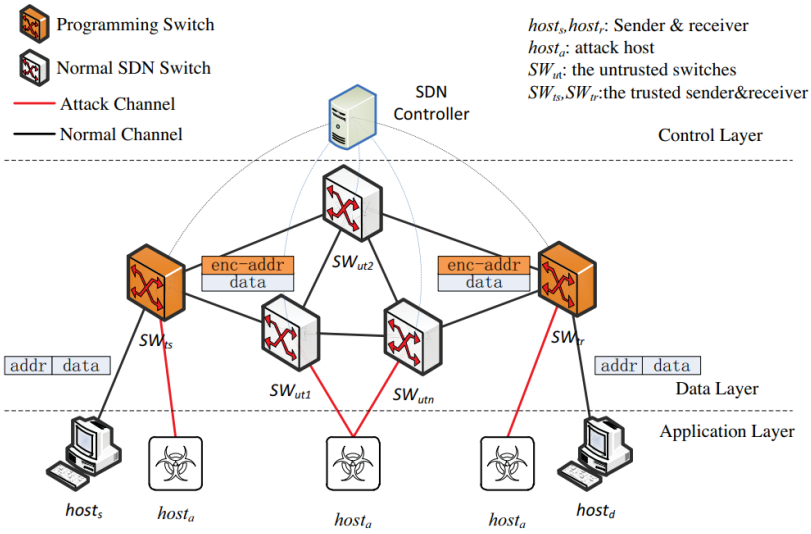

Figure 11.2: Proactive Defense Architecture SDN[12]

untrusted entity can read the source and target IP address of packets. Providing this information to another attacker might be harmful. To mitigate this risk, cryptographic schemes can be used to mask the IP address of the packets. In order to encrypt and decrypt the IP address, cryptographic parameters in the form of keys are used. The SDN control plane generates and distributes the cryptographic parameters to the sending and receiving switch. A set of temporary flow rules is generated alongside based on the cryptographic parameters and pushed down to the data layer. It is important to note here that hosts are not responsible for address transformation. This task is performed by trusted switches. The sending switch encrypts the IP address with the received key and sends the packet. Intermediary switches forward the packet according to the received, temporary flow rules. The receiving switch decrypts the IP address and forwards it to the respective host. To mitigate malicious switches form analyzing traffic patterns, cryptographic keys must be changed frequently.
This type of IP address randomization ensures the non-linkability of IP addresses which proactively mitigates scan detection attacks that could potentially lead to DoS attacks [12]. On the other hand, the cryptographic mechanism introduced comes at a price of higher complexity.

11.3.4 Side channel attack

The separation of the control plane and data plane in SDNs requires that network policies and rules are distributed among network participants. The distribution of this information itself can be an attack vector as malicious parties are able to gather relevant information about the network configuration and topology. That information could in turn provide support for further attacks. These kind of attacks are referred to as side channel attack.
A side channel attack specific to SDN is the Know Your Enemy (KYE) attack. It builds upon the assumption that the attacker is able to learn or infer the flow rules dictated by the controller and stored in the flow table of an SDN switch. The KYE attack itself is independent from how the flow rules are obtained. The goal of an KYE-attacker is to collect information about a network without being detected. More specifically, an attacker is only able to:

1. Send packets through the target network

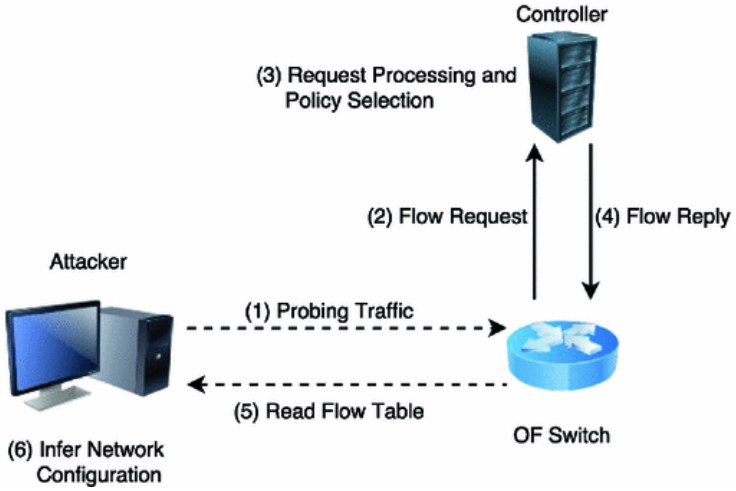

Figure 11.3: KYE attack [16]

2. Use the side channel to access the flow rules installed on attacked switch

Thus, the attacker uses the side channel on the flow table only to read data, not to modify or erase it.
The KYE attack consists of two phases - a probing and inference phase. During the probing phase, the attacker's goal is to trigger the installation of additional flow rules on the compromised switch. In the second phase, the inference phase, the attacker attempts to obtain the flow rule through the side channel. The attacker then uses the obtained information to plan further attacks. Hence the name "Know Your Enemy".
Let's look at an example how a KYE attack can provide support for a DoS attack. As presented in sections 11.3.1 and 11.3.3, SDN adopt IDS to detect and mitigate DoS attacks. During the probing phase, a KYE-attacker starts to penetrate the network in the way that is similar to the behaviour of legitimate users. Gradually, the attacker ramps up the penetration while observing the flow rules installed on the switch. Ultimately, the DoS attack will be detected by the SDN controller and new flow rules will be put in place to instruct the switch how to handle messages of the intruder. The attacker can now learn the parameters and load factor when the IDS will eventually be triggered. This information can be used to plan further DoS attacks that may not be detected by the IDS. Further, the network is not aware that it was under a KYE attack [16].
In [17] the authors propose a countermeasure called flow obfuscation to mitigate the KYE attack. The measure takes advantage of OpenFlow's ability to alter packets in transit such that the attacker is not able to distinguish its own packets from those of other users. When the attacker sends a probe message to the compromised switch, the SDN controller installs a new flow rule to modify the packet's header and forward it to another switch. The process is repeated for a predefined number of switches. At the last switch, the SDN controller installs a new flow rule enforcing the network policy destined for initial probe flow sent by the attacker. This measure prevents the attacker from learning the correlation between the issued probe messages and the resulting network policies as they are installed on a switch beyond the attacker's control.

11.3.5 Packet Injection Attack

Packet Injection Attack aims at crashing or disturbing network applications. In this scenario, a malicious user continuously sends messages with forged header fields. Specifically, the user sends messages with random MAC addresses in the header field. As a consequence, the receiving switch sends packet-in messages to the SDN controller as there aren't any flow rules present in the switch's flow table. The SDN controller processes the

incoming in-packet messages. While processing the in-packet messages, the network topology is extended with a new host mapped to the forged MAC address. Repeating these steps causes the network topology to grow continuously, resulting in a 'ghost' network of non-existing devices. The deceived network topology can eventually mislead network applications or even crash them. In an experiment, the authors were able to successfully conduct a Packet Injection Attack and cause a network configuration application to crash.
In [18] the authors propose a countermeasure against the packet injection attack that prohibits malicious users from spamming the network with forged messages. Packet-in messages shall be checked on their integrity and then either be dropped or processed regularly. To distinguish between forged packet-in messages and legitimate ones, SDN controller maintains a mapping table with information on connected hosts. The mapping table stores the DPID, an ID that identifies connected switches uniquely, the MAC address of the host and the port number on which the host is connected. When the network starts, the SDN controller collects the information and stores it in the mapping table. When a host leaves the network the controller receives a Port-Status message. The SDN controller then removes the mapping for the leaving host from the mapping table. Checking whether a packet-in message is forged or not becomes a trivial task. The SDN controller verifies whether the MAC address and port number of the packet-in message exist in the mapping table. If the entry is found the packet-in message is processed or dropped in case there exists no such entry.
In their evaluation, the prototype was able to mitigate nearly 100% of the forged messages with negligible overhead [18].

