



University of
Zurich^{UZH}

*Burkhard Stiller, Alberto Huertas, Bruno Rodrigues, Christian Killer,
Eder John Scheid, Eryk Schiller, Jan von der Assen, Katharina O.
E. Müller, Muriel Franco, Rafael Ribeiro
(Edts).*

Communication Systems XV

TECHNICAL REPORT – No. IFI-2022.06

August 2022

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



Contents

1	A Virtualized/Containerized Radio Access Network as a Foundation for Multi-Radio Access Technology Communication Systems	3
	<i>Florian Herzog</i>	
2	A Technical Overview Of Blockchain Retroactive Public Funding Schemes	19
	<i>Dario Gagulic and Lynn Zumtaugwald</i>	
3	A Survey of Non-Fungible Token Use-Cases and Their Technical Standardization	47
	<i>Tumen Dambiev and Marek Gajewski</i>	
4	Brain-Computer Interfaces: Overview and Application Scenarios	73
	<i>Zeen Wang, Szymon Modrzynski</i>	
5	Data Plane Programmability: Overview, Abstractions, and Use Cases	105
	<i>Basler D. B., Browne J. I.</i>	
6	Digital Sovereignty and Digital Colonization: A Technical Discussion Regarding Challenges and the Digital Future of Countries	131
	<i>Jiaming Tong</i>	
7	An Overview of Mitigation Techniques for Spoofing Attacks in IoT	147
	<i>Geyu Meng, Zi Ye</i>	
8	Threat Hunting: An Overview of Proactive Cybersecurity Methods	173
	<i>Heman Tanos, Karin Brunner</i>	

Chapter 1

A Virtualized/Containerized Radio Access Network as a Foundation for Multi-Radio Access Technology Communication Systems

Florian Herzog

1.1 Introduction

The fifth-generation mobile communication technology standard (5G) promises a future-proof framework for not only cellular networks but various radio network applications, including IoT [2] [11]. With the ability to adjust its properties with respect to an application's needs by amplifying the network in the three major domains of ultra reliable low latency communication (URLLC), enhanced mobile broadband (eMBB), and massive machine type communication (mMTC), it offers diverse solutions to satisfy countless use cases [3]. To achieve these ambitious goals, 5G makes use of the latest technology. One core idea that 5G wants to implement in the foundation of radio access network (RAN) architectures is virtualization, leading to the concept of a virtualized radio access network (vRAN). vRAN introduces radical changes in the network design. The essential computational part of this previously distributed network is now centralized and cloudified. This does not only help solving issues of the previous design standard but comes with considerable potential and at the same time introduces new challenges for both latency and throughput.

1.2 Objectives of vRAN

1.2.1 Decoupling of Proprietary Systems

The main objective of a virtualized Radio Access Network is to decouple the network software from proprietary hardware. In 4G, the previous generation of mobile communication networks, the standard architecture has been a Distributed Radio Access Network (D-RAN). A brief overview to D-RAN will be given in sec:dran. One characteristic property of D-RAN is that the computational work on the traffic is done by sophisticated hardware composing the base band unit (BBU) directly at the Base Station (BS). While this practice allows an effective optimisation of the utilized circuits, it also excludes D-RAN from the benefits of general-purpose processors. In contrast to those, sophisticated hardware cannot capitalise on mass production. Moreover, these proprietary systems are hard to evolve and offer little flexibility.

1.2.2 Flexibility by Virtualization

Through means of containerisation, vRAN wants to divide the whole software system into an ecosystem of interconnected microservices in a virtual environment. These microservices can then be expanded, replaced, or removed individually. Any functional modification of the network can be introduced merely through a software command, without having to replace any hardware. Thanks to the detachment of proprietary systems and the use of general-purpose processors, all parts of the BBU should be able to run any computational service. Besides adding and removing individual parts of the system, a virtual environment allows an entire running process to be captured in its current state and duplicated or conveyed to a different BBU over the network with only a short downtime; A possibility that can come in handy especially when devices travel between the reachable ranges of different base stations. With these tools virtualization offers a powerful framework for an evolvable and dynamic system.

1.2.3 Better Manageability through Cloudification

With location-independent services vRAN allows cloudification, meaning that the system can run on a dynamically scalable and reliable infrastructure. Naturally, the base stations still need to be equipped with an RRH to translate between electromagnetic waves and electronic signals. However, while the physical antenna still must be present on the base station, the reconstruction of the IP package from a sampled wave (or vice versa) can be computed in the cloud. This leaves the base stations requiring only an RRH with a direct fronthaul to the cloud, without its own BBU on site. Instead, the RRH will communicate with a dynamic BBU pool (the cloud) over the fronthaul. This is particularly attractive for internet service providers as it reduces the capital expenditure for a RAN infrastructure, requiring significantly less hardware to provide the same connectivity and bandwidth to its users because of the ability to distribute fluctuating workloads over the whole cloud,

instead of one local BBU having to be able to handle any level of utilization. Such a scalable and efficiency-aware design is crucial in order to realize the envisioned future of the internet, which expects the amount of data traffic to explode: See figure 1.1

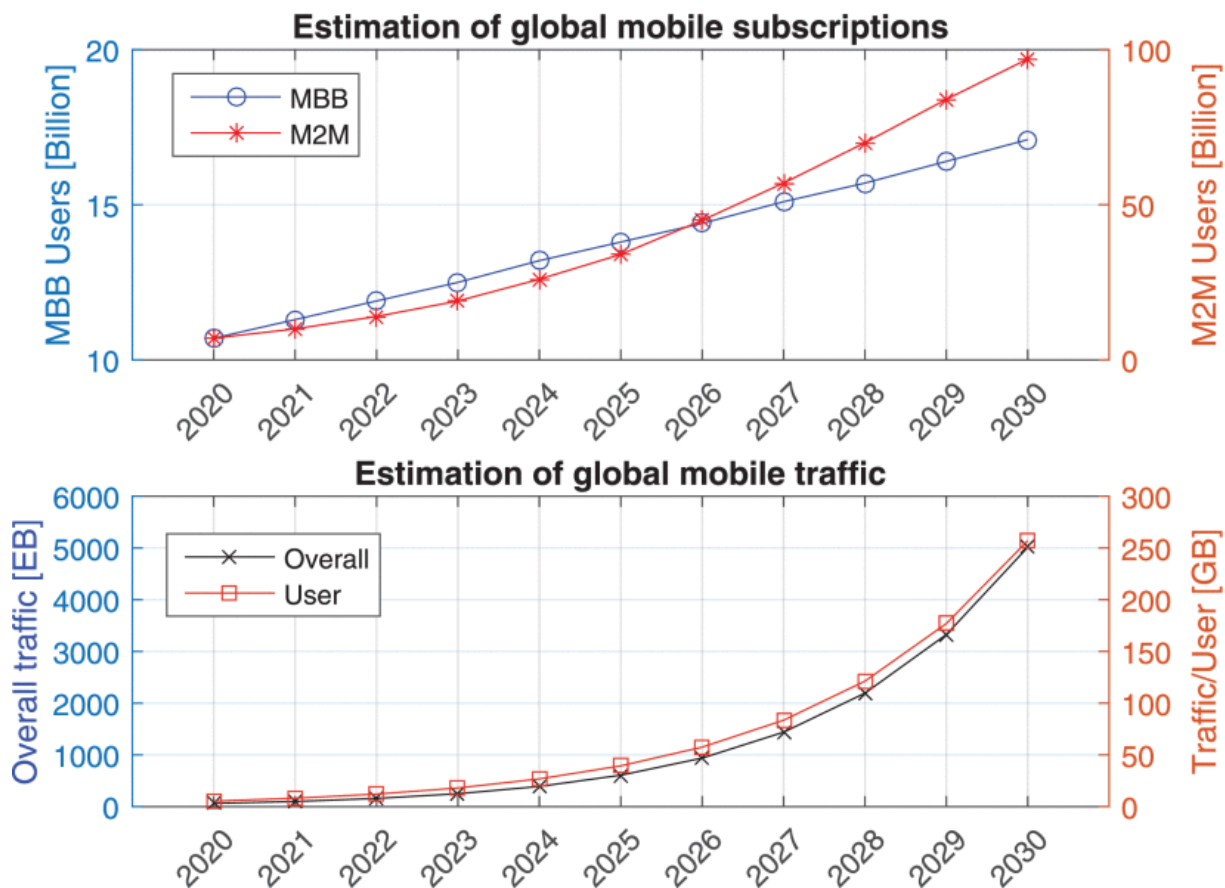


Figure 1.1: estimated development of cellular network load (source: [16])

1.2.4 A Robust and Adaptive Framework

By virtualizing the computational part of a radio access network, most parts of the network are now controlled by software. This allows for a sophisticated system architecture appropriate to best-practice designs, in a way that autonomous and independent features can be placed on top of this framework. This way, future enhancements in the form of changes or additional applications of the network can be implemented easier and in a modular manner. With that, a virtualized RAN lays the foundation for many visions of the next generation of cellular networks. [16]

1.2.5 Optimized Resource Management

With an overall improved monitorability and controllability, resources can be allocated more efficiently. New models are shown to operate with significantly less power consumption than legacy systems: [15]

1.2.6 Network Slicing

A virtualized network can be divided into logically independent parts that still operate on top of one underlying physical infrastructure. This concept is known as network slicing [18]. 1.2

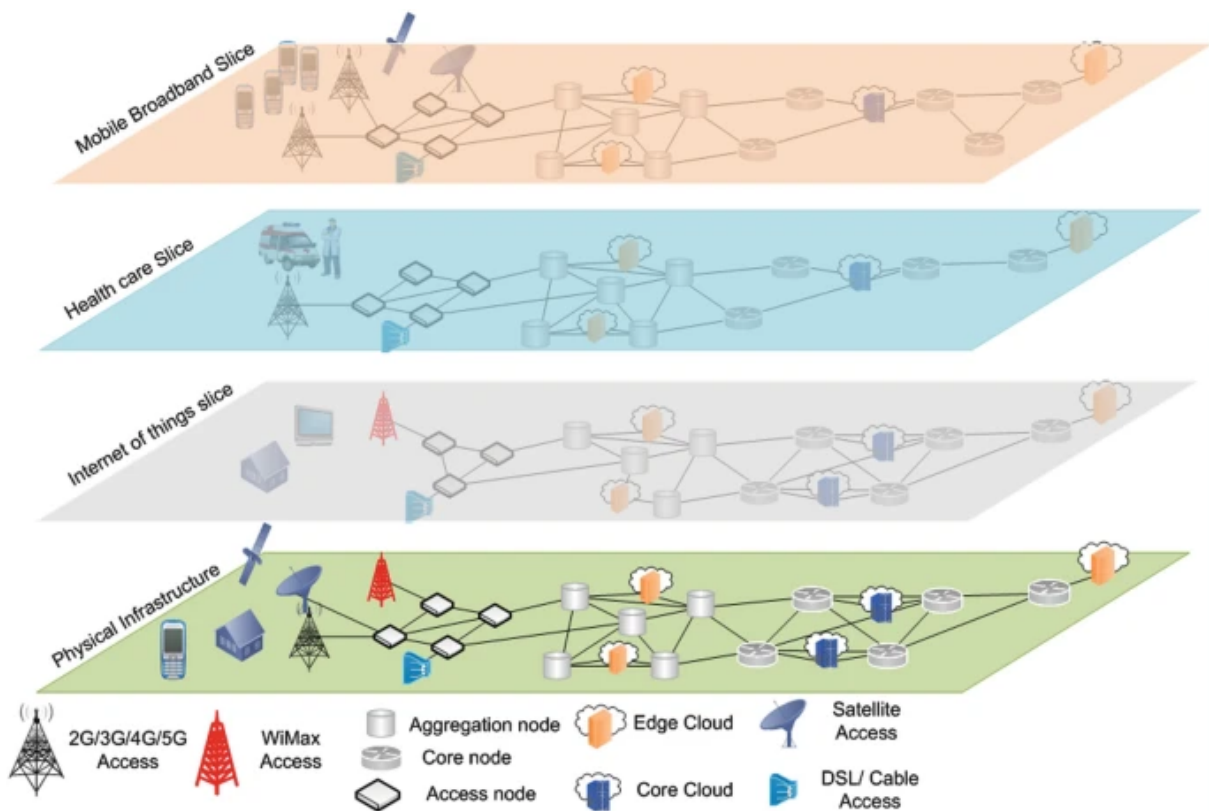


Figure 1.2: multiple network slices over the same infrastructure(source: [18])

In 5G, this mostly finds application in the way that networks are split into three major domains, each one tuned to maximize one particular property of the network's capabilities:

- ultra reliable low latency communication (URLLC)

- enhanced mobile broadband (eMBB)
- massive machine type communication (mMTC)

The users of the network can then subscribe to the slice that best fits their preferences. E.g. an autonomous vehicle that strongly relies on real-time data would chose the URLLC slice. Note that in theory the network can be sliced arbitrarily.

1.3 Challenges of vRAN

1.3.1 Latency

Time is often critical in cellular networks. Existing cellular protocols postulate certain requests to be answered within a tight timeframe. While the proprietary hardware processing the traffic in D-RAN made sure that this time budget could be met, both virtualization and cloudification have performance drawbacks as a cost of abstraction compared to a primitive, low-level system. [10] [9]

1.3.1.1 Higher Physical Distance.

While in the distributed design of D-RAN the BBU was located directly at the base station, the centralized architecture of C-RAN has the base station connected to a cloud infrastructure over a fronthaul that might have several kilometers length for messages to travel before being processed and forwarded to the core network (or vice versa).

1.3.1.2 Cloud Hardware Allocation

The main characteristic of a cloud infrastructure, that a workload can be assigned to any machine in the cloud, brings another performance issue. Since a message may be processed by any node in the cloud, this node does not necessarily have to be the one closest to the base station holding the connection to the user. To minimize the travelling time between the base station and the core network, the cloud should implement edge computing [14] to distribute its work in a location-aware manner.

1.3.1.3 Additional Effort from Abstraction.

A virtual environment has by design more data layers than a native setup to be able to provide its abstraction functionalities. Virtual hardware as well as a hypervisor are essential components of a containerized system which vRAN incorporates. These two additional layers bring their own overhead and hence cost to the system's performance.

1.3.1.4 Latency by Scheduling

In the virtual BBU, a newly arrived package may remain idle for some time, waiting for its responsible process to be scheduled by the hypervisor. In conventional virtual platforms this can take up to several milliseconds. Even when being scheduled immediately, precious time can be lost when the machine must go through a context switch. In a real-time application where a request must be answered within a couple of milliseconds in order to keep the connection alive, this is unacceptable and must be circumvented.

1.3.2 Throughput

While the cloudified BBU pool can easily grow to handle any desirable data rate, the fronthaul connecting the base station to the cloud can now become a bottleneck. 5G envisions to use multiple input multiple output (MIMO) systems on a high frequency channel. This results in immense requirements for the fronthaul capacity [8]. It is highly recommended to use compression algorithms to decrease the size of the data passing through the fronthaul. On that effect a compression of 2x was found to be lossless and a compression of 3x was found to have a corruption degree that is practically acceptable [1]. However, even with this heavily reduced data volume the desired capacity of 5G can easily go beyond the limits of modern ethernet capabilities as they are used for the fronthaul in a standard C-RAN design. In order to further release the pressure on the fronthaul, a heterogeneous cloud radio access network (H-CRAN) proposes the use of HetNets for an enhanced network design [12]. Furthermore, Bartelt illustrated: the combination of utilization-dependent functional splits and statistical multiplexing can additionally reduce aggregated transport traffic by up to almost one order of magnitude as compared to today's CPRI-based networks [8].

1.3.3 Geographical Complexity

With 5G's newly allocated frequency spectrum a cellular network gains significant bandwidth capacity. However, this has the drawback that high-frequency waves can travel shorter distances. Therefore, the network topology requires an additional density of base stations in order to serve the high-frequency spectrum. This does not only mean a geographical increase of infrastructure, it also magnifies the challenge of planning a network. In many cases a complete coverage of high frequency radio access is either not profitable, very complicated, or simply impossible for various reasons.

note: this is a challenge of the 5G frequency spectrum rather than RAN virtualization, but it is mentioned because the topics are closely correlated and there is a common solution to this and vRAN-specific problems.

1.4 RAN architecture

1.4.1 From D-RAN to C-RAN

A distributed RAN (D-RAN) base station (BS) mainly consists of a remote radio head (RRH) and a base band unit (BBU). See figure 1.3. The RRH acts as a gateway between the two different types of signals used in mobile communication: Electronic traffic on the wire to communicate over the internet, and electromagnetic waves closing the connection to the user equipment over the air. Once the RRH produced a digital sampling of an incoming wave, the common public radio interface (CPRI) carries this sampling down to the BBU. The BBU then is responsible for the computational work that comes with this change of medium. As both communication over a wire and over the air bring their own challenges and opportunities, the two technologies have been optimized in their own ways and it takes computational effort to translate data from one format into the other. The

BBU translates sampled waves to IP packages and vice versa.

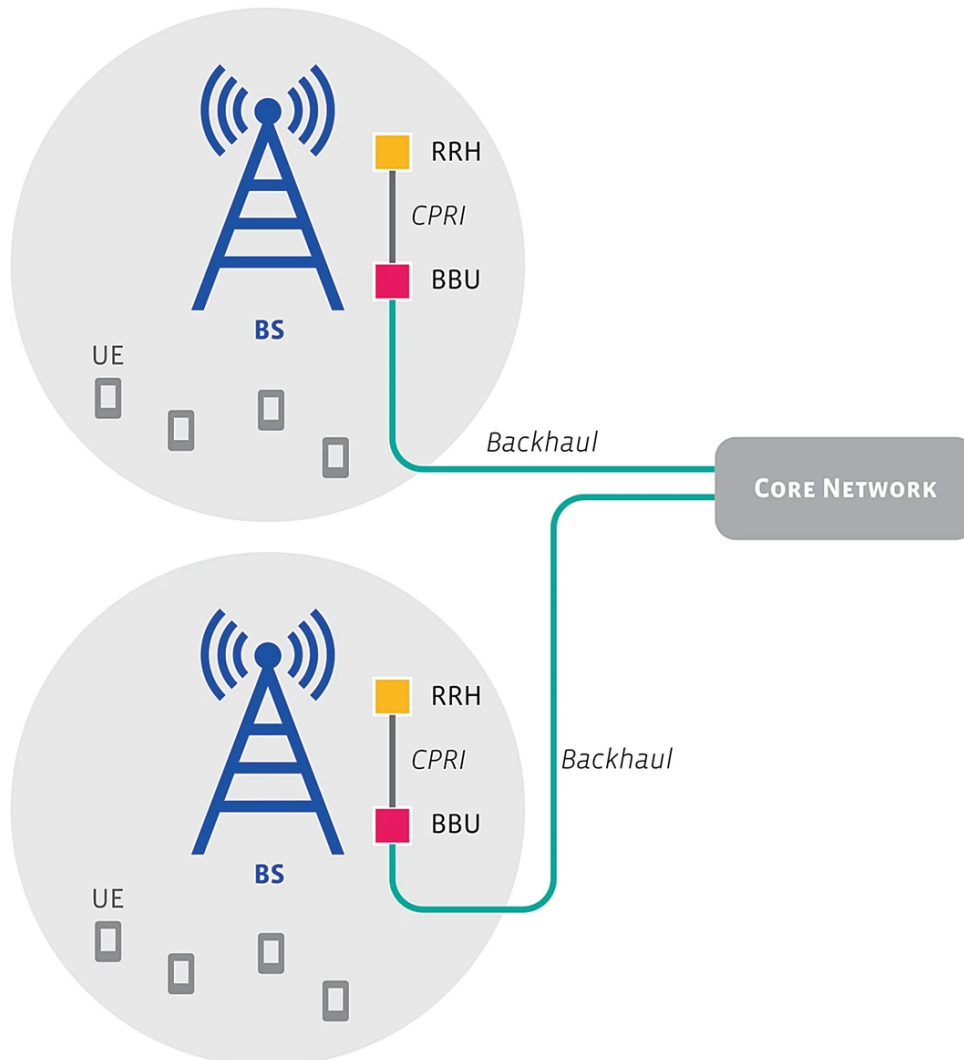


Figure 1.3: basic D-RAN architecture (source: [7])

In contrast to D-RAN where each base station has its own BBU, allowing the BS to connect directly to the core network and provide a fully functional access point to its users, a C-RAN base station no longer has a local BBU and can hence only convert between electromagnetic waves and digitally sampled waves but has no way to handle real internet traffic. Instead, the BBU handling the conversion between digitally sampled waves and IP packages is virtualized. The RRH at the BS now is connected to a cloud infrastructure of virtualized BBUs running on a general-purpose processor platform. The CPRI that previously connected the RRH and the BBU is replaced by an ethernet based fronthaul. The choice of ethernet is based on the maturity of the technology [6]. See

figure 1.4

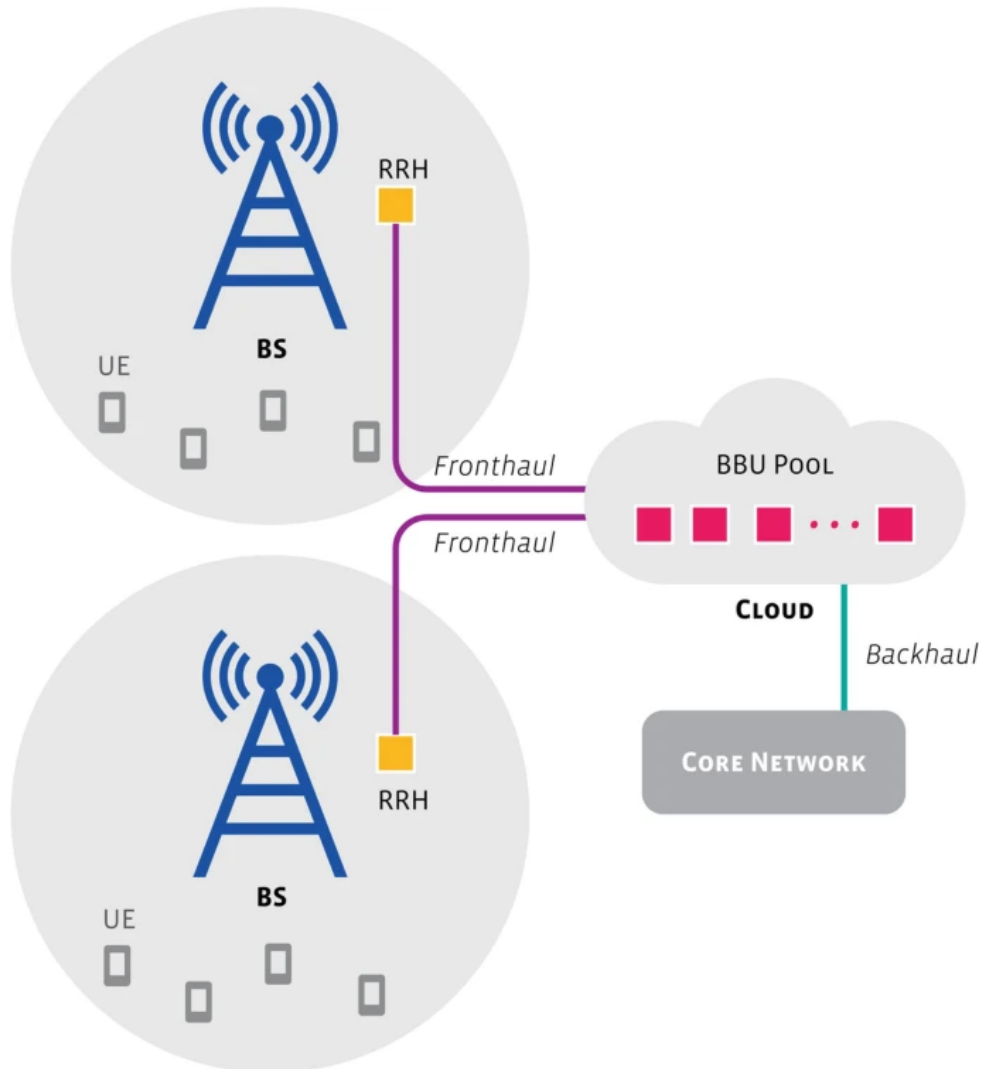


Figure 1.4: basic C-RAN architecture (source: [7])

1.4.2 H-CRAN for heterogeneity

The problem of sec:geocomplex can be solved with Heterogeneous Networks (HetNets)[7]. As the name suggests, HetNets use heterogeneous base stations. The traditional macro cell network is supplemented with a variety of small cells transmitting low power signals within their subarea. These small cells are heterogeneous in the sense that they exist in different classifications: Femto cells, Pico cells, and Micro cells. They differ in their radius, capacity, and output power. Femto cells are also incompatible with an integrated

Distributed Antenna System (DAS); a distributed MIMO system that seeks to prevent route loss and shadowing by providing spatial diversity: See table 1.5.

	Femto cells	Pico cells	Micro cells	Macro cells
Output power	1–250 mW	250 mW–1 W	1–10 W	10–50+ W
cell radius	10–100 m	100–200 m	0.2–2 km	8–30 km
Users	1–30	30–100	100–2000	2000+
DAS integration	No	Yes	Yes	Yes

Figure 1.5: Cell specification (source: [19])

While the macro cell base stations, or High Power Nodes (HPN), are still responsible for wide area coverage, the small cell base stations (SBS) are equipped with a low power node (LPN) and are placed more densely and with respect to demand. They provide high bandwidth to hot spots. This results in Heterogeneous Cloud Radio Access Networks (H-CRAN), a backward compatible design with seamless coverage that can adapt to the network's demand: See figure 1.6

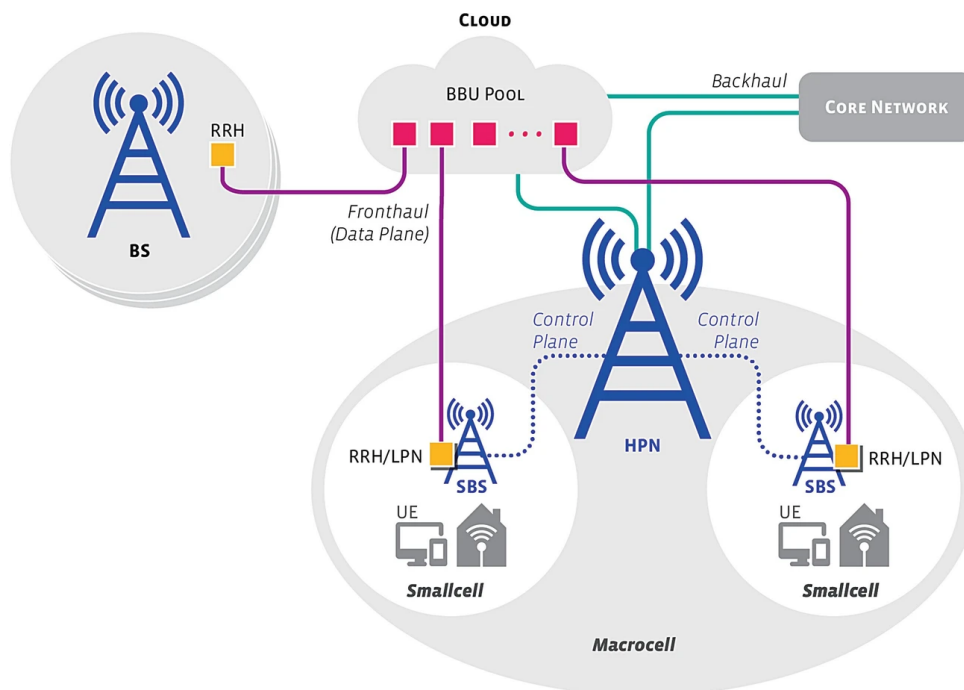


Figure 1.6: basic H-CRAN architecture (source: [7])

1.4.3 F-RAN for Low Latency and Better Load Balancing

Addressing the problem of sec:throughput, Fog Computing Radio Access Networks (F-RAN)[7] use fog computing (also known as edge computing) technology on top of an H-CRAN to release pressure from the front-haul. F-RAN only introduces edge-computing-technology adaption of two components of the H-CRAN architecture: The Access Point and the User Equipment both require the capability of radio resource management (RRM), caching, and cooperative signal processing. This results in the LPN turning into a fog computing-based access point (F-AP) and the user equipment turning into fog user equipment (F-UE): See figure 1.7

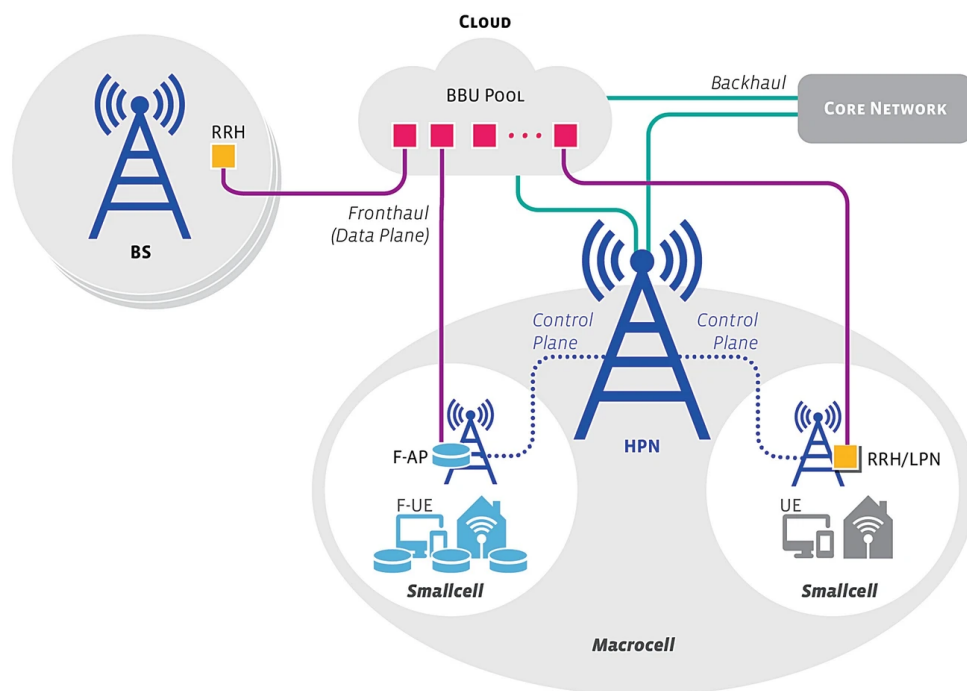


Figure 1.7: basic F-RAN architecture (source: [7])

This allows a significant amount of computational effort to be resolved within the small cell itself, such that the data does not even have to travel through the front-haul. This does not only mitigate the problem of the front-haul being a potential bottleneck, it also reduces the average latency of regular mobile traffic.

1.4.4 O-RAN for a future-proof design

An interesting expansion to vRAN is the concept of an open RAN (O-RAN). O-RAN envisions to establish industry-wide standards for all interfaces between parts of a C-RAN system [13]. This further adds to the flexibility of a RAN system, giving the network operator the freedom to choose from the whole available market for each module. Not only does this inter-compatibility create vendor-independence for a network operator, but also the guaranteed possibility to upgrade only a small part of the network without having to substitute the whole system.

With such an overarching standardization and well-defined interfaces, an O-RAN network system would also develop the reliable structure needed for large-scale applications built on top of the network, such as intelligent optimization algorithms to further improve device collaboration or to maximize the efficiency of resource usage. 1.8

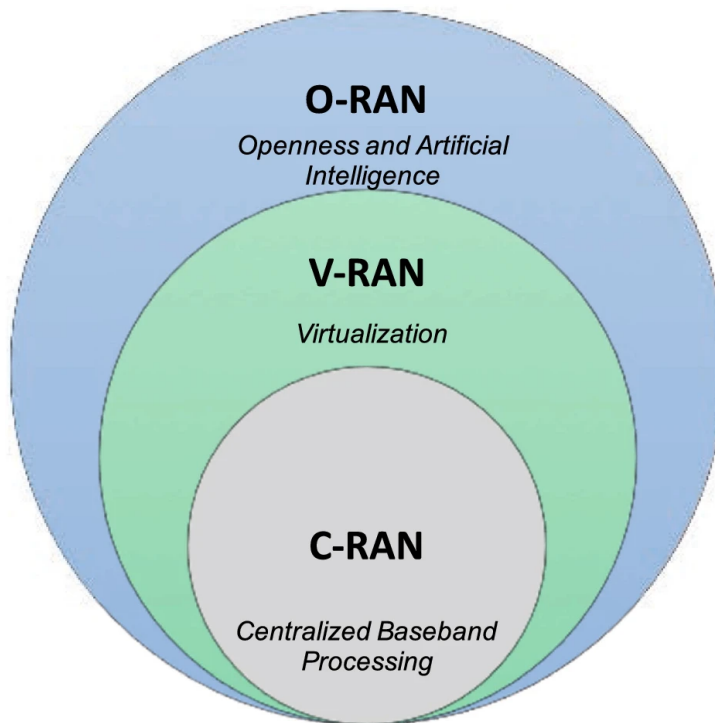


Figure 1.8: O-RAN on top of V-RAN (source: [13])

1.5 Discussion

Cloudification. Moving from a distributed and tailored infrastructure towards a centralized and dynamic one allows a more cost- and energy efficient use of resources. The use

of general-purpose processors instead of proprietary hardware also enables quick adaptation of technological advancements from the software area. Thanks to refined network designs and smart algorithms a cloudified radio access network can even outperform the conventional setup in many cases. *Virtualization.* This concept is by no means new. In computer science, operating systems in particular, virtualization, abstraction, and encapsulation have been used as a guide to simplistic and robust system design for a long time. Undoubtedly there is always some performance overhead when adding new layers to a system, but with modern optimization algorithms it should be possible to keep this cost at a minimum, allowing developers to fully flesh out the newly found potential of a virtualized system. Seeing how the discussed concepts have already proven their power and revolutionized the software world in the past, it is very well possible that vRAN will unlock new possibilities and act as a catalyzer for the next generation of mobile network applications.

1.6 Summary

vRAN helps creating more dynamic and efficient radio access networks. Virtualization and cloudification are the two core concepts it incorporates. The architectural changes leading from a distributed RAN to a centralized RAN are substantial and come with several challenges in terms of throughput and latency that need to be addressed in a market-ready vRAN. Ultimately, with vRAN as a basis, RAN interfaces can now be standardized corresponding to the vision of O-RAN.