11.3.6 ARP Spoofing

In computer networks, the Address Resolution Protocol (ARP) is responsible to translate IP addresses from the network layer to MAC addresses in the data link layer of the OSI model. When a new host joins a Local Area Network (LAN), it's assigned a unique IP address and a mapping from IP address to MAC address is generated and cached for further lookup. To save space and avoid stale data, ARP cache is frequently updated. If an ARP mapping for a particular IP address is missing, ARP sends a broadcast message to all clients asking them to respond with their MAC address if they recognize the IP address as their own. A malicious host can respond to the broadcast message to map its own MAC address to an IP address of another host [13]. This attack is known as ARP spoofing and can be used as an attack vector to support other attacks [14]. To mitigate ARP attacks, multiple countermeasures exist and have been deployed. They include host/server software, static configurations, and proprietary systems and induce challenges in scalability, maintenance cost and vendor neutrality [14].
To ensure network compatibility, OpenFlow adapts the ARP protocol. In fact, hosts are not aware whether they are in an OpenFlow based network or not. Thus, threats associated to ARP protocol apply to SDN as well. The authors of [15] simulated ARP spoofing in an OpenFlow based network to analyze how the network is affected. Their results show that ARP spoofing can be effectively conducted and that the victim host's ARP cache can be contaminated. Further, they concluded that ARP spoofing on OpenFlow switches can potentially lead to a greater impact as traffic between the control plane and data plane could be intercepted by a succeeding man-in-the middle attack. Thus, critical information, such as cryptographic parameters (c.f. section 11.3.3), could be disclosed [15].
To mitigate ARP spoofing in SDN, the authors propose a mechanism called Active ARP Inspection (AAI), which is deployed in the control plane. OpenFlow switches are configured to forward all ARP related messages to the control plane where they are processed by the AAI. The advanced processing of ARP messages goes through several stages to

identify potential conflicts where multiple hosts claim the same IP or MAC address. The system then updates the ARP cache only with validated and non-conflicting ARP response messages.
The implementation was evaluated under the same simulation as mentioned above. The system was able to identify all forged ARP messages without a considerable decrease in network throughput rates. It, thus, shows that SDN can be leveraged to mitigate ARP spoofing [15].

11.4 Discussion and the Future of SDN

The security of SDN networks and their components requires carefully and holistically designed SDN architectures. Especially the SDN controller imposes a single-point-of-failure and, thus, is a high-value target to attackers. Security is only one aspect to be considered in the design of an SDN network. Reliability and scalability are two other important dimensions that need to be included in the network design.
For more reliable and scalable networks, a cluster consisting of multiple SDN controllers could be leveraged. The cluster of SDN controllers could work in parallel to create a certain degree of desired redundancy. An attack against the SDN controller would no longer yield a complete failure of the network, but could be mitigated by redundant SDN controllers to take over. In [21] the authors propose a cluster of redundant and heterogeneous SDN controllers. The proposed architecture improves security against attacks attempting to exploit known vulnerabilities in the SDN controller. The heterogeneous cluster of SDN controllers enhances security by lowering the chance that a particular SDN controller type becomes an attack vector [20].
The discussed multi-controller architectures may enhance network security and reliability but also introduce an overhead in complexity. An important research question in this field is thus the orchestration and management of multiple SDN controllers. New components will be needed to coordinate the actions between the controllers, and to maintain a consistent state. These components should be designed with security in mind as they potentially introduce new vulnerabilities [20].
Security is also an aspect in other current research areas of SDNs. The integration of SDN networks with cyber threat intelligence (CTI) systems yields promising results. These systems collect data on cyber attacks including attack type, indication, attackers, duration and countermeasures [20]. Leveraging this information can help network operators to include it in the design of their defense strategies. While human interpretation of collected cyber threat information is effective, information could be utilized to generate actual flow rules out of it. To this end, ML and DL techniques are leveraged in processing the collected threat information and transform them into actionable flow rules [20].

11.5 Conclusions

SDN architecture has gained traction and adoption in industry for its numerous benefits in network management. Despite the benefit of easier network provisioning, SDN architecture offers opportunities related to cyber security for both attackers and network defenders [20]. With the increasing adoption of SDN networks, security aspects become more critical and must be at the center of the network design.
The literature review in this report showed that challenges and opportunities of SDN in the field of cyber security is a popular research area. Most of the reviewed challenges and opportunities are related to threat mitigation or detection of a particular cyber attack. To this end, proposed solutions often include modifications or extensions to the OpenFlow

standard. While this makes sense in an academic setting, it is not clear how valuable these contributions are in practice without them being adopted by the OpenFlow standard.
This report presents a selection of cyber security opportunities and challenges for SDN networks. It focuses on specific attack types and highlights attack vectors and vulnerabilities specific to SDN networks. On the other hand, we present opportunities in the SDN architecture attempting to mitigate cyber threats. At the end of the report an outlook on the development of SDNs is given. We believe this report is useful to readers interested in getting an overview of cyber security in SDN networks.

Bibliography

[1] Feamster, N., Rexford, J., Zegura, E. (2014). The road to SDN: an intellectual history of programmable networks. *ACM SIGCOMM Computer Communication Review, 44(2)*, 87-98.

[2] Neghabi, A. A., Navimipour, N. J., Hosseinzadeh, M., Rezaee, A. (2018). Load balancing mechanisms in the software defined networks: a systematic and comprehensive review of the literature. *IEEE Access, 6*, 14159-14178.

[3] Blial, O., Ben Mamoun, M., Benaini, R. (2016). An overview on SDN architectures with multiple controllers. *Journal of Computer Networks and Communications*, 2016.

[4] Kreutz, D., Ramos, F. M., Verissimo, P. E., Rothenberg, C. E., Azodolmolky, S., Uhlig, S. (2014). Software-defined networking: A comprehensive survey. *Proceedings of the IEEE*, 103(1), 14-76.

[5] SDxCentral Studios. What Is OpenFlow? Definition and How it Relates to SDN, https://www.sdxcentral.com/networking/sdn/definitions/what-is-openflow, visisted April 17, 2021

[6] Tourrilhes, J., Sharma, P., Banerjee, S., Pettit, J. (2014). The evolution of SDN and OpenFlow: a standards perspective. *IEEE Computer Society, 47(11)*, 22-29.

[7] Jain, S., Kumar, A., Mandal, S., Ong, J., Poutievski, L., Singh, A., ... Vahdat, A. (2013). B4: Experience with a globally-deployed software defined WAN. *ACM SIGCOMM Computer Communication Review*, 43(4), 3-14.

[8] Dridi, L., Zhani, M. F. (2016, October). SDN-guard: DoS attacks mitigation in SDN networks. *In 2016 5th IEEE International Conference on Cloud Networking (Cloudnet)* (pp. 212-217). IEEE.

[9] Kuerban, M., Tian, Y., Yang, Q., Jia, Y., Huebert, B., Poss, D. (2016, August). FlowSec: DOS attack mitigation strategy on SDN controller. *In 2016 IEEE International Conference on Networking, Architecture and Storage (NAS)* (pp. 1-2). IEEE.

[10] Hande, Y., Muddana, A. (2019, November). Intrusion Detection System Using Deep Learning for Software Defined Networks (SDN). *In 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT)* (pp. 1014-1018). IEEE.

[11] Sultana, N., Chilamkurti, N., Peng, W., Alhadad, R. (2019). Survey on SDN based network intrusion detection system using machine learning approaches. *Peer-to-Peer Networking and Applications*, 12(2), 493-501.

[12] Chang, D., Sun, W., Yang, Y. (2019, November). A SDN Proactive Defense Mechanism Based on IP Transformation. *In 2019 2nd International Conference on Safety Produce Informatization (IICSPI)* (pp. 248-251). IEEE.

[13] M. Zydyk, Address Resolution Protocol (ARP), TechTarget. `https://searchnetworking.techtarget.com/definition/Address-Resolution-Protocol-ARP`, visited April 22, 2021.

[14] Cox, J. H., Clark, R. J., Owen, H. L. (2016, March). Leveraging SDN for ARP security. *In SoutheastCon 2016* (pp. 1-8). IEEE.

[15] Xia, J., Cai, Z., Hu, G., Xu, M. (2019). An active defense solution for ARP spoofing in OpenFlow network. *Chinese Journal of Electronics*, 28(1), 172-178.

[16] Conti, M., De Gaspari, F., Mancini, L. V. (2017, May). Know your enemy: Stealth configuration-information gathering in SDN. *In International Conference on Green, Pervasive, and Cloud Computing* (pp. 386-401). Springer, Cham.

[17] Conti, M., De Gaspari, F., Mancini, L. V. (2018). A novel stealthy attack to gather sdn configuration-information. *IEEE Transactions on Emerging Topics in Computing*, 8(2), 328-340.

[18] Deng, S., Gao, X., Lu, Z., Gao, X. (2017). Packet injection attack and its defense in software-defined networks. *IEEE Transactions on Information Forensics and Security*, 13(3), 695-705.

[19] Lara, A., Kolasani, A., & Ramamurthy, B. (2013). Network innovation using openflow: A survey. *IEEE communications surveys & tutorials*, 16(1), 493-512.

[20] Yurekten, O., & Demirci, M. (2021). SDN-based cyber defense: A survey. *Future Generation Computer Systems*, 115, 126-149.

[21] Hu, H., Wang, Z., Cheng, G., & Wu, J. (2017). MNOS: a mimic network operating system for software defined networks. *IET Information Security*, 11(6), 345-355.

Chapter 12

Detection and Analysis Methods for Fake News

Dario Akhavan Safa

Public misinformation, disinformation, and fake news: recent political events popularized these terms that are often mistakenly used interchangeably. This report aims to consolidate and review the status quo of scientific research conducted on detection, analysis, and mitigation methods for fake news. The core principles and different approaches of fake news detection will be analyzed and explained. Finally, limitations and open challenges, such as the potential impact on censorship and society in general will be discussed.

Contents

12.1 Introduction

Although the usage of the term 'fake news' has gained popularity on a large scale only relatively recently, the means and history of fabricating and spreading alternative facts within a society date back to the ancients. One of the earliest known examples is denoted by Darnton [1], which describes the work of the Byzantine historian Procopius. In his now-famous *Anecdota*, he smeared the public reputation of the reigning Emperor Justinian with fabricated stories, portraying him as cruel, demonic, and incompetent. It remains unclear whether Procopius' motivations resided in a personal grudge against the emperor, or an attempted coup in which he possibly took part of - the only thing that is clear, is the fact that his stories have nowadays been confirmed to be exaggerated, inconsistent or just downright false.
In today's day and age, a glimpse and an analysis of Google Trend's search data reveals that the usage of the term 'fake news' has emerged and gained rapid attraction during the U.S. presidential election in November of 2016, and remained popular since[1].

12.2 Background

12.2.1 Definition and Taxonomy

Searching for a definition of fake news makes it clear that there exists no universal answer. This chapter aims to consolidate these different definitions and distinguish between the different types of fake news. In general, we can state that fake news relates to the fabrication or propagation of factually incorrect information. On a high level, the authors of [2] differentiate between three main categories that fulfill these properties. [2] distinct between misinformation, disinformation, and fake news, mentioning that the latter overlaps with the first two categories. Public *misinformation* is false, inaccurate, or misleading information that is communicated no matter regardless of the intention to deceive people. Examples can include false rumors, insults, and pranks. *Disinformation* is regarded as a subset of misinformation; on top of false information, this subset aims to be deceptive, misleading, or biased deliberately. Often a clear motivational background is recognizable. Hoaxes, public smear campaigns, and political propaganda fall into this category. Lastly, fake news can be defined as a hybrid between misinformation and disinformation, with the distinction being that false information tries to mimic mainstream news media content.

[1] https://trends.google.com/trends/explore?date=all&q=fake%20news

12.2.2 Definition approaches in Related Works

Media outlets and journalists are using the term *fake news* and scientific research across various interdisciplinary areas, and thus definitions vary considerably. Since there is no overarching definition at hand for this specific term, this section will provide an overview of different approaches of specifying the characterization of *fake news*. The Cambridge dictionary [3] defines fake news as:

> "false stories that appear to be news, spread on the Internet or using other media, usually created to influence political views or as a joke."

From a journalism point of view, the New York Times [4] defines fake news as follows, while including an emphasis on the often sensationalist aspect of fake news articles:

> "Narrowly defined, 'fake news' means a made-up story with an intention to deceive, often geared toward getting clicks"

Lastly, after the events that followed the U.S. presidential campaigns of 2016, Facebook retired the term 'fake news' altogether in favor of 'false news' [5]:

> "The term 'fake news' has taken on a life of its own. False news communicates more clearly what we're describing: information that is designed to be confused with legitimate news, and is intentionally false."

In the end, it is debatable whether the term fake news can be universally defined. It is a broad umbrella term with no exact date of origin, that can have many different meanings and interpretations within different contexts. The authors of [6] identified seven different types of online content under the label of fake news, while clearly defining the characteristics of each one of them - (*cf.* Fig. 12.1)

Types of content	Fact-checked	Emotionally charged	Source verification	Registration inconsistency	Site pedigree	Narrative writing	Humor
Real news	Yes	No	Yes	No	Yes	No	No
False news	No	Yes	No	Yes	No	Yes	No
Polarized content	No[a]	Yes	No	No	No	Yes	No
Satire	No	No[a]	No	No[a]	No	No	Yes
Misreporting	No	No	No	No	Yes	No	No
Commentary	Yes	Yes	Yes	No	Yes	Yes	No
Persuasive information	No	No[a]	No	No[a]	No[a]	Yes	No
Citizen journalism	No	No[a]	No	No[a]	No[a]	No[a]	No

[a]Feature not available, tagged as "No" for the purpose of this exercise.

Figure 12.1: Seven Types of news content as described by Molina *et al.* [6].