Bibliography

- [1] Guo, B., Cao, W., Tao, A. & Samardzija, D. LTE/LTE-A signal compression on the CPRI interface. *Bell Labs Technical Journal*. **18**, 117-133 (2013)
- [2] Ghosh, A., Maeder, A., Baker, M. & Chandramouli, D. 5G Evolution: A View on 5G Cellular Technology Beyond 3GPP Release 15. *IEEE Access*. **7** pp. 127639-127651 (2019)
- [3] Nikaein, N., Schiller, E., Favraud, R., Katsalis, K., Stavropoulos, D., Alyafawi, I., Zhao, Z., Braun, T. & Korakis, T. Network Store: Exploring Slicing in Future 5G Networks. *Proceedings Of The 10th International Workshop On Mobility In The Evolving Internet Architecture*. pp. 8-13 (2015), <https://doi.org/10.1145/2795381.2795390>
- [4] Liu, L., Yang, F., Wang, R., Shi, Z., Stidwell, A. & Gu, D. Analysis of handover performance improvement in cloud-RAN architecture. *7th International Conference On Communications And Networking In China*. pp. 850-855 (2012)
- [5] Sousa, A., Melo, D. & Monteiro, P. A worst case analysis of C-RAN fronthaul coverage length with Ethernet based technologies. *2017 19th International Conference On Transparent Optical Networks (ICTON)*. pp. 1-4 (2017)
- [6] Pliatsios, D., Sarigiannidis, P., Goudos, S. & Karagiannidis, G. Realizing 5G vision through Cloud RAN: technologies, challenges, and trends. *EURASIP Journal On Wireless Communications And Networking*. **2018** (2018,5), <https://doi.org/10.1186/s13638-018-1142-1>
- [7] Gonçalves, G., Santos, G., Ferreira, L., S. Rocha, Souza, L., Moreira, A., Kelner, J. & Sadok, D. Flying to the Clouds: The Evolution of the 5G Radio Access Networks. *The Cloud-to-Thing Continuum*. pp. 41-60 (2020), https://doi.org/10.1007/978-3-030-41110-7_3 Akhtar, T., Tselios, C. & Politis, I. *Radioresource management : approaches and implementations from 4G to 5G and beyond*. *Wireless Networks*. **27**, 693–734 (2020, 11), <https://doi.org/10.1007/s11276-020-02479-w>
- [8] Bartelt, J., Vucic, N., Camps-Mur, D., Garcia-Villegas, E., Demirkol, I., Fehske, A., Grieger, M., Tzanakaki, A., Gutierrez, J., Grass, E., Lyberopoulos, G. & Fettweis, G. 5G transport network requirements for the next generation fronthaul interface. *EURASIP Journal On Wireless Communications And Networking*. **2017** (2017,5), <https://doi.org/10.1186/s13638-017-0874-7>

- [9] Vakili, S. & Elbiaze, H. Latency Control of ICN Enabled 5G Networks. *Journal Of Network And Systems Management*. **28**, 81-107 (2019,4), <https://doi.org/10.1007/s10922-019-09497-w>
- [10] Nikaein, N., Schiller, E., Favraud, R., Knopp, R., Alyafawi, I. & Braun, T. Towards a Cloud-Native Radio Access Network. *Studies In Big Data*. pp. 171-202 (2016,11), https://doi.org/10.1007/978-3-319-45145-9_8 Mesogiti, I., Lyberopoulos, G., Setaki, F., Giglio, A., Pelcelsi, A., Serra, L., Zou, J., Tzanakaki, A. *economicanalysehighlightingaspectsof5Gtransportnetworkdeployments.PhotonicNetworkComm* 268(2020, 10), <https://doi.org/10.1007/s11107-020-00912-w>
- [11] Sabella, D., Serrano, P., Stea, G., Viridis, A., Tinnirello, I., Giuliano, F., Garlisi, D., Vlacheas, P., Demestichas, P., Foteinos, V., Bartzoudis, N. & Payar³, M. *Designingthe5Gnetworkinfrastructure : aflexibleandreconfigurablearchitecturebasedoncontextandcontentinformation.EURASIPJournal* [//doi.org/10.1186/s13638-018-1215-1](https://doi.org/10.1186/s13638-018-1215-1) Yang, C., Chen, Z., Xia, B. & Wang, J. *WhenICNmeetsC-RANforHetNets : anSDNapproach.IEEECommunicationsMagazine*. **53**, 118 – 125(2015)
- [12] Sousa, A., Melo, D. & Monteiro, P. A worst case analysis of C-RAN fronthaul coverage length with Ethernet based technologies. *2017 19th International Conference On Transparent Optical Networks (ICTON)*. pp. 1-4 (2017)
- [13] Gavrilovska, L., Rakovic, V. & Denkovski, D. From Cloud RAN to Open RAN. *Wireless Personal Communications*. **113**, 1523-1539 (2020,3), <https://doi.org/10.1007/s11277-020-07231-3>
- [14] Peng, M., Yan, S., Zhang, K. & Wang, C. Fog-computing-based radio access networks: issues and challenges. *IEEE Network*. **30**, 46-53 (2016)
- [15] Wang, K., Yang, K. & Magurawalage, C. Joint Energy Minimization and Resource Allocation in C-RAN with Mobile Cloud. *IEEE Transactions On Cloud Computing*. **6**, 760-770 (2018)
- [16] Jiang, W., Han, B., Habibi, M. & Schotten, H. The Road Towards 6G: A Comprehensive Survey. *IEEE Open Journal Of The Communications Society*. **2** pp. 334-366 (2021)
- [17] Song, C., Zhang, M., Zhan, Y., Wang, D., Guan, L., Liu, W., Zhang, L. & Xu, S. Hierarchical edge cloud enabling network slicing for 5G optical fronthaul. *Journal Of Optical Communications And Networking*. **11**, B60-B70 (2019)
- [18] Kazmi, S., Khan, L., Tran, N. & Hong, C. Network Slicing: The Concept. *Network Slicing For 5G And Beyond Networks*. pp. 13-24 (2019), https://doi.org/10.1007/978-3-030-16170-5_2 Mishra, A. *Fundamentals of Network Planning and Optimisation 2G/3G/4G : Evolution to 5G*

Chapter 2

A Technical Overview Of Blockchain Retroactive Public Funding Schemes

Dario Gagulic and Lynn Zumtaugwald

Public goods can not be efficiently priced due to being non-excludable and creators of public goods are often not rewarded for their important contribution to the community. Open source software being a public good suffers from the same circumstances and depends on high level of volunteerism of developers which implement it in their free time. Widespread concerns about the sustainability of this open source community have been expressed, and it has been shown that indeed contributions to and number of open source projects have decreased and that the vast majority of open source projects fails due to different reasons. In summary, incentives for open source projects are needed and one possibility is through public funding. Considering this, this work looks at the landscape of public funding platforms and their different implementation of public funding schemes. Parameters of funding schemes researched in this work go from proactive to retroactive funding, from linear to quadratic funding, from one time donations to recurring money transfer in form of subscriptions, from high to non-existing fees, from FIAT to cryptocurrency payment, blockchain and non-blockchain based and many more. All these funding schemes and their parameters show advantages and disadvantages when applied to specific use cases and that is why one can not conclude for an optimal funding scheme that fits everyone. The role of blockchains currently is money transfer without the need of an intermediary but may become important for transparent voting or creation of project tokens in the future. Despite having all these public funding platforms to support the open source community, it is still unclear if this is enough to help this community to become sustainable again.

Contents

[19]	2.1 Introduction	21
	2.2 Definitions And Related Work	22
	2.2.1 Definitions	22
	2.2.2 The Decline Of Open Source Contributions And Sustainability	23
	2.2.3 Failure Reasons Of Open Source Projects	24
	2.2.4 What Motivates Developers To Accept Monetary Rewards	25
	2.2.5 The Impacts Of Monetary Rewards On Open Source Projects	26
	2.2.6 The Different Parameters Of Funding	27
	2.3 Approaches	27
	2.4 Solutions	29
	2.4.1 Platforms In A Timeline	29
	2.4.2 Kickstarter	30
	2.4.3 Flattr	30
	2.4.4 Goteo	31
	2.4.5 Patreon	32
	2.4.6 Liberapay	32
	2.4.7 GitCoin	33
	2.4.8 BuyMeACoffee	34
	2.4.9 GitHubSponsors	34
	2.4.10 Optimism Retroactive Public Goods Funding	35
	2.4.11 Flatfeestack	36
	2.5 Evaluation And Discussion	36
	2.5.1 One Time Donation VS Recurring Money Transfer	37
	2.5.2 Proactive VS Retroactive	37
	2.5.3 Linear VS Quadratic Funding	38
	2.5.4 Experts VS Everyone	38
	2.5.5 The Role Of Blockchain, Fees And Transparency	39
	2.5.6 The Optimal Platform(s)	40
	2.6 Summary And Conclusion	41

2.1 Introduction

There are plenty of shared resources among the world that one can use for free. Such resources, once created or served by nature, individuals, or a group of individuals and then shared to the open community for free use, are often referred to as public goods. As examples, one can name research published by universities researchers, the rule of law, national defense, access to clean air and drinking water, knowledge sharing, open-source software, and many more. Basically, it is every shared resource that individuals can or do receive a benefit from, which is greater than their individual contribution to this resource. Public goods are extremely important for human civilization since they benefit everyone in society and build the core of human flourishing [1]. Imagine Alexander Fleming would not have shared his discovery of Penicillin in 1928 [2] with human civilization, would we still live in a world without antibiotics treatments where minor bacterial infections would cost hundreds of thousands of human lives every year?

The public goods problem is well known in economics and refers to the fact that these goods can not be efficiently priced due to the impossibility to exclude individuals from access. Often, creators of public goods are the ones getting the least benefit from it and are not rewarded according to the value they created for society [1].

In the 21st century, public goods more and more took the form of digital content. From content and knowledge shared on free use internet platforms to open source software and tools, everyone can download and use for free. Recent studies have shown, that the open-source community is contributing to a high reduction of costs to build and deploy software [3]. And a survey done with 1313 companies showed that 65% of these companies rely on open source to speed up application development [4].

While the rise of version controlling and code sharing platforms like GitHub, GitLab, and Bitbucket created a renaissance of the open-source movement, recent concerns about the sustainability of the open source movement were expressed in the community. The concern is the question of whether open source innovation can be sustained if developers are not making private gains due to a high level of volunteerism [5]. While some developers may be ideologically motivated and code open source for their joy, some may wish to earn a reasonable livelihood from their contribution [6]. Further, recent studies have shown a plateau in contribution to open source in 2016 while growing exponentially until 2013 [8] and expressed the need and presence of external rewards [9][10].

A possibility to reward humans for their contribution to the open-source community may be public funding schemes. These schemes can be and are realized in many different forms and shapes in the digital world. Therefore, this report explores and discusses in a first part the landscape of public funding platforms, their specifications and functionalities. Then a discussion of the funding parameters is introduced, showing their advantages and disadvantages, and their fit for specific use cases. Also, the impact and usefulness of blockchain technologies in such platforms are exposed. Finally, a conclusion about the perfect public funding platform and its funding parameters is addressed.

2.2 Definitions And Related Work

In the following, basic definition that are important to understand the report, such as different funding types are provided in subsection 2.2.1. The decline of open source contributions and sustainability is evaluated in subsection 2.2.2, following by a deep look on failure reasons of open source projects in subsection 2.2.3. What motivated developers of open source projects to accept money is discussed in subsection 2.2.4 and research to the impacts of monetary rewards is shown in subsection 2.2.5. Then an overview of different parameters of funding is shown in subsection 2.2.6.

2.2.1 Definitions

This subsection introduces the notion of public funding schemes, different points in time when funding can happen, the different types of funding and the special case of quadratic funding.

A **public funding scheme** is a scheme to fund people, projects, organizations, or activities. Typically, the fund comes from the public, for example, the government or a community of humans.

Proactive funding happens ahead of time, typically a project from which the sponsor believes will be successful in the future.

Retroactive funding happens behind the time, typically projects that have already been shown to be successful in the past get funded.

A **donation** is a voluntary transfer of property (often money) from the transferor (donor) to a transferee (donee). Typically, this is a one-time transfer of money.

A **subscription** is a voluntary transfer of money which occurs repetitively (*e.g.* weekly, monthly, yearly). In exchange, the donator may get value in different forms, for example, online content, or it may also be with no value in exchange.

Sponsoring is supporting a person, organization, or activity by giving money, encouragement, or other help.

Crowdsourcing is the practice of obtaining information or input into a task or project by enlisting the services of a large number of people, either paid or unpaid, typically via the internet.

Quadratic voting/funding describes an optimized voting scheme with a nice property, that the amount of influence you purchase is proportional to how much you really care. This effect prevents the issue of *buying influence*. In contrast to a linear voting/funding scheme, in a quadratic funding scheme the total matched amount is calculated using a quadratic formula of each vote's square root summed together (see figure 2.1). For instance, if project A gets a total funding of 100 from exactly one contributor and project B gets the same total funding of 100 from ten different contributors, then linear funding

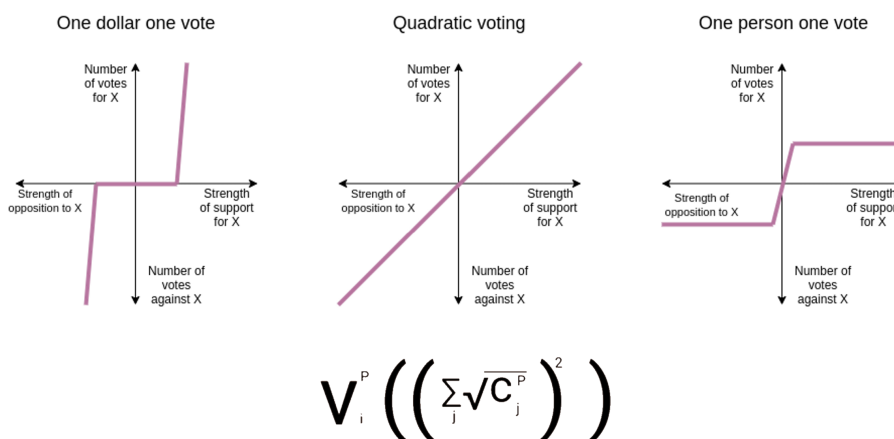


Figure 2.1: Illustration Of Linear Proportionality In Quadratic Funding "Strength Of Support For A Project X" vs "Number Of Votes For Project X" And Formula For Quadratic Funding.

would distribute the total matching pool equally among the two projects. Differently, quadratic funding would weight the total funding of project B higher because the funding came from more distinct contributors.

2.2.2 The Decline Of Open Source Contributions And Sustainability

The open source literature measures the growth and sustainability of open source software projects for many years. The first study, conducted in 2003 showed that open source grows with respect to byte size [11] by analyzing 406 software projects. A later study in 2007 claims that open source grows quadratic with respect to lines of code [12] by analyzing 4047 projects and a third study shows that in 2008 open source grows exponentially with respect to lines of code and also grows exponentially with respect to projects by analyzing 5122 projects [13].

A more recent study conducted in 2020 took these measurements of growth defined in the previous studies and investigated if these results are still correct or did change in one way or the other. They analyzed 224'844 open source projects in the main growth attributes: Lines of code, Commits, life cycle state, and contributors. They could show that in terms of lines of code open source software did grow exponentially through 2010, but the growth has continuously slower down since 2011. Similarly, the number of commits grew exponentially until March 2010, peaking in 1.5 Mio. commits, but declined ever since. In 2018, between 600'000 and 800'000 commits have been made to open source projects, which is a clear reduction compared to 2010. Also, the mean and median of monthly commits per project is decreasing [8]. Figure 2.2 shows open source projects life cycle phases over time. When all projects are included, also inactive and abandoned ones, the exponential curve fits until 2011. More interesting is the fact, that the number of abandoned projects relative to active projects continues to grow massively and overall

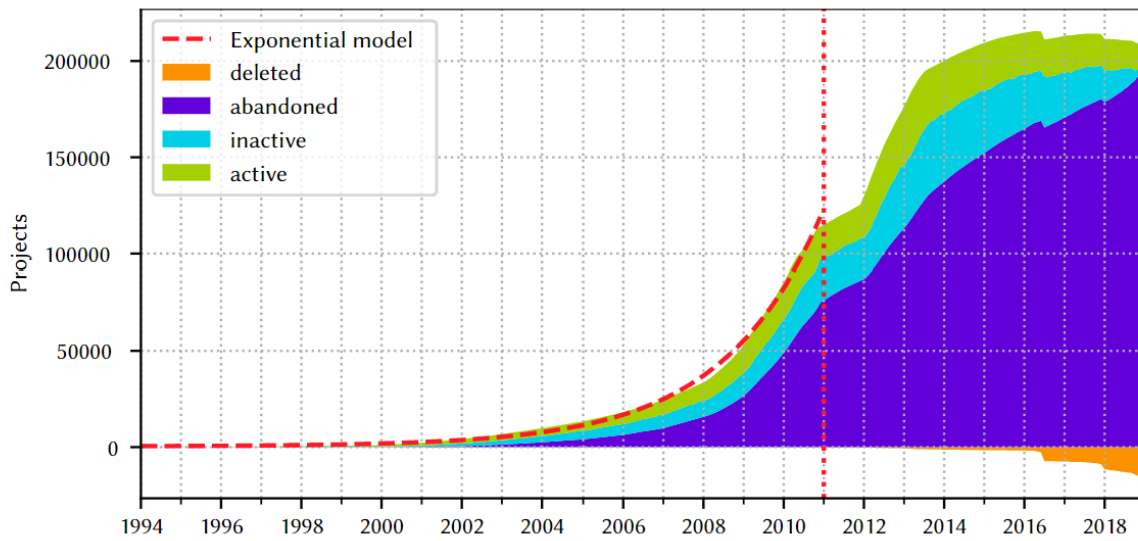


Figure 2.2: Available Open Source Projects By Life Cycle Phase, Where The Time for Implicit State Changes Is One Month [8].

the vast majority of open source projects are abandoned [8]. The next sub-chapter looks at the reasons why modern software projects fail.

2.2.3 Failure Reasons Of Open Source Projects

A study conducted in 2016 looked at 618 projects on GitHub to find reasons why modern software projects fail. Failed projects in their sense mean projects that had no contribution in the last year or have been clearly marked as abandoned or failed in their README. Surveying the developers of these projects showed the following reasons for failure (figure 2.3): One major reason is that the project has been usurped by competitors, developers

Reasons	Group	Projects	
Usurped by competitor	Environment	27	■
Obsolete	Project	20	■
Lack of time	Team	18	■
Lack of interest	Team	18	■
Outdated technologies	Project	14	■
Low maintainability	Project	7	■
Conflicts among developers	Team	3	
Legal problems	Environment	2	
Acquisition	Environment	1	

Figure 2.3: Why Open Source Projects Fail [3].

reported that there is strong competition in open source software, and often, a better

solution pops up that makes the own project not being used anymore. Most of the time these competitors are big tech companies, which pick up on an idea. Further, projects failed because they have become obsolete. Reasons for that are that there have been technical advances or shifts that make certain technologies useless or not applicable.

Lack of time and lack of interest make the other two major reasons why projects fail. Developers answered that they just do not find the time to maintain the projects anymore, because they had to go back to work or that they lost interest in the project because they found interest in other projects. Another important reason for failure is that the projects maintainability has become to complex. The study did not research a lack of funding as a direct source of failure for open source software projects, but 69% reported that they did not receive any rewards to maintain the project and thus acted completely voluntary. It also shows that developers work on open source software in their free time and that this is not their source of income [3].

2.2.4 What Motivates Developers To Accept Monetary Rewards

Funding developers to contribute to open source is great, but one questions is whether they actually accept money, and what influences their willingness to accept monetary rewards for their work. Financial incentives in the open source can be of many types. Developers might be paid by a non-profit organization by monthly salaries or by a profit corporation by bounty to solve a particular task [14, 15]. Another possibility is to be paid through voluntary contributions [16]. Interestingly, not all developers in past accepted these voluntary contributions [7]. Figure 2.4 shows the influence of different parameters research on the intention of developers to accept monetary rewards. Considering the independent

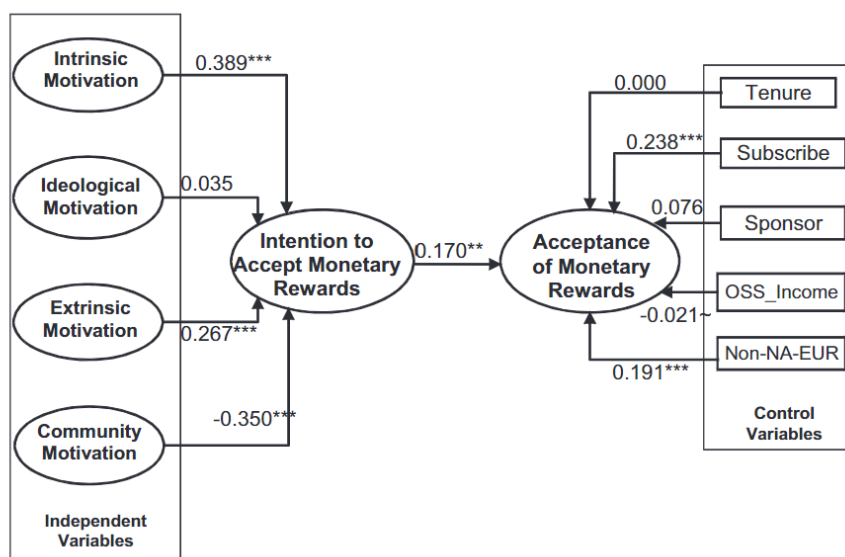


Figure 2.4: What Influences Developers To Accept Monetary Rewards [7].

variables, counter-intuitively, the intrinsic motivation of developers has the highest positive impact on the intention to accept monetary rewards with 0.389. It is argued that since

developers voluntarily choose to accept their rewards, they will not perceive such rewards as controlling, rather as supportive. That is why "crowding in" is occurring. Further, extrinsic motivation has also a high positive impact with 0.267. Ideological motivation does not have a reasonable impact on the intention to accept monetary rewards. The only researched variable that has a negative impact is the community motivation of the developers. With 0.35 it can be considered on having a significant impact. A reason for that can be that accepting monetary rewards can be viewed as giving more importance to individual gain than benefiting the community. Other developers in the community might see accepting monetary rewards as selfish. This fact could cause financial instability in the open source community. Still, it is unclear how many developers share which motivations [7]. Other studies researched the impacts of monetary rewards on open source projects further.

2.2.5 The Impacts Of Monetary Rewards On Open Source Projects

Open-source software is ubiquitous, but sustaining it is a challenge. No matter if the software is open- or closed source, it requires continuous effort such as fixing bugs and vulnerabilities or adapting and evolving technical and non-technical environments and requirements in order to remain relevant. It seems difficult for an external observer to understand why developers are willing to invest their valuable time in continuously maintaining such open source software, long after the software is finalized, and other parties start using it too, without any form of monetary reward. Of course, there are other factors (*e.g.* intrinsic motivation) that may influence the willingness to contribute to a project. However, in this section, we want to compare what related work found out about the extent, to which monetary rewards impact firstly the behavior of the open source software developers and secondly the quality of open source projects. Even though donations seem to gain traction to support open source activities, little is known about their prevalence, success, and impact [53]. Nakasai et al. measured the impact of donation badges on the contributors as well as the organizers. With the following reporter metrics, they measured the impacts on the contributors [19]:

- Number of commits, which captures both coding and noncoding activities
- Issue resolution speed, which captures the efficiency of maintenance and community support activities

Their findings show that the appearance of donation badges is appealing for both contributors and organizers. Further, Overny et al. identified open source projects requesting donations and analyzed observable characteristics of projects receiving donations. By conducting a time-series analysis of donations' effects on project activity, they found some evidence, but not strong support, for the hypothesis that donations lead to higher levels of development or maintenance activity. However, a clear decline in activity can be recognized for projects which were asking for donations but did not receive any [53]. Other open source software literature documented clearly the need and presence of external rewards [9, 20]. It has been argued that intrinsic motivation (joy, pride of contributing to a better world) are important factors, but many developers may "simply wish to earn a reasonable livelihood from their efforts" [6].

2.2.6 The Different Parameters Of Funding

Regardless of all available funding schemes having the same interest to support open source software development, there are multiple parameters in which certain platforms are differentiating themselves from the rest. Probably the most obvious parameter is the selection criteria of the projects to be considered. While some platforms allow the donator to freely choose who they want to support, other funding schemes focus only on projects with a major contribution in a specific field. Depending on the number of projects available, some funding schemes apply filtering mechanisms by experts in order to preselect high-quality projects. Another important parameter is the timing of the financial support in relation to the project's life-cycle.

It takes capital to make anything happen in the business world. Therefore, proactive funding schemes can support projects in advance, provide financial resources and ensure that the plans get moving. Conversely, retroactive funding focuses on projects which have proven to be more or less successful, and the outcome of the project is already known. An important fact is, that it is easier to judge quality retroactively than ahead of time [48]. While in some funding schemes individuals contribute to projects by themselves through donations or subscriptions, other funding schemes divide a pot of available money into several projects through the use of votes. Hereby, one can differentiate between a linear and a quadratic voting design. In a quadratic scheme, the receiver of the vote or fund gets the quadratic amount of that funded amount. This benefits projects which are supported by a broader range of communities and help to identify high-quality projects.

Also, Buterin mentions the importance of transparency and a potential conflict of interest rules by the badge holders [48]. One could indirectly benefit by supporting nonprofit organizations with a vote, while at the same time the badge holder is also part of the nonprofit organization. Not only direct self-beneficial behavior can be detected by making the voting publicly available, but also benefiting personal relations in a close community can be identified. On the other side, he also mentions downsides of transparent voting, such as vulnerability to bribery or supporting projects in part with the subconscious motivation of winning favor with them. The technology used by the funding scheme is another important parameter. For instance, payments based on blockchain technology could result in a more transparent setting.

2.3 Approaches

This paper is mainly generated by combining and transferring knowledge that already existed in various sources, classical literature research. Additionally, a collection of different public funding platforms has been gathered and ordered. It includes funding platforms that are still active today, because they are clearly more important to the future of the open-source community. A further requirement is that they in some way or another enhance and support the open-source community, since this is the main purpose of this report. Further, to compare the funding platforms, the inclusion of platforms with different metrics has been chosen. These different metrics are their technical underlying

infrastructure (blockchain-based, non-blockchain), Pay-in and/or pay-out in FIAT money or crypto, one-time funding/recurring funding, possibility to add extra content for subscribers on the platform/no extra content, proactive/retroactive funding scheme and time of founding has been kept as diverse as possible.

The following platforms have been chosen:Kickstarter; Flattr; Goteo; Patreon; Liberapay; GitCoin; BuyMeACoffee; GitHubSponsors; Optimism and Flatfeestack

The following parameters are discussed for each chosen platform. The first four parameters give an overview about the size and the founding year of a platform in order to classify the reach of the platforms. The main focus of the platform is chosen to evaluate how important the platform is to support open source software projects. The parameters six to eleven are chosen to introduce the offerings a platform has and to show how a sponsor or a creator can use and interact with the platform to provide a basic overview and also to support readers to find the most interesting platforms for themselves. Fees and points of payment are included to compare the platforms and their dealings with users. The parameters fourteen and fifteen are important to reason about how decision of money flow are made on the platforms. The last chosen parameter is the technical infrastructure of the platform to show and reason about different possible implementations, their advantages and disadvantages, and their differentiation's through the technical details.

1. Founding year
2. Activity status
3. Number of users
4. Number of projects/creators who received money, depending on the platform
5. Main focus of the funded projects (software, online content, arts etc.)
6. The offering of the platform
7. The functionality of becoming a sponsor
8. The functionality of becoming a creator
9. Who is eligible to become a creator
10. Types of funding possible (donations, subscriptions etc.)
11. Funding scheme (proactive/retroactive)
12. Fees
13. Point of payment and payment possibilities (FIAT money/cryptocurrencies)
14. Who decides to whom the money flows and how (linear/quadratic funding and voting)
15. Transparency of the platform
16. Underlying technical infrastructure (blockchain/non-blockchain) and technical specialties of the platform

2.4 Solutions

2.4.1 Platforms In A Timeline

The following figure 2.5 shows the chosen platforms in a timeline and their most important features. In following sub chapters, more details to each platform are provided. The platform Flatfeestack is not shown in this graph because it is not implemented fully yet, but still discussed in this report.

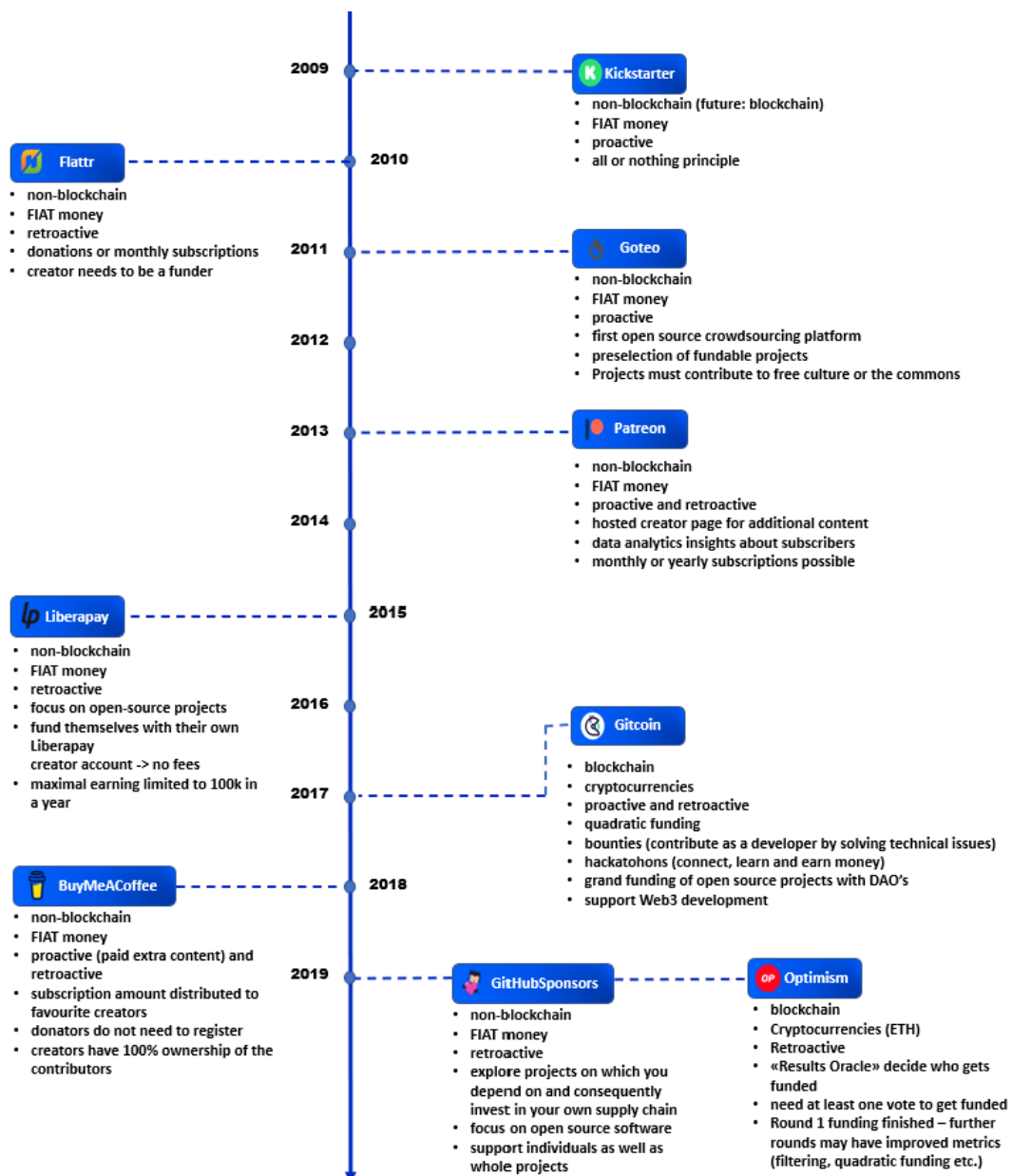


Figure 2.5: Chronological Self Illustration Of The Chosen And Already Implemented Funding Platforms

2.4.2 Kickstarter

Kickstarter is a platform founded in 2009 that has actually more than 200'000 creative projects [21]. It has pledged over 6 billion USD till today and supports one-time projects of all kinds. It is a crowdfunding platform that acts proactively. It follows an all-or-nothing principle, which means that the payment method will only be charged if the project reaches the funding goal before the campaign deadline. As a creator, everyone can list their project on Kickstarter. They need to subscribe and provide a description of the project. As a sponsor, one can fund projects they like through one-time donations. Kickstarter does not have special requirements for projects, they do not need to contribute to the open source community, but still can. Kickstarter can also be used to fund a company or a business idea. Money transfer on Kickstarter is done via FIAT money only and happens after a campaign deadline [22]. Despite Kickstarter being a non-blockchain-based platform, they recently expressed plans to develop an open source protocol that will essentially create a decentralized version of Kickstarter's core functionality. The public blockchain they choose to deploy to is called Celo. This protocol will be available for collaborators, independent contributors, and even Kickstarter's competitors from all over the world to build upon, to connect, and to use [21].

2.4.3 Flattr

Flattr was founded in 2010 and is still active today with over 100'000 active users and an estimated 30'000 creators funded [25]. The platform is implemented in a centralized manner and supports different creators like musicians, artists, writers, craftspeople, video makers, storytellers, teachers, photographers, cooks, guides, developers, and gurus [24]. It offers two different ways of funding: one-time donations to a single creator and monthly subscriptions, in which your donation gets split to your favorite creators [25]. Since creators only created the content before the fund, or are continuously creating content, this is considered a retroactive funding scheme. Everyone who wants can get on Flattr, as a creator, you must sign up and create a short description of your work, and then you are already ready to receive money [23].

One needs to note that for being a creator, you also need to be a sponsor, otherwise you will not receive money [27]. Also as a sponsor, you have to sign in and then choose the creator you want to pay money to. Further, you have to choose whether you want a one-time donation or a subscription. Hence, sponsors themselves decide who gets their money and it is a linear funding scheme[25]. The currency used on this platform is USD and the creators receive money always after a 30-day cycle to their bank account [26]. The fee structure is as follows: 90% goes to the creators, 5% is charged as a transaction fee, and 5% goes to Flattr itself [23]. Transparency-wise everyone can see the individual posts of creators and also the number of clicks but this does not apply to the revenue of the creators. A specialty of this platform is that you have to be a sponsor to be able to be a creator, as mentioned above. This, according to Flattr, encourages active participation [27].

2.4.4 Goteo

Goteo is a platform founded in 2011 and claims to be the first free/open source crowdfunding platform. Till 2021, Goteo raised 2'711'964 Euro for projects. Goteo received by then 1005 projects and published 252 of them [28]. This also means that not every project that applies gets listed on Goteo. A strong requirement for projects who want to raise money on Goteo is that they contribute to free culture or the commons in some way. With respect to the openness requirements, Goteo supports a broad diversity of projects. According to their statistics, the main categories of projects are social (20%), cultural (15%), education (15%), environmental (13%), technological (10%), entrepreneurship (9%), communication (8%) and scientific (7%) [29].

For a project to get listed on Goteo, several steps must be completed. First, one needs to sign up and create a detailed description of the project. Then the selection process by Goteo starts. In case of being accepted, a revision of the project and its description is performed. Then the project is published on Goteo and two rounds (40 + 40 days) of co-financing start. After these rounds, the distribution of rewards, communication of results, and shared materials is done. In case of rejection, the project does not get listed on Goteo. The Goteo Foundation charges a commission of 5% of the total funds collected over the two rounds for services provided. These include publication, use of management tools, project promotion, consultancy on all aspects of project presentation. Further commissions for financial transactions are charged that depend on the transaction type. For payments through PayPal, the commission is 3,4% + 0.35 Euro for each payment. Payments on Goteo only happen in FIAT money. As discussed, Goteo itself decides which projects get published on the platform. This selection is carried out by a Goteo team in close collaboration with a community of experts in a wide range of fields. Following parameters influence the decision of the Goteo team [30]:

- Aims
- Theme
- Background
- Pertinence
- Degree of innovation
- Estimate of collective return
- Abilities and experience of promoter/s

As technical specialties of the Goteo platform are first that it claims to be the first free/open source crowdfunding platform. Because its code is open source, it has been forked in Japan [28]. Second, its strong requirements for the openness of the projects, that need to contribute to free culture or the commons in some way. And third, it also enables participants to contribute as volunteers for the projects, and not just with money [29].