12.3 Detection and Analysis Methods

There are a lot of different approaches to detect fake news, and the area of research is rapidly evolving. In a recent work published by Zhou et al. (2020), multiple state-of-the-art models are consolidated and compared to each other, while the authors propose a

new theory-driven to detect fake news [7]. The authors of [8] make a distinction between content-based and feedback-based identification. Content-based approaches analyze the textual content in fake news and try to separate authentic and fake news by looking at several linguistic cues and features in the text of the news article. The second category focuses on the secondary information surrounding a fake news article. Feedback-based detection looks at comments and user feedback (*e.g.,* likes, retweets, subscriptions) on the specific news articles that are verified. In the end, all of these approaches share one characteristic: they all rely on machine learning algorithms and deep learning to analyze and process large, mostly noisy, unstructured, and semi-structured data sets. The differences between the approaches ultimately fall into the specific patterns and features fed into these algorithms to detect fake news.

12.3.1 Deep Learning with Neural Networks

The published work from Nielsen [9] explains the relation and core concepts behind deep learning and neural networks in detail. Deep learning is a subset of machine learning, which primarily focuses on learning with neural networks. Neural networks are considered a new programming paradigm that enables a computer to reenact the learning process of neurons in biology. This paradigm makes it possible for a computer to learn from observational data and draw conclusions about complex correlations in the data that often are too complex for humans to understand. Therefore, Deep Learning (DL) is a term referred to as the robust set of techniques and frameworks for learning in neural networks. Due to recent breakthroughs in Natural Language Processing (NLP) using neural networks, and their frequent usage within fake news detection models, a high-level summary of three of the most popular used network structures is given in this section.

12.3.1.1 Artificial Neural Network (ANN)

The detailed definition of a neural network is given by [9], whereas a summary is given in this section. The most general definition of a neural network is given by an Artificial Neural Network (ANN). An ANN consists of a collection of nodes, so-called neurons, which loosely correspond to the idea of neurons in a brain. Each neuron can communicate with a certain set of other neurons, through synapses (*i.e.,* edges) between them. An artificial neuron receives an input signal, processes it, and outputs its computed signal to the next neuron(s) (if any are available). Inputs can be feature values of a data set, whereas in the detection of fake news, the input is usually a set of documents, words, or phrases. If the neuron is located on an in-between network layer, the input of a neuron is the collection of outputs of other neurons in the preceding layer of the network. The last layer of a network is the output layer, which determines the final outcome of a task, such as determining whether a news article is fake or genuine. Both neurons and their connections (*i.e.,* edges) generally have a weight that is assigned and automatically adjusted during the learning process. Neurons can have a threshold so that a signal is only propagated if its input crosses a pre-defined threshold level. Each layer of a neural network performs a different type of computation or transformation on its inputs.

12.3.1.2 Dense Neural Network (DNN)

A Dense Neural Network (DNN) is a simple subset of an ANN and consists of three parts - an input layer, hidden layers in-between, and an output layer. A *dense* neural network, refers to the characteristic that layers are fully connected in all network layers, *i.e.,* a neuron (depicted as a circle in figure 12.2) receives an input from all neurons of the previous layer - therefore, the network is considered to be densely connected [10; 9; 11].

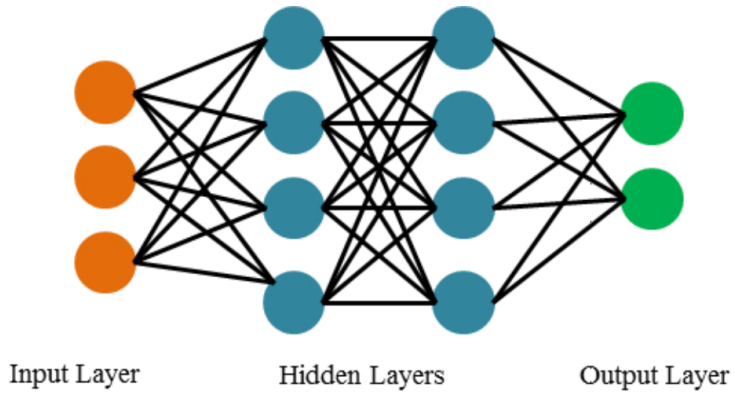

Figure 12.2: Visual representation of a DNN [11]

12.3.1.3 Convolution Neural Network (CNN)

A Convolutional Neural Network (CNN) differs from an ANN in that it operates over a set of inputs. CNN's are widely used in image processing and recently gained popularity in text processing [12]. A convolution is mathematically defined as a combination of two functions to produce a third relationship - thus, it joins two sets of information. In a CNN the neural network layers contain different dimensions. The convolution is formed on the input data by applying a filter/kernel, which produces a feature map. After several pooling operation steps, the multi-dimensional output is reduced to a single vector of probability scores, which are then fed into a 'regular' ANN [13].

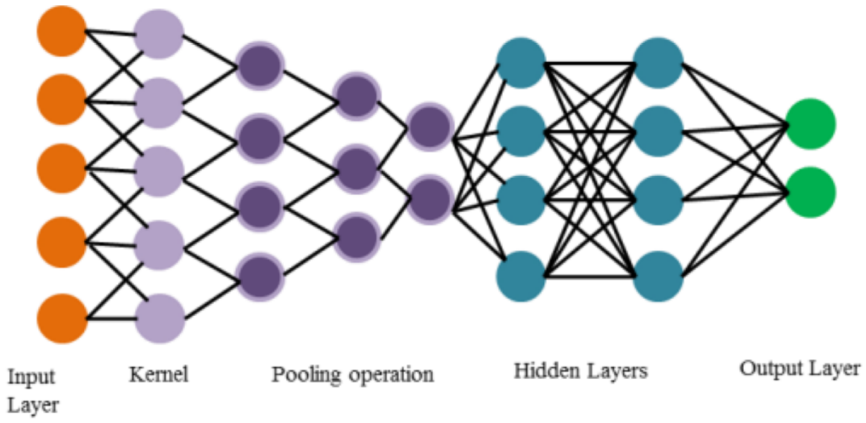

Figure 12.3: Visual representation of a CNN [11]

12.3.1.4 Recurrent Neural Network (RNN)

Recurrent Neural Networks (RNN) are typically applied when working with sequential data, such as text, genomes, handwriting, or numerical times series data [11]. As explained by Sharma *et al.* [8] each network layer stores a memory about the state of the preceding layer. At each step, the information from the currently processed word updates the layer's hidden state. This memory model is useful in the context of predicting the next word of a text sequence.

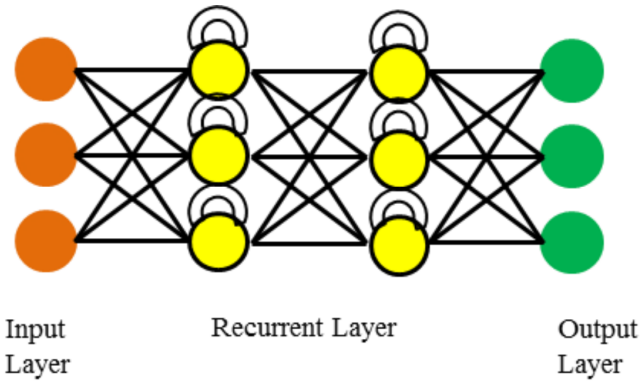

Figure 12.4: Visual representation of a RNN [11]

12.3.2 Content-based Detection

As the name suggests, content-based approaches rely on the analysis of qualitative news content, with the help of different frameworks. Zhou *et al.* categorize three forms of representation of news content [7]. These three perspectives include *(i)* Knowledge, *(ii)* Style, and *(iii)* Latent features.
(i) First, we consider the **knowledge** framework, that aims to process sentences out of a given text, *e.g.,* news articles or social media posts, and extracts sets of tuples, containing three elements: subject, predicate, object - also known as the semantic triple [14]. An example for the usage of this framework would be the SPO tuple (ElonMusk, Profession, CEO) for the sentence "Elon Musk is the CEO of Tesla, Inc.". The resulting tuples, *i.e.,* the to-be-verified knowledge extracted from the news sources are then compared with a trusted source in the form of an oracle - typically a knowledge graph [15]. Paulheim [16] defines the term knowledge graph as follows:

> "A knowledge graph mainly describes real-world entities and their interrelations, organized in a graph, defines possible classes and relations of entities in a schema, allows for potentially interrelating arbitrary entities with each other, and covers various topical domains."

A well-known use case of the implementation of a knowledge graph is Google's knowledge panel, see figure 12.5. The panel captures and displays the most frequently associated values in regards to the search query

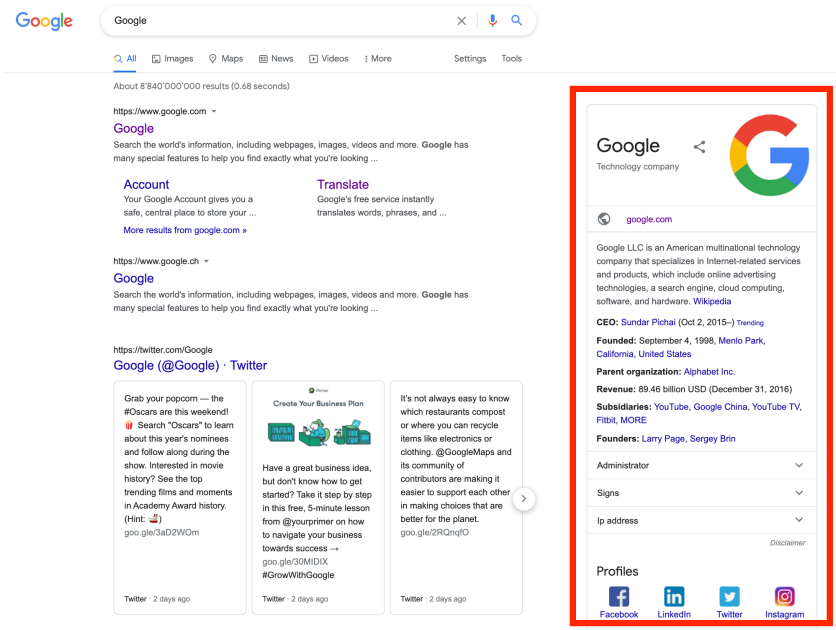

Figure 12.5: Google knowledge panel

Knowledge graphs contain a large amount of manually processed relational knowledge from various sources, most often from the internet. There are, however, multiple challenges when working with knowledge graphs. First, when applying any form of mathematical or statistical analysis and processing in the form of machine learning algorithms, knowledge graphs are assumed to be incomplete and noisy, and entities and relations may contain duplicates [17]. Since knowledge graphs are often impressively massive in their size, this requires a large amount of post-processing and impacts the performance of fake news detection algorithms. Another challenge that arises when working with knowledge graphs, is that the comparison between to-be-verified information and the knowledge graph requires an always timely, *i.e.,* up to date knowledge data set, especially when evaluating news of recent events. Third, the knowledge framework can only determine whether the verified news article is factually false (*i.e.,* content-wise) and not whether the source of the false information was intentionally misreporting a news event. This makes for an essential distinction between the formerly described classification of misinformation and the intentionally deceptive nature of fake news.
(ii) **Style** is another approach to detect fake news in a content-based matter with machine learning. The focus in this method relies on the analysis of a text with the help of certain self-defined sets of machine learning features. These algorithms map characteristics that regularly appear in fake news data to the data that the user would like to analyze. An overview of commonly used machine learning techniques, tools, and features in this category is listed below:

- Term Frequency - Inverse Document Frequency (TF-IDF)
  Often used as a weighting factor, TF-IDF reflects the importance of a keyword in a text corpus [18]. The TF-IDF value proportionally increases to the number of times a word appears in a given document. The value is decreased by the number of documents in the corpus that contain the word, to take into account that certain words appear more frequently in general (*e.g.,* preposition words). Tf-idf can be used to identify text segments with similar importance A concrete use case of tf-idf can be seen in the automatic fake news detection described by [19], where it was utilized to encode their feature sets of their model as a form of pre-processing.

- Bag of Words (BoW)
  In this model, a text (can be a sentence or a full document) is represented as the multiset (*i.e.,* bag) of its words, disregarding grammar and word order but retaining multiplicity.

- word2vec, doc2vec
  word2vec, introduced by Mikolov *et al.* [21] solves the problem of representing words numerically to generate representation vectors in a machine learning context. Words can be encoded as numeric values, while relationships between words are still persisted. These representations encapsulate the relationship between words, like synonyms, antonyms, or analogies. As seen in figure 12.6, even though the words will be encoded in numerical values, the relationship between certain words has still prevailed within the algorithm.
  doc2vec is an extension of word2vec with the goal to represent documents (in place of words) numerically, regardless of the length of the document [22].

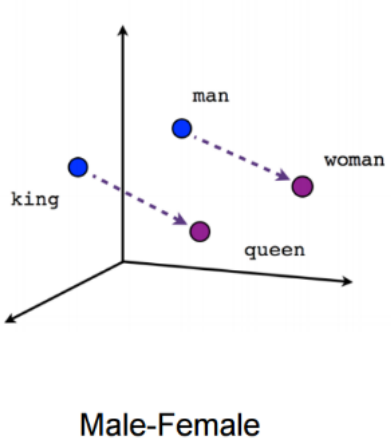

Figure 12.6: word2vec internal representation - king to queen is like man to woman [20].

- n-grams
  Mostly used in computational linguistics and probability, n-grams are contiguous sequences of n items from a given sample of text or speech. In fake news detection models, n-grams are used to extract unigrams and bigrams derived from a data set of representative words of a news article [23]. These extracted features are often encoded as tf-idf values.

- Linguistic Inquiry and Word Count (LIWC)
  LIWC is a transparent text analysis program and is used to process the psychological meaning of words, which can be used to analyze emotions and sentiments that are conveyed through text [24].

(iii) Lastly, **latent features** are obtained by deep learning techniques (such as the TI-CNN described in[12]) and matrix/tensor factorization executed on news articles. Fake news detection through the means of latent features is often quite effective. However, due to the high complexity and hard-to-follow computations, latent features are difficult to comprehend, making it difficult to explain the exact reasoning behind the classification of a fake news article.