2.4.5 Patreon

Patreon is a platform founded in 2013 that in 2022 over 250'000 creators and 8 Mio. users. On Patreon, 100'000'000 CHF monthly funds and over 3.5 billion funds overall are raised. On Patreon, a lot of different people and projects are supported. It includes podcasters, video creators, musicians, visual artists, communities, authors and journalists, gaming creators, non-profit organizations, tutorials and guides, and creatives of all kinds [31]. As a creator, everyone can use Patreon. Creators can choose between three different subscriptions, Lite, Pro and Premium. Each of these has different additional functionalities like analytics of the Patreons (subscribers) and many more. It is free to get started, Patreon only receives a small amount percentage once a creator starts earning on Patreon. Plus fees for payout 5% + 10 cents for amounts under 3 Euro and 2.9% + 30 cents for amounts over 3 Euro [31]. The creator can either be payout monthly or can also trigger payout whenever they want. Pay-in and pay-out are done via FIAT money only, depending on which country the users live in. As a contributor, one also needs to sign up for a creator and choose a monthly or yearly subscription. Mostly, creators offer additional content for their subscribers. Technical specialties of Patreon are the hosted creator page: the creator page on Patreon is where creators can invite fans and provide additional content. Further, there are Patreon workshops where one can learn from the creator success team and the data analytics where creators can get more information on who their subscribers are [32].

2.4.6 Liberapay

Liberapay is a platform funded in France in the year of 2015. Until now, the platform counts more than 61'000 users in total, of which only last week (23.02.2022) 9'691 sponsors were donating a total amount of 13'722 CHF [34]. The sponsors themselves decide who they want to fund, and it is a linear funding scheme. Liberapay mainly focuses on open source project such as software, knowledge transfer and others [33]. A remarkable feature worth mentioning is, that Liberapay does not raise any fees by themselves, except the external payment service provider fees of 5% depending on the sponsors' payment method (Stripe and PayPal are supported), but rather they fund themselves with their own Liberapay creator account [36]. This demonstrates their main focus and willingness to support open source projects rather than acting profit oriented. Creators are paid Wednesdays in FIAT money, usually in their local currency (33 currencies available for payout). Currently, no cryptocurrency payouts are supported. When creators are signing up and want to create a profile, they have to explain why they are asking for donations and what they plan to do with the received money. After setting the preferred payment method, they can reach out to their target audience and contact people who will benefit from their work and ask them for their financial support.

Similarly, a contributor creates a profile and has the full control of their donations by setting the frequency (weekly, monthly or yearly) and selecting a recurrence (automatic or manual). With the automatic recurrence, the contributor can keep the donation running, otherwise he/she will be notified when a donation needs to be renewed. On Liberapay, payments are made in advance. The contributor is in control of how much money and

when they donate. Paying a larger amount once will result in a lower share of transaction fees [35]. Technical specialities of Liberapay are that they only support recurrent funding (subscriptions), no one time donations, with the goal of ensuring the creators a stable income [33]. Further, the maximal earning for each creator is limited to 100'000 CHF per year to dampen undue influence [36] and creators do not know who their sponsors are [35]. Liberapay is financed by donations only and transparently run by a non-profit organization. The source code is publicly available [33].

2.4.7 GitCoin

GitCoin was founded in 2017 and is still active today with 65'000 sponsors and 287'500 earners. It is a blockchain-based funding scheme which provided an exceptional amount of 53'500'000 USD of for open source software in over 1'700'000 complete transactions until today. They provide quadratic funding and follow the mission to build and fund the open web together [40]. They are supporting projects which create community and infrastructure for Web 3 - a diverse range of tools, technologies, and networks that enable people to work for the open internet [38]. It offers three different ways of funding:

- **Proactive:** earn through bounties as a hunter

Bounties are certain technological development challenges or issues which are crowd-sourced to a large number of developers, so-called hunters. Hunters who have time to contribute and increase speed of project development decide to earn money by solving these bounties. Posters can easily create an issue on GitCoin and let developers work on their project. Every issue requires a fund issue form to be filled out such that a contributor (hunter) can start working on it. The poster approves the hunter's request via the email notification. Once work has been completed, the hunter will submit a PR for review. [37].

- **Proactive:** earn through sponsored hackathons

By participating in hosted hackathons, developers have the opportunity to learn something and simultaneously earn money by solving real-world open source problems. They can collect rewards and tokens as they hack your way through creative solutions on a variety of protocols. Team work is asked as developers connect with other Web3 builders to advance the technology that will shape our future [37].

- **Retroactive:** earn through grant funding

GitCoin Grants makes funding for open source projects and public goods possible and allow for smaller donations from individuals to have an exponential impact through the matching fund. The Grants Rounds take place quarterly, with the opportunity to earn a portion of the generous matching pool of the Ethereum Foundation that has helped to launch thousands of projects. The community and sponsors have the possibility of funding pool categories of OS projects DAO and community (*e.g.* funding Ethereum public goods, support for Ukraine, climate change, UniSwap etc.) [37].

The site and the site services operate with Ether or ERC20 compatible tokens for the payment from posters to hunter. Payment to hunters will occur upon poster's acceptance of the hunter's work product or, in the case of a Grant, immediately upon a poster agreeing to contribute a Grant to the hunter and upon each payment cycle thereafter. Acceptance of a hunter's work product is at the sole discretion of the poster. GitCoin charges a 10% listing fee of the total task value (the "listing fee") payable to GitCoin. The listing fee is paid by the poster and due upon a poster submitting a task to the site [39].

2.4.8 BuyMeACoffee

BuyMeACoffee is a funding scheme created in 2018 with over 300'000 creators (state: April 2021). The platform is not blockchain-based and supports different creators like video creators, artists, writers, musicians, developers and gaming [41].

Everyone can register as a creator and make a profile with a very short description of their creation. They have the possibility to extend the profile with paid extra content. Once everything is set up, they can receive payments. Further, contributors have the possibility to buy a coffee (make a donation) either one time or monthly without having the need to register themselves. Contributors have access to the paid extra content of the supported creators [44]. Donors decide themselves who they want to support, and both one time donations and subscriptions (monthly, yearly) to specific creators are possible.

Some special technical aspects of BuyMeACoffee worth mentioning are that they do not force donators to register with their personal information, which makes donating much simpler and lowers the barrier. Another attractive feature is that creators have 100% ownership of the contributors, since BuyMeACoffee never emails contributors directly. Donations can stay anonymous or contribute by name. Transparency is provided by making numbers accessible on the website about how many supporters a creator has. Lastly, it claims to be a better alternative to Patreon, especially because of the lower fees (5% transaction, 95% to creators) [44, 45].

2.4.9 GitHubSponsors

GitHubSponsors is a platform created in 2019, which still is active. Anyone who wants to contribute to an open source project and lives in a supported region is eligible to become a sponsored developer. Contributions include but are not limited to bug reports, issue triage, code, documentation, leadership, business development, project management, mentorship, and design [51].

Everyone resided in supported regions can be a creator or a sponsor. Sponsors decide themselves who they want to sponsor. Remarkably, GitHubSponsors does not charge any fees for sponsorships from user accounts, so 100% of these sponsorships go to the sponsored developer or organization [51]. Creators can manage payments, *e.g.*, pay out in FIAT money depending on region.

One special idea about GitHubSponsors is that one can explore projects on which they depend on and consequently invests in their own supply chain [51]. GitHubSponsors is a retroactive funding scheme. Related work showed that the activity and popularity of a project on GitHub are positively correlated with the likelihood of receiving donations, and projects that received donations showed a short-term increase in the number of commits and a reduction in the time to resolve issues [53]. Another technical characteristic is, that one can sponsor a project or a single developer: While most OSS donation services in the past have targeted projects, GitHub Sponsors is unique in that it allows users to donate to individual OSS developers. As of now, GitHub matches contributions of up to 5k during a developer's first year in GitHub Sponsor [50].

2.4.10 Optimism Retroactive Public Goods Funding

Optimism is an Ethereum layer two blockchain company which started a blockchain based retroactive public good funding in the year 2019, where a "Results Oracle", which is the process and people deciding, commits to rewarding projects recognized to have delivered value. In round 1 a value of 1'000'000 USD was donated to 58 public goods projects [47], where the money was donated by Optimism itself.

Instead of DAO's where people select promising projects ahead of time and give them money, the Optimism retroactive public good funding gives money to open source project which already have shown to be valuable and have an impact in the past, implementing the retroactive funding scheme. A smaller but more skilled number of badge holders (decision-making board) can make decisions on which projects to support. They can decide to send money to a single individual or organization that was primarily responsible for making the project happen or to a smart contract representing a fixed allocation table, splitting funds between multiple individuals and/or organizations who had contributed time and/or funding to the project.

The team still wants experiment with the results oracle and change parameters of it in retrospective to every funding round. Like this, they want to ensure that the decision process on who gets funded gets better every round. In round one, they found several issues they want to improve for round to. For example, they want to diverse badge holders such that not only tech projects get funded and they want to implement a quadratic funding scheme such that money gets distributed less evenly.

Another idea of the Optimism team is that an open-source project could create and distribute project tokens to contributors of the open-source project, which then can be traded on a market in exchange to Ethereum or other coins. Creating the possibility for investors to invest in the project and, thereby, also allowing project builders to use this money to pay off development costs. This idea can close the cycle of retroactive and proactive funding and create a prediction market for what the results oracle will fund. If the results oracle decides to fund this project retroactive, they would just buy these coins and hence investors would receive money for their investment.

2.4.11 Flatfeestack

Flatfeestack is a flat-fee platform where donations are distributed based on different git metrics. Interestingly, the donation payment is done with credit cards while the payout happens with cryptocurrency. The contributor can decide to transfer a certain amount, which will be equally distributed based on one's selected repositories. The total amount

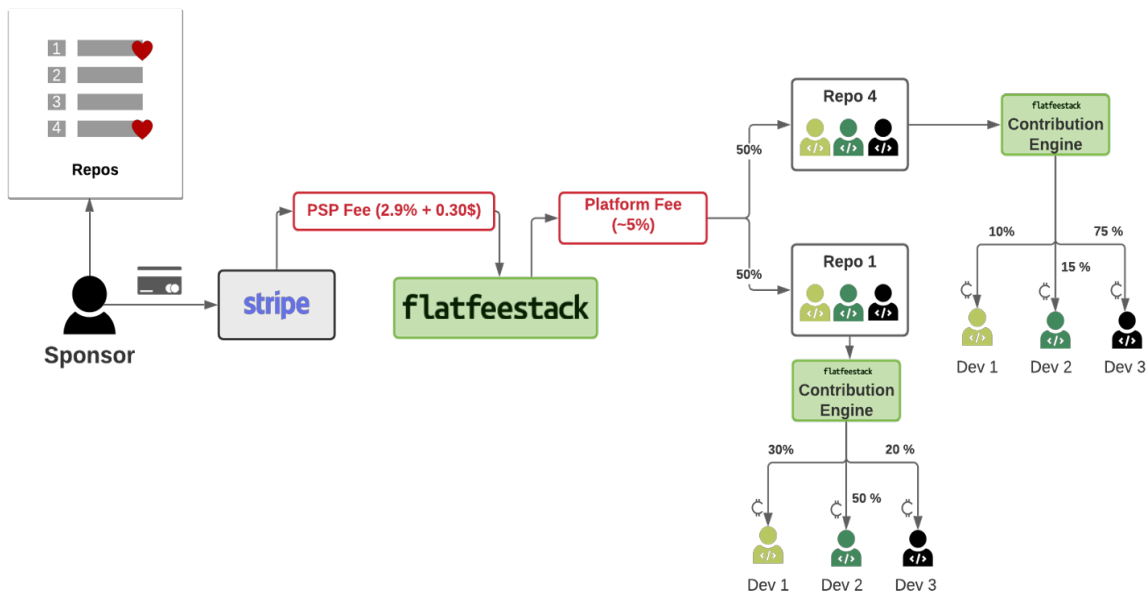


Figure 2.6: Sponsoring Flow of Flatfeestack With A Contribution Engine.

for each repository is distributed by a contribution engine, which computes the contribution relevance for each developer based on git-based data, using a special formula with weighting factors (cf. figure 2.6). This computation can be implemented on a smart contract. However, this approach faces some challenges and limitations:

- Who/how to determine weights?
- Who/how to determine code quality?

These questions are not so easy to answer, since every answer has their benefits and disadvantages. One possible suggestion for describing code quality could be fostering *clear, simple and bug-free code* rather than just looking at metrics such as *added lines of code* and so on.

2.5 Evaluation And Discussion

This section discusses the information showed before in this report. The parameters of platforms, their advantages and disadvantages for specific use cases and their importance

of creating a sustainable open source environment. What optimal platforms for specific use cases should look like, and what role blockchains already play and may play in the future. These aspects are discussed in respect to how they can motivate people to contribute to the open source community and thus stabilize the open source community.

2.5.1 One Time Donation VS Recurring Money Transfer

Both one time donations and recurring rewarding (for example in the form of salaries or subscriptions) can be useful. Most of the analyzed platform offer both types of funding in order to adapt to different use cases (Flattr, Patreon, Liberapay, BuyMeACoffee) [23, 31, 33, 41]. One time donations are especially advantageous in cases where a project needs a high amount of money in advance, where it needs the money to have the resources to start developing the project. For example, if a project need high technical, human or physical resources like hardware, buildings etc. to be able to develop in the first place. Further, one time donations are useful for retroactive funding a project that has been shown to be important. Another use case is also small donations, where everyone can tip the creators as a sign of thanking. For example, the platform BuyMeACoffee implements this thought, where you can donate small amounts like 5 USD to a creator [41].

On the opposite, recurring funding is more advantageous in use cases where the project development is still ongoing and will need maintenance over a period of time. For example, a developer that maintains a plugin a lot of other people depend on, could lose interest in doing that after receiving a one time donation but will keep interest if receiving recurrent funding as long as the work is performed by him. Further, recurring funding can act as a stable, previsible and predictable source of income. Hence, serve as a main and secure source of income for creators and thus motivate them to contribute full- or part-time to the open source community [14].

2.5.2 Proactive VS Retroactive

The timing of the funding plays a crucial role and affects many other correlated factors. As Buterin [48] mentioned, it is easier to decide on the success of projects retroactively, as the software is already built and has proven to be of high quality. Undeniably, not all projects that are believed to become successful and are proactively funded have a guarantee to make it. The *optimism retroactive public goods funding round 1* selected higher quality project compared to the *Bitcoin round 11 (tech only)* projects selection. Another great benefit of retroactive funding is the possibility of developers getting recognized and financially supported, even after their service/product has already been delivered. This psychologically motivates the developers to ship more high quality content, nevertheless of the financial reward depending on possible future donations. Some great examples for such retroactive donations are platforms like Flattr, Patreon or Liberapay. In the case of Patreon, the creator also has the possibility to host additional content, available for the contributor only by supporting the creator proactively [31]. This is a great trade-off, since retroactive funding also has a downside. The retroactive funding model may work well for maintaining an already-built codebase, but the beginning phase of kicking off a project

is still extraordinarily difficult [46]. The beginning of a project is often a slow-moving labor of love from a small group of highly dedicated individuals, where proactive funding definitely helps to build the financial foundation of a project.

2.5.3 Linear VS Quadratic Funding

The idea of quadratic funding has been gaining popularity rapidly over the last few years. As purposed by Buterin [48] it works well as a way to allocate the distribution of funds since it tackles the challenge of aggregating people's preferences. He phrases the problem as such: "Ultimately, we want a scheme where how much influence you 'buy' is proportional to how much you care" [49]. Gitcoin is a perfect example of having implemented the quadratic funding mechanism. Conversely, in a linear funding mechanism, each vote is equally weighted and considered to have the same constant cost. This can lead to individuals with a strong preference (or large budget) buying as many votes as they want in order to support their project.

This brings up the question about fairness on whether quadratic funding considered as always being better compared to linear funding. When considering the mathematically optimal way of funding public goods in a democratic community, one can clearly argue so, since it correctly amplifies the donations made by a large community over the contributions made by a small group with big pockets [49]. However, there exists one case where actions of a malicious actor in control can strategically contribute to a minority quickly burning down their votes. For instance, the malicious actor can repeatedly support projects where the majority weakly approves and the minority strongly disapproves. By adding its vote to a majorly weak supported project, one additional vote has a comparably high impact and compensates for the strongly disapproving minority because of the quadratic effect. In contrast to Gitcoin, BuyMeACoffee is a good example where linear funding is being implemented, by distributing the total subscription amount to one's favorite creators [44].

2.5.4 Experts VS Everyone

The discussion about who decides whether a project gets funded, clearly influences the outcome. In platforms like Flattr, Patreon, Liberapay, BuyMeACoffee GitCoin and Kickstarter, which we have previously discussed, every contributor has the chance to support their preferred project based on their preference. Contributors, using the GithubSponsors platform, may have an affinity of supporting projects they depend on (support your supply chain). However, these platforms have in common that everyone has the possibility to decide whether they want to support a project and consequently influence the outcome of the project's funding. The *Optimism's retroactive funding round 1* practiced a new kind of governance through badge holders. These badge holders consist of 22 known experts, which are entitled to make decisions on the board [48]. Therefore, the decision is non-public and depends strongly on the opinion of these badge holders. One benefit of this setting is that a smaller, but more skilled, number of experts can make better decisions

than "the crowd". This can lead to high quality projects getting recognized by the experts and supported with an increased probability.

Badge holders may take smarter decisions, but the crowd is more diverse and diversity plays a crucial role in funding [47]. Also, since the decision-making board consists of only a medium-sized group, taking influence on someone's preference may have a greater impact on the end result than manipulating a single vote in a public voting setting, where everyone is entitled to participate. To counter this incentive of influencing a badge holder's decision, the decision process could be made more transparent. For instance, badge holders could be required to explain their decisions, *e.g.* writing a post or a paragraph for each project they made votes on [48].

2.5.5 The Role Of Blockchain, Fees And Transparency

At the moment, only two of the analyzed nine platforms use blockchain technology (Bitcoin and Optimism), but there are other examples like Giveeth [54] that are not addressed in this report. The major use case of them is to perform transactions on the blockchain. This aspect may lower the transaction fees or also may increase them. Transaction fees on blockchains may be volatile, depending on the traffic (*e.g.*, the amount of people that want to perform a transaction at that time). This is because miners of a block can decide which transactions they write in the block and if there are many people wanting that space in that block at that time, then miners can rise the transaction fee or simply pick the candidates that offer the highest transaction fees. The fee is also dependent on the blockchain protocol itself. On the Bitcoin blockchain for example, a transaction costs in the mean between two and five USD. On the Ethereum blockchain, the transaction fees are about 35 USD. These transaction costs can also rise up to 50 USD in some cases. Both of these platforms currently use the proof-of-work algorithm. On proof-of-stake blockchains, transaction fees can be much lower [56]. For example, on the Cardano blockchain, a transaction costs about 0.4 USD, on the Solana Blockchain 0.00025 USD and on the Harmony ONE blockchain around 0.000001 USD [57] [58].

For platforms which mostly provide one time donations with big amounts of money transferred, *e.g.*, thousands or more, also a high fee blockchain can be applicable. In case of Optimism, where high amounts of money are donated to a project, it is not weighty to pay these 30 USD or even 50 USD transaction fee. In fact, this can be cheaper than using PayPal or other non-blockchain transaction systems, which demand a percentage of the transaction amount as a fee. For platforms like BuyMeACoffee for example, where a lot of users only donate 5 USD (a coffee), it becomes costly to use high fee blockchains, but low fee blockchains like Solana or Harmony One could still come in handy. Further, transactions on blockchains can be really fast or also take some minutes, depending on the use case of the platform, there can be advantages in using blockchains as a mean of transaction because the money is directly transferred to the creators, and they do not have to wait to the end of the month or some other period to receive their earned money. It can offer a way of automatizing money treason. If developers and creators want to receive their reward in cryptocurrencies needs yet to be researched and does not become clear out of actual research. One problem that may arise is the current volatility of such currencies.

Another aspect where blockchains could be useful in the public funding landscape is the voting process for projects. Blockchains can offer a fully transparent way to vote for projects to fund, where everyone can see who voted for which project(s). This full transparency however, comes with advantages and disadvantages. On one hand, the full transparency can be considered as fair, since everyone knows everything. This may discourage expert cycle of voters for voting for their own projects, or for projects that benefit their own projects, hence it could prevent misuse of their power. Further, one could easily detect expert voters that continue to vote for projects that do not show to be successful in future. Therefore, it could act as an assessment for the experts. On the other hand, this transparency could prevent experts for voting for risky or bold developments, since they do not want to be criticized for their decision afterwards. This behavior may hinder innovation. Another disadvantage of the full transparency can be that expert voters feel pressure to vote for projects of owners they know or of which they depend on in some way, and this could foster lobbying in the voting process. The optimism project wants to play around with these transparency parameters in future funding rounds but the results still remain to be seen [48]. Further, blockchains may support the creation and use of project tokens for open-source projects and thus, close the cycle between proactive and retroactive public funding as discussed in the Optimism part [48].

2.5.6 The Optimal Platform(s)

It is not straightforward to conclude the parameters for an optimal public funding platform for every use case. The metrics a platform should optimally have is dependent on the users and use cases of the platform. Nevertheless, there are some parts that are similar for all kinds of use cases. Users definitely benefit from usable interfaces of the platform and low barriers. Further, not having the need to register as a sponsor may also lower the barrier to fund projects or people, since less work is needed to do so. Additionally, users benefit from low or non-existing fees, since the total amount then goes to the creators of the content and not to the platform providers.

A good solution for the platform to finance itself is to create a funding account on a platform for itself, such that it can be donated by users and therefore implement the whole idea of public funding itself and also act as a good role model. Liberapay applies this idea of self funding through the own public funding platform very well and this can be considered as a strong advantage of Liberapay as a platform [33].

High-cost projects often need a big amount of money in advance to have the ability to be developed. They need monetary, physical and human resources, which need to be paid. Therefore, for this specific use case, a funding scheme using proactive funding is necessary and advantageous over retroactive funding. Often, these projects need more money at the beginning than during the further development, therefore a one-time donation can be advantageous. To this idea, the platform Kickstarter fits well [21]. Conversely, not all high cost projects need the biggest fund at the beginning. Some may also need continuously a constant amount of money. Projects with high human or electricity resource need, but low physical resource needs, for example. In these cases, a recurrent monetary donation may offer better conditions, which are supported by platforms such as BuyMeACoffee, Patreon

and Flattr. High cost projects may profit from platforms where an expert cycle decide which projects are funded, since only some projects get money, but typically a bigger amount. Here, the platform Goteo and Optimism offer advantages [28]. Nevertheless, also on platforms where every user decides which projects he/she wants to fund can raise big amounts of money, like Kickstarter. In terms of money type, both FIAT and cryptocurrencies can have benefits and drawbacks, as discussed above.

Low cost projects profit from platforms where everyone decides which projects are to be funded since more projects can get the money, it is better distributed and the project does not need to compete in selecting phases with other projects that much. It is also less time consuming and the application is easier for creators. Both retroactive and proactive funding is applicable, as well as FIAT money transfer and cryptocurrencies. The volatility problem of cryptocurrencies may be less weighty in this use case than with high cost projects. Platforms applicable for these types of use cases are BuyMeACoffee, Patreon, GitHubSponsors, GitCoin, Liberapay, Flattr and Kickstarter.

For funding projects, every discussed platforms is applicable. If the goal is to fund individual people however, the platforms Kickstarter, Flattr, Patreon, Liberapay, GitHubSponsors and BuyMeACoffee provide better characteristics. These platforms provide fast application processes and offer a choice of funding schemes.

Overall, the optimal platform would let space for all of these different parameters and funding schemes. One big platform where every type of funding is possible would lower the barrier for creators and sponsors, since it erases the need to register and use a different platform. There are some platforms that allow different use cases [50, 44, 37]. However, there is none that combines all of them. If an introduction of all different funding schemes on one single platform would create advantages or results in an overload and lowers the usability by confusing users, is still to be researched.

2.6 Summary And Conclusion

Public goods are extremely important for human civilization since they benefit everyone in society and build the core of human flourishing. The public goods problem is well known in economics and refers to the fact that these goods can not be efficiently priced due to the impossibility to exclude individuals from access [1]. Open source software is one of these public goods that have shown to be extremely important for the community by highly reducing the cost to build and deploy software [3]. Contributions to the open source community have been grown exponentially until 2013, reached a plateau in 2016 and have ever been decreasing since. Various resources expressed the need of external rewards in the open source community [9, 10] and researched the impacts of monetary rewards of the open source community.

Many platforms exist to raise money for open source projects. All of them implement different funding schemes and ideas. From proactive to retroactive funding, from linear to quadratic funding, from expert decision cycles to everyone deciding which projects to fund, from money transfers in FIAT money to various cryptocurrencies, from high fees to no

fees, from one time donations to recurring money transfer, from no transparency to high transparency. As discussed, all of these parameters have advantages and disadvantages in different use cases, and it is not straightforward to conclude on a single optimal public funding scheme. Retroactive funding is a fairly new idea that shows the advantage of rewarding projects and individuals that already have shown to be successful in the past. Vitalik Buterin, the founder of Ethereum, also expresses that it is easier to interpret the importance of already build projects than guessing if a project will have influence in the future. However, it is unclear if possible future rewards motivate developers enough to contribute to the open source community, since there may also be no rewards [48]. The role of blockchain is at the moment only the instant transfer of money [37, 46], which has the possibility to lower fees but also can be problematic due to volatility of cryptocurrencies. It is unclear if creators want to receive their rewards in cryptocurrencies or if FIAT money is preferred.

While studies showed positive influences of monetary rewards on the motivation of developers to contribute to the open source community [9, 20, 6] others were not able to find strong support but recognized a clear decline of project activity in projects requesting, but not receiving any financial support [53]. Clearly, the open source community is important for human civilization, and actions are needed to help this community to be sustainable. Despite all these different possibilities to reward open source projects and individuals contributing to this community, it is questionable if this alone is enough to help the open source community to again become sustainable. This however, and also how further actions to achieve this goal should look like, remains to be researched in the future.

Bibliography

- [1] Paul M. Romer: *Increasing Returns in Long-run Growth*; Journal of Political Economy, (94,5), 1986, 1002-1037.
- [2] S. Y. Tan and Y. Tatsumura: *Alexander Fleming (1881-1955): Discoverer of penicillin*; Singapore Med J, (56,7), 2015, 366-367.
- [3] J. Coelho and M. T. Valente: *Why Modern Open Source Projects Fail*; in Proceedings of the 2017 11th Joint Meeting on Foundations of Software Engineering, ser. ESEC/FSE 2017. New York, NY, USA: Association for Computing Machinery, 2017, 186-196. <https://bit.ly/30yaadz>
- [4] Stephanos Androutsellis-Theotokis, Diomidis Spinellis, Maria Kechagia, Georgios Gousios, et al. : *Open source software: A survey from 10,000 feet*. Foundation-sand Trends in Technology, Information and Operations Management 4, (3,4), 2011, 187-347.
- [5] E. Von Hippel and G. Von Krogh: *Open source software and the private collective model: issues for organization science*, Organization Science, (14,2), 2003, 209-223.
- [6] B. Fitzgerald: *The transformation of open source software*, MIS Quarterly, (30,3), 2006, 587-598.
- [7] S. Krishnamurthy, S. Ou and A. K. Tripathi: *Acceptance of monetary rewards in open source software development*, Research Policy, (43,4), 2014, 632-644. <http://dx.doi.org/10.1016/j.respol.2013.10.007>
- [8] D. M. M. Carparo and A. Barcomb: *Quo Vadis, Open Source? The Limits of Open Source Growth.*, arXivpreprint arXiv:2008.07753, 2020.
- [9] J. Lerner and J. Tirole: *Some simple economics of open source*, Journal of Industrial Economics, (50,2), 2002, 197-234.
- [10] C. Okoli, O. Wonseok: *Investigating recognition-based performance in an open content community: a social capital perspective*, Information and Management, (44,3), 2007, 240-252.
- [11] A. Capiluppi, P. Lafo and M. Morisio: *Characteristics of open source projects*, Sevents European Conference on Software Maintenance and Reengineering 2003 Proceedings, 2003, 317-327. <https://doi.org/10.1109/CSMR.2003.1192440>

- [12] S. Koch: *Software evolution in open source projects - a larger-scale investigation*, Journal of Software Maintenance and Evolution: Research and Practice, (19,6), 2007, 361-382. <https://doi.org/10.1002/smr.348>
- [13] A. Deshpande and D. Riehle: *The Total Growth of Open Source*, Open Source Development, Communities and Quality, (275), 2008, 197-209. https://doi.org/10.1007/978-0-387-09684-1_16
- [14] J. Roberts, I. H. Hann and S. Slaughter: *Understanding the motivations, participation, and performance of open source developers: a longitudinal study of the apache projects*, Management Science, (52, 7), 1998, 984-999.
- [15] S. Krishnamurthy, A. Tripathi: *Bounty programs in free/Libre/open source software (floss): an economic analysis.*, The Economics of Open Source Software Development, 2006.
- [16] S. Krishnamurthy, A. Tripathi: *Monetary donations to an open source software platform*, Research Policy, (38,2), 2009, 404-414.
- [17] Rodrigues, B., Eisenring, L., Scheid, E., Bocek, T., Stiller, B. (2019, April). Evaluating a blockchain-based cooperative defense. In 2019 IFIP/IEEE Symposium on Integrated Network and Service Management (IM) (pp. 533-538). IEEE.
- [18] Bruno Rodrigues, Eder John Scheid, Jonas Brunner, Calvin Falter, Guilherme Sperb Machado, Thomas Bocek, Burkhard Stiller: FlatFeeStack: a Blockchain-based Sustainable Public Funding of Open Source Projects; UZH, IFI-TecReport No. 2022.05, Zürich, Switzerland, June 2022, URL: <https://files.ifi.uzh.ch/CSG/staff/rodrigues/extern/publications/TR-FlatFeeStack.pdf>.
- [19] K. Nakasai, H. Hata and K. Matsumoto: *Are donation badges appealing?: A case study of developer responses to eclipse bug reports*, IEEE Software, 2018, 36(3), 22-27.
- [20] J. Lerner and J. Tirole: *Economic perspectives on open source in Intellectual Property and Entrepreneurship*, Emerald Group Publishing Limited, 2004.
- [21] Kickstarter Articles; <https://www.kickstarter.com/articles/the-future-of-crowdfunding-creative-projects?ref=section-protocol-promo-the-future-of-crowdfunding-creative-projects>, March, 2022.
- [22] Kickstarter Projects; <https://www.kickstarter.com/projects/3dprintmill/creality-cr-scan-lizard-capturing-fine-details-of-view?ref=b1nmn4&gclid=Cj0KCQiAmpyRBhC-ARIsABs2EArDnNqbqjd8DC1-2rTshchlEqj2TmRUo6z0cxyLn-mTEIyYKn-Y8AwcB>, March, 2022.
- [23] Flattr About; <https://flattr.com/about>, March, 2022.
- [24] Flattr Creators; <https://flattr.com/creators>, March, 2022.
- [25] Flattr; <https://flattr.com/>, March, 2022.

- [26] Flattr FAQ; <https://flattr.com/faq>, March, 2022.
- [27] F. Loll, C. Mumme and N. Pinkwart: *Flattr this! Explorative Evaluation von Social (Micro-)Payments als alternatives Bezahlmodell*, IfI Technical Report Series Clausthal University of Technolgoy, (10,10), 2010, 9-17.
- [28] Goteo About; <http://en.goteo.org/about>, March, 2022.
- [29] Goteo; <http://en.goteo.org>, March, 2022.
- [30] Goteo Sponsor; <http://en.goteo.org/faq/sponsor>, March, 2022.
- [31] Patreon; <https://www.patreon.com>, March, 2022.
- [32] Patreon Product; <https://www.patreon.com/product>, March, 2022.
- [33] Liberapay; <https://de.liberapay.com/>, March, 2022.
- [34] Liberapay Stats; <https://de.liberapay.com/about/stats>, March, 2022.
- [35] Liberapay About; <https://de.liberapay.com/about/>, March, 2022.
- [36] Liberapay FAQ; <https://de.liberapay.com/about/faq>, March, 2022.
- [37] GitCoin Earn; <https://gitcoin.co/earn>, March, 2022.
- [38] GitCoin Quickstart; <https://gitcoin.co/grants/quickstart>, March, 2022.
- [39] GitCoin Terms; <https://gitcoin.co/legal/terms>, March, 2022.
- [40] GitCoin Stats; <https://gitcoin.co/results#:~:text=Since%20its%20launch%20in%20November,an%20audience%20of%20287%2C601%20earners>, March, 2022.
- [41] Miller, T.: *This Startup Wants To Make It Much Easier For Creators To Get Paid*, Forbes, 2021, 10-06.
- [42] Rodrigues, Bruno, et al. "A blockchain-based architecture for collaborative DDoS mitigation with smart contracts." IFIP International Conference on Autonomous Infrastructure, Management and Security. Springer, Cham, 2017.
- [43] Rodrigues, Bruno, Thomas Bocek, and Burkhard Stiller. "Enabling a cooperative, multi-domain DDoS defense by a blockchain signaling system (BloSS)." Semantic Scholar (2017).
- [44] Buy Me A Coffee; <https://www.buymeacoffee.com/>, March, 2022.
- [45] Buy Me A Coffee - Patreon Alternative; <https://www.buymeacoffee.com/patreon-alternative>, March, 2022.
- [46] Retroactive Public Goods Funding; <https://medium.com/ethereum-optimism/retroactive-public-goods-funding-33c9b7d00f0c>, March, 2022.
- [47] Optimism About; <https://www.optimism.io/about>, March, 2022.

- [48] Review of Optimism retro funding round 1; <https://vitalik.ca/general/2021/11/16/retro1.html>, March, 2022.
- [49] Vitalik Buterin on quadratic funding; <https://vitalik.ca/general/2019/12/07/quadratic.html>, March, 2022.
- [50] T. Miller: *Exploring a New Way to Contribute to Open Source*, arXiv preprint, 2021, 10-06.
- [51] GitHubSponsors Sponsors; <https://github.com/sponsors>, March, 2022.
- [52] GitHubSponsors Setup; <https://docs.github.com/en/sponsors/receiving-sponsorships-through-github-sponsors/setting-up-github-sponsors-for-your-user-account>, March, 2022.
- [53] C. Overney, J. Meinicke, C. Kästner, and B. Vasilescu: *How to Not Get Rich: An Empirical Study of Donations in Open Source*, In Proceedings of ACM/IEEE 42nd International Conference on Software Engineering (ICSE '20), 2020, 1209-1221.
- [54] Giveeth; <https://giveth.io/>, March, 2022.
- [55] Bocek, Thomas, et al. "Blockchains everywhere-a use-case of blockchains in the pharma supply-chain." 2017 IFIP/IEEE symposium on integrated network and service management (IM). IEEE, 2017.
- [56] Blockchain Transaction Fees: Why Do They Matter; <https://learn.bybit.com/blockchain/blockchain-transaction-fees-explained>, March, 2022.
- [57] Cardano Transactions; <https://solberginvest.com/blog/cardano-fees>, March, 2022.
- [58] Harmony ONE Transactions; <https://docs.harmony.one/home/general/technology/transactions>, March, 2022.