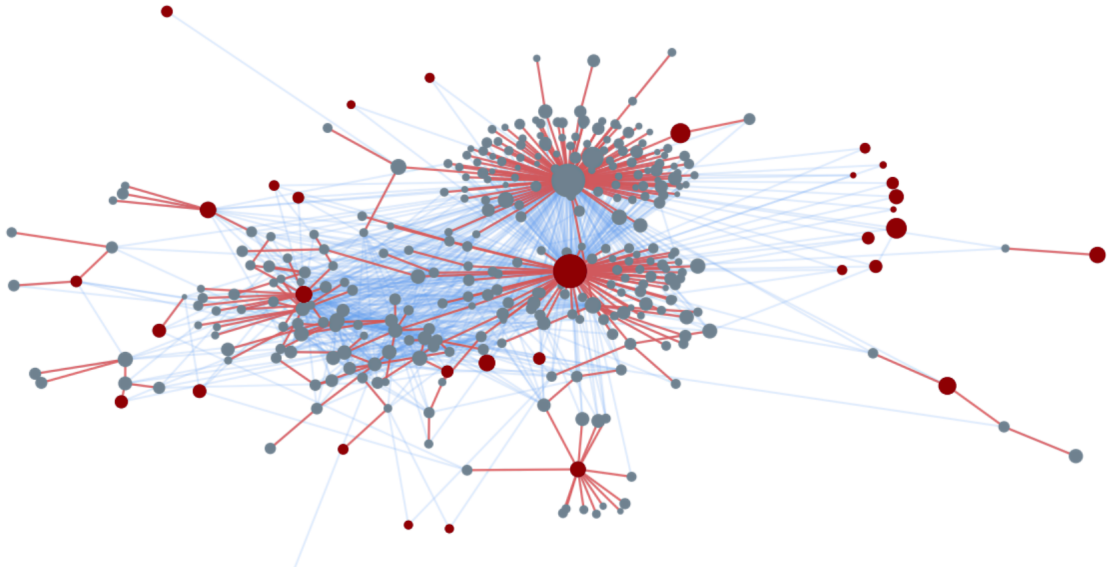

Figure 12.7: Example of a news cascade visualization [20]

12.3.3 Feedback-based detection

Instead of analyzing the actual content of a news article, feedback-based approaches aim to look at the network interactions and events between users and news articles and interaction between the users themselves.

12.3.3.1 Propagation Pattern Analysis

Due to the massive usage of social media networks in recent years, a modern approach to detect fake news utilizes social context information, such as propagation, spreading, and connectivity between fake news spreaders. One example of this approach is the published model from Monti *et al.* [25] which takes into account the *user profiles* sharing news content (geolocalization, language, etc.), *user activity* (engagement with the network), the user's *network, and spreading* (connections, followers, etc.), and *content* (published social media posts). To see how a news article is spread on social media, data can be presented using a news cascade tree structure. The cascades are defined as the news diffusion tree of a source object (*e.g.,* the first tweet referencing a certain URL) and its children (*e.g.,* retweets and reactions). In figure 12.7, Monti *et al.* visualized the spreading of a single news story on a subset of Twitter accounts. Users that share the news article's URL are termed cascade roots and are pictured in red. The cascades are shown with red edges. Social connections between users are displayed with blue edges. Wu *et al.* [26] have modeled news cascades as multivariate time series, and detect fake news by using deep learning techniques, such as recurrent and convolutional neural networks. Their model is based on the assumption that the structure of fake and authentic news cascades look and behave fundamentally different and can therefore determine a probability for a given news article to be classified as fake news.

12.3.3.2 Temporal Pattern Analysis

Temporal modeling methods rely on the differences in temporal dynamics of user feedback when comparing fake news to trustworthy news. In the modeling of temporal methods, a sequence of social media posts regarding a news article is ordered chronologically and divided into a fixed number of time intervals. Ma *et al.* [27] consider three types of temporally relevant features that can be applied in the analysis of each interval:

- Text Features
  A collection of text style features related to the text of the posts. Examples include: proportion of posts with exclamation or question marks, the proportion of posts with user mentions, the proportion of posts that include a hashtag, number, and type of emoticons used in the posts

- User Features
  A collection of user features related to the owner of the posts. Examples include: the amount of user followers, registration date and age of a user, the verification status of a user

- Propagation Features
  A collection of metadata related to the propagation of the posts. Examples include: average number of shares, the average number of comments

The variation of features between two adjoining time intervals is made up of the number of differences in the feature values divided by the length of the interval. In the end, a final vector of features from each interval can be used to compare and differentiate fake news from actual news. Ma *et al.* observe a significant performance improvement with the usage of temporal pattern analysis over methods that solely rely on measured propagation features (such as the number of shares or likes of a post) without taking into consideration the variations over time [27]. One shortcoming of this model is the fact that it requires a manual selection of handcrafted features that are further restricted to numerical features. This limitation can be overcome by combining the temporal pattern analysis model with deep learning techniques, so that intricate temporal variation differences between genuine and fake news can be automatically extracted, as described in the model of Ruchansky *et al.* [28].

12.3.4 Mitigation Methods

In this following section, it is discussed how the effects of exposure to fake news articles can be mitigated, *i.e.,* how news and social media platforms can ensure that fake news is detected early and contaminated to block the spreading to other users or at least how to inform the users of the platform of the risk of viewing a fake news article. This chapter discusses several categories of mitigation efforts and discusses the rationale and limitations of such approaches.

12.3.4.1 Automated Mitigation

In today's digitally connected world, news articles can spread at astonishingly fast levels, through the vast amount of social media channels. These spread events, or news cascades as defined before, are of chaotic nature and cannot be tracked efficiently by a single individual. Therefore, we need to seek help and assistance from artificial intelligence to mitigate the spread of fake news. This section presents the most promising approaches to tackle this problem.

- Decontamination
  Proposed by Nguyen *et al.* [29], decontamination is a strategy where users that have been exposed to fake news are decontaminated by being confronted with genuine news. A greedy algorithm is implemented that selects the best set of users, that will propagate the genuine news cascade. The aim is to reach as many contaminated users as possible with the intention to confront their initial interaction with a fake news article with truthful, verified facts, in the hopes of mitigating their fake news consumption. *Limitations of this model*: This form of mitigation is only a response to an already spread-out fake news event. Therefore this model can only be regarded as a means of improvement of an already escalated situation.

- Competing Cascades
  Here, a genuine news cascade is introduced to compete against a detected fake news cascade. Budak *et al.* [30] introduced an 'influence blocking maximization objective, in which a strategic selection of a user set is made, so that as many users as possible will be activated (*i.e.,* confronted) with genuine news before a fake news cascade could reach them. This model assumes that the first activation of a user with a news source will determine whether a user belongs to a genuine news or a fake news cascade. Therefore, if a higher amount of users get activated by a genuine news cascade first, the amount of people that are affected by fake news stays minimal.
  Limitations of this model: The genuine news cascade will only be propagated after a fake news cascade has been detected. Also, it can be argued that the assumption made by this model about the activation of users is too simplified and does not reflect reality.

- Multi-stage Intervention
  Farajtabar *et al.* [31] considered a multistage intervention framework based on multivariate point processes. The model assumes that news-sharing events of a user in the past can trigger and influence future news-sharing events based on the intensities of influence and time delay between events [31; 8], therefore addressing the limitations of competing cascades, since users that are first confronted with fake news do not necessarily have to share the same views as the media they were exposed to. The end goal of this model is to keep the propagation of fake news in control and incentivize users to counteract fake news themselves by sharing and spreading trustworthy news.
  Limitations of this model: Multi-stage intervention requires the exact tracking of a fake news cascade and its propagation through the network, which is not a trivial task. Furthermore [8] argue that since we assume the detection of fake news as a requirement for this model, it is possible that the previously explained models of decontamination and competing cascades are more effective when it comes to mitigating the negative effects of fake news.

12.3.4.2 Early Identification

In the previously discussed mitigation methods, we have only looked at methods that mitigate the effects of fake news *after* an article has already been publicized and shared. Moreover, the published research on information cascades by Friggeri *et al.* [32] shows that information is often readily and rapidly transmitted over the network, even if the credibility or veracity of the source is dubious at best. To counteract against this drawback, we will look at three early identification strategies, presented in [8], which can be then combined with a mitigation strategy to create a quick and effective mitigation plan against fake news.