Chapter 3

A Survey of Non-Fungible Token Use-Cases and Their Technical Standardization

Tumen Dambiev and Marek Gajewski

In this chapter we will discussing Non-Fungible tokens, which showed a substantial increase in value , indicating signs of a speculative bubble that could have an impact beyond their native ecosystem. We will present the technical aspects of standardization of such technology and the impact it will have.

Contents

3.1	Technical Keywords	50
3.2	What exactly is a Non-Fungible Token?	50
3.2.1	Size of the industry	50
3.2.2	Fungibility	50
3.2.3	Bitcoin and Ethereum	51
3.2.4	Origin of monetising digital goods	51
3.2.5	Bulk of the NFTs are procedurally generated	52
3.2.6	Network generated NFTs	52
3.3	Ethical problems with NFTs	53
3.3.1	Wild west of digital world	53
3.3.2	We have only one planet	54
3.4	Standardization of the technology	54
3.4.1	New technology, old challenges	54
3.4.2	Smart contracts	55
3.4.3	Protocols	56
3.5	Token Standards	57
3.5.1	Desirable traits of Non-Fungible Tokens	57
3.5.2	ERC-20	58
3.5.3	ERC-721	58
3.5.4	ERC-998	59
3.5.5	ERC-1155	59
3.6	Technical Security Concerns	60
3.7	Copyright implications of NFTs	61
3.8	NFT system design flaws	62
3.8.1	Concentration of Power	62
3.8.2	Phishing	62
3.8.3	Money Laundering	62
3.9	Potential of NFTs	63
3.9.1	Gaming Industry	63
3.9.2	Virtual Events	63
3.9.3	Protection of digital artworks	64
3.9.4	Metaverse	64
3.10	Legal Hurdles	64
3.11	The NFT communities	65
3.11.1	A wide range of motivators for NFT creation (Q1)	65
3.11.2	Bit by bit engagement with NFTs and the community (Q2)	67

3.11.3 Challenges on the way for NFT creators (Q3)	67
3.12 Conclusion	67

3.1 Technical Keywords

Blockchain: Blockchain were proposed alongside Bitcoin by Satoshi Nakamoto after the market crash of 2008. "Blockchain is defined as a distributed and attached-only database that maintains a list of data records linked and protected using cryptographic protocols" [9]. Blockchain is visible by every user and changes can be made to the blockchain only after the majority of nodes are in agreement. And once the information is stored on the blockchain, it becomes permanent.

Address: Each person has a unique identifier on the blockchain. It is similar to bank account, it is also used to send or receive cryptocurrencies. It is created using a public key and a private key. Users can only interact with their own address on the blockchain (send and receive funds). And users cannot even decline the funds going their way.

Data Encoding: This process is used to convert data into another form for the purpose of saving memory (compression) or creating a high resolution data (decompression). Ethereum uses hex values to encode transaction elements such as the function names, parameters and return values. Therefore, this implies that an owner claiming ownership of a certain property, instead claims an ownership of this original hex value signed by the original creator.

3.2 What exactly is a Non-Fungible Token?

3.2.1 Size of the industry

In recent years, NFTs have attracted an enormous amount of attention from general populace and scientific communities. Reportedly, the daily trading volume of the NFT industry is 4,592,146,914 USD [9], while the daily trading volume of cryptocurrencies is 341,017,001,809 USD. The NFT liquidity is accounting for 1.3% of the whole cryptocurrency market and it occurred in a short span of time. The NFT market has grown incredibly fast compared to 2020 and early adopters are receiving thousandfolds returns on their investments. Total number of NFTs sales has exploded from 12 million to 340 million in a single year, which was followed by a media craze and talks of the future of digital assets.

3.2.2 Fungibility

The Non-Fungible Tokens (NFT) can be compared to cryptocurrencies in a number of ways. Both hold a certain value and can be exchanged upon an agreed value, but the difference comes in fungibility. In real world, people can swap 100 dollar bill for another 100 dollar bill and nobody would disagree. Because dollar is fungible and same bills hold same value. If somebody have tried to swap one baseball card for another, then there would be some problems. Baseball cards are non-fungible, since they cannot be

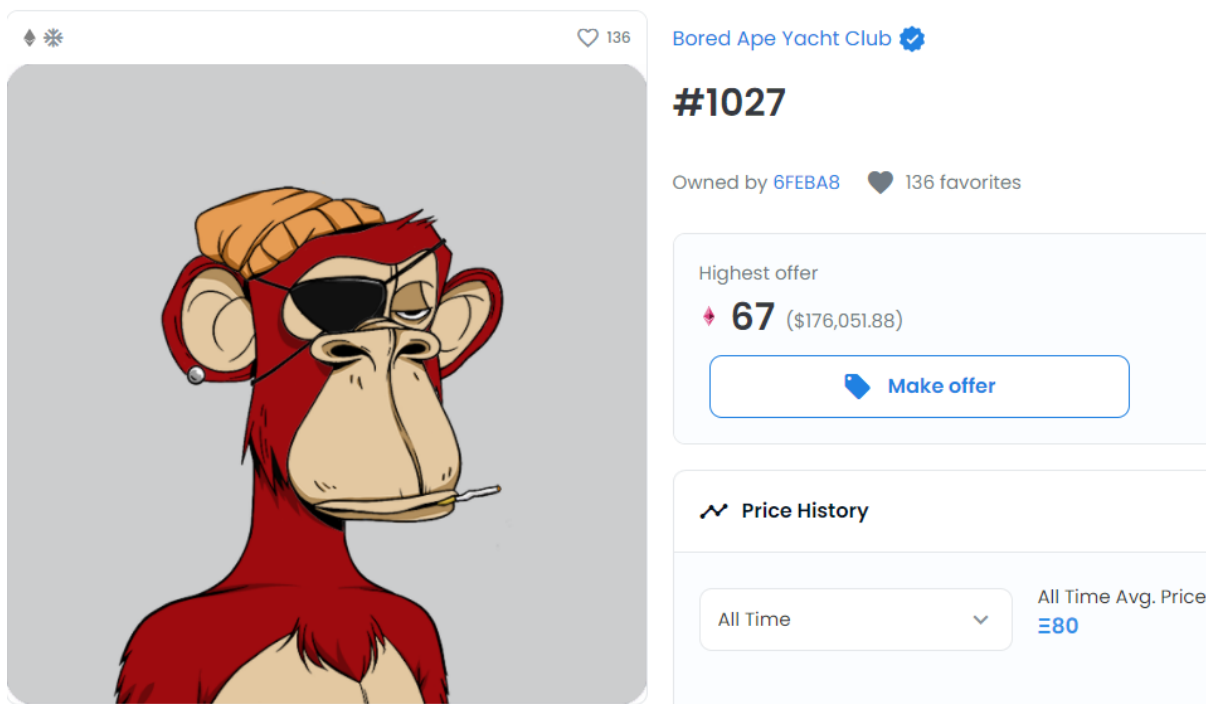


Figure 3.1: NFT for 175,000\$ [1]

exchanged freely like currency. Same applies to cryptocurrencies and NFTs. Ethereum (the primary cryptocurrency for trading NFTs) is fungible, but NFTs (just like baseball cards) themselves are non-fungible by design.

3.2.3 Bitcoin and Ethereum

Comparing Bitcoin to Ethereum is like comparing a calculator to a computer. Calculators are accomplishing a single function and computers can be programmed to accomplish multiple functions. Bitcoin acts as a primary cryptocurrency, while Ethereum can be a cryptocurrency and something else if programmed. For this reason, on a technical level, Ethereum is a primary cryptocurrency used for NFT trading.

3.2.4 Origin of monetising digital goods

The Non-Fungible Token is a creation born for the purposes of monetizing the digital goods by assigning a unique owner to each product. When the idea was first aired in 2014 on the seven on seven conference, nobody would have predicted where this technology would go afterwards. Originally the technology was developed to monetize the graphical goods (digital art) and hence the name of the project "monegraph", created by Anil Dash and Keven McCoy.

The technology was in place ever since 2014 and hasn't changed much ever since. To this day, the URLs are still being used to store the picture. NFTs gained traction recently



		
Design vision	Digital money, decentralized virtual currency	Smart contracts, decentralized applications
Founder	Satoshi Nakamoto (<i>anonymous</i>)	Vitalik Buterin et al.
Release date	2009	2015
Consensus	Proof-of-Work	Proof-of-Work, switching to Proof-of-Stake
Programming Language	Prescriptive	General Purpose (Turing Complete), Smart Contracts
Market Cap (at time of writing)	~\$120 Bil.	~\$41 Bil.

Figure 3.2: Bitcoin and Ethereum comparison [14]

due to speculations. The creators have hoped, that the technology would not become "another method of exploiting creative professionals".

When purchasing the NFT, the buyer does not acquire the exclusive right to a certain digital good. But rather a line of code on a blockchain which verifies that this person has bought the right to display an NFT. In real world, it would look like printing a picture of a celebrity and getting a signature. And all of sudden, a worthless piece of paper appreciated in value and became non-fungible, because an original artist has endorsed this piece of paper. Same applies to NFTs on blockchains. There are infinite copies of certain works, but only the buyer has an exclusive right to claim that he got an endorsement from an original artist.

3.2.5 Bulk of the NFTs are procedurally generated

The vast majority of the NFTs on open market are generated through algorithms which change the baseline model by a little bit, as shown on figure 3.3. And such a hefty price on some of the NFTs collections come from their artificially inflated value due to speculations and promises of increase in value. NFTs are being seen as risky investments, which can multiply oneself wealth. These algorithms are not necessarily complex or involve machine learning and practically anyone can make collections of NFTs with slightly variable features.

3.2.6 Network generated NFTs

There were some work done to see whether the general populace would be able to differentiate between the Generative Adversarial Network and the actual NFTs [2]. The results have shown, that there is not a significant difference between network generated NFTs and the actual ones. Some of the examples can be seen on figure 3.4



Figure 3.3: Various mass produced NFTs showing little variability

3.3 Ethical problems with NFTs

3.3.1 Wild west of digital world

Nowadays NFTs are associated with scams, thievery and speculations. The result of such bad reputation can be attributed to inherit anonymous nature of transactions, through usage of cryptocurrency and almost instantaneous transferal of funds with little ability to repeal the purchase. However, by far the worst aspect of NFTs is the failure to fulfill its initial objective as a powerful tool to monetize digital goods.

In the current world of digital art, the ownership of the digital art cannot be sold to a specific buyer. Due to nature of the internet, not a single file containing art is free from being copied and presented as an original. The NFT technology could help digital artists to acquire means of selling ownership to their art (just like in auction houses), unfortunately the bulk of art created for selling as NFTs hold little to none artistic value and are mass produced to fuel the speculative nature of this business model (figure 3.1 and 3.3).

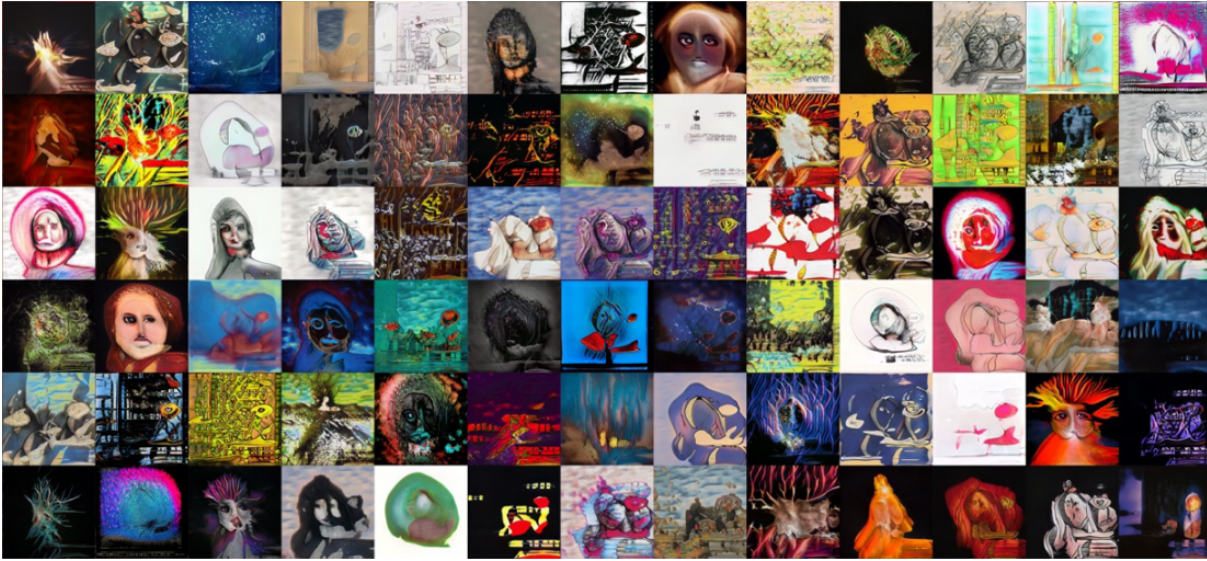


Figure 3.4: Network generated NFTs [2]

3.3.2 We have only one planet

An enormous boom in the NFT industry is closely correlated to the usage of primary NFT cryptocurrency, Ethereum. Due to the nature of the blockchains, the enormous amount of energy is required for any transaction to occur. In 2021, Ethereum has clocked in at 100 TWh per year [3]. To put this into perspective, London, the Capital City of United Kingdom had a demand of 330 TWh per year. The Three Gorges Dam in China, the most powerful Dam in the world produces 100 TWh per year. And the GPUs, which are being used to run these operations wear themselves out significantly quicker, becoming e-waste in enormous amounts.

Such a massive energy consumption is unwarranted for such an uninfluential function. The Ethereum company has recognized this and are working on new ways to mitigate this.

3.4 Standardization of the technology

3.4.1 New technology, old challenges

Any new technology has a tendency to be exploited for the short-term gains, for example internet. At first, internet was notorious for numerous ways of getting scammed. But at the end, new channels have been established for more secure experience.

Nowadays, less people are clicking on the pop up ads and give away their sensitive data to the unidentified people online. And the majority of people are now using proper procedures to lower the risk. Businesses were built around the internet model and government implemented new regulations to lower the exploitation of general populace. Same could

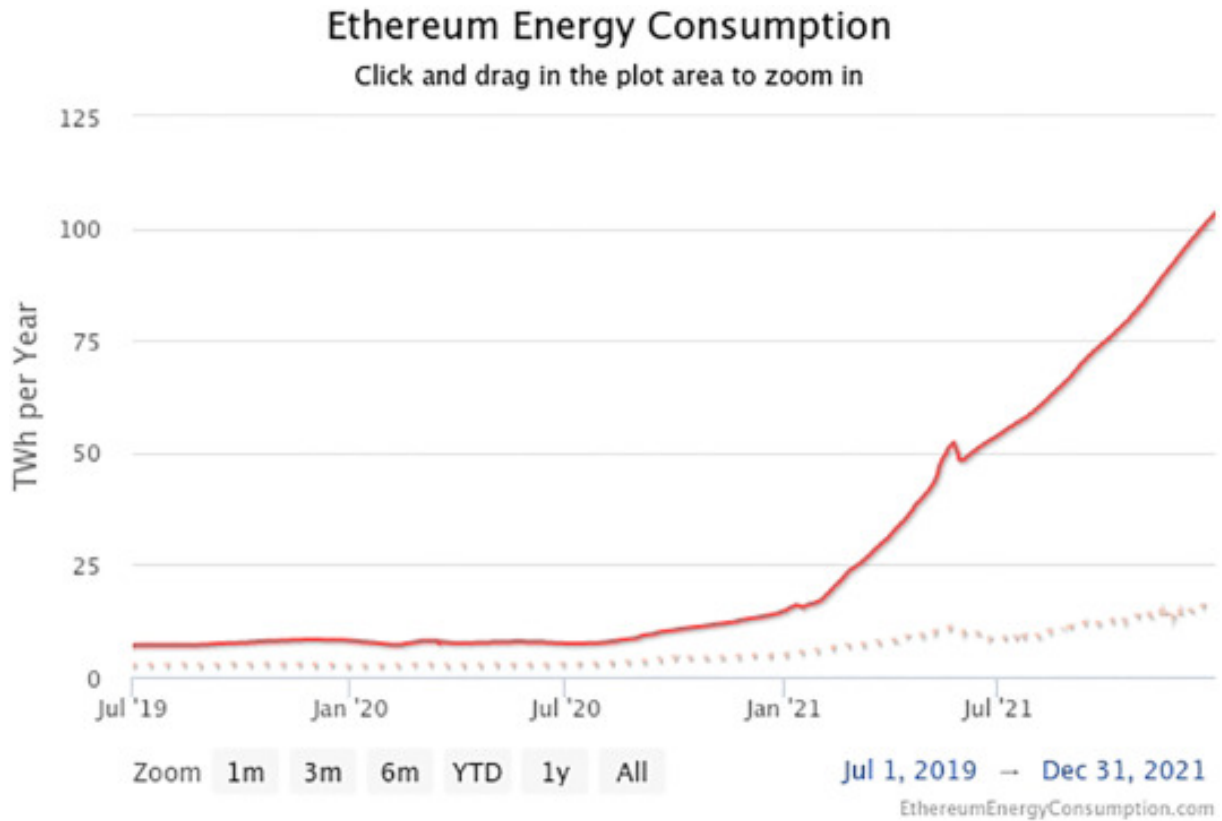


Figure 3.5: Ethereum Energy Consumption [11]

happen to NFTs, not on such a massive scale, but for the digital art community, this could be a saving grace.

Using blockchains for digital art could be just the beginning for digitization of all sorts of assets. Some companies are experimenting with digital real estate, where users can buy a plot of land in a virtual reality. NFT boom could pave the way for WEB 3.0, where people could be able to interact without any third parties.

3.4.2 Smart contracts

Behind the NFTs lies code that is on the blockchain. This code describes how those objects should be utilized and in what way the other users of the internet can interact with tokens. This code on the blockchain is called Smart Contract and is responsible not only for NFTs but for cryptocurrencies and other aspects of distributed ledgers as well. NFTs could be described as a sort of service certain blockchain offer with the collaboration of the initial creator of the token. NFTs come in a number of protocols that are being used in many ways, depending on the purpose of a token.

Because NFTs are really the code stored on the blockchain they use the concept of Smart Contracts to be used and interact among internet users. Smart contracts are simply programs stored on a blockchain that run when predetermined conditions are met. They

typically are used to automate the execution of an agreement so that all participants can be immediately certain of the outcome, without any intermediary's involvement or time loss. They can also automate a workflow, triggering the next action when conditions are met. Because there's no third party involved, and because encrypted records of transactions are shared across participants, there's no need to question whether information has been altered for personal benefit. Blockchain transaction records are encrypted, which makes them very hard to hack. Moreover, because each record is connected to the previous and subsequent records on a distributed ledger, hackers would have to alter the entire chain to change a single record. Thanks to these mechanics of Smart Contract, it is easy to create new NFTs on the blockchain and help internet users interact with them.

3.4.3 Protocols

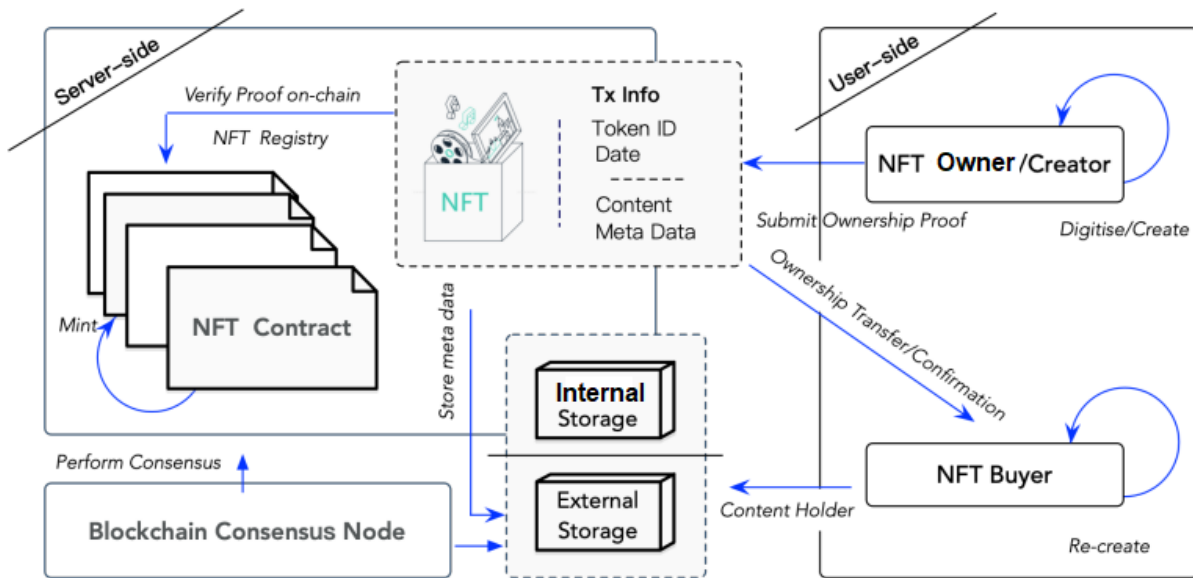


Figure 3.6: NFT system schematic [9]

Assuming that there is a secure, complete, consistent and available ledger with records, we can identify two design patterns for the NFTs: **Top to Bottom** and **Bottom to Top**.

Top to Bottom: An initiator creates the NFT and sells it directly to the buyer.

- Initiator checks that the file corresponds to the contract specifications and all details are correct. And then converts raw data into an appropriate format.
- Initiator stores the raw data outside of the blockchain, for the cos efficiency. Storing pictures on the blockchain is very energy consuming and expensive.
- Initiator signs the transaction and the hash leading to the raw data, and then sends the transaction to the smart contract of his/her own choosing.
- Smart contract receives the NFT raw data with transaction and starts minting.

- After the transaction is verified, the minting is over and NFT will be linked forever to the unique blockchain address.

Bottom to Top: An initiator creates an NFT template which is used by other users.

- Initiator creates a template via a smart contract.
- Users bid on NFTs and acquire additional feature on top of basic template. The ontop features are randomly chosen.
- Minting process starts, once the smart contract is triggered.
- Once the transaction is verified, the minting is complete and NFT is forever linked to the unique blockchain address.

Such system preserves the history of all transactions occurring on the blockchain. And once the block reaches its capacity, other blocks are added, which are linked to original data.

3.5 Token Standards

Certainly, NFTs follow some sort of a standardization, so that they are being used in an ordered manner and no further complications occur while interacting with them. That is why several protocols have been created to ease the use of NFTs. Protocols can be described as blueprints for anyone that wants to create a new token, which will be used on the blockchain. It must be mentioned that currently (Q2 2022) most of NFTs are being handled by Ethereum Network that, so the protocols and Smart Contracts are native to this network.

3.5.1 Desirable traits of Non-Fungible Tokens

NFTs are based on the decentralized ledger system which benefit from following properties [9]:

- **Verification.** Everything is on the public blockchain and can be easily verified.
- **Transparency.** All activities including in transaction and minting are publicly accessible.
- **Availability.** The system never shuts down, and the market is always open for transactions.
- **Resistance to tampering.** The records in the blockchain across multiple nodes restrict anyone from tampering with the confirmed transaction data.

- **Usability.** Every NFT is updated and information is clearly displayed.
- **Atomic transactions.** Transactions are completed in atomic, consistent, isolated and durable (ACID) process.
- **Tradability.** All NFTs can be traded and exchanged upon agreed value.

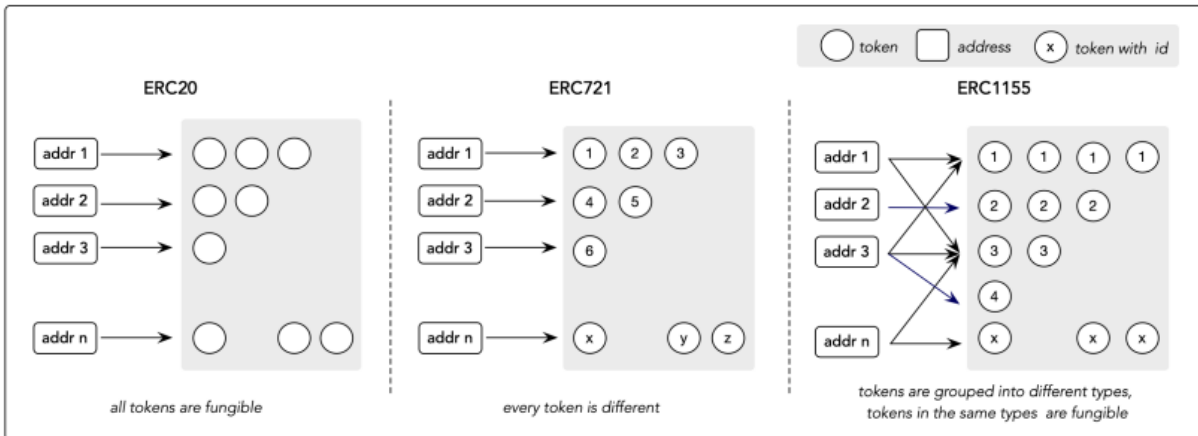


Figure 3.7: NFT token standards [9]

3.5.2 ERC-20

One of the most significant Ethereum tokens is known as ERC-20. ERC-20 has emerged as the technical standard; it is used for all smart contracts on the Ethereum blockchain for token implementation and provides a list of rules that all Ethereum-based tokens must follow. ERC-20 is similar, in some respects, to bitcoin, Litecoin, and any other cryptocurrency; ERC-20 tokens are blockchain-based assets that have value and can be sent and received. The primary difference is that instead of running on their own blockchain, ERC-20 tokens are issued on the Ethereum network. The main aspect of this protocol is that it handles the fungible tokens.

3.5.3 ERC-721

ERC-721 is the token standard that started it all, and it remains the most popular, widely used NFT standard to this day. ERC-721 is a free, open token standard that describes how to build NFTs on the Ethereum platform. The ERC-721 token standard was developed to standardize NFTs. As a result, a new dawn of digital content, games, and applications has emerged from this protocol. In contrast to the previous protocol this one is designed to handle solely non-fungible items on the blockchain.

```

pragma solidity ^0.4.20;

/// @title ERC-721 Non-Fungible Token Standard
/// @dev See https://github.com/ethereum/EIPs/blob/master/EIPS/eip-721.md
/// Note: the ERC-165 identifier for this interface is 0x80ac58cd
interface ERC721 /* is ERC165 */ {
    /// @dev This emits when ownership of any NFT changes by any mechanism.
    /// This event emits when NFTs are created (`from` == 0) and destroyed
    /// (`to` == 0). Exception: during contract creation, any number of NFTs
    /// may be created and assigned without emitting Transfer. At the time of
    /// any transfer, the approved address for that NFT (if any) is reset to none.
    event Transfer(address indexed _from, address indexed _to, uint256 indexed _tokenId);

    /// @dev This emits when the approved address for an NFT is changed or
    /// reaffirmed. The zero address indicates there is no approved address.
    /// When a Transfer event emits, this also indicates that the approved
    /// address for that NFT (if any) is reset to none.
    event Approval(address indexed _owner, address indexed _approved, uint256 indexed _tokenId);

    /// @dev This emits when an operator is enabled or disabled for an owner.
    /// The operator can manage all NFTs of the owner.
    event ApprovalForAll(address indexed _owner, address indexed _operator, bool _approved);

    /// @notice Count all NFTs assigned to an owner
    /// @dev NFTs assigned to the zero address are considered invalid, and this
    /// function throws for queries about the zero address.
    /// @param _owner An address for whom to query the balance
    /// @return The number of NFTs owned by `_owner`, possibly zero
    function balanceOf(address _owner) external view returns (uint256);

    /// @notice Find the owner of an NFT
    /// @dev NFTs assigned to zero address are considered invalid, and queries
    /// about them do throw.
    /// @param _tokenId The identifier for an NFT
    /// @return The address of the owner of the NFT

```

Figure 3.8: Part of the ERC-721 Protocol that handles non fungible tokens. [4]

3.5.4 ERC-998

ERC-998 tokens are non-fungible, just like ERC-721 tokens. However, ERC-998 tokens are composable as well. ERC-998 tokens can be organized into complex digital assets and valued, traded, or sold as one entity. The ERC-998 token standard can hold various Non-Fungible Tokens, like the ERC-721, and fungible tokens, such as the ERC-20. Think of the ERC-988 token as a unique portfolio of digital assets. People use these NFTs to organize their digital assets in a single place.

3.5.5 ERC-1155

This protocol can be described as a mixture of ERC-20 and ERC-721 as ERC-1155 tokens allow users to register fungible tokens and Non-Fungible Tokens in the same smart contract. This token standard was written primarily with gaming in mind. Fungible tokens, such as in-game currency, are often used to purchase NFTs, such as in-game items and other digital collectibles. This protocols further widens the possibilities that internet users can utilize to their needs and liking.

Algorithm 1: NFT Standard Interfaces (with selected functions)

```

interface ERC721 {
    function ownerOf(uint256_tokenId) external view returns (address);
    function transferFrom(address_from, address_to, uint256_tokenId)
    external payable; ...
}
interface ERC1155 {
    function balanceOf(address_owner, uint256_id) external view returns
    (address);
    function balanceOfBatch(address calldata _owners, uint256 calldata
    _ids) external view returns (uint256 memory);
    function transferFrom(address_from, address_to, uint256_id, uint256
    quantity) external payable; ...
}

```

Figure 3.9: ERC-721 and ERC-1155 algorithm [6]

3.6 Technical Security Concerns

There are six security concerns when it comes to NFTs which can be abbreviated as **STRIDE** [9], this method is employed through risk evaluation taking into account: authenticity, integrity, non-repudiability, availability and access control:

- **Spoofing (Authenticity):** An ability of impersonation of another individual, may lead to theft of the private key and illegal transferal of NFTs, therefore additional verification is required.
- **Tampering (Integrity):** An ability to manipulate NFT data, although the data on the blockchain cannot be tampered with. The data outside can be manipulated and lead to loss of sensitive information. Therefore, both hash and original data should be sent when trading.
- **Repudiation (Non-repudiability):** Users cannot deny NFTs, therefore hash data can be bound with an attacker's address. Therefore, the contract should be confirmed by more than one participant.
- **Information Disclosure (Confidentiality):** Confidential information can become exposed if NFT owner is not using the privacy preserving smart contracts. The attacker may exploit the linkability of the hash.
- **Denial of Service (DoS) (Availability):** DoS attacks break basic functions of the service. Luckily, the blockchain is very resilient to such attack. But the outside web service is still susceptible. Therefore a hybrid blockchain architecture should be used.

- **Elevation of Privilege (Authorization):** An attacker may gain higher authority, which will grant some undesirable rights.

STRIDE	Security Issues	Solutions
Spoofing <i>(Authenticity)</i>	<ul style="list-style-type: none"> • An attacker may exploit authentication vulnerabilities • An attacker may steal a user's private key. 	<ul style="list-style-type: none"> • A formal verification on the smart contract. • Using the cold wallet to prevent the private key leakage.
Tampering <i>(Integrity)</i>	<ul style="list-style-type: none"> • The data stored outside the blockchain may be manipulated. 	<ul style="list-style-type: none"> • Sending both the original data and hash data to the NFT buyer when trading NFTs.
Repudiation <i>(Non-repudiability)</i>	<ul style="list-style-type: none"> • The hash data may bind with an attacker's address. 	<ul style="list-style-type: none"> • Using a multi-signature contract partly.
Information disclosure <i>(Confidentiality)</i>	<ul style="list-style-type: none"> • An attacker can easily exploit the hash and transaction to link a particular NFT buyer or seller. 	<ul style="list-style-type: none"> • Using privacy-preserving smart contracts instead of smart contracts to protect the user's privacy.
Denial of service <i>(Availability)</i>	<ul style="list-style-type: none"> • The NFT data may become unavailable if the asset is stored outside the blockchain. 	<ul style="list-style-type: none"> • Using the hybrid blockchain architecture with weak consensus algorithm.
Elevation of privilege <i>(Authorization)</i>	<ul style="list-style-type: none"> • A poorly designed smart contract may make NFTs lose such properties. 	<ul style="list-style-type: none"> • A formal verification on the smart contracts.

Figure 3.10: Potential Security Concerns and their Possible Solutions [9]

3.7 Copyright implications of NFTs

It is no secret that the sale of an NFT does not necessarily transfer the underlying copyright in the work which exists "off-chain" to the purchaser. Such is the case when selling a physical copy of nearly any type of creative work - the transfer of the underlying copyright is up to the creator or most recent copyright owner. Unfortunately, many people still do not know that what they are really buying, what is described as NFTs, are not the actual pictures but the internet links to those pictures, videos or music that is hooked on some page not on the blockchain.

3.8 NFT system design flaws

3.8.1 Concentration of Power

It is certain that the blockchain is exploitable and many frauds happen, often involving millions of dollars of stolen assets. While many developers do their best to secure the blockchain from further breaches still many people that invested heavily into the world of non-fungible tokens are being robbed off their wealth. This sparks an underlying question. Should there be some sort of surveillance that could freeze assets or stop fraudsters from their deeds. However, the main idea behind the decentralized world of finances is that there is no human interaction when it comes to regulating or surveilling anything that is happening in the blockchain, because the only authentication is needed is the one that is provided by distributed ledger mechanism. However, as the number of frauds, scams and breaches taking place some interventions have been made. Most notably recently (Q1 2022) OpenSea, the leading marketplace for NFTs, had to freeze a few million dollars worth of stolen assets. Other examples may include OpenSea's security flaws has been successfully exploited by hackers. The hacker attack happened in January 2022 and resulted in the theft of about 1 million dollar worth of NFTs. The hackers were able to buy at least 8 NFTs for much less than what was thought to be their "fair market value".

3.8.2 Phishing

Phishing is another cybercrime that is widespread in the NFT space and beyond. One of the latest phishing attacks happened in January 2022 when scammers tricked supporters of CryptoBatz - Ozzy Osbourne's NFT project. Of course, a collection of 9,666 digital bats issued by a famous rock star attracted the public's attention. Scammers took advantage of this opportunity. They used an old URL of the project and created a fake Discord server. Since the old tweets from CryptoBatz and Ozzy himself contained the previous URL and were not deleted, they unintentionally directed users to a server taken over by scammers. As a result, users clicked on a message asking them to verify their crypto assets. They were redirected to a phishing site, where they connected their crypto wallets and ended up losing money. Selling fake works attributed to famous artists as non-fungible tokens is also a popular NFT scam.

3.8.3 Money Laundering

It is also worth mentioning that traditional art has been used for money laundering for years. NFTs could make this process even easier. For the time being, NFTs can be anonymous, and you may not know who is behind the artist's nickname and avatar. There are also no laws or regulations for NFTs. These facts make NFTs ideal for hiding illegally earned money. NFT marketplaces, even the largest ones, are not required to comply with Anti-Money Laundering (AML) and Counter Terrorist and Proliferation Financing (CTF/PF) standards, and users don't need to go through Know Your Customer (KYC)

procedures. However, there is more and more discussion about the possibility that NFT platforms will soon have to adopt KYC, AML, and CTF/PF solutions.

”The PrivacyHQ survey spoke to 1,008 people in the U.S. who are actively investing in and own NFTs. And according to the report, there are some horror stories and great lessons to be learned. The key takeaways from the survey are:

- Less than half of NFT owners feel their NFTs are secure
- Two out of 3 respondents said they had panic sold NFTs in the past
- Nine out of 10 respondents had experienced an NFT scam
- Half of the respondents had lost access to their NFTs at some point

When it comes to NFT scams there were multiple ways in which buyers were scammed. Topping the list of the most common scams experienced by these respondents starts out with the NFT provider shutting down or changing their URL at 44.8%. Next is investing in an NFT project that disappeared at 43.8% followed by buying an NFT from a fake marketplace at 43.4% to round off the top three.”

3.9 Potential of NFTs

3.9.1 Gaming Industry

There are already games which use in game mechanics to create NFTs, which can be sold for profit. Some games possess mechanics which allow to mix certain attributes and create unique NFTs. Such games can earn a substantial amount of money for the developers. However, the majority such games are mobile and relatively small. We are yet to see a major gaming publisher to implement NFTs into their games, although there are plans brewing. Such models allow to have a mutually beneficial relationship between the consumers and developers, where both parties can profit from NFT market[9].

3.9.2 Virtual Events

Real world tickets can be fraudulent, cancelled or resold multiple amount of times. With virtual events and ability to sell tickets on the blockchain, it could be possible guarantee a spot on the event. Traditional events require trust into third party, that everything will go smoothly. With smart contracts, such trust is not required, since everything is as much as transparent as possible.

3.9.3 Protection of digital artworks

The problem with real world art, is the problem with monetizing. Artists in real world have limited amount of pathways for making a living using their talent. Therefore, only the most prestigious and well-known artists make it to the comfortable levels of living. This problem could be alleviated with digital artists, if we give tools for propelling crypto-artists talent across the barrier of obscurity. Smart contracts can be adapted to pay royalties to the original artists, whenever a new transaction occurs. Digitized artworks could yield income to artists passively, whenever their work is being used anywhere (videos, advertisements or articles). This level of control is unimaginable in real world.

3.9.4 Metaverse

A virtual reality, where people can communicate, collect, interact and do all sorts of things. NFTs could be used to display ones emotion or show the state of their finances. They could be used in games or traded like collectibles. At this point, virtual land can be leased to other individuals and there can be virtual landlords. And all of this can earn money, which can be used in the real world. However it is a bit too early to speculate about it, since technology still does not allow for immersion to the metaverse in the first place.

3.10 Legal Hurdles

NFTs are confronting legal basis of most countries from all sides. And government restrictions vary from country to country. India and China have strict rules around cryptocurrencies and NFTs sale. Therefore, the NFT traders have to overcome the challenges of governance to be able to turn the profit. It is important to know rules and scrutiny of different governments and be able to tolerate such requirements.

Governments would like to be able to tax capital gains on NFT market. At this point, only USA taxes cryptocurrencies, while other countries haven't considered it [9]. This makes it easier to conduct financial crimes under cover of NFT trading.

The main aspects of legal barriers include[12]:

- **Copyright:** Purchasing an NFT does not transfer the ownership of the digital work. The artwork creator or the third-party seller still retain the right copy, distribute, modify, and publicly display or perform the work.
- **Privacy and data protection laws:** Data protection laws allow for individuals to erase their personal information. Having personal information on the blockchain makes it impossible to execute. Therefore NFTs containing personal information may violate data protection laws.

- **Property law:** The asset's location can determine the property law, therefore it is important to know the which legal system governs the asset. However the NFTs represent the unique hash towards the digital work on the external database, but not the actual work itself.
- **Money laundering:** The sheer value of NFT transaction and a widespread of NFTs is raising concerns, whether this system is being widely used for purposes of breaking anti-money laundering laws.
- **Regulatory:** Since NFTs are non-fungible and cannot be traded away, they are not securities and are not subject to securities regulations. Therefore, an unregulated NFT transactions and "wash trading" can create an artificial impression of demand for an asset.
- **Taxation issues:** It is very hard to pin point, whether NFTs are subjects to tax regulations.
- **Security:** While the blockchain itself is immutable, the digital works are not and can be subjects of cyber-hacking. Therefore investors should be wary of such possibility
- **Estate and succession planning:** What happens to NFTs after owner's death. Leaving digitally stored assets to the next generation can lead to difficulties for executors.

3.11 The NFT communities

We should also consider the human factor, and try to understand why would people invest heavily into such insecure form of an asset. For that we need to answer three questions[13]:

- **Q1-**Why do NFT creators create NFTs?
- **Q2-**How do NFT creators engage with NFTs and their communities?
- **Q3-**What challenges do NFT creators encounter?

For the questions to be reasonably answered, a group of volunteers (15) were used with experience in NFT trading.

3.11.1 A wide range of motivators for NFT creation (Q1)

- **Uniqueness:** *"Fungible tokens are more for speculation, and people trade them in the same way as stocks. But non-fungible tokens, most of them are kind of art. [...] They have some special traits, very different from each other. If you own one of them, you have some personal emotion for it, and would use it as an Avatar."*

Each NFT can be unique and connect to the somebody on the personal level, this factor is not driven by the goal to enrich oneself, but by the satisfaction of having something special.

ID	Gender	Country	Role	Year(s) of exp	Main platform(s)	Total created	Price range(ETH)	Specific art/media
P1	M	UAE	Creator	<1	Rarible	28(r)	0.1-0.15	AI tools, Illustration
P2	M	Canada	Creator	<1	Rarible	27(r)	0.003-0.05	3D, VR with Tiltbrush
P3	M	USA	Creator	<1	Twitter NFT group	108(r)	0.1-3	2D, 3D art
P4	F	Germany	Creator	<1	Rarible	98(r)	0.05-0.1	3D ducks, 2D Linework
P5	M	USA	Creator	<1	Rarible, Decentraland	29(r)	0.3-4	3D game, holographic
P6	M	Argentina	Creator	1	Opensea, Rarible	205(os)	0.005-5	3D abstract porn art and tattoo
P7	M	China	Creator+Collector	<1	Opensea	na	na	na
P8	M	USA	Creator	1	Rarible	160	0.013-5	Piet Mondrian Vector Mash-Ups
P9	F	China	Creator	<1	Own platform	na	na	na
P10	M	China	Creator+Collector	<1	Own platform, Rarible	na	na	na
P11	M	Taiwan	Creator+Collector	2	Opensea, Foundation	na	na	na
P12	M	USA	Creator	1	Opensea, Rarible, Tezos	66(tz)	0.01-5	Experimental 4D/AV
P13	M	Switzerland	Creator	<1	Foundation, Opensea, NFTB	75(os)	0.35-1.5	Luminescence, Hyperreal Landscape
P14	M	Iran	Creator	1	Foundation, Rarible, Opensea	205(os)	0.14-2	Geometric shapes of faces, sculptures
P15	M	Belarus	Creator	<1	Rarible, Foundation, Opensea	12(os)	0.09-0.5	Natural things and digital apocalypse

Figure 3.11: Basic information of study participants. For NFT creation, three major marketplaces were used. r:Rarible, os: Opensea, tz: Tezos. Price range is given in ETH (1ETH=3200USD at the time of writing) [13]

- **Proof of Ownership:** NFTs are valued for its ability to verify the ownership with 100% certainty, whereas more traditional works might be disputed. This trait bring an unprecedented level of transparency to the art market. As P7 has put it:

Before NFT was a thing, I would post paintings on Instagram, and offer them for sale. 99.9% of people would just say, why should I buy it? I can just copy the image and download it for myself. There's no real way to prove that you created it or that you've bought it. Now every time it is sold, I get the percentage

- **Expressing Personal Emotions:** Perhaps a more strong motivator would be the most human one. People just create NFTs from personal emotional feelings, whether they want to support their favourite NBA team or keep a memory of one the closest relatives. P11 noted:

"I was thinking of creating NFT because my child will be born in December. I'm thinking of making some connections between the baby and the blockchain, and even create a trend on the market. I think I'm fine with Photoshop and simple art, as well as uploading it to the Ethereum blockchain. I want to do this because I want to keep some memories for him, like funny stories and his dreams."

- **Expressing Creativity:** Minting can be used as a way to express oneself, especially if the person is artistically talented and has a lot to show. P4 was one these talents:

"This was a way to step out of my comfort zone of normal art a little bit, start something new, learn and create something new. In my previous [tattoo] business, I created tattoo designs for my clients' ideas. But NFT are just my own ideas, and that's pretty cool to make something new."

- **Leveraging New Models for Art Business:** NFTs mean a great deal for the unknown artists who are trying to make the name for themselves. This system provides low barrier of entry, since there is no dependency on event organizers or auctioneers. The autonomous platform empowers the independent artists by making their work easily exportable. It was even more crucial during the pandemic. P3, the owner of the NFT company mentioned:

"It empowers individuals to create their own business model. Your business transactions would always be recorded [...] benefit freelancers to automate their workflow. You don't have to worry about tracking receipts. They're all stored on blockchain [...]in theory, if you use it right, if you trust the platform you work with, you would have unlimited control."

- **Being A Part of "The New Renaissance"**: And for many, being part of the great movement is enough. Many participants consider NFT system revolutionary. P3 explained:

"Almost everyone that I've interacted with believes this [NFT] to be one of the most important technologies and advancements in our lifetime. A lot of people are calling it the new Renaissance for artists[...]. A lot of new inspiration and new art is coming out, and people are very willing to collaborate in network, across countries, across the world."

3.11.2 Bit by bit engagement with NFTs and the community (Q2)

Each volunteer has started the route to creating NFTs differently, but the structure they followed can be generalized in fig 3.12. The pathway can be split into two routes of a beginner (early stage) and experienced. During the early stage, people encounter problems due to lack of knowledge of IT concepts. It takes a while to learn new terminologies and understand why and how is everything is working. And usually these people tend to help themselves though research (google, youtube and etc.). On the other hand, experienced people (someone what has been on the market for months) travel though different route. These people discover and find new information through community hubs and acquire direct information first hand from other people.

3.11.3 Challenges on the way for NFT creators (Q3)

Due to the fact that NFT is a new phenomenon, the major hurdles are learning new material, vast amount of scams and phishing and lack of metrics to assess NFT creators. The summary can be seen on fig 3.13. The NFT community at this point is a wild west and there are no concrete rules. Unfortunately, this paints a bad picture for the NFTs as a whole. But with a healthy and motivated community, these problems can be dealt with as time passes on.

3.12 Conclusion

The NFT industry is booming and does not show the signs of slowing down. It certainly looks like a bubble, but with possibly enormous potential. The market is ripe with scams, but the base schematic is robust enough. Ethereum is switching to proof-of-stake, which reduces the energy consumption by 99.95%, so the environmental aspect might be tolerable

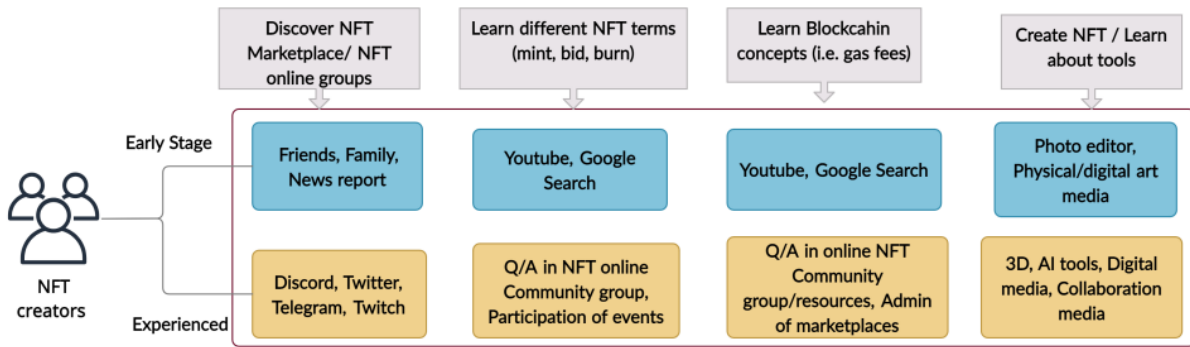


Figure 3.12: Two stages of how our participants created NFTs: an early stage (when they initially joined the NFT marketplace), and an experienced stage (when they stayed in the marketplace for a few months). The top row shows the steps taken to create NFTs, whereas the bottom rows (blue and orange boxes) show the corresponding resources used for each step, for early stage and experienced stage, respectively. During the early stage, our participants tended to seek information and create NFTs on their own. In contrast, during the experienced stage, they tended to discover and participate in relevant online groups and communities, and create NFTs more collaboratively. [13]

in the future. The technology is still young and not mature enough. People using this technology are still not used mature enough to use it. NFT will take the same process as Internet, where everyone slowly started understanding how to properly communicate online over time. This report has compiled history, ethics, legal and technical part of Non-Fungible Coins and maybe shed light on its future.

Challenges	Current strategies to address the challenges
Difficulty in understanding new and complex concepts & technologies	<ul style="list-style-type: none"> • Seek help from community members, e.g., sharing tools/resources, solving bugs • Seek help from NFT project/platform administrators upon request • Search answers online by themselves (e.g., Google, YouTube)
Lack of assessment metrics for (new) NFT creators & NFTs	<ul style="list-style-type: none"> • Look at prior reputation and how well-known the creator is • Look at the number of their social media followers (mainly Twitter accounts) • Estimate the amount of work/time devoted to creating the NFT • Assess aesthetics of the NFT
Prevalence of phishing NFTs and scams	<ul style="list-style-type: none"> • Manually inspect NFTs and compare them with existing NFTs • Report scams to NFT project administrators • Alert/Remind other community members (mostly via Discord channels and Twitter) • (Some NFT platforms) Use a KYC-like process and/or an “invite-only” mechanism to restrict community membership

Figure 3.13: Challenges faced by NFT creators and their current strategies to address these challenges. [13]



Figure 3.14: A virtual NFT gallery in Decentraland, owned by P5. Galleries are interactive media in the game, designed for people to sell and buy NFTs. Each user has her own avatar/character in the virtual world. [13]

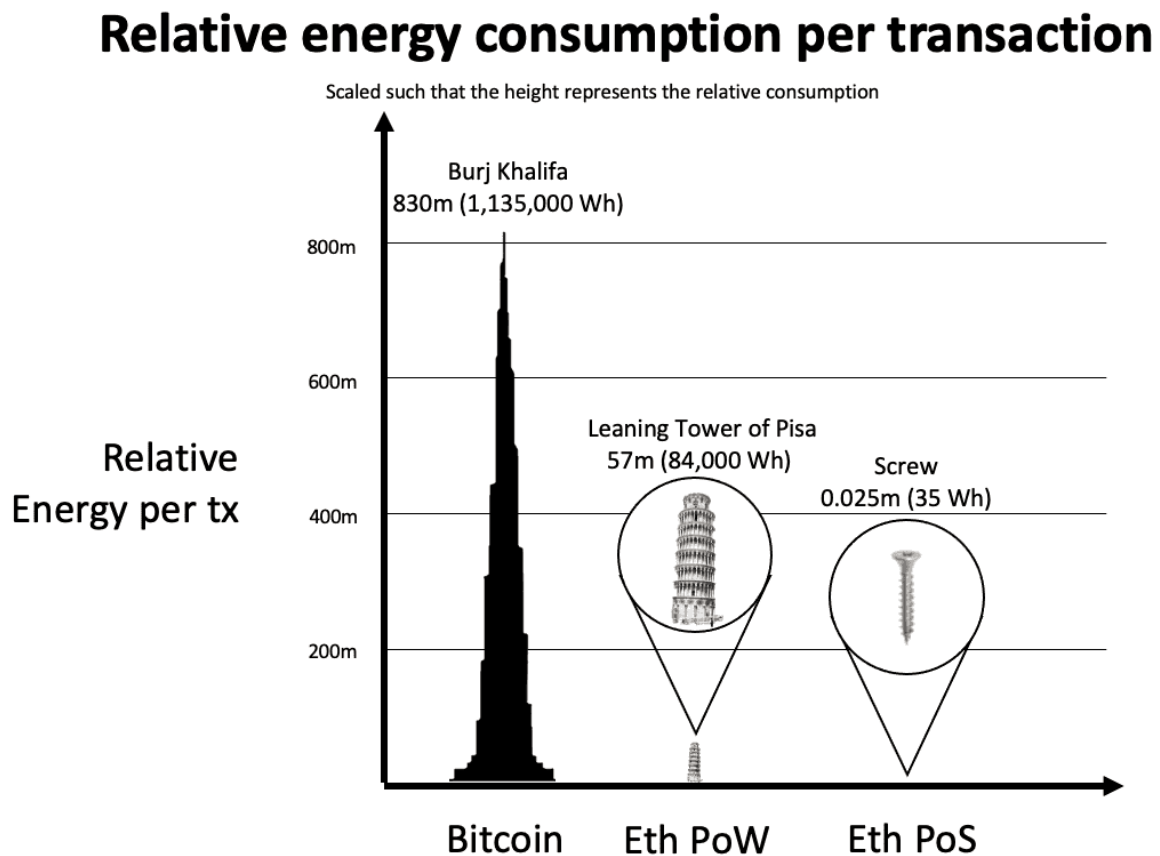


Figure 3.15: Estimate of PoW energy consumption per tx used in figure based on May 2021 data [11]

Bibliography

- [1] OpenSea <https://opensea.io/>
- [2] Sakib Shahriar, Kadhim Hayawi: *NFTGAN: Non-Fungible Token Art Generation Using Generative Adversarial Networks*, International Conference on Machine Learning Technologies, December 2021. <https://arxiv.org/abs/2112.10577?context=cs>.
- [3] Jon Truby, Rafael Dean Brown, Andrew Dahdal, Imad Ibrahim: *Blockchain, climate damage, and death: Policy interventions to reduce the carbon emissions, mortality, and net-zero implications of non-fungible tokens and Bitcoin*, Energy Research & Social Science, 2022. <https://tinyurl.com/memhnsy2>
- [4] OpenZeppelin|DOCS: *ERC721*. <https://tinyurl.com/mr2xyf8s>
- [5] Ethereum.org: *ERC20*. <https://ethereum.org/en/developers/docs/standards/tokens/erc-20/>
- [6] LCX.com: *ERC-721 vs ERC-998 vs ERC-1155: The Difference*, January 2022. <https://www.lcx.com/token-standards-erc-721-erc-998-and-erc-1155-how-are-they-different/>
- [7] Gregory J. Chinlund, Kelley S. Gordon: *What are the copyright implications of NFTs?* October 2021. <https://www.reuters.com/legal/transactional/what-are-copyright-implications-nfts-2021-10-29/>
- [8] Michael Guta: *Only 1 in 10 NFT Owners Have Never Experienced a Scam*, March 2022 <https://smallbiztrends.com/2022/03/1-in-10-nft-owners-never-experienced-scam.html>
- [9] Qin Wang, Rujia Li, Qi Wang, Shiping Chen: *Non-Fungible Token (NFT): Overview, Evaluation, Opportunities and Challenges*, October 2021. <https://arxiv.org/pdf/2105.07447.pdf>
- [10] Johnny Harris: *How Crypto will Change the World (or Not)* https://www.youtube.com/watch?v=v0V_zkng4go
- [11] Ethereum.org: *Ethereum energy consumption* <https://ethereum.org/en/energy-consumption/>
- [12] Hamed Ovaisi: *What are the legal issues around NFTs?* <https://tinyurl.com/4ahmmsvr>

- [13] Tanusree Sharma, Zhixuan Zhou, Yun Huang, Yang Wang: *"It's A Blessing and A Curse": Unpacking Creators' Practices with Non-Fungible Tokens (NFTs) and Their Communities*, University of Illinois at Urbana-Champaign, USA, January 2022. <https://arxiv.org/pdf/2201.13233.pdf>

- [14] cameronfitchett: *Bitcoin vs Ethereum*, September 2018. <https://golucidity.com/bitcoin-and-ethereum-whats-the-difference/>

Chapter 4

Brain-Computer Interfaces: Overview and Application Scenarios

Zeen Wang, Szymon Modrzynski

The brain-computer interface (BCI), also called a brain-machine interface (BMI), is a direct communication and control channel established between the human brain and a computer or other electronic devices. Through this channel, people can express ideas or manipulate devices directly through the brain without language or action. Brain-computer interface technology is an interdisciplinary technology involving neuroscience, signal detection, signal processing, pattern recognition, etc. Many scientists from all over the world have contributed to the technology in this scientific domain. They developed brain signal collection and processing techniques and tools. With scientists' help, BCI is no longer empty talk. In this paper, we present the most relevant aspects of the BCI and all the major achievements that have been made over the history of this research domain. We mention people who were pioneers in area of BCI and groundbreaking technological breakthroughs. We expect readers to have a general understanding and knowledge of the BCI domain by reading this paper.

Contents

4.1	Introduction and Problem Statement	75
4.1.1	Introduction	75
4.1.2	Problem Statement	76
4.2	Related Work	77
4.2.1	Brain and nervous	77
4.2.2	Brain waves	77
4.2.3	Event-related Potential	78
4.2.4	Machine Learning	78
4.2.5	Our Contribution	78
4.3	Types and Designs of BCI	79
4.3.1	Reasons for Categorization	79
4.3.2	Passive, Active and Reactive	79
4.3.3	Motor, Sensory, Sensorimotor, Cognitive and Brainets	80
4.3.4	Non-Invasive, Semi-Invasive and Invasive	80
4.4	BCI Technology	81
4.4.1	BCI System Structure	81
4.4.2	Signal Acquisition	82
4.4.3	Device Specifications	82
4.4.4	Non-Invasive BCI	82
4.4.5	Semi-Invasive BCIs	85
4.4.6	Invasive BCIs	85
4.5	BCI Application Scenarios	86
4.5.1	Medical Applications	86
4.5.2	Military Applications	88
4.5.3	Entertainment Applications	89
4.6	Risks Associated with BCI Usage	90
4.6.1	Technical Security Risks	90
4.6.2	Privacy Risks	91
4.6.3	Decision-making Autonomy Risks	92
4.6.4	Social Equity Risks	92
4.7	Future Development of BCI	93
4.7.1	Major Hurdle	93
4.7.2	Future trends and possible development directions	95
4.8	Conclusion	96

4.1 Introduction and Problem Statement

4.1.1 Introduction

The brain-computer interface (BCI) is a direct and two-way communication link between the brain and external device or a computer. BCI has long been recognized as a potential technology to assist people with disabilities, primarily deaf and blind. For an EEG-based BCI system, it is important to start with the short characteristic of the EEG, which was first recorded by Hans Berger in 1924 [3] and that's what has led to the identification of the alpha and beta waves [13]. The first systematic attempt to implement an electroencephalogram EEG-based BCI was made by J. J. Vidal in 1973 [2], who recorded evoked electrical activity in the cerebral cortex in intact skulls. Later, the researchers used the P300 to establish direct communication between the computer and the brains of people with severe movement disorders (Farwell and Donchin, 1988 [4]). As an alternative to traditional treatments for movement disorders, BCI technology helps to artificially enhance or re-energize synaptic plasticity in affected neural circuits. By harnessing undamaged cognitive and emotional function, BCI aims to re-establish connections between the brain and damaged peripheral parts (Vansteensel et al., 2016 [5]).

As the first experiment to be carried out on animals, Vladimir Vladimirovich Pravdicz-Nieminski was the first to demonstrate the bioelectricity activity of the animal brain in 1913. He recorded up to seven different types of changes and recorded alpha and beta frequencies. [15]. This kind of bioelectricity was later called "electroencephalography (EEG)".

At the same time, the application of BCI technology has gradually become a reality and enriched, such as brain fingerprinting for lie detection (Farwell et al., 2014 [6]), detecting drowsiness for improving human working performances (Wei et al., 2018 [7]), estimating reaction time (Wu et al., 2017b [8]), controlling virtual reality (Vourvopoulos et al., 2019 [9]), quadcopters (LaFleur et al., 2013 [10]) and video games (Singh et al., 2020 [11]), and driving humanoid robots (Choi and Jo, 2013 [12]).

The core of BCI technology is a conversion algorithm that converts EEG signals input by users into output control signals or commands. A very important part of BCI research work is to adjust the mutual adaptation relationship between the human brain and the BCI system, which is to find suitable signal processing and conversion algorithms. This enables the neural electrical signals to be converted into command or operation signals that can be recognized by the computer through the BCI system.

Due to the different application occasions and the required signal characteristics, the corresponding signal processing methods and conversion algorithms are also very different, and even each BCI system has its own algorithm. This makes the evaluation of BCI performance difficult. Therefore, establishing a unified evaluation standard that can be accepted by most people is an indispensable and important link in the development of BCI technology. This standard should be able to fairly and objectively evaluate the vast majority of BCI technologies.

BCI technology is currently a very active research field in the world. As an interdisciplinary technology that integrates multiple disciplines, there are still many problems in the development of BCI technology. This requires more scientific and technological workers to devote to in-depth research. In order to promote the development of BCI technology, this paper makes a detailed review of the principle, concept, key technology, and development of BCI on the basis of consulting relevant materials, and discusses the existing problems and possible development directions in this field.

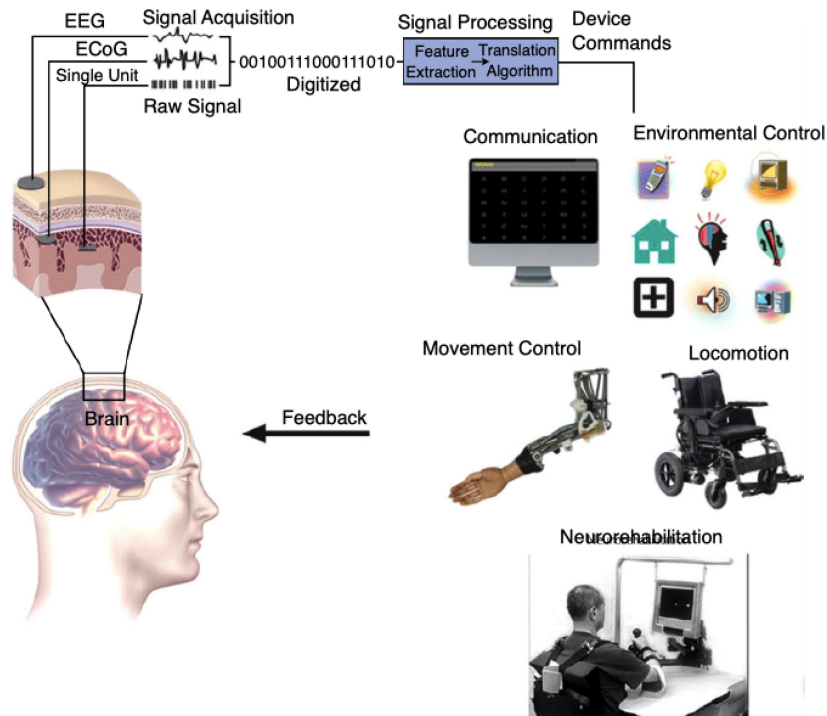


Figure 4.1: The brain computer interface [42]

4.1.2 Problem Statement

4.1.2.1 The potential of BCI in medicine

BCI is currently used in multiple therapeutic areas, including epilepsy, Parkinson's disease, depression, Alzheimer's disease, etc. At the same time, BCI has been well applied in the field of medical rehabilitation and has become an important way to improve the quality of life of patients. As a new medical treatment technology, BCI has become the most rapidly developing discipline in the field of medical science in the past two decades. Along with technological breakthroughs, the industry has developed very rapidly. It is estimated that the global BCI market will exceed 9 billion US dollars by 2024.

4.1.2.2 The potential of BCI in non-medical fields.

The concept of the metaverse is very popular recently, and BCI will be an integral part of the metaverse. The BCI uses the mind to control the game, which may realize more free operation. In the metaverse, players can freely move every part of the body of their own will, the software no longer needs rigid preset actions, and players can interact with the virtual world as they wish. In terms of interaction, using BCI can not only get rid of the "shackles" of preset actions but also the feedback of multiple senses will become possible through the bidirectional transmission of brain signals.

4.2 Related Work

This section reviews the background of the BCI. By reading this section, readers will better understand the subsequent articles.

4.2.1 Brain and nervous

The brain is the central organ of the nervous system of all vertebrates and most invertebrates. In life, we often compare the brain to a general-purpose computer. However, this concept is far beyond the truth because of the complexity of the brain. The brain is a set of sub-systems that cooperate to control the whole human body and its functionalities.

The brain could be generally divided into two main parts - the cerebral cortex and sub-cortical regions. The cerebral cortex is the outermost layer of the brain. It is related to our mental activity. The subcortical structures are a group of diverse neural formations deep within the brain which include the diencephalon, pituitary gland, limbic structures, and the basal ganglia, which control the basic and vital functions such as heart rates, body temperature respiration.

The nervous system is your body's command center. It is divided into two main parts which are central and peripheral systems. Originating from your brain, it controls your movements, thoughts, and automatic responses to the world around you. It also controls other body systems and processes, such as digestion, breathing, and sexual development (puberty). The nervous system uses specialized cells called neurons to send signals, or messages, all over your body. These electrical signals travel between your brain, skin, organs, glands, and muscles. Different kinds of neurons send different signals. For example, sensory neurons take information from your senses and send signals to your brain.

4.2.2 Brain waves

The brain is composed of two types of cells, i.e., neurons and neuroglia, which have different functions. The neuroglia's function is to fix neurons and provide neurons with nutrients and oxygen. The neurons are responsible for the information transfer through

chemical and electrical impulses, called nerve impulses. This constant flow of electrical current in the brain is called brain waves.

The simplest way to measure brain waves in EEG, which is non-invasive and simple in nature. However, the captured signals are very weak and not effective, because they need to cross several layers of tissues (the skull, the meninges, and the scalp). To solve this problem, it is often necessary to use several electrodes to have a higher spatial and a more precise system.

There are five kinds of the band of brain waves, i.e., delta (1-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (13-30 Hz), and gamma (30-150 Hz) waves, classified according to its frequency. Each band is associated with certain mental states.

4.2.3 Event-related Potential

The event-related potential is a special type of brain-evoked potential that intentionally impart special psychological meaning to stimuli. It reflects the neuro-electrophysiological changes (cognitive potentials) of the brain during cognitive processes. Cognitive potentials are brain potentials recorded from the surface of the skull when people are cognitively processing a topic. The main components of classical ERP include P1, N1, P2, N2, and P3, of which the first three are called exogenous components, and the latter two are called endogenous components. The main characteristics of these components are: firstly, they are not only the embodiment of the pure physiological activity of the brain, but also reflect some aspects of the psychological activity; secondly, their elicitation must have a special stimulus arrangement, and more than two stimuli Or a change in stimulus. Among them, P3 is one of the most concerned and researched endogenous components in ERP, and it is also the most important indicator for polygraph detection.

4.2.4 Machine Learning

Machine learning is an approach that empowers machine learning to do things that cannot be done by direct programming. But in a practical sense, machine learning is a way of using data, training a model, and then using the model to make predictions. Machine-learning techniques can eventually identify how the brain's neural network maps correlate with those specific thoughts or emotions.

4.2.5 Our Contribution

In this section, we compare our work with other similar work. The results are shown in the table below.

	Our Paper	Brain-Computer Interface Software: A Review and Discussion	Brain-Computer Interface Review	Brain Computer Interfaces, a Review
History	Yes			
Type	Yes	Yes		Yes
Technology	Yes	Yes	Yes	Yes
Application Scenarios	Yes	Yes	Yes	Yes
Bibliometric Analysis		Yes		
Risk	Yes			
Future	Yes		Yes	Yes

4.3 Types and Designs of BCI

4.3.1 Reasons for Categorization

As a result of various technologies used and techniques applied, the field of BCI is quite diverse. To avoid chaotic nomenclature, different schemes for the classification of the tools have been proposed. These are based on the distinction between their' usage cases, functions, or technical design. In this article, we will mention three such categorizations.

4.3.2 Passive, Active and Reactive

One of the ways to categorize a BCI system is to capture in which way it uses the brain (Simanto et al., 2021 [13]). This classification separates BCIs into passive, active, and reactive.

Passive BCI relies on decoding the brain's unintentional affective/ cognitive states (Zander et al., 2009 [14]). Such a system works partially autonomous, as it analyses the user's emotional and mental state, to provide feedback (Kawala-Sterniuk et al., 2021 [16]). A great example of a passive BCI would be the application in detecting driver's drowsiness to prevent road accidents (Lin et al. [18], 2008; Gao et al., 2019 [19]). It is also a rather emerging trend in the BCI field, as traditional BCI systems would be predominantly active or reactive [16].

Active BCI (also called independent as in (Lebedev MA, Nicolelis MAL., 2016 [17])), on the other hand, is controlled directly through a specific mental task performance [16]. It requires the user to visualize an actual movement so the system can then transcribe it accordingly, or incorporate the act such as blinking to provide the system with input [16].

The third type would be the reactive BCI (also called dependent [17]). It is based on brain signals that the human brain produces in response to external stimuli such as visual or auditory stimuli. These stimuli create a natural response to an event, called Event-Related Potential [16], a phenomenon greatly used in the research.

4.3.3 Motor, Sensory, Sensorimotor, Cognitive and Brainets

Another possible distinction is the one referring to the physiological function BCI is intended to emulate [17]. The ones aiming at providing control over a given limb or general body navigation are the motor BCIs. Sensory BCIs provide the reproduction of sensations and accordingly, sensorimotor BCIs couple these two functions. The role of cognitive BCIs is to enable higher-order brain functions, such as memory (Berger TW et al., 2011 [20]), attention (Fuchs T et al., 2003 [21]; Lubar JF, 1995 [21]), and decision-making (Hasegawa RP et al., 2009 [20]; Musallam S2 et al., 2004 [21]) [17]. Another BCI class proposed in [17] is the Brains, which requires the participation of multiple subjects, as such a system operates on a combination of the electrical activity of multiple brains concurrently.

The major caveat of the aforementioned classification is its binding to the traditional brain labeling, which divides its parts into the motor, sensory, or higher-order [17]. Such classification is, according to present knowledge, outdated as the brain areas actually serve as multi-function units, and there is an abundance of literature supporting this claim (Brovelli A et al., 2004 [25]; Fetz EE, 1992 [26]; Lilly JC, 1956 [27]; Nicoletis MA, Lebedev MA, 2009 [28]; Rolls ET, 1989 [29]).

4.3.4 Non-Invasive, Semi-Invasive and Invasive

Classification in terms of signal acquisition divides BCIs depending on the level of invasiveness they demonstrate. The distinction is rather straightforward. Systems that require surgery to open up the scalp of the subject would be classified as invasive, with the separation between semi-invasive and invasive being the further penetration of the brain tissue in case of a fully invasive approach, e.g. placing an electrode recording grid on the brain's surface - electroencephalography (ECoG) - in the case of semi-invasive approach. The rest of the systems, i.e. the ones that do not invade the user's body, would be classified as non-invasive BCIs. This quite substantial group of BCIs would employ a range of technologies such as functional near-infrared spectroscopy (fNIRS), functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG), positron emission tomography (PET), functional transcranial Doppler (fTCD) and most predominantly electroencephalography (EEG).