- Network Monitoring
  Network monitoring keeps track of a list of dubious sources that are susceptible of propagating fake news. Through an early fake news detection filter, the network could block users spreading intentionally deceiving false news [8].
  Limitations of this model: Network monitoring can become very expensive, the larger the network grows and malicious users could fake their identity, *i.e.,* create new accounts to appear as neutral users [8].

- Crowd-sourcing
  Users in the network are given the right to report and flag suspicious content on the platform [8].
  Limitations of this model: Reporting functionality can potentially be abused by malicious users. Further, there is a trade-off between the number of already contaminated users by fake news with the number of reports needed to draw an accurate conclusion on a potential fake news post [8].

12.3.4.3 Mitigation Through Education

Lastly, another innovative mitigation method would be to raise awareness and educate society on fake news as a whole - *e.g.,* through the use of the gamification theory in education. The authors from [33] have conducted a randomized field study, where they let the participants play a game about the creation of a fake news article. This game aims to familiarize the players with different viewpoints, motivations, and techniques that the narrative of a news topic can be manipulated by. The aim of the study was to determine whether games can be used to educate users to detect fake news more reliantly.

The setup of the game looked like this: First, players were divided into different groups of player personalities, that aim to imitate the behavior and characteristics of common journalism sectors. This included a group of *deniers*, who tried to make the news topic look as insignificant and unimportant as possible. *Alarmists* acted the opposite of deniers. *Clickbaiters* tried to generate sensationalist news headlines that gained lots of attraction, and finally *conspiracy theorists*, who tried to derail from the mainstream narratives. The groups of players were then asked to produce news articles about the overarching topic of a highly relevant news event at that time (in regards to immigration reception in Europe). While formulating the news article, the participants had to choose between various predefined answer blocks that formed a news story that best reflected their group character's motivation and views. In the end, the groups were judged and scored by the number of correct answers they had given, *i.e.,* the group with the highest accordance of answers that matched their characteristics, won. After playing the game, participants were asked to read two randomly assigned fake news articles. The researchers then conducted a survey where they collected the participant's views of these unknowingly made-up news stories on various aspects, such as perceived reliability of the articles, persuasiveness, and personal agreement with the articles. Comparing the survey's answer with a control group, who have not played the game before, the results showed significant differences in judgments about the reliability of the fake news articles. The control group rated the reliability of the fake news articles significantly higher than the treatment group. Persuasiveness and personal agreement with the articles exposed effects in a similar direction, albeit not statistically significant (Both persuasiveness and personal agreement resulted in lower measurements for treatment group compared to control group). Summarized, this experiment shows potential in media education and handling of fake news. Since the setup of this particular experiment had a sample size of just n=95 and an median age of 16 years, it would be interesting to replicate this research with bigger sample sizes and broader demographics. One could argue that young students still develop their believes and experiences with news outlets, and therefore may be better suited to empower against fake news exposure.

12.4 Discussion

Fake News, disinformation and misinformation are big problems we experience in today's form of communication. Through the rapid rise of social networks and fast-flowing information exchange, it is inevitable that we cannot avoid the problem of spreading false information. Fake news is a modern-day weapon of propaganda and can be used to alter a person's perception and opinion on a certain topic. At its worst, it can be used as a weapon to manipulate public opinion on political issues. However, the measures and methods that are currently researched to counter against the fake news movement also come with their own limitations and downsides. [8] identified a set of several challenges. These challenges are discussed the following sub chapter.

12.4.1 Open Challenges

- Potential impact of automated fake news detection on censorship and societies:
  One problem that arises with the methods explained in this report stems from the fact that automation-based fake news detection systems are not perfect. Shu *et al.* [34] analyzed the performance of several different methods on two popular data sets from PolitiFact and GossipCop. The maximal accuracy of detection was measured at 69% and 79.6% respectively. On top of that, automated systems are not only prone to false negatives, they can also produce false positives. Both error types are very problematic. If a fake news article is not detected, we may run into the issue that the false information spreads rapidly and could risk that the contaminated users will be manipulated. On the other hand, false positives are also dangerous, as a truthful verified news article could be blocked or removed from a network for no reason. This is undesirable as it can lead to censorship and suppression.

- High Stakes and Many Players:
  As stated by [8], the World Economic Forum ranked the spread of fake news amongst the top threats the world is facing today [35]. Due to the rapidly expanding nature of propagation of fake news in social media, it becomes a very difficult challenge to mitigate false information before it can cause damage. To give an indication of the propagation speed of fake news articles nowadays, Kim *et al.* [36] state that the fact-checking organizations Snopes and Politifact, which specialize in human verification of news data, cannot keep up with the vast amount of data that needs to be fact-checked every day.

- Malicious Intent
  Propagating fake news in a network can attract a lot of different malicious forms of motivation. From rather harmless 'internet trolling up to events of geopolitical warfare that aim to manipulate a nation's election campaigns - unfortunately, it is inevitable, that malicious attacks will always play a role when it comes to the spread of fake news. In turn, this increases the complexity of fake news as a whole and makes it harder for both humans and machines to find effective methods to detect and mitigate fake news [8].

- Constant Change
  As we are living in a fast-paced globally connected world, it becomes increasingly more difficult to track all news events. Content-based detection methods, such as fact-checkers, suffer as the required knowledge base has to expand and update itself automatically, while simultaneously ensuring correctness, integrity, and availability of data [8].

12.5 Summary and Conclusions

This report summarizes recent trends in fake news detection and mitigation. In recent years, there has been a lot of progress in the research field of automated fake news detection. The different models and approaches can be categorized into two major categories. First, we have looked at content-based detection, where the focus relies on the textual content of the news article itself. The textual content of a news article can be analyzed with three major focus points. First, *(i)* the content's factual correctness can be analyzed and verified using knowledge graphs. Second, *(ii)* we can analyze certain hand-picked style features of the text, such as sentence structures, word relationships, n-grams, or the psychological meaning of words. Lastly, *(iii)* latent features through the usage of neural networks can be determined and used for detection algorithms.
The second way to detect fake news is by looking at a feedback-based approach. The interaction of users with news articles and between themselves can reveal useful information that can be facilitated to detect fake news.
Finally, this report concatenated the most popular mitigation techniques and stated open challenges for the future of fake news.

Bibliography

[1] R. Darnton, "The True History of Fake News," 4 2017. [Online]. Available: https://www.nybooks.com/daily/2017/02/13/the-true-history-of-fake-news/

[2] D. M. J. Lazer, M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, M. J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, M. Schudson, S. A. Sloman, C. R. Sunstein, E. A. Thorson, D. J. Watts, and J. L. Zittrain, "The science of fake news," *Science*, vol. 359, no. 6380, 3 2018.

[3] Cambridge Dictionary, "Fake News definition by Cambridge Dicitonary," 4 2021. [Online]. Available: https://dictionary.cambridge.org/us/dictionary/english/fake-news

[4] S. Tavernise, "As Fake News Spreads Lies, More Readers Shrug at the Truth," 4 2016. [Online]. Available: https://www.nytimes.com/2016/12/06/us/fake-news-partisan-republican-democrat.html

[5] W. Oremus, "Facebook Has Stopped Saying "Fake News"," 4 2017. [Online]. Available: https://slate.com/technology/2017/08/facebook-has-stopped-saying-fake-news-is-false-news-any-better.html

[6] M. D. Molina, S. S. Sundar, T. Le, and D. Lee, ""Fake News" Is Not Simply False Information: A Concept Explication and Taxonomy of Online Content," *American Behavioral Scientist*, vol. 65, no. 2, 2 2021.

[7] X. Zhou, A. Jain, V. V. Phoha, and R. Zafarani, "Fake News Early Detection: A Theory-driven Model," *Digital Threats: Research and Practice*, vol. 1, no. 2, 7 2020. [Online]. Available: https://dl.acm.org/doi/10.1145/3377478

[8] K. Sharma, F. Qian, H. Jiang, N. Ruchansky, M. Zhang, and Y. Liu, "Combating Fake News: A Survey on Identification and Mitigation Techniques," *ACM Transactions on Intelligent Systems and Technology*, vol. 10, no. 3, 5 2019. [Online]. Available: https://dl.acm.org/doi/10.1145/3305260

[9] M. A. Nielsen, *Neural networks and deep learning*. Determination press San Francisco, CA, 2015, vol. 25.

[10] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[11] A. Thota, P. Tilak, S. Ahluwalia, N. Lohia, S. Ahluwalia, and N. Lohia, "Fake News Detection: A Deep Learning Approach," Southern Methodist University, Tech. Rep. 3, 2018. [Online]. Available: https://scholar.smu.edu/datasciencereview/vol1/iss3/10

[12] Y. Yang, L. Zheng, J. Zhang, Q. Cui, Z. Li, and P. S. Yu, "TI-CNN: Convolutional Neural Networks for Fake News Detection," *arXiv*, 6 2018. [Online]. Available: http://arxiv.org/abs/1806.00749

[13] Y. Kim, "Convolutional Neural Networks for Sentence Classification," *EMNLP 2014 - 2014 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*, pp. 1746–1751, 8 2014. [Online]. Available: http://arxiv.org/abs/1408.5882

[14] O. Lassila and R. R. Swick, "Resource Description Framework (RDF) Model and Syntax Specification," 4 1995. [Online]. Available: https://www.w3.org/TR/PR-rdf-syntax/

[15] X. Dong, E. Gabrilovich, G. Heitz, W. Horn, N. Lao, K. Murphy, T. Strohmann, S. Sun, and W. Zhang, "Knowledge vault: a web-scale approach to probabilistic knowledge fusion," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. New York, NY, USA: ACM, 8 2014.

[16] H. Paulheim, "Knowledge graph refinement: A survey of approaches and evaluation methods," *Semantic Web*, vol. 8, no. 3, 12 2016.

[17] M. Nickel, K. Murphy, V. Tresp, and E. Gabrilovich, "A Review of Relational Machine Learning for Knowledge Graphs," *Proceedings of the IEEE*, vol. 104, no. 1, 1 2016.

[18] A. Rajaraman and J. D. Ullman, "Data Mining," in *Mining of Massive Datasets*. Cambridge: Cambridge University Press, 2011.

[19] V. Perez-Rosas, B. Kleinberg, A. Lefevre, and R. Mihalcea, "Automatic Detection of Fake News," in *Proceedings of the 27th International Conference on Computational Linguistics*. Santa Fe, New Mexico, USA: Association for Computational Linguistics, 8 2018, pp. 3391–3401. [Online]. Available: https://www.aclweb.org/anthology/C18-1287

[20] G. Shperber, "A gentle introduction to Doc2Vec - Wisio," 4 2019. [Online]. Available: https://medium.com/wisio/a-gentle-introduction-to-doc2vec-db3e8c0cce5e

[21] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," in *1st International Conference on Learning Representations, ICLR 2013 - Workshop Track Proceedings*. International Conference on Learning Representations, ICLR, 1 2013. [Online]. Available: http://ronan.collobert.com/senna/

[22] Q. Le and T. Mikolov, "Distributed Representations of Sentences and Documents," in *Proceedings of the 31st International Conference on International Conference on Machine Learning - Volume 32*, ser. ICML'14. JMLR.org, 2014, pp. 1188–1196.

[23] G. Sidorov, F. Velasquez, E. Stamatatos, A. Gelbukh, and L. Chanona-Hernández, "Syntactic N-grams as machine learning features for natural language processing," *Expert Systems with Applications*, vol. 41, no. 3, pp. 853–860, 2 2014.

[24] J. W. Pennebaker, R. L. Boyd, K. Jordan, and K. Blackburn, "The development and pychometric properties of LIWC2015," LIWC, Tech. Rep., 9 2015. [Online]. Available: https://repositories.lib.utexas.edu/handle/2152/31333

[25] F. Monti, F. Frasca, D. Eynard, D. Mannion, and M. M. Bronstein, "Fake News Detection on Social Media using Geometric Deep Learning," *arXiv*, 2 2019. [Online]. Available: http://arxiv.org/abs/1902.06673

[26] K. Wu, S. Yang, and K. Q. Zhu, "False rumors detection on Sina Weibo by propagation structures," in *Proceedings - International Conference on Data Engineering*, vol. 2015-May. IEEE Computer Society, 5 2015, pp. 651–662.

[27] J. Ma, W. Gao, Z. Wei, Y. Lu, and K.-F. Wong, "Detect Rumors Using Time Series of Social Context Information on Microblogging Websites," in *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*. New York, NY, USA: ACM, 10 2015.

[28] N. Ruchansky, S. Seo, and Y. Liu, "CSI: A Hybrid Deep Model for Fake News Detection," in *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*. New York, NY, USA: ACM, 2017.

[29] N. P. Nguyen, G. Yan, M. T. Thai, and S. Eidenbenz, "Containment of misinformation spread in online social networks," in *Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci '12*. New York, New York, USA: ACM Press, 2012.

[30] C. Budak, D. Agrawal, and A. E. Abbadi, "Limiting the spread of misinformation in social networks," in *Proceedings of the 20th International Conference on World Wide Web, WWW 2011*. New York, New York, USA: ACM Press, 2011, pp. 665–674. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1963405.1963499

[31] M. Farajtabar, J. Yang, X. Ye, H. Xu, R. Trivedi, E. Khalil, S. Li, L. Song, and H. Zha, "Fake News Mitigation via Point Process Based Intervention," *34th International Conference on Machine Learning, ICML 2017*, vol. 3, pp. 1823–1836, 3 2017. [Online]. Available: http://arxiv.org/abs/1703.07823

[32] A. Friggeri, L. Adamic, D. Eckles, and J. Cheng, "Rumor cascades," in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 8, 2014.

[33] J. Roozenbeek and S. van der Linden, "The fake news game: actively inoculating against the risk of misinformation," *Journal of Risk Research*, vol. 22, no. 5, pp. 570–580, 5 2019. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/13669877.2018.1443491

[34] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake News Detection on Social Media: A Data Mining Perspective," *ACM SIGKDD Explorations Newsletter*, vol. 19, no. 1, 9 2017.

[35] L. Howell, "Global risks 2013," World Economic Forum, Tech. Rep., 2013.

[36] J. Kim, B. Tabibian, A. Oh, B. Schölkopf, and M. Gomez-Rodriguez, "Leveraging the crowd to detect and reduce the spread of fake news and misinformation," in *WSDM 2018 - Proceedings of the 11th ACM International Conference on Web Search and Data Mining*, vol. 2018-Febuary. Association for Computing Machinery, Inc, 2 2018, pp. 324–332. [Online]. Available: http://arxiv.org/abs/1711.09918