Classification according to invasiveness is important for a variety of reasons, the major one being a safety concern [17]. As the well-being of the patient is of utmost importance, the risks associated with each application have to be clearly outlined. The prevalence of non-invasive BCIs stems mainly from the (mostly) devoid of risk approach. For example, in the case of EEG recordings, the electrodes are simply placed on the scalp surface (Niedermeyer E et al., 2005 [30]) through an easy and safe procedure, particularly if dry sensors are used (Chi YM et al., 2010 [31]; Fonseca C et al., 2007 [32]; Gargiulo G et al., 2010 [33]; Guger C et al., 2012 [34]; Taheri BA et al., 1994 [35])[17].

In the case of an invasive approach, there are multiple factors that influence the safety of the application. First is the specific procedure of opening up the skull and inserting the BCI, which in itself carries the risk of tissue damage or infection, specifically when

the implant consists of extracranial parts (Bouton CE et al., 2016 [36]; Collinger JL et al., 2013. [37]). After the application of the system, there still exists a possibility of its rejection by the subject's body and the development of conditions impairing the brain functions [17]. Extended use of the implant also introduces additional risk wired to the movement of the subject's body, which in turn causes irritation through micromovements of the inserted BCI [17].

With each of the solutions come its advantages and disadvantages. In the case of non-invasive BCI prioritized is the safety and ease of application over precision. Due to the lack of direct contact with the brain, we sacrifice the clarity and quality of the signal (Lebedev MA, Nicolelis MA, 2006 [38]). In the case of EEG, we have to deal with a signal that comes as a filtered brain activity of many millions of neurons [17], which is highly lacking in spatial resolution and is subject to many interferences originating from other parts of the body and the environment.

Invasive BCIs, in contrast, provide neural-level recording and stimulation, at the cost of in-brain intervention. Such granularity enables monitoring ranging from single neurons to large neuronal populations, which makes it possible to record the action and local field potentials (LFPs) [17]. Another major improvement over non-invasive BCI is the ability to stimulate the desired area to influence sensory, motor, or cognitive processing (41,45,48,146,704,705,719,770). Although the solution provides the best possible interaction with the brain, due to its cost both in financial and health terms, we see the application mostly in patients with specific disorders or animals.

4.4 BCI Technology

4.4.1 BCI System Structure

The BCI cycle begins by observing the user's behavior in regards to a provided stimuli, or its lack when trying to monitor passively. The sensors gather the signals collected from the brain activity and these signals are then adequately pre-processed to remove noise before further processing [16] [40]. The succeeding step is to extract relevant features for the classifier to use. Classification is a critical step in any BCI system [16], therefore has to be carefully suited for a given task. Some of the techniques used are linear discrimination analysis (LDA) (F. Lotte et al. 2018 [41]) in the case of P300 and MI-based BCIs, canonical correlation analysis (CCA) for SSVEP, and fuzzy neural networks (FNNs) for analyzing the mental state of the subject [16]. The outcome can then be used as input for controlling external devices or collecting data on the user's state. The cycle is closed by providing the user with feedback, which may come in different forms, and is generally tailored to the specific task at hand. The most common way to present it is through a computer monitor, due to its common availability and convenience [16] [41]. One thing worth noting is that further cycles of the interaction of the user with the computer present the possibility of improved performance of the system through machine learning and the user getting familiar with the system.

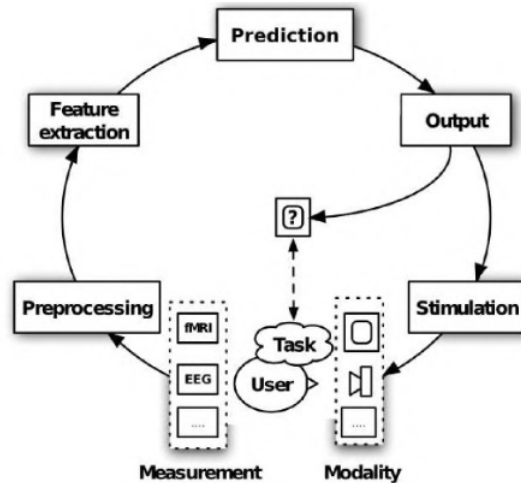


Figure 4.2: The BCI Cycle [40]

4.4.2 Signal Acquisition

The action of acquiring a signal from the subject's brain is available thanks to the various existent measuring techniques. Depending on the measuring technique used, different BCI paradigms are applied and different designs are used. To further investigate differences between devices used for signal acquisition, technical features of the most common devices will be presented and analyzed.

4.4.3 Device Specifications

Although major characteristics affecting the popularity of a device regard the safety and cost, there are also multiple other features affecting the device's attractiveness. Amongst the most important technical factors are spatial and temporal resolution. Spatial resolution is a term to describe the precision with which the measurements are taken with regard to the area covered. Accordingly, the temporal resolution is the description of the frequency with which the device collects data. These aspects are crucial to the outcome of the measurement, as they determine the degree of precision and reliability of extracted data. Factors like portability, amplitude, and noise resistance also play a significant role when considering the value of a given device [16].

4.4.4 Non-Invasive BCI

4.4.4.1 EEG

EEG-based BCIs are the most prevalent systems that have been thoroughly studied on a range of human subjects [17] [16]. It provides the ability to record the brain activity from the surface of the scalp by placing electrodes in a grid to record electrical current

TABLE. Summary of Current Sensory Modalities Used in BCI							
	Invasiveness	Spatial resolution	Temporal resolution	Depth of brain	Maintenance	Portability	Procedure needed
Implanted micro-electrode	Invasive (intracortical)	Single unit: 0.05 mm Multi-unit: 0.10 mm Local field potential: 0.50 mm	3 ms	0-5 neurons around electrode (1 mm-3 cm)	Moderate/high: periodic recalibrations or replacement; LFP recordings are most stable	Moderate: single neuron electrode recordings sensitive to changes in position	Craniotomy
ECoG	Moderately invasive (subdural or epidural)	1.0 mm	5 ms	0.5 mm-3 mm	Low/moderate: periodic recalibrations; inflammation still a long-term consideration	Good: fully implanted wireless systems available	Craniotomy
Intra-vascular electrode	Minimally invasive	2.4 mm	5 ms	260-680 um	Low: robust signal stability over time and in moving subjects	Good: fully implanted wireless systems available	Groin puncture
EEG	Noninvasive	10 mm	50 ms	1 mm-3 cm	Low/moderate: requires frequent gel reapplication	Good: fully implanted wireless systems available	None (or skin incision if implanted on skull)
fMRI	Noninvasive	1 mm or better	1 s	Whole brain	High: frequent machine maintenance	Poor: equipment is bulky	None
fNIRS	Noninvasive	1 cm	1 s	1-3 cm	Low: robust signal acquisition	Good: wearable systems available	None
ftCD	Noninvasive	1-3 cm	1-5 s	80 mm	Low: robust signal acquisition	Good: wearable systems available	None
MEG	Noninvasive	1 mm or better	1-5 ms	Whole brain	High: frequent machine maintenance	Poor: equipment is bulky	None
PET	Noninvasive	3-5 mm	50-100 s	Whole brain	High: frequent machine maintenance	Poor: equipment is bulky	None

Figure 4.3: Summary of the most popular sensors applied in BCIs [42]

fluctuations generated by the brain (Niedermeyer E., da Silva F.L., 2004 [44]). Although the signals collected by EEG are prone to noise contamination and provide lower quality recordings than the intracranial solutions, its portability, ease of use, low cost, and still decent temporal ($\approx 0.05s$) and spatial ($\approx 10mm$) resolutions lead to the popularity of this solution. EEG-based BCIs often rely on one of the following paradigms [16]:

- ERP - event-related potentials (P300 and other components),
- ERD - associated with motor imagery (MI),
- SSVEP - steady-state visual evoked potentials,
- ASSR - auditory steady-state response,
- SCP - slow cortical potentials,
- SMR - sensorimotor oscillations,
- hybrid systems (based on more than one input signal)

The most prominent out of the aforementioned would be the ERP, with the P300 being specifically popular due to its low requirement in terms of training the subject and its high sensitivity to reaction in "oddball" scenarios (Alexander JE et al., 1995 [67]). Another popular design would be the employment of SSVEP, in which the subject would focus on one of the flickering objects, typically displayed on the screen, to provide the system with adequate input [17].

The convenience of use also lead to the expansion of the EEG headsets outside the scientific field and onto the general public, where they would find use cases in the entertainment and educational industries.

4.4.4.2 MEG

Magnetoencephalography employs magnetic induction to measure the magnetic activity of the brain (Chin-Teng Lin et al., 2020 [68]). It detects magnetic fields of magnitude in the order of picoTeslas, which are generated by populations of cortical neurons [17]. The required sensitivity of measurements implies the usage of precise magnetometers, alongside appropriate shielding to diminish the impact of external magnetic signals. This at the same time acts as one of the major disadvantages of MEG, as the required equipment is expensive, large, and not portable (Jerbi K. et al., 2011 [69]). To combat this issue, there has been intensive progress in the development of portable brain scanners - mobile MEG [70] - which are promising to become consumer-available in the next years (Yang et al., 2019 [71]; Boto et al. 2018 [72]).

4.4.4.3 fMRI

Another signal acquisition method is functional Magnetic Resonance Imaging. The data is collected from the blood oxygen level-dependent (BOLD) activity of the brain (detecting changes in blood flow) by using the MRI scanner (Sitaram R et al., 2007 [73]). Similarly to MEG, the fMRI presents itself with a lack of portability and high cost, alongside challenging data analysis and even patient discomfort due to the operational noise (Yoo, S.S. et al., 2004 [74], Sorger, B. et al., 2020 [75]). It also has a low temporal resolution (1-2s), with the additional 3-6s lag between neuronal activity and BOLD response [17]. The spatial resolution on the other hand is high and allows for monitoring of the whole brain activity, with an option to target a specific area as the source of the signal. Additionally, as the data from BOLD and EEG seem to be correlated, there are systems in which those technologies get combined to obtain better results (Weiskopf N. et al., 2004 [76]).

4.4.4.4 fNIRS

Functional near-infrared spectroscopy measures the brain's hemodynamic activity by detecting changes in oxyhemoglobin (HbO) and deoxyhemoglobin (HbR) in the blood flow [16]. To extract this data it applies near-infrared light in the range of 600-1000 nm through the skull, but can only measure regions in close proximity to the cortical surface, as opposed to fMRI. Though spatial (≈ 1 cm) and temporal (100-1000 ms) resolutions are not the best, with additional several-second delays between neural activity and blood oxygenation, its low cost and portability make it an appealing solution. Similar to fMRI, fNIRS is also employed in hybrid systems like EEG-fNIRS BCIs (Hasan, M.A. et al., 2020 [77], Nazeer, H. et al., 2020 [78]).

4.4.5 Semi-Invasive BCIs

4.4.5.1 ECoG

Electrocorticographic grids are an approach in which the scalp is penetrated, but the brain tissue remains undamaged. They often contain several tens to several hundred electrodes, while allowing minimally invasively recordings of multichannel field potentials from large cortical territories (Crone NE et al., 2006 [64]; Hill NJ et al., 2012 [65]). An emerging trend in this approach has been to miniaturize the electrodes while increasing their density [17]. Additional electronics hardware embedded in the grids also proved to increase the quality of recordings (Viventi J et al., 2011 [66]).

4.4.6 Invasive BCIs

Invasive BCIs are the most accurate type of devices thanks to their proximity to the brain tissue and technological advancement. They generally consist of sets of micro-wires organized in a specific manner, tailored to the task. These micro-wires are generally composed of silicon, platinum, tungsten, or iridium and coated with glass, teflon or polyamide [17]. Each wire's diameter is generally on a micrometer scale and the length is determined by the location desired to target. One of the main problems associated with the usage of micro-wire technologies is its susceptibility to tissue growth and material deterioration (Rousche PJ et al., 1998. [50]; Saxena T et al., 2013 [51]).

The most prevalent invasive technologies using the micro-wires are the micro-wire recording cubes, microelectrode arrays (MEAs), and deep brain stimulation (DBS).

4.4.6.1 Utah Array

Utah array is a microelectrode array that consists of 100 rigid micro-electrodes in a fixed 10x10 arrangement (Campbell PK et al., 1991[45]). The needles' shafts are coated with polyamide and their sharpened tips are coated with platinum. The distance between neighboring needles is 0.4mm. It currently is the only United States Food and Drug Administration (FDA) approved micro-electrode array implant for human use [16]. It has been employed in the number of clinical studies (Normann RA et al., 2009 [46]; Jaroch DB et al., 2009 [47]; Hochberg LR et al., 2012 [48]). Biocompatibility issues have been reported with regard to the usage of Utah arrays. They indicate that implants may cause tissue reactions, such as microhemorrhages, microglia activation, and long-term inflammation with the level of severity depending on the tissue damage during the implantation (Fernandez E et al., 2014.[49]).

4.4.6.2 Michigan Probe

Another interesting micro-electrode design is the Michigan probe, which allows for recordings alongside the shaft of the planar electrodes, as opposed to only at the tip. The size

of the grid is configurable and allows recordings from local populations as well as from different cortical and hippocampal layers at the same time (Mizuseki K, Buzsaki G, 2014 [52]).

4.4.6.3 Micro-wire Recording Cubes

Recording Cubes differ slightly from previous examples due to their implementation of guiding tubes. The design is based on a 10x10 grid of tubes with 1mm spacing between each. Each guiding tube contains 3-10 micro-electrodes, with possible variations in length. Therefore the total number of micro-wires might amount to 1000. With 1-2 individual neurons being recorded by each of the micro-wires, a sample of 1000 to 2000 neurons can be monitored by a single cube [17].

4.4.6.4 Deep Brain Stimulation

Deep Brain Stimulation is a design involving long one or two electrode wires inserted into the brain, alongside a computer(neurostimulator) implanted on the chest and wired to the electrodes. It's mostly used in the treatment of movement disorders (Kringelbach ML et al., 2007 [53]) and conditions such as obsessive-compulsive disorder (OCD) and epilepsy. Currently, the primary manufacturer of the DBS systems is Medtronic (Christine A. et al., 2017 [54]). These devices are capable of dual-channel stimulation with a frequency range of 2-250Hz, a pulse width of 60 to 450 ms, and an amplitude of 0.0 to 10.5 V [54]. DBS is also approved by the United States Food and Drug Administration (FDA).

4.4.6.5 Neural Dust

Neural dust is a recording method that utilizes small (10-100 micrometer) sensors ("dust") that detect extracellular neuronal potentials and communicate them via an ultrasonic link to an interrogator placed under the skull (Seo D et al., 2015. [63]). Each of the sensors consists of a set of electrodes to record neuronal activity, metal-oxide-semiconductor (CMOS) circuitry for the amplification of the signal, and a piezoelectric transducer for the conversion of electrical potentials into ultrasound. Interrogator-produced ultrasound is used to power the dust, as well as to examine its state. The interrogator is also the component that communicates with the extracranial components [17].

4.5 BCI Application Scenarios

4.5.1 Medical Applications

Because the BCI technology can directly realize the interaction between the brain and external equipment and span the conventional brain information output pathway, there

is a broad application prospect in the medical and health field. This field is the largest market application field in the current brain interface and the fastest-growing field.

4.5.1.1 The application of BCI in the diagnosis and treatment of disabled people

The goal of BCI in the diagnosis and treatment of disabled people is to improve the current state and improve the quality of life through the auxiliary treatment of this technology.

Specifically, there are two main ways to apply the BCI. One is the auxiliary BCI, which refers to obtaining the patient's movement intention through the BCI devices to realize the control of external devices such as prostheses or exoskeletons. For example, the BrainRobotics prosthetic hand was selected as the best invention of 2019 by TIME. This prosthesis based on BCI is different from the traditional sleeve prosthesis. The new system is a nerve-muscle-bone prosthesis. It can directly connect with the nerves and muscles of the limbs, and users can control it with their brains. The second is the rehabilitation BCI. Because the central nervous system has plasticity, the BCI device can directly stimulate the brain, and then restore the connection between neuronal synapses. Italian Percro Laboratory [81] proposed a full upper-limb exoskeleton for motor rehabilitation of reaching, grasping, and releasing in post-stroke patients. They show the potential of the proposed system for being introduced in a rehabilitation protocol.

4.5.1.2 The application of BCI in the diagnosis and treatment of consciousness and cognitive impairment

A chronic disorder of consciousness includes two levels: vegetative state and micro-conscious state. In recent years, the research on BCI technology has gradually increased in the field of chronic disorder of consciousness diagnosis and treatment. Obtaining and analyzing the patient's brain-electrical signal through the BCI devices can grasp the patient's conscious state, realize the diagnosis and evaluation of consciousness disorder, and judgment, and even communicate with patients with conscious disorders. The P300 paradigm is often used. Specifically, the patient's name, photo, and other information are usually used as target stimulation [80]. The BCI devices obtain the patient's stimulus of the brain electrical signal and analyze the patient's status. Some patients may have a specific reaction to the target stimulus. This "brain and electrical communication" helps doctors to determine whether the patient has the possibility of awakening recovery.

In terms of cognitive impairment, the early symptoms of Alzheimer's disease were found through the test of the patient's cerebral waves. A paper [79] by Professor Li-Huei Tsai's team at the Massachusetts Institute of Technology found that light and sound stimulation produced beneficial brain waves in the brains of mice, leading to improved cognition and memory. This non-invasive therapy offers the possibility of treating Alzheimer's disease with brain waves.

4.5.1.3 The application of BCI in the diagnosis and treatment of mental illness

Compared to other physiological signals, EEG signals can provide more in-depth, realistic emotional emotions than other physiological signals. Through learning algorithms, EEG features can be extracted to discriminate a wide range of emotions (such as pleasure, sadness, calmness, anger, fear, surprise, anger, etc.). sadness, calmness, anger, fear, surprise, anger, etc.). Therefore, EEG-based emotion recognition studies can be used to assist with depression. EEG-based emotion recognition research can be used to assist in the study and treatment of the pathogenesis of depression, anxiety disorders, and other psychiatric disorders. In addition, it can be used in the rehabilitation of mental diseases. In addition, in the rehabilitation of mental disorders, neurofeedback training based on the brain-computer interface can be used in the treatment of depression and anxiety disorders. In addition, brain-computer interface-based neurofeedback training can play an active role in the treatment of depression and anxiety disorders. Although neurofeedback predates brain-computer interfaces, it is essentially the earliest application of brain-computer interfaces. Although neurofeedback predates brain-computer interfaces, it is essentially one of the earliest applications of brain-computer interfaces.

4.5.2 Military Applications

The military has often been the driving force in technology development and it certainly will play a role in the BCI development. Although the military sector generally remains hidden from the general public, an assessment of the potential direction in which the advance will steer has been conducted by The RAND Center for Global Risk and Security [55]. Through a game convening technical and operational experts, tested was the potential utility of BCI application in the military field against the arising obstacles.

The results have suggested that BCI technologies will likely be of practical use on the future battlefield, as the volume of human-machine interaction increases. The major advantages of the application of BCI seem to be connected to the existent bottleneck between a soldier and his ability to communicate and operate technological interfaces faster, and improving situational awareness during the mission. The report has also outlined the potential risks associated with BCI technology deployment. These would not only consist of operational risks, but also new areas of an ethical and legal risk. One of the most appealing use cases shows to be the incorporation of the brain to brain interfaces (BTB), but concurrently it also carries the greatest operational and institutional risks. Another obstacle to the implementation could be a possible lack of trust in the BCI technology amongst personnel, both lower and higher in the hierarchy, stemming from an aversion towards used technology, or in fear of possible exploitation.

There has also been conducted a study in which participants would enact a military mission scenario, with real-time stress-level acquisition provided to the commander via telephone application and emotional levels detected by Neurosky EEG headset (Md Mostafizur Rahman, 2021 [60]). Results showed that the BCI system almost correctly (98%) classified the stress status of the participants. Additionally, field interviews showed consensus regarding the usefulness, ease of use, and portability of the applied solution.

In regards to improved machine operational skills, there have been attempts at controlling a drone with a consumer-grade EEG headset (S Rosca et al., 2018 [56]; Peining et al., 2017 [57]). Attempts showed the possibility to control drones in up to 2 dimensions with efficacy ranging from 54% (Muse ICE) to 71% (Emotiv ECE) depending on the headset used and the user operating the device [57]. Taking into account the fact that these were consumer-grade headsets, there certainly is a potential for applications in this field.

4.5.3 Entertainment Applications

With the rising availability of consumer-grade BCIs, there has been increasingly more interest in BCI outside of the medical field, which it had mainly been applied in. One particular area that is gaining more evidence due to the arrival of consumer-grade devices is that of computer games, as it allows more user-friendly applications of BCI technology for the general public (Gabriel Alves et al., 2020 [59]). Commercial BCIs' use reaches back to the late 2000s [59] when predominantly EEG headsets would become relatively popular equipment for the game development. It was mostly due to the low-cost, ease of acquisition, and the abstraction they would present. Developers did not have to interact with low-level implementations of the system, but rather build onto existing functionality. Moreover, the emergence of wireless devices would upgrade comfort, an element of significant importance in the consumer market [16] [59]. The current price of available headsets starts from as low as 99\$, which makes it even more approachable [16].

Manufacturer	Wearable	Sensors Type	Channels Amount	Sampling Rate	Data Transfer
Neurosky	YES	Dry	1	500 Hz	Bluetooth
Emotiv	YES	Wet/Dry	5-32	500 Hz	Bluetooth
OpenBCI	YES	Wet/Dry	8-21	250-500 Hz	Bluetooth
ANT Neuro	YES	Dry	32-256	<16 kHz	Wi-Fi
g.tec	YES	Wet/Dry	8-256	500 Hz	Cable/Wi-Fi
Cognitionics	YES	Dry	8-128	>2 kHz	Bluetooth
CREmedical	YES	Wet	20	500 Hz	Cable
interaXon	YES	Wet	4-7	250 Hz	Bluetooth
Cognitionics	YES	Wet	8-128	<2 kHz	Bluetooth

Figure 4.4: Some of the most popular inexpensive, off-the-shelf BCI systems-summary table [16]

The implementation of BCI into computer games mostly refers to changes in input. Instead of a traditional mouse and keyboard, the player is required to use one of the BCI control paradigms, or in some games player's mental or emotional state is examined [58]. The design of BCI games is mainly restricted by factors such as omnipresent noise, especially in the case of VR games where body movement is required, low transfer rate, and most evidently, relatively slow decision making when compared to traditional input controls (Ferreira et al., 2014 [61]).

An analysis [59] of the 87 articles on BCI games has provided insight into statistical game characteristics. The majority of games implemented attention (40%) as the main control signal, with the use of meditation at 21%. Action dominated the game genre category

with 48% (puzzle 12%, simulation 9%). The major purpose occurred to be training with 38% (research 17%, entertainment 11%).

One of the notable applications outside the game industry could be its appearance in the cosplay space [62]. A collaboration between a Dutch designer Anouk Wipprecht, Institute for Integrated Circuits at Johannes Kepler University and neurotechnology company, and g.tec medical engineering GmbH would be the origin of the first BCI dress, namely Pangolin Scales BCI+[62]. The creation would make a presentation at the Ars Electronica Festival bringing the BCI even closer to the general public.

4.6 Risks Associated with BCI Usage

Risk is an important part that cannot be ignored in the development of BCI technology. At present, the risks of BCI technology mainly focus on technical security risks, privacy risks, autonomy risks, and fairness risks. In response to these risks, in the research, development, and promotion of brain-computer interface technology, we should place risk issues in a more prominent position, evaluate risks scientifically and objectively, and avoid unnecessary harm to individuals and society.

4.6.1 Technical Security Risks

The application of brain-computer interface technology should first pay attention to the safety of the technology itself, especially when surgery may be involved. Safety must be placed in the highest position. Figure 4.5 presents a summary of the classification studies, also including the different families included in each classification. The most prominent security risk of current BCI technology is that invasive BCI technology may cause harm to people's physical and mental health. These risks include surgical trauma to the human brain, bleeding, infection, etc [88]. In addition, some potential risks also require our high attention. For example, implants may lead to some unintended consequences, such as changes in people's emotions, personalities, etc. The long-term use of implanted devices such as brain chips may also cause rejection by the body and cause unnecessary harm. At the same time, these devices may gradually fail over long periods of use., and the weakened function may lead to operation errors in the BCI system. In addition, the mutual movement between the brain tissue and the implant may also cause changes and damage to the function of the brain tissue [87].

Compared with invasive BCI technology, non-invasive BCI technology is easier to accept, but its technical security issues cannot be ignored either. First, the risk of injury from the misreading of EEG signals has not received enough attention. Due to the blocking effect of the skull on nerve signals, it is difficult for non-invasive BCI devices to obtain continuous and stable EEG signals, and the possibility of indirect harm to the human body due to misinterpretation increases [86], especially for some devices that require high Brain-controlled actions of precision operation, such as controlling a wheelchair to cross the road, driving a car, etc.. Second, the long-term risks of wearing electrode caps are

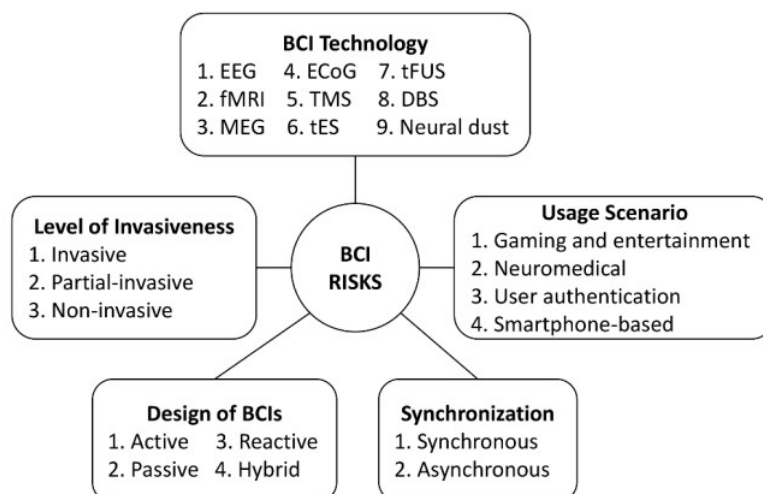


Figure 4.5: BCI classifications studied and their associated families[82]

easily ignored. People may have headaches, fever, and blurred vision after using the electrode caps [85], and the long-term wear of electromagnetic radiation is also worthy of attention. Although the safety of some devices has been verified in medical treatment, there is a big difference between medical use and daily life use. A patient needs to use few times for medical treatment, but for a person who uses non-invasive BCI in daily life, the damage of electromagnetic radiation cannot be ignored. In addition, due to the plasticity of the human brain, some scholars also worry that the long-term use of brain-computer interface devices by teenagers and even adults who are still in the developmental stage will cause irreversible changes in the human brain, resulting in unknown negative effects [86].

4.6.2 Privacy Risks

Since the BCI devices need to extract information from the brain, the extracted information may be abused or leaked, so there is a huge risk of privacy violation. Compared with other emerging technologies, the potential invasion of privacy by BCI technology may be manifested in the following two aspects. First, collecting and processing neural information without permission. Neural information stored in digital form may be more likely to be quietly collected by hackers or illegally authorized companies [84]. In this Internet age, the consequences of this neural information privacy leak are even more serious. In theory, others can directly obtain information in the human brain remotely through the BCI system connected to the network, and it can be widely and rapidly disseminated directly through the Internet, which is potentially harmful to privacy.

If man-made information leakage and propagation can be prevented and controlled by improving relevant work systems, then information leakage caused by insecurity defects in the system itself is more difficult to prevent and control. These problems will lead to a new phenomenon: BCI devices may make wireless control or remote monitoring of the brain become a reality. Second, it leaks people's thinking information. Thinking

information may be the most private of all people's information. If neural information can be accurately interpreted by brain-computer devices, then human thinking process and thinking results may be inferred, and human thinking may be leaked or stolen. Collecting and analyzing human thinking information brings more worries. Although brain wave signals cannot be directly represented by the human brain. The changes in brain waves are likely to show the user's emotional reactions such as emotions, and various tendencies and preferences. After interpretation and analysis, it is possible to infer personal interests, hobbies, emotions, etc., especially in marketing. On the other hand, merchants may use BCI technology to place advertisements in a more targeted manner, judge customers' consumption preferences and purchase intentions, and then analyze their consumption behavior and consumption ability [83], which may lead to new unfair transactions or intentional inducement of consumption and so on.

4.6.3 Decision-making Autonomy Risks

Whether we will lose self-control by using BCI devices in the future is a question of great concern. In theory, in a BCI system, the human brain can transmit command information to the outside, and it is also possible to receive command information from the outside. Such a design could serve as an "error correction mechanism" for the device, protecting the user's safety in times of danger. For example, a smart wheelchair may adopt a similar design to avoid danger. When a person sits in a wheelchair and uses his mind to control a left turn, it just detects that there is an obstacle or danger on the left, and this signal is sent back to the brain to prevent the brain from making a left turn. However, the controversy here is, for such a device, how can human autonomy be manifested? Is the device controlling the human or the human controlling the device?

Generally speaking, an autonomous person should be able to decide actions according to his or her desires, values, emotions, and plans. If these are inconsistent with the person's actions, then the person may lack autonomy. In this sense, the BCI device changes people's thinking commands by transmitting signals in reverse, to avoid wrong and dangerous operations. Such actions may cause users to be unable to act according to their wishes, thus causing people to lose self-control. However, if the device simply blocks the transmission of signals from the human brain to the external device, then the user is still autonomous in a general sense. In addition, the BCI may also be used by third parties to manipulate people, such as through forced shopping and psychological suggestions. These seem to indicate that the BCI device may potentially violate the user's autonomy.

4.6.4 Social Equity Risks

The application of BCI technology may pose the risk of social injustice. This unfair problem is mainly manifested in two aspects, one is how to use BCI technology fairly, and the other is how to obtain BCI technology fairly. In terms of how to use BCI technology fairly, people are mainly worried that the use of BCI will increase the unfairness in the competitive field. For example, in competitive sports, some people may use external

devices of BCI to obtain better physical fitness, endurance, etc.. Users can achieve better results than non-users, resulting in unfair competition. In education, BCI devices may have the potential to enhance cognitive function. If users use BCI devices to improve their level of concentration and cognitive load, they may gain advantages. But more people worry that such equipment may allow students to be monitored everywhere, and bring more unfairness to the competition. Users will gain advantages that others do not have, and non-users can only make up for this by being forced to use them.

In terms of equitable access to BCI technology, the main concern is the accessibility of the device. Not everyone in society has the ability to pay for these devices, which may make some people, especially low-income people, unable to use BCI devices in the beginning, and unable to share the results and benefits of BCI technology. Therefore, the use of BCI technology may eventually become a special treatment by some rich people and cannot benefit the public. At the same time, BCI is not similar to other high-tech products. The knowledge and skill level of potential users will also prevent the BCI device from being accessible and universally beneficial to more people, such as the elderly. Some elderly groups may be affected by the emergence of BCI. Those without BCI devices may be at a competitive disadvantage in life, education, etc., creating social stratification or exacerbating disparities between people. In the absence of government regulation, our society will face more equity challenges.

4.7 Future Development of BCI

4.7.1 Major Hurdle

4.7.1.1 Bandwidth

The human cerebral cortex contains about 14-16 billion neurons, while the cerebellum contains about 55-70 billion neurons. At present, there have never been more than a couple of hundred electrodes in a human brain, which means very low data volume. When it comes to vision, that equals a super low-res image like last century's pixel games. In this case, higher bandwidth is needed.

As the graph shows in [89] called Stevenson's law. This research suggests that, in good approximation, the number of recorded neurons has grown exponentially since the 1950s, doubling every 7 years. Just like the Moore's Law. Ian Stevenson and Konrad Kording came to an approximate conclusion [90]. If this pattern continues, it will take us till the end of this century to reach a million and until 2225 to record every neuron in the brain, which is unbearable.

One potential strategy is to shrink the size of electrodes so that bioengineers can cram more of them into the brain at one time, the other is to think of completely new approaches. In conclusion, this will be a problem that scientists need to solve in the future.

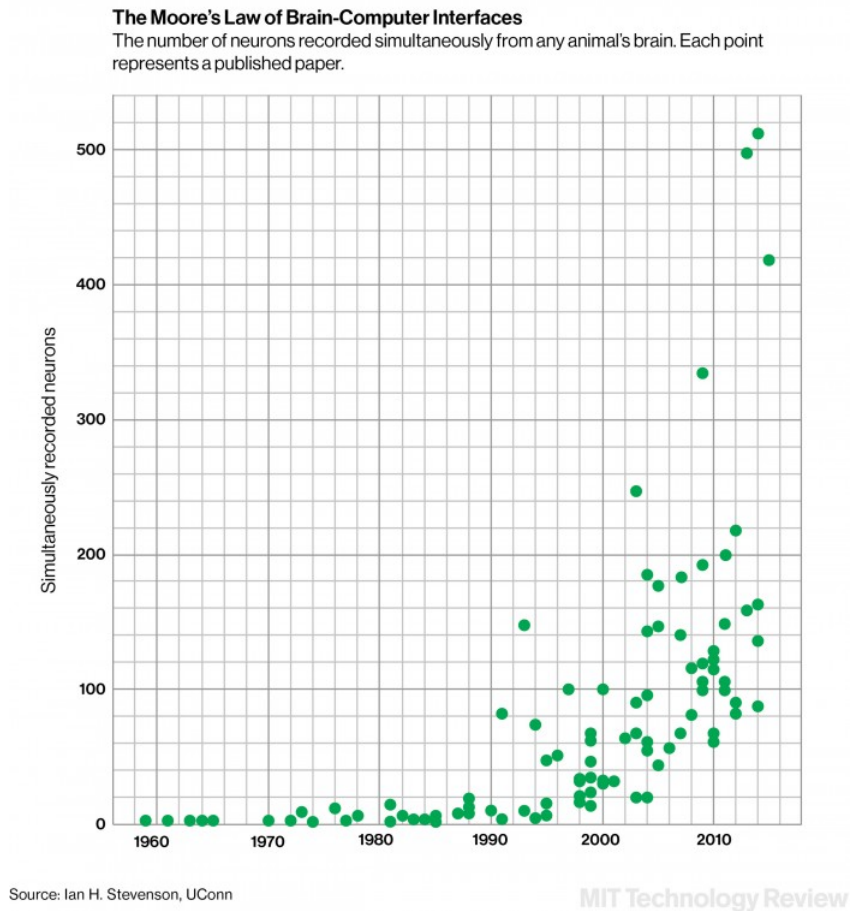


Figure 4.6: Stevenson's law[89]

4.7.1.2 Implantation

Nowadays, if you want to have an invasive BCI, you have to have a skull-opening surgery. On top of being both a major barrier to entry and a major safety issue, invasive brain surgery is expensive and in limited supply. Elon Musk, Founder of Neuralink, came up with an eventual BCI implantation process that could be automated like Lasik (Laser-assisted in situ keratomileusis). According to the paper published by Neuralink [91], their current system can implant up to 96 electrode threads, each bundled with 32 electrodes (3072 electrodes in total). The electrode wires they made were only 10 times as thick as human hair. There is no need to open the skull during the operation, but a special robot uses a very thin needle to implant the electrode wire into the brain, minimizing damage to blood vessels and other tissues.

4.7.2 Future trends and possible development directions

4.7.2.1 Intrusive and non-intrusive technologies will still co-exist and may even be combined

When invasive BCI technology is gradually put into practical application, it has extremely high requirements for basic theoretical research, engineering implementation, a long investment return period, high dependence on compound talents, high development cost, and requires clinical trials to prove its safety. It takes a long time for the company. Therefore, from a market perspective, the invasive BCI application will be very narrow in a short period of time.

Non-invasive BCI products are expected to lead the consumer market. At the same time, non-invasive BCI products have the characteristics of many fields of application and a large number of people, such as public life, entertainment, and education. In the future, one trend may be that invasive and non-invasive methods co-exist and combine for a long time, and the other trend is that minimally invasive BCI between invasive and non-invasive methods may become a hot research and development direction.

4.7.2.2 Invasive brain-computer interface technology drives innovation in robotics, software, and hardware

In the future, the innovation of invasive BCI products may focus on the development of low power consumption, miniaturization, and wireless implantation of chips/ sensors, as well as neurosurgical surgical robots that can quickly perform surgery for many people. Implanted chips will put forward new requirements for semiconductor design and manufacturing, ultra-high-bandwidth data transmission technology, two-way communication, high-density integration, digital signal processing modules (such as DSP, ASIC, FPGA), and new material development. The hardware design and development of implanted chips put forward new requirements for integrated technology, micro-fabrication manufacturing technology, testing technology, and inductive power supply. The software system implanted in the chip put forward higher requirements for digital module architecture and logic design, modeling, driver, signal amplification, analog-to-digital conversion, data compression, coding, and decoding. To process terabytes of neural data, it is necessary to develop efficient, accurate, and robust encoding and decoding algorithms, and artificial intelligence technologies such as feature learning, machine learning, and deep neural networks may play a role.

4.7.2.3 Non-invasive methods for detecting brain activity tend to diversify

Non-invasive BCI technology can also be carried out in parallel with other indicators and detection methods, for example, based on electric fields, ultrasound, magnetic fields, optogenetics, etc. Optogenetics involves injecting a virus into brain cells, which can respond to light stimuli. There are also studies that use the measurement of blood oxygen

content to determine the state of neuronal activity, which is used for the detection of brain activity.

4.7.2.4 Centralized data aggregation analysis and management platform may appear

The design, manufacture, and application of BCI in the future will generate massive amounts of valuable data. Centralized data aggregation and analysis platform can be used to collect multiple types of data such as aggregated neural signal recordings, brain tissue images, and microfabricated manufacturing data, and to realize data visualization, data analysis, data evaluation, and assist in the generation of reports for industry supervision and approval, survey results.

4.8 Conclusion

BCI is a multi-disciplinary emerging technology, which involves neuroscience, signal detection, signal processing, pattern recognition, and other disciplines. The research on BCI technology has important theoretical significance and broad application prospects. Due to the late development of BCI technology, the corresponding theory and algorithm are very immature, the research on its application is very imperfect, and more scientific and technological workers need to be devoted to the research work in this field. With the continuous improvement and maturity of technology, BCI will gradually be applied to reality and open up new application fields for bionics. After referring to a large number of literature, this paper gives a detailed review of the working principle of BCI technology and the key technologies involved, points out the main problems in this field, and looks forward to the development direction of this field.

Bibliography

- [1] Martin Waldburger, Patrick Poullie, Burkhard Stiller: *Guideline for Seminar Reports*, Communication Systems Group, Department of Informatics, University of Zurich, January 2013. <http://www.csg.uzh.ch/teaching/guideline-seminar-report-v05.pdf>.
- [2] Vidal JJ. Toward direct brain-computer communication. *Annual review of Biophysics and Bioengineering*. 1973 Jun;2(1):157-80.
- [3] Berger H. Uber das elektroenkephalogramm des menschen. *Archiv fur psychiatrie und nervenkrankheiten*. 1929;87(1):527-70.
- [4] Farwell LA, Donchin E. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and clinical Neurophysiology*. 1988 Dec 1;70(6):510-23.
- [5] Vansteensel MJ, Pels EG, Bleichner MG, Branco MP, Denison T, Freudenburg ZV, Gosselaar P, Leinders S, Ottens TH, Van Den Boom MA, Van Rijen PC. Fully implanted brain-computer interface in a locked-in patient with ALS. *New England Journal of Medicine*. 2016 Nov 24;375(21):2060-6.
- [6] Farwell LA, Richardson DC, Richardson GM, Furedy JJ. Brain fingerprinting classification concealed information test detects US Navy military medical information with P300. *Frontiers in neuroscience*. 2014 Dec 23;8:410.
- [7] Wei CS, Wang YT, Lin CT, Jung TP. Toward drowsiness detection using non-hair-bearing EEG-based brain-computer interfaces. *IEEE transactions on neural systems and rehabilitation engineering*. 2018 Jan 5;26(2):400-6.
- [8] Wu D, Lance BJ, Lawhern VJ, Gordon S, Jung TP, Lin CT. EEG-based user reaction time estimation using Riemannian geometry features. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 2017 Apr 28;25(11):2157-68.
- [9] Vourvopoulos A, Pardo OM, Lefebvre S, Neureither M, Saldana D, Jahng E, Liew SL. Effects of a brain-computer interface with virtual reality (VR) neurofeedback: A pilot study in chronic stroke patients. *Frontiers in human neuroscience*. 2019:210.
- [10] LaFleur K, Cassady K, Doud A, Shades K, Rogin E, He B. Quadcopter control in three-dimensional space using a noninvasive motor imagery-based brain-computer interface. *Journal of neural engineering*. 2013 Jun 4;10(4):046003.

- [11] Singh AK, Wang YK, King JT, Lin CT. Extended interaction with a BCI video game changes resting-state brain activity. *IEEE Transactions on Cognitive and Developmental Systems*. 2020 Apr 2;12(4):809-23.
- [12] Choi B, Jo S. A low-cost EEG system-based hybrid brain-computer interface for humanoid robot navigation and recognition. *PloS one*. 2013 Sep 4;8(9):e74583.
- [13] Saha S, Mamun KA, Ahmed K, Mostafa R, Naik GR, Darvishi S, Khandoker AH and Baumert M (2021) Progress in Brain Computer Interface: Challenges and Opportunities. *Front. Syst. Neurosci.* 15:578875. doi: 10.3389/fnsys.2021.578875.
- [14] Zander, T. O., Kothe, C., Welke, S., and Rotting, M. (2009). Utilizing secondary input from passive brain-computer interfaces for enhancing human-machine interaction, in *International Conference on Foundations of Augmented Cognition (Las Vegas, NV: Springer)*, 759-771. doi: 10.1007/978-3-642-02812-0_86.
- [15] Kulak W, Sobaniec W. Historia odkrycia EEG. *Neurologia Dzieciecea*. 2006;15(29):53-6.
- [16] Kawala-Sterniuk, A.; Browarska, N.; Al-Bakri, A.F.; Pelc, M.; Zygarlicki, J.; Sidikova, M.; Martinek, R.; Gorzelanczyk, E.J. Summary of over Fifty Years with Brain-Computer Interfaces-A Review. *Brain Sci.* 2021, 11, 43. <https://doi.org/10.3390/brainsci11010043>.
- [17] Lebedev MA, Nicolelis MAL. Brain-Machine Interfaces: From Basic Science to Neuroprostheses and Neurorehabilitation. *Physiol Rev* 97: 767-837, 2017. Published March 8, 2017; doi:10.1152/physrev.00027.2016.
- [18] Lin, C.T., Chen, Y.C., Huang, T.Y., Chiu, T.T., Ko, L.W., Liang, S.F., et al. (2008). Development of wireless brain computer interface with embedded multitask scheduling and its application on real-time driver's drowsiness detection and warning. *IEEE Trans. Biomed. Eng.* 55, 1582-1591. doi: 10.1109/TBME.2008.918566
- [19] Gao, Z., Wang, X., Yang, Y., Mu, C., Cai, Q., Dang, W., et al. (2019). EEG-based spatio-temporal convolutional neural network for driver fatigue evaluation. *IEEE Trans. Neural Netw. Learn. Syst.* 30, 2755-2763. doi: 10.1109/TNNLS.2018.2886414
- [20] Berger TW, Hampson RE, Song D, Goonawardena A, Marmarelis VZ, Deadwyler SA. A cortical neural prosthesis for restoring and enhancing memory. *J Neural Eng* 8: 046017, 2011.
- [21] Fuchs T, Birbaumer N, Lutzenberger W, Gruzelier JH, Kaiser J. Neurofeedback treatment for attention-deficit/hyperactivity disorder in children: a comparison with methylphenidate. *Appl Psychophysiol Biofeedback* 28: 1-12, 2003.
- [22] Lubar JF. Neurofeedback for the management of attention-deficit/hyperactivity disorders. In: *Biofeedback: A Practitioner's Guide*, edited by Schwartz MS. New York: Guildford, 1995.
- [23] Hasegawa RP, Hasegawa YT, Segreaves MA. Neural mind reading of multi-dimensional decisions by monkey mid-brain activity. *Neural Networks* 22: 1247-1256, 2009.

- [24] Musallam S, Corneil BD, Greger B, Scherberger H, Andersen RA. Cognitive control signals for neural prosthetics. *Science* 305: 258-262, 2004.
- [25] Brovelli A, Ding M, Ledberg A, Chen Y, Nakamura R, Bressler SL. Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by Granger causality. *Proc Natl Acad Sci USA* 101: 9849-9854, 2004.
- [26] Fetz EE. Are movement parameters recognizably coded in the activity of single neurons? *Behav Brain Sci* 15: 679-690, 1992.
- [27] Lilly JC. Distribution of motor functions in the cerebral cortex in the conscious intact monkey. *Science* 124: 937, 1956.
- [28] Nicolelis MA, Lebedev MA. Principles of neural ensemble physiology underlying the operation of brain-machine interfaces. *Nat Rev Neurosci* 10: 530-540, 2009.
- [29] Rolls ET. Parallel distributed processing in the brain: implications of the functional architecture of neuronal networks in the hippocampus. In: *Parallel Distributed Processing: Implications for Psychology and Neurobiology*, edited by Morris RGM. New York: Clarendon, 1989.
- [30] Niedermeyer E, Lopes da Silva FH. *Electroencephalography Basic Principles, Clinical Applications, and Related Fields*. Philadelphia, PA: Lippincott Williams & Wilkins, 2005.
- [31] Chi YM, Jung TP, Cauwenberghs G. Dry-contact and noncontact biopotential electrodes: methodological review. *IEEE Rev Biomed Eng* 3: 106-119, 2010.
- [32] Fonseca C, Silva Cunha JP, Martins RE, Ferreira VM, Marques de Sa JP, Barbosa MA, Martins da Silva A. A novel dry active electrode for EEG recording. *IEEE Trans Biomed Eng* 54: 162-165, 2007.
- [33] Gargiulo G, Bifulco P, McEwan A, Nasehi Tehrani J, Calvo RA, Romano M, Ruffo M, Shephard R, Cesarelli M, Jin C, Mohamed A, van Schaik A. Dry electrode biopotential recordings. *ConfProc IEEE Eng Med Biol Soc* 2010: 6493-6496, 2010.
- [34] Guger C, Krausz G, Allison BZ, Edlinger G. Comparison of dry and gel based electrodes for p300 brain-computer interfaces. *Front Neurosci* 6: 60, 2012.
- [35] Taheri BA, Knight RT, Smith RL. A dry electrode for EEG recording. *Electroencephalogr Clin Neurophysiol* 90: 376-383, 1994.
- [36] Bouton CE, Shaikhouni A, Annetta NV, Bockbrader MA, Friedenber DA, Nielson DM, Sharma G, Sederberg PB, Glenn BC, Mysiw WJ, Morgan AG, Deogaonkar M, Rezai AR. Restoring cortical control of functional movement in a human with quadriplegia. *Nature* 533: 247-250, 2016.
- [37] Collinger JL, Wodlinger B, Downey JE, Wang W, Tyler-Kabara EC, Weber DJ, McMorland AJ, Velliste M, Boninger ML, Schwartz AB. High-performance neuroprosthetic control by an individual with tetraplegia. *Lancet* 381: 557-564, 2013.

- [38] Lebedev MA, Nicolelis MA. Brain-machine interfaces: past, present and future. *Trends Neurosci* 29: 536-546, 2006.
- [39] Martini, M.L.; Oermann, E.K.; Opie, N.L.; Panov, F.; Oxley, T.; Yaeger, K. Sensor modalities for brain-computer interface technology: A comprehensive literature review. *Neurosurgery* 2020, 86, E108-E117.
- [40] Marcel van Gerven¹, Jason Farquhar, Rebecca Schaefer, Rutger Vlek, Jeroen Geuze, Anton Nijholt, Nick Ramsey, Pim Haselager, Louis Vuurpijl, Stan Gielen and Peter Desain The brain-computer interface cycle 2009 *J. Neural Eng.* 6 041001.
- [41] F. Lotte et al., "A review of classification algorithms for EEG-based brain-computer interfaces: A 10 year update," *J. Neural Eng.*, vol. 15, no. 3, 2018, Art. no. 031005.
- [42] Michael L. Martini, Eric Karl Oermann, Nicholas L. Opie, Fedor Panov, Thomas Oxley, Kurt Yaeger, *Sensor Modalities for Brain-Computer Interface Technology: A Comprehensive Literature Review* 86:E108-E117, 2020
- [43] Lopez Bernal, Sergio & Huertas, Alberto & Martinez Perez, Gregorio. (2021). *Cybersecurity Risks Associated With Brain-Computer Interface Classifications*. 10.4018/978-1-7998-7789-9.ch013.
- [44] Niedermeyer E.; da Silva F.L. (2004). *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Lippincott Williams & Wilkins. ISBN 978-0-7817-5126-1.
- [45] Campbell PK, Jones KE, Huber RJ, Horch KW, Normann RA. A silicon-based, three-dimensional neural interface: manufacturing processes for an intracortical electrode array. *Biomed Eng IEEE Trans* 38: 758-768, 1991.
- [46] Normann RA, Greger B, House P, Romero SF, Pelayo F, Fernandez E. Toward the development of a cortically based visual neuroprosthesis. *J Neural Eng* 6: 035001, 2009.
- [47] Jaroch DB, Ward MP, Chow EY, Rickus JL, Irazoqui PP. Magnetic insertion system for flexible electrode implantation. *J Neurosci Methods* 183: 213-222, 2009.
- [48] Hochberg LR, Bacher D, Jarosiewicz B, Masse NY, Simeral JD, Vogel J, Haddadin S, Liu J, Cash SS, van der Smagt P, Donoghue JP. Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature* 485: 372-375, 2012.
- [49] Fernandez E, Greger B, House PA, Aranda I, Botella C, Albisua J, Soto-Sanchez C, Alfaro A, Normann RA. Acute human brain responses to intracortical microelectrode arrays: challenges and future prospects. *Front Neuroeng* 7: 24, 2014.
- [50] Rousche PJ, Normann RA. Chronic recording capability of the Utah Intracortical Electrode Array in cat sensory cortex. *J Neurosci Methods* 82: 1-15, 1998.
- [51] Saxena T, Karumbaiah L, Gaupp EA, Patkar R, Patil K, Betancur M, Stanley GB, Bellamkonda RV. The impact of chronic blood-brain barrier breach on intracortical electrode function. *Biomaterials* 34: 4703-4713, 2013.

- [52] Mizuseki K, Buzsaki G. Theta oscillations decrease spike synchrony in the hippocampus and entorhinal cortex. *Philos Trans R Soc Lond B Biol Sci* 369: 20120530, 2014.
- [53] Kringelbach ML, Jenkinson N, Owen SL, Aziz TZ (August 2007). "Translational principles of deep brain stimulation". *Nature Reviews. Neuroscience*. 8 (8): 623-35. doi:10.1038/nrn2196
- [54] Neurostimulation Devices for the Treatment of Neurologic Disorders Christine A. Edwards, MS; Abbas Kouzani, PhD; Kendall H. Lee, MD, PhD; and Erika K. Ross, MS, PhD n, 2017 <http://dx.doi.org/10.1016/j.mayocp.2017.05.005>
- [55] Binnendijk, A., Marler, T., & Bartels, E. M. (2020). *Brain-Computer Interfaces: U.S. Military Applications and Implications, An Initial Assessment*. RAND Corporation.
- [56] S Rosca et al 2018 IOP Conf. Ser.: Mater. Sci. Eng. 294 012048
- [57] Peining, Pan & Tan, Gary & Aung, Aung & Phyo wai, Aung aung. (2017). Evaluation of Consumer-Grade EEG Headsets for BCI Drone Control.
- [58] A Review of Brain-Computer Interface Games and an Opinion Survey from Researchers, Developers and Users Minkyu Ahn, Mijin Lee, Jinyoung Choi and Sung Chan Jun 2014, 14, 14601-14633; doi:10.3390/s140814601
- [59] Gabriel Alves Mendes Vasiljevic & Leonardo Cunha de Miranda (2020) Brain-Computer Interface Games Based on Consumer-Grade EEG Devices: A Systematic Literature Review, *International Journal of Human-Computer Interaction*, 36:2, 105-142, DOI: 10.1080/10447318.2019.1612213
- [60] Md Mostafizur Rahman Brain-Computer Interface (BCI) is a Direct Communication Pathway between a Brain and a Device: A Comprehensive Analysis 2021 ISSN 2582-7421
- [61] Ferreira, A. L. S., Marciano, J. N., de Miranda, L. C., & de Miranda, E. E. C. (2014). Understanding and proposing a design rationale of digital games based on brain-computer interface: Results of the admiralmind battleship study. *SBC Journal on Interactive Systems*, 5(1), 3-15.
- [62] Aysha M. A 3D printed dress that can extract data from the brain 2020 <https://www.3dnatives.com/en/pangolin-3d-printed-dress-310820205/#!>
- [63] Seo D, Carmena JM, Rabaey JM, Maharbiz MM, Alon E. Model validation of untethered, ultrasonic neural dust motes for cortical recording. *J Neurosci Methods* 244: 114-122, 2015.
- [64] Crone NE, Sinai A, Korzeniewska A. High-frequency gamma oscillations and human brain mapping with electrocorticography. *Prog Brain Res* 159: 275-295, 2006
- [65] Hill NJ, Gupta D, Brunner P, Gunduz A, Adamo MA, Ritaccio A, Schalk G. Recording human electrocorticographic (ECoG) signals for neuroscientific research and real-time functional cortical mapping. *J Vis Exp pii*: 3553, 2012.

- [66] Viventi J, Kim DH, Vigeland L, Frechette ES, Blanco JA, Kim YS, Avrin AE, Tiruvadi VR, Hwang SW, Vanleer AC. Flexible, foldable, actively multiplexed, high-density electrode array for mapping brain activity in vivo. *Nature Neurosci* 14: 1599-1605, 2011.
- [67] Alexander JE, Porjesz B, Bauer LO, Kuperman S, Morzorati S, O'CONNORSJ, Rohrbaugh J, Begleiter H, Polich J. P300 hemispheric amplitude asymmetries from a visual oddball task. *Psychophysiology* 32: 467-475, 1995.
- [68] C. -T. Lin and T. -T. N. Do, "Direct-Sense Brain-Computer Interfaces and Wearable Computers," in *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 1, pp. 298-312, Jan. 2021, doi: 10.1109/TSMC.2020.3041382.
- [69] Jerbi, K.; Vidal, J.; Mattout, J.; Maby, E.; Lecaiguard, F.; Ossandon, T.; HamamÃ©, C.; Dalal, S.; Bouet, R.; Lachaux, J.P.; et al. Inferring hand movement kinematics from MEG, EEG and intracranial EEG: From brain-machine interfaces to motor rehabilitation. *Irbm* 2011, 32, 8-18.
- [70] g.tec Medical Engineering | Brain-Computer Interfaces and Neurotechnology. <https://www.gtec.at/>
- [71] Yang, M., Z. Yang, T. Yuan, W. Feng, and P. Wang. 2019. A systemic review of functional near-infrared spectroscopy for stroke: current application and future directions, *Frontiers in neurology* 10, 58.
- [72] Boto, E., N. Holmes, J. Leggett, G. Roberts, V. Shah, S. S Meyer, L. D. Munoz, K. J Mullinger, T. M Tierney, S. Bestmann, et al. 2018. Moving magnetoencephalography towards real-world applications with a wearable system, *Nature* 555, no. 7698, 657-661.
- [73] Sitaram R, Caria A, Veit R, Gaber T, Rota G, Kuebler A, Birbaumer N. FMRI braincomputer interface: a tool for neuroscientific research and treatment. *Comput Intell Neurosci* 25487, 2007.
- [74] Yoo, S.S.; Fairmeny, T.; Chen, N.K.; Choo, S.E.; Panych, L.P.; Park, H.; Lee, S.Y.; Jolesz, F.A. Brain-Computer interface using fMRI: Spatial navigation by thoughts. *Neuroreport* 2004, 15, 1591-1595.
- [75] Sorger, B.; Goebel, R. Real-time fMRI for brain-computer interfacing. In *Handbook of Clinical Neurology*; Elsevier: Berlin/Heidelberg, Germany, 2020; Volume 168, pp. 289-302.
- [76] Weiskopf, N.; Mathiak, K.; Bock, S.W.; Scharnowski, F.; Veit, R.; Grodd, W.; Goebel, R.; Birbaumer, N. Principles of a braincomputer interface (BCI) based on real-time functional magnetic resonance imaging (fMRI). *IEEE Trans. Biomed. Eng.* 2004, 51, 966-970.
- [77] Hasan, M.A.; Khan, M.U.; Mishra, D. A Computationally Efficient Method for Hybrid EEG-fNIRS BCI Based on the Pearson Correlation. *BioMed Res. Int.* 2020, 2020, 1838140.

- [78] Nazeer, H.; Naseer, N.; Khan, R.A.; Noori, F.M.; Qureshi, N.K.; Khan, U.S.; Khan, M.J. Enhancing classification accuracy of fNIRS-BCI using features acquired from vector-based phase analysis. *J. Neural Eng.* 2020, 17, 056025
- [79] Martorell AJ, Paulson AL, Suk HJ, Abdurrob F, Drummond GT, Guan W, Young JZ, Kim DN, Kritskiy O, Barker SJ, Mangena V. Multi-sensory gamma stimulation ameliorates Alzheimer's-associated pathology and improves cognition. *Cell.* 2019 Apr 4;177(2):256-71.
- [80] Cruse D, Chennu S, Fernández-Espejo D, Payne WL, Young GB, Owen AM. Detecting awareness in the vegetative state: electroencephalographic evidence for attempted movements to command. *PloS one.* 2012 Nov 21;7(11):e49933.
- [81] Barsotti M, Leonardis D, Loconsole C, Solazzi M, Sotgiu E, Procopio C, Chisari C, Bergamasco M, Frisoli A. A full upper limb robotic exoskeleton for reaching and grasping rehabilitation triggered by MI-BCI. In 2015 IEEE international conference on rehabilitation robotics (ICORR) 2015 Aug 11 (pp. 49-54). IEEE.
- [82] Bernal SL, Celdrín AH, Pérez GM. Cybersecurity Risks Associated With Brain-Computer Interface Classifications. In *Advances in Malware and Data-Driven Network Security 2022* (pp. 236-259). IGI Global.
- [83] Jebari K. Brain machine interface and human enhancement-an ethical review. *Neuroethics.* 2013 Dec;6(3):617-25.
- [84] Vlek RJ, Steines D, Szibbo D, Kubler A, Schneider MJ, Haselager P, Nijboer F. Ethical issues in brain-computer interface research, development, and dissemination. *Journal of neurologic physical therapy.* 2012 Jun 1;36(2):94-9.
- [85] Drew L. Agency and the algorithm. *Nature.* 2019;571(7766):S19-21.
- [86] Tamburrini G. Brain to computer communication: ethical perspectives on interaction models. *Neuroethics.* 2009 Nov;2(3):137-49.
- [87] Clausen J. Man, machine and in between. *Nature.* 2009 Feb;457(7233):1080-1.
- [88] Burwell S, Sample M, Racine E. Ethical aspects of brain computer interfaces: a scoping review. *BMC medical ethics.* 2017 Dec;18(1):1-1.
- [89] Adam Piore. Government Seeks High-Fidelity 'Brain-Computer' Interface. 2016. <https://www.technologyreview.com/2016/02/02/162473/government-seeks-high-fidelity-brain-computer-interface/>.
- [90] Stevenson IH, Kording KP. How advances in neural recording affect data analysis. *Nature neuroscience.* 2011 Feb;14(2):139-42.
- [91] Musk E. An integrated brain-machine interface platform with thousands of channels. *Journal of medical Internet research.* 2019 Oct 31;21(10):e16194.

Chapter 5

Data Plane Programmability: Overview, Abstractions, and Use Cases

Basler D. B., Browne J. I.

With the emergence of software defined networking, the controlling functionality (control plane) could be separated from the devices in networks. The respective forwarding devices, such as switches, remained responsible for the data plane functionalities while being controlled by a centralized control instance. In order to establish a communication between the forwarding plane and the control instance, OpenFlow was introduced, which is one of the early APIs in software defined networks and enabled the management of match-action flow tables. However, OpenFlow came with limitations such as protocol dependence and a number of successor techniques followed in an attempt to solve these limitations. These introduced techniques can be labeled data plane programmability. In particular four different approaches were highly relevant in recent years, including the following ones: P4, Domino (both data plane programming languages), protocol oblivious forwarding, and SRv6 programmability. It is important to understand these different approaches and compare them with each other, to gain an understanding of why data plane programmability improves the limitations of the traditional network setup with OpenFlow. After providing a brief overview of the most notable terms necessary to understand concepts related to network programmability and SDN, the main part dives into data plane programmability in more detail and covers the mentioned approaches. Each of these include an overview, a context, a description of the architecture and forwarding model, and a display of some use cases. Finally, all are evaluated and compared to each other in a discussion, followed by a conclusion regarding the current state and an potential outlook of future developments.

Contents

5.1	Introduction To Network Programmability	107
5.2	Definitions And Terms	107
5.2.1	Data, Control And Management Plane	107
5.2.2	Software-Defined Networking (SDN)	108
5.2.3	SDN Over Legacy Protocols	109
5.2.4	Programmability In OpenFlow	110
5.2.5	Programmability In PCEP	111
5.2.6	Programmability In hSDN	112
5.3	Data Plane Programmability	113
5.3.1	Motivation	113
5.3.2	Existing Solutions	114
5.4	Comparison Of Techniques And Discussion	124
5.5	Conclusion	125

5.1 Introduction To Network Programmability

The fast growth of internet over the last decade has caused the scales of networking applications, and network elements to rise tremendously. Software defined networks (SDN) provides network operators with programmatic autonomy over their networks. Meanwhile, the network architecture has gotten increasingly sophisticated as a result of the competitive tussle among many stakeholders [1]. As a result of these concerns, software-defined networking has evolved as a new paradigm that alters how the Internet operates [2]. The control plane and forwarding plane are kept separate in SDN, and one control plane controls numerous forwarding devices. While forwarding devices can be configured in a variety of ways, having a standard, open interface allows a control plane to operate forwarding devices. As a result of the control plane's forwarding plane abstraction, new networking protocols can be simply implemented. Being one of the first SDN deployments, OpenFlow offers network operators with a strong tool set for programming and managing their networks flexibly [3]. Nevertheless, the forwarding plane's programmability is still limited due to its protocol dependence. OpenFlow, in particular, defines its matching fields in flow tables in accordance with existing network protocols [4]. As a result, in order to parse packets and execute flow matching, OpenFlow switches must comprehend the protocol headers, which may pose major compatibility concerns when attempting to add new protocols or delete header fields [1]. As a result, it would be preferable if network programmability could be improved further such that the forwarding plane is protocol-independent and can be dynamically reprogrammed to support new protocols seamlessly. Following this concept, current research suggest a few new SDN technologies, such as protocol-oblivious forwarding (POF) [2], programming protocol-independent packet processors (P4) [5], Segment Routing IPv6 (SRv6) [6] and Domino [7]. They share the basic principle in that they all attempt to detach network protocols from packet forwarding and make the forwarding plane customizable and programmable. This article explores how to use the mentioned techniques to improve the programmability of the forwarding plane in SDN. We begin by reviewing some of the basic concepts of SDN and explain why P4, POF, SRv6 or Domino was introduced.

5.2 Definitions And Terms

5.2.1 Data, Control And Management Plane

Before we go into the specifics of what a control, data or management plane is we must first clarify what a plane in networking in general is. A plane is an abstract term in networking that represents the location of particular operations [8]. The data plane, the control plane, and the management plane are the most widely used planes in networking. The core DNA in today's networking, gear to transfer IP packets from A to B, are the Control Planes and Data Planes. The other important component is the management plane, which is also referred as the user-to-hardware interaction [9]. These operational planes serve as the foundation for the layered architecture to which network systems have expanded to.

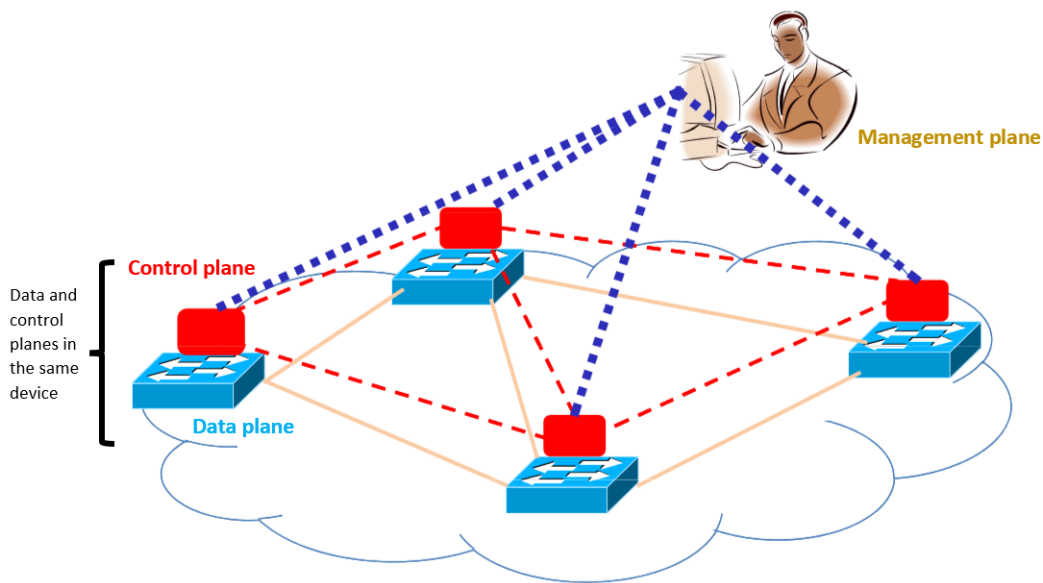


Figure 5.1: Traditional Networking with modifications taken from Rexford, J. (2012), Software-Defined Networking Course.

Control plane: The control plane in routing refers to those operations and procedures that identify the route to take to transport a packet or a frame [8]. The control plane's duty is to populate the routing table, modelling network topology and therefore activating the data plane capabilities [8]. Simply said, it is responsible of how packets are routed and forwarded.

Data plane: In routing, the data plane corresponds to all activities and functions that use the control plane logic to transfer packets from one interface to another. The data plane function is mainly composed of the forwarding table, routing table and routing logic [8]. In a nutshell, it is responsible for transporting packets from source to destination. It is therefore also known as the forwarding plane.

Management plane: The management plane processes communication traffic going to the network device that is intended to monitor, setup or administer the device [10]. In other words, the management plane traffic is characterized in the same manner as the control plane traffic, but the communication's purpose is to setup, administer, or monitor the network device [10].

5.2.2 Software-Defined Networking (SDN)

Traditional networks are vertically integrated, meaning that the control plane and the data planes are situated together network devices. This reduces flexibility and a solution which emerged to this limitation is software-defined-networking (SDN). SDN breaks the vertical integration by taking the control plane out of the network devices and thus, making the control logic more independent. The control logic is now in a centralized controller also called network operating system, which is situated in between network

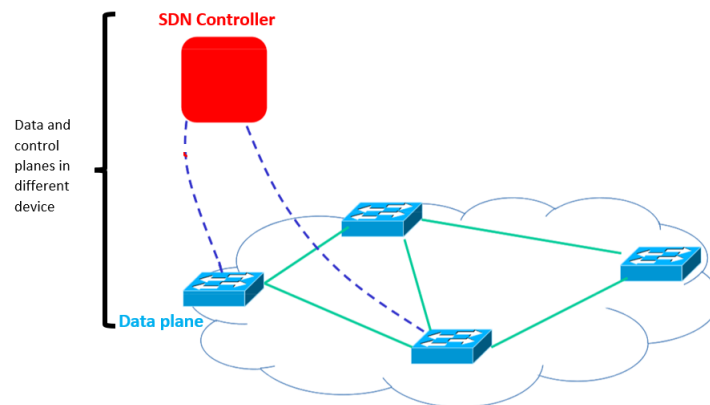


Figure 5.2: Software-defined Networking with modifications taken from Rexford, J. (2012), Software-Defined Networking Course.

applications and the data plane including the network devices [4]. The application plane contains network applications such as firewalls, intrusion detection systems, and network virtualization. Further it includes mobility management in order use the northbound API to interact with the controller or control plane below. The centralized controller offers services including routing computation, network monitoring, and load balancing. It receives high-level instructions from the application plane and translates them into sets of instructions in the form of flow entries. These entries are installed on the devices which are located on the data plane. In SDN the data plane devices make no decisions themselves, but are rather used as simple forwarding devices [11].

By separating the control plane from the data plane, it became more feasible to run the control plane or controller in software on standard servers. This enabled the creation of new virtualized controllers and custom-made services. Using well defined APIs, applications on the application plane are able to program the underlying network to control the underlying data plane. Doing this enables precisely defined actions and treatment for packets down to the data plane. Network administrators are not required to resort to manual configuration of low-level interfaces anymore, but can use the higher layers. The value proposition of SDN includes a number of new capabilities including rapid introduction of new network functions at software speed rather than hardware or firmware product cycles, and more seamless integration of the network with IT processes in the enterprise through programmable service-oriented APIs. Further, SDN provides a new method for applications to interact with the network. Abstract APIs are used for direct configuration and operation of networks and a query API to ask networks for information. This information can then be used to plan and optimize operations. Finally, SDN can decouple the network service API from the underlying implementation, which enables the infrastructure to evolve with reduced impact to applications [12].

5.2.3 SDN Over Legacy Protocols

Before the existence of APIs like OpenFlow, the network configuration protocol (NETCONF) was introduced. The first version was distributed in 2006 [13] and a revision of it

in 2011 [14]. It emerged because operators previously relied on protocols like the simple network management protocol (SNMP) which were too difficult to use and were mainly used for monitoring activity [15]. NETCONF allows for a network device to be managed in a simple way. Further information corresponding to the configuration can be retrieved and new configured. Thus, the protocol allows devices to expose an API. It makes use of remote procedure calls (RPC) to communicate between user and device and XML documents for the configuration information which is exchanged. NETCONF can be split into the following four layers from top to bottom: content with the configuration data, operations to get and edit data, RPC to establish communication, and transport protocol as the underlying layer [13].

As NETCONF sessions are the logical connections between network configuration applications and network devices [13], we can see how this concept is related to SDN. Like OpenFlow, NETCONF is located at the southbound interface between controller and data plane.

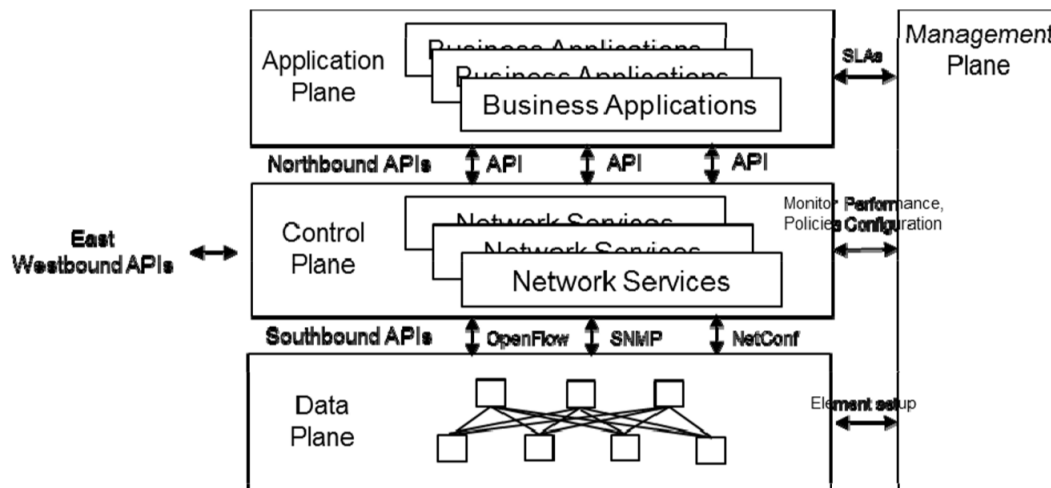


Figure 5.3: SDN Architecture Overview, Open Networking Foundation (2013)

5.2.4 Programmability In OpenFlow

In the year 2008, McKeown et al. [3] proposed OpenFlow. Their main motivation for developing OpenFlow was the need for programmable networks at the time due to the lack of ways to test and experiment new network protocols with existing networks. While their proposal was intended for academic experiments on campus networks, OpenFlow became much more widespread in the following years and is one of the most notable Southbound API implementation in the industry [11]. It became so widespread, that most vendors of commercial switches support the OpenFlow API today [4]. In summary, OpenFlow basically acts as an API between a SDN controller and the data plane including switch and routing devices. SDN controllers use the API to get access to the data plane elements. Switches in the context of OpenFlow, have flow tables with packet-handling rules. Each such rule can match with certain traffic elements based on different characteristics. Thanks to this, users can program a device via the controller to behave as desired [4]. OpenFlow

switches can also handle forwarding rules locally without depending on a remote SDN controller [11].

OpenFlow Switches consist of three basic parts: a flow table, a secure channel, and an OpenFlow protocol. The flow table tells the switch how to handle a flow of traffic depending on certain characteristics. Then, the secure channel connects the controller with the switch using the OpenFlow protocol to communicate the different commands and packets. Through this communication, users do not have to program the switch anymore. Controllers install the rules in the devices which match particular patterns such as port numbers to an action. These flow tables with the rules are built using Ternary Content Addressable Memory (TCAM) [3]. TCAM is a type of memory that can perform parallel search at high speeds [16], which is ideal for the use case of matching traffic flows and flow-rules. It can provide a lookup of flow entries within a single clock cycle [11]. Since the beginning of SDN, a large issue is scalability. When large volumes of traffic from many flows is passed through a network, the problem gets worse. If flow entries in switches increase very fast, TCAM can become exhausted, which results in inferior forwarding performance overall [4].

Flow tables can typically store sets of flow entries which are made up of 15 field tuples. Some of the fields are optional. However in many cases matching fields, action, statistical counter, priority, and timeout mechanism are used. Among the meta information of IP packets used in the matching process are MAC addresses (source and destination), IP addresses (source and destination), and ethernet types. These fields can either be matched in an exact or wildcard manner. In wildcard-matching multiple traffic flows can be matched as some values are marked with an asterisk (*), symbolizing that that field can have any value. Whenever a packet matches with the flow, the corresponding action is taken. OpenFlow supports multiple flow table pipelining processing using more than just one flow table. In both single and multiple table scenarios, when there is no match, the packet can either be dropped or the packet-in message is sent back to the controller which requests a flow setup request [11]. The multiple flow table pipeline was introduced to extend OpenFlow's initial limitations of the header field sets. With a large array of predefined header fields this was achieved [17]. Currently the Open vSwitch is the most widely used OpenFlow software switch. It uses a universal flow-caching based datapath for implementing match-action pipelines [17]. Open vSwitch is a virtual switch with multiple layers which was introduced by Pfaff et. al. in 2015. Its purpose was to help networking users to carry out their work in virtual environments [18].

5.2.5 Programmability In PCEP

The Path Computation Element (PCE) was originally designed to alleviate the path computation tasks from routers in an Multiprotocol Label Switching (MPLS) traffic network [19]. Path computations in MPLS are done by the provider's head-end router which obtains path requests from the customers [20]. However, processing constraints-based pathways for varied flows might overload the router's CPU, decreasing overall routing efficiency. Path calculation, in general, gets even more complicated when inter-domain and as well as multi-layer networks, are involved [19]. This necessitates the deployment of

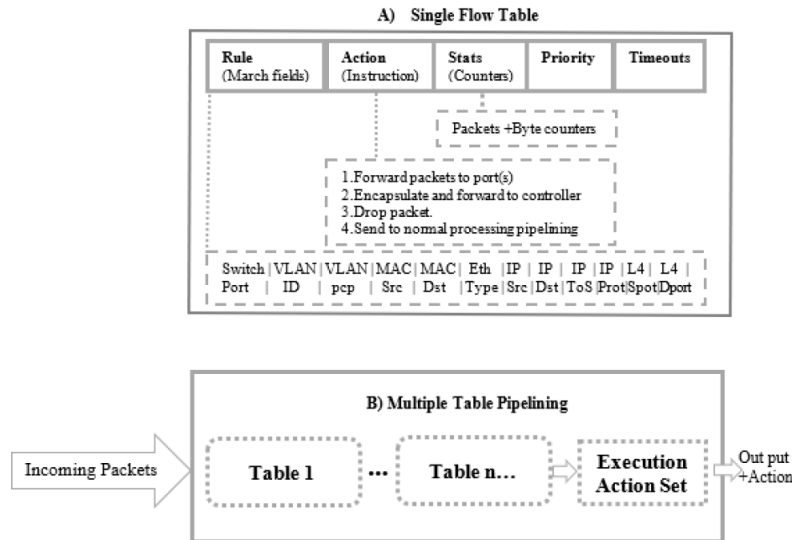


Figure 5.4: Single vs. multiple flow table taken from Isyaku, B. et al. (2020), Software defined networking flow table management of openflow switches performance and security challenges: A survey

a dedicated path computation server capable of computing pathways within big complex, meshed networks as well as across domains, regardless of the underlying network layer.

PCE is able to compute best pathways for traffic over a network and can adjust the paths to address traffic demand changes [19]. Ever since, the PCE's function and purpose have expanded to include a variety of different use cases such as in usage with Generalized Multiprotocol Label Switching [20] or to provide delegated control [21].

As described in Subsection 5.2.2, an element that specifies how network packets will be utilized and how the switches will be configured is an essential feature in an SDN design. In the context of PCEP, this element may be viewed as doing particular computations to relocate traffic flows throughout the network which is precisely what a PCE does and demonstrates how a PCE fits into an SDN system [22]. As a result, it is legitimate to investigate PCEP as a control protocol respectively as an Southbound Interface for usage in these settings in order to fully empower the PCE as a central controller [22].

5.2.6 Programmability In hSDN

A hybrid software-defined network (SDN) is a network consisting both traditional networking as well as software-defined networking protocols [23]. Network administrators may use a hybrid SDN to deploy newer SDN technologies such as OpenFlow to legacy systems without having to completely restructure the system infrastructure [23]. In a hybrid software-defined network technicians can operate SDN technologies and traditional switching protocols along the same physical hardware. While standard distributed network protocols remain to partly direct the data traffic, a network administrator is able to configure the software-defined network control plane to control particular traffic flows [23].

5.3 Data Plane Programmability

5.3.1 Motivation

Most network components, such as routers and switches, are developed bottom up [12]. Switch providers who provide products to their customers typically depend on external components from third-party manufacturers [5]. The chip is a system's brain, determining how the system OS is implemented and what capability it can provide. Because the chip is a fixed-function element with an internal packet forwarding process that cannot be conveniently altered at runtime, introducing an additional set of features is a difficult operation that may take several months [4]. This is due to the fact that a chip redesign is frequently necessary. Programming languages like P4 [5], POX [24] or Domino [7] have an entirely different method, comparable to how Graphics Processing Units or Central Processing Units operate [25]. These functional units run executable code in a particular programming language such as C++ or OpenCL. The code is therefore compiled before being loaded into these functional units such as the CPU [5]. P4, POX or Domino on the other hand, is centered on the same concept, however for network devices. This is a typical top-down strategy. In contrast to the bottom-up paradigm mentioned above, with these programming languages the network structure can be described by a programmer.

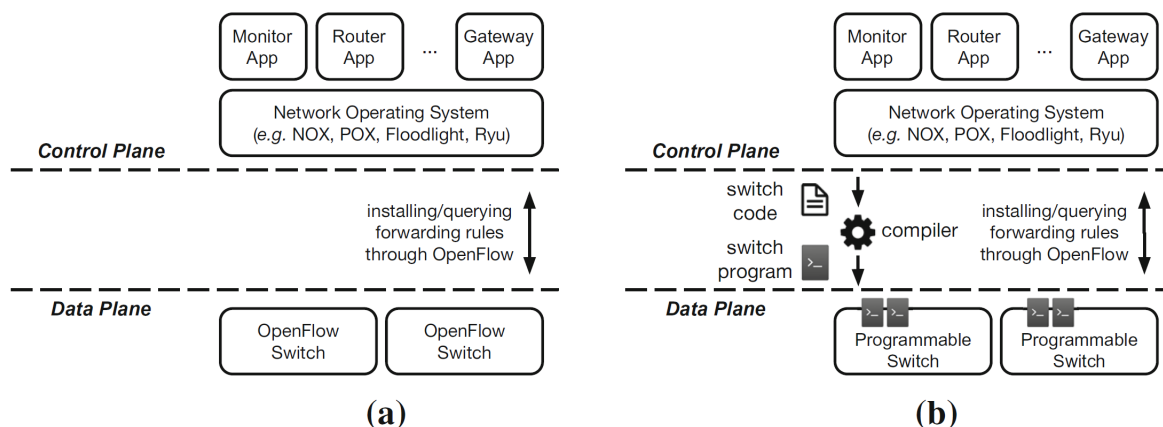


Figure 5.5: Traditional SDN (a) and SDN with a programmable data plane (b) taken from da Costa Cordeiro, W. L. et al. (2017), Data plane programmability beyond openflow: Opportunities and challenges for network and service operations and management

Traditionally, the term programming was associated with populating switch tables with rules on how to forward flows in the data plane. However, with SDN solutions, a more complex problem arose. Network operators can get constrained by the set of features and protocols supported by fixed-function ASIC switches. They are not able to design custom packet headers and packet parsing routines. Further, the support of novel L1, L2, and L3 protocols requires a complex and time-consuming pipeline, which includes protocol standardization and the release of an OpenFlow version that supports the protocol header fields. This process can take a long time (even years in some cases) [26]. In order to solve these existing problematic aspects, data plane programmability became

more prominent after 2010. It goes beyond the low-level SDN protocols such as OpenFlow and allow operators to satisfy the need of more complex SDN applications [17]. The new high-level programming languages enable network operators to reprogram forwarding devices in-field to deploy novel networking protocols, customize network behavior, and consequently develop innovative services [26]. Furthermore, with modern hardware, such as flexible switches, languages specify the switch processing architecture. This includes the specification of the layout of match-action tables, the protocols understood by the parser, and the supported actions [17]. In a programmable forwarding plane, operators can write switch code that specifies custom header fields and define packet header matching and parsing semantics. By using a vendor supplied compiler, operators can compile the code and deploy the resulting program into the switch. OpenFlow can be used to populate forwarding rules built upon programmed switch features [26].

5.3.2 Existing Solutions

In the following sub chapters, four different approaches are introduced in more detail showing advancements in the field of data plane programmability throughout the last years. All approaches emerged from the traditional SDN networks with OpenFlow and attempt to eliminate the limits that approach has. First, two programming languages, P4 and Domino, are introduced which are designed to enhance the data plane functionalities. Then, two similar approaches, POF and SRv6, are introduced which are intended replacements of OpenFlow and are therefore southbound APIs in the SDN architecture.

5.3.2.1 P4

P₁₄ was the acronym of the very first P4 language deployment and P₁₆ is the current language standard [5]. There are significant differences between P₁₄ and P₁₆, and it is important to keep in mind that P₁₆ includes certain backwards-incompatible modifications to P₁₄ in respect of semantics and syntax [27]. The primary reason for moving from P₁₄ to P₁₆ was to minimize the language's complexity and to provide a solid base of the language to ensure that current P₁₆ programs will be syntactically valid in the long term when evaluated against later releases of the language [27]. Simultaneously, a considerable number of native P₁₄ features have been incorporated into libraries for developing efficient P4 software [25].

P4 was initially introduced in the publication "Programming Protocol-Independent Packet Processors" [5]. P4 enables developers to completely control how packets passing configurable dataplane blocks will be treated. Target is a common term in P4 that refers to a range of devices that may be programmed using P4, such as routers, switches or Network Interface Cards [27]. P4 has a significant benefit in that it can process not just conventional well-known protocol headers such as Ethernet or TCP as typical routers or switches do, but also entirely customized ones [5] as shown in Figure 5.6.

The P4 code must specifically describe all header types and how they are processed [5]. It is also the job of the coder to construct a match-action pipeline within the specified

```

header ethernet {
    fields {
        dst_addr : 48;
        src_addr : 48;
        ethertype : 16;
    }
}

header vlan {
    fields {
        pcp : 3;
        cfi : 1;
        vid : 12;
        ethertype : 16;
    }
}

Custom protocol header {
    header customProtocol {
        fields {
            up1 : 8;
            up2 : 8;
            down1 : 8;
            down2 : 8;
            ethertype : 16;
        }
    }
}

```

Figure 5.6: P4 Header Formats based on Bosshart, P. et al. (2014), P4: Programming Protocol-Independent Packet Processors

dataplane block [27]. Such a block in turn can be made up of one or more tables that are matched against processed header fields [27] as can be seen in Figure 5.7.

One or more actions could be assigned to each table and executed at runtime [28]. These actions specify what should be done with a packet such as changing the entries of certain header fields, drop packets or forward packets to the specified physical port [5]. A programmer is responsible for all elements of the tables and their inner structural properties such as the match-action fields, number of tables or action behavior which, in turn, makes P4 a powerful approach [28].

The PSA architectural model as shown in the upper half of Figure 5.7 is often used with the P₁₆ language definition. As such, the P4 architectural model specifies the function blocks for a specific dataplane target. As depicted in Figure 5.7 both programmable blocks and fixed-function blocks are applicable. With that being said, the target manufacturers dictate the behavior of the fixed-function blocks, which is beyond the control of the programmer [5]. However, P4 is then used to configure the programmable blocks. It's also worth noting that two two fixed-function blocks and six programmable blocks are specified in the architecture model [28]. The Packet Buffer and Replication Engine (Block four in Figure 5.7) and the Buffer Queuing Engine (Block eight in Figure 5.7) are target specific components [28]. The format of these two blocks may alter depending on the device and manufacturer. However, P4 is entirely configurable for the remaining blocks. In addition to the PSA architecture model, there are other ones such as the V1Model, the PISA or BMv2 architecture model. The V1Model is the architecture commonly used with the BMv2 Simple Switch [28] and the PISA concept dates back to when P₁₆ was the primary language version.

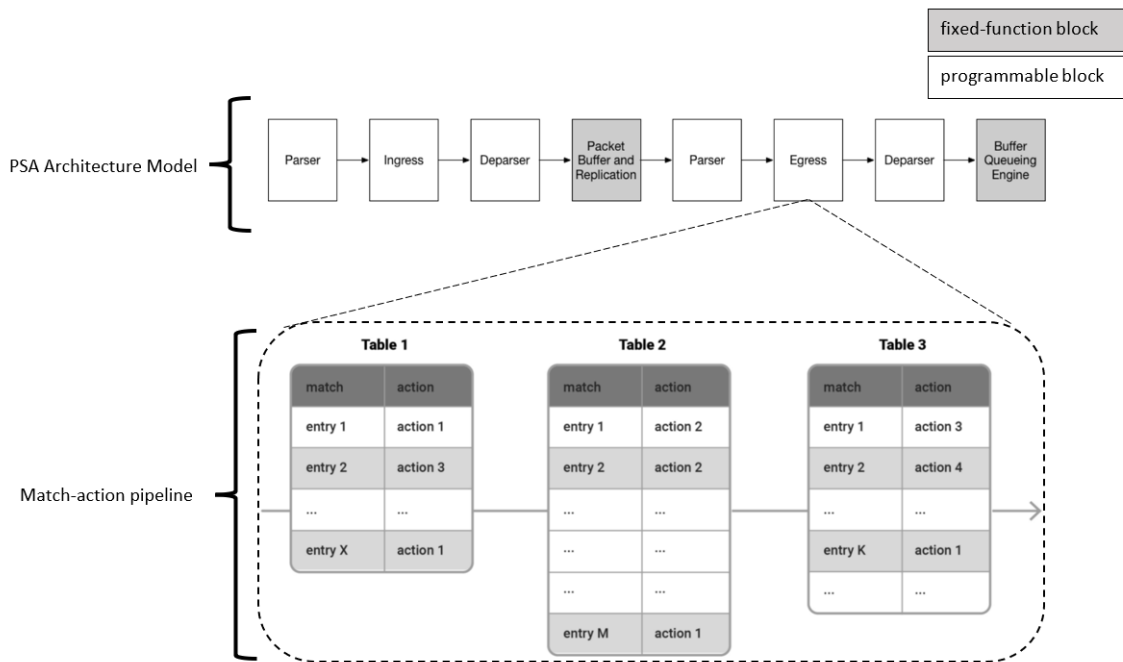


Figure 5.7: PSA Architecture Model and match-action pipeline within a given programmable block based on P4.org Architecture Working Group (2021), P₁₆ Portable Switch Architecture (PSA) and Pawel, P. (2020), P4 programming Language - Introduction to network programming with P4 Course

Regarding P4, a lot of research is being done at the moment. Fattaholmanan and Carzangia, for example, introduced P4 Weaver, a new way for incremental programming [29]. P4 Weaver is intended to integrate new data plane code features into a core program in a systematic and controlled manner in order to maintain the switch’s reliability. P4 Weaver enables switch users to adopt and deploy new concepts and algorithms while guaranteeing that the switch’s essential traffic-handling functionality remains functional. P4 Weaver adds so called annotations to a P4 applications which enable to control how a program may be altered without breaking it.

5.3.2.2 Domino

In 2016 Sivaraman et al. developed Domino which is a domain-specific and imperative programming language to express data-plane algorithms. It is similar to the C programming language but has several constraints, which are described below [7]. The main goal of the Domino concept is to express data-plane algorithms and make the respective hardware flexible enough, while maintaining the performance of high speed switches. Traditionally, data-plane algorithms were implemented using dedicated hardware, making investing in new and expensive hardware necessary every time a new algorithm came along. Efforts to build programmable switches meant to have worse performance in order to achieve programmability. However, in recent years programmable switching chips such as Intel’s Tofino 3 [30] or Marvell’s Teralynx [31] have emerged which are attractive as they don’t compromise data rates [32]. In order to always run at line-rate of a switch,

the data-plane algorithms traditionally had to be implemented directly in the hardware, and therefore, not being modifiable. This leads to static and non-flexible solutions. With Domino, it is possible to formulate data-plane algorithms conveniently which are compiled into low-level microcode which runs on line-rate switching chips. Further, the concept of packet transactions was introduced which enables line-rate programmability for stateful algorithms [7]. In contrast to the stateful algorithm support of Domino, the P4 language, discussed in the previous sub-chapter, is more suitable for packet processing tasks which don't manipulate states on switches [33].

As mentioned the Domino language comes with some constraints. For instance, it doesn't support for loops, access to packet payloads, unstructured control flow and dynamic memory allocation. The reason for these constraints is that the line-rate must be guaranteed in all cases. By making these constraints required, deterministic performance can be achieved. In the case of loops, it would be impossible to guarantee the algorithm finishing in time to satisfy the line-rate. Payload access also requires too much time for switches to process [7].

Domino comes with the following three main concepts [7]:

- Packet transactions (abstraction for programming algorithms) and the Domino language itself
- The Banzai machine model for line-rate switches
- The Domino compiler

```

#define N 30
struct Packet {
    int sample;
}
int count = 0;

void func(struct Packet pkt){
    if(count == N - 1){
        pkt.sample = 1;
        count = 0;
    } else{
        pkt.sample = 0;
        count = count + 1;
    }
}

```

Figure 5.8: Domino code taken from Sivaraman, A. et al. (2016) Packet transactions: High-level programming for line-rate switches

A packet transaction in Domino is an abstraction to implement data plane algorithms. It is a sequential code block which has the property of being atomic, meaning that it is isolated from other packet transactions. Therefore, developers can focus on the operations of each packet without thinking about concurrent packets. While the actual hardware (switches) has several pipelines working on the packets, developers view the transactions

as if they were executed in a serial matter. This means, any visible state is equivalent to serial execution of the transactions in the order of arrival of the packets [7]. In order to write code to express packet-processing algorithms, a packet transaction is used. The transaction captures the logic of the processing of a packet [33]. In figure 5.8 a code example is shown to demonstrate a very simple packet transaction which samples packets only every Nth time. The code is written in the Domino language.

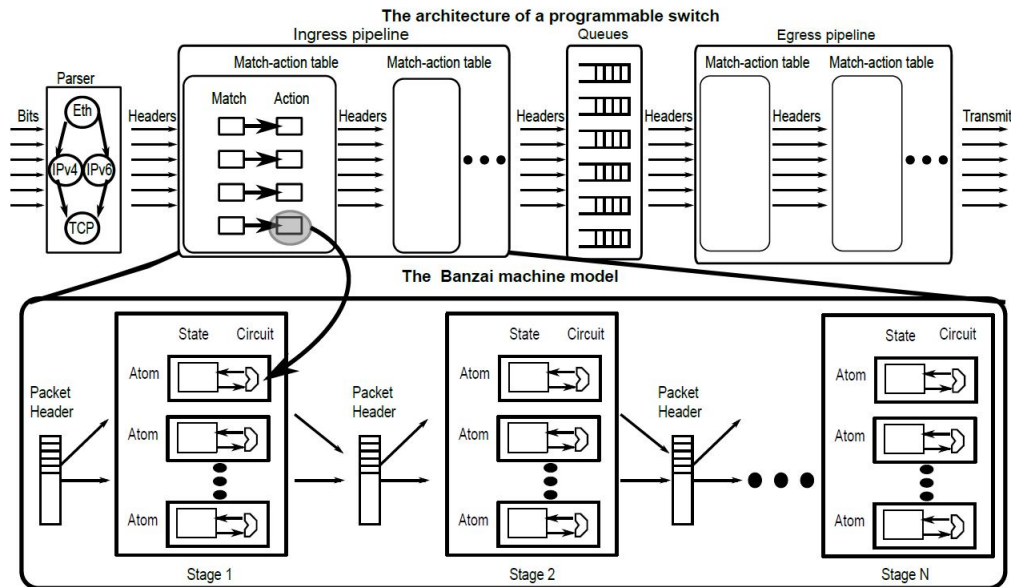


Figure 5.9: The Banzai machine model pipeline of a programmable switch taken from Sivaraman, A. et al. (2016) Packet transactions: High-level programming for line-rate switches

The second concept of Domino is the Banzai machine model which is depicted in figure 5.9. A goal of this machine model is to represent switches and experiment with different ones [33]. Banzai is a C++ simulator of programmable switches and can model either the ingress or egress switch pipeline. As seen in graphic, the machine model models only the action phase of the match-action table. Further, the model assumes the packets are already parsed on arrival. Banzai is essentially a feed-forward pipeline which executes different stages on every clock cycle. At the passing of every clock cycle, exactly one packet is processed at a stage. An important property of Banzai is, that it always satisfies line-rate and is deterministic [7]. In the context of Banzai, a key element was introduced called the atom. Atoms are located in every pipeline stage of the machine model in form of a vector and represent atomic operations which are supported by the switches natively. Each atom in pipeline stage modifies mutually exclusive sections of one packet's header in parallel, allowing for experiments with different switches. Atoms can also modify the actual switch device, making stateful data plane algorithms possible. In order for these algorithms to be executed at line-rate, Banzai uses an atomic modifying operation which enables an atom to terminate the the read, modify, and write operation in one clock cycle. The modify part can be any stateful operation [7]. Banzai does however come with some constraints and limitations. As every atom's state is local, this state can't be shared with other atoms. This limits the memory capacity. Computation limits arise due to the fact that Banzai requires atoms to terminate within one-clock cycle. Finally,

resources are limited because the amount of atoms in each stage and the amount of stages in the pipeline are limited. Therefore, Banzai is not a feasible solution for algorithms which modify larger sets of packet headers and require larger amounts of computational resources. Examples include WAN optimization or deep packet inspections which require switches to process the payloads of packets as well [7]. These tasks, especially at line-rates over 1GHz, should rather be run on hardware such as the CPU [34].

The last Domino concept is the Domino compiler which translates Domino code to Banzai targets. This compiler is a bit special as it only compiles the Domino code if it can be run at line-rate (all-or-nothing policy). Therefore, line-rate is guaranteed. The compiler consists of three phases: preprocessing, pipelining, and code generation [7]. In a first step, the compiler takes the packet transactions and outputs a pipeline of codelets. These are small code fragments which guarantee the transaction's semantics. Secondly, every codelet is mapped an atom in the switching hardware [33]. For more detailed information about the compiling process, the paper on Domino [7] can be studied.

5.3.2.3 POF

Protocol Oblivious Forwarding is an improvement to the SDN forwarding architecture based on OpenFlow as it increases the SDN programmability [2]. POF allows forwarding devices to adopt any new protocol without altering the devices' software [2]. To achieve this objective, a protocol abstraction is utilized. As previously stated, existing OpenFlow-based SDN increases network device programmability by loading flows into devices from the controller, however it cannot dynamically support additional protocols. The programmability is limited to established protocols. If a new protocol-based service must be introduced, the operator must request that the device manufacturer adjust the code of the devices to accommodate the new service.

POF essentially adopts OpenFlow's network topology, in that a centralized controller exists in the control plane to regulate the forwarding activities of the switches in the forwarding plane using Flow-tables. POF's advances, on the other hand, are the protocol-oblivious definition for flow-matching fields and a collection of generalized flow operations, with which protocol-independent packet forwarding in switches may be efficiently accomplished [35].

POF represents any protocol field with the following format where offset indicates the starting point of the field relative to the current protocol head and length the field's length in bits [2]. As an example, the figure below depicts the Ethernet protocol header format with its corresponding fields. It is apparent that any current or new protocol may be represented by the offset, length tuple.

To achieve packet parsing and flow matching, the switches employ a series of basic search keys and table lookup instructions. POF essentially specifies a matching field's search key as a tuple "offset, length." All of the directives in the tables use the tuple "offset, length" as well, to find the data on which to act. This allows one to alter any packet at any time, which is significantly more adaptable than the OpenFlow method. As mentioned above in the OpenFlow switch specifications, an action remains protocol-specific, requiring

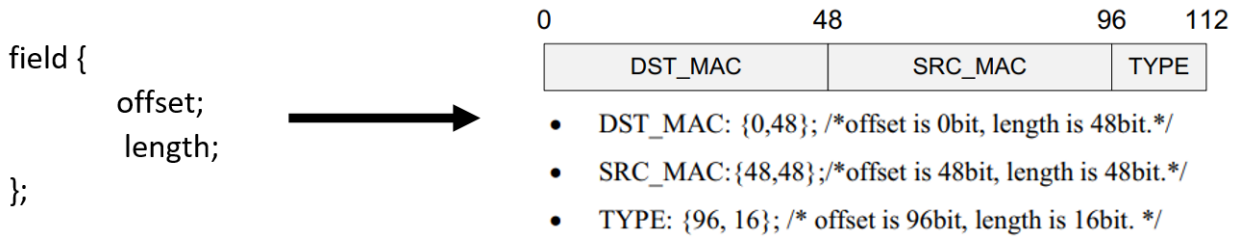


Figure 5.10: POF - protocol field representation based on Yu, J., Wang, X., Song, J., Zheng, Y., Song, H., (2014) Forwarding programming in protocol-oblivious instruction set

numerous actions to be provided for each outdated protocol. Flow tables in a POF-enabled switch are categorised into four types: masked match (MM), longest-prefix-match (LPM), extract-match (EM), and direct tables (DT). These tables have varying memory capacities and can be searched using a variety of table lookup techniques [1]. To be more specific, Yu, J. et al. (2014), outline the packet structure recommended for POF-based source routing where the source routing-related header fields are placed between the IP and Ethernet and IP header.

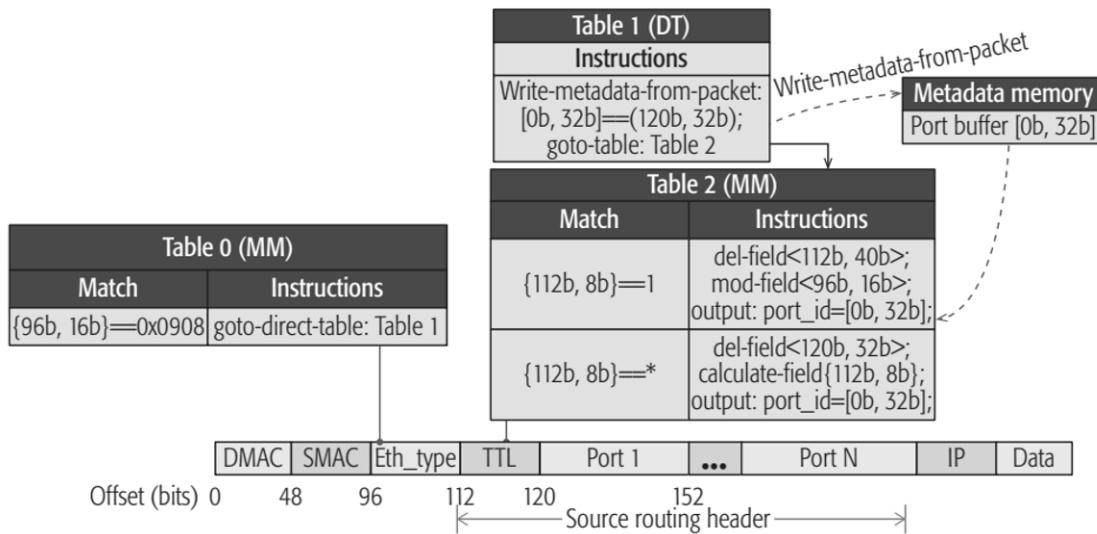


Figure 5.11: POF-based source routing taken from Yu, J., Wang, X., Song, J., Zheng, Y., Song, H., (2014) Forwarding programming in protocol-oblivious instruction set

The next section provides an overview of the different processes that occur during source routing with protocol-oblivious forwarding, which as well can be read in more detail in [1][2][35].

- When the first packet reaches the ingress edge switch, with the help of the POF controller, the source routing related header field is inserted in between the Ethernet header respectively the IP header in a first step, and the Ethernet header’s type field is changed to 0x0908 in a second step, to signal that the Ethernet frame includes a POF-based source routing packet.

- In Figure 5.11, Table 0 contains a check to see whether the type field of the packet's Ethernet header matches 0x0908.
- If the Ethernet frame comprises a POF-based source routing packet, the switch at table 1 performs the write-metadata-from-packet action to replicate the Port field to its metadata memory to encode the output port.
- If the time-to-Live field in the source routing header matches 1 (the specific switch is the packet's final hop), the switch executes the del-field operation to erase all source routing related fields in the packet and reset the type field in the Ethernet header to its initial value. Otherwise, the switch does nothing but delete the Port field. The packet is subsequently sent to its selected output port using the output instruction.

As we have seen, POF's main idea is to decouple the defined network protocols from the packet forwarding in order to make the forwarding plane adaptable and programmable. POF, in particular, presents a protocol-independent instruction set, allowing a network operator to specify the protocol stack and packet forwarding mechanism in a considerably more modular fashion than the OpenFlow standards.

5.3.2.4 SRv6 programmability

The notion of SRv6 programmability is defined as segment routing over IPv6 (SRv6) network programming [6]. It is a framework, introduced by the IETF, which enables network operators to encode sequences of instructions in IPv6 packet headers and therefore, being able to achieve processing programmability. Every instruction can be implemented in different network nodes. To identify these instructions, SRv6 segment identifiers (SID) in the packets are utilized [6]. Basically, SRv6 is a combination of the traditional segment routing and IPv6, where flexible IPv6 extension headers are used to implement network programming [36]. Before looking closer at SRv6, the concept of segment routing must be introduced.

The term segment routing refers to an architecture which is flexible and scalable. It was introduced in 2013 by IETF [37] and emerged as networks evolved more and more towards application-centric platforms requiring more flexibility. The basic technique of segment routing is to allow an edge router to steer packets through a network by making use of segments. Such segments are defined as identifiers for topological instructions which enables packets to be steered over certain paths. This omits the need for intermediate routers or nodes to contain the state for the different paths [38]. As the only the ingress node maintains per-flow state information, this enables better network scalability. The TCAM memory requires less space to store flow table. Overall, segment routing is turning out to be an effective alternative for the traditional destination based routing schemes. It can solve challenges of traditional networks by using an ordered list of instructions, which is called the segment list, to navigate through the paths. Segment lists are encoded in the packet headers as either MPLS labels or as IPv6 addresses [39]. The latter case is the one which is relevant for SRv6.

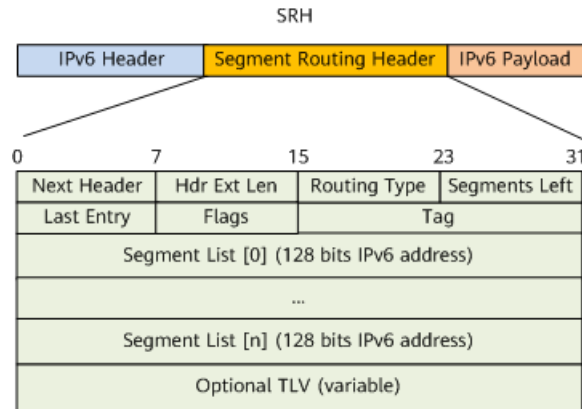


Figure 5.12: The format of a segment routing header (SRH) in SRv6 taken from Lanjun, L. (2021) SRv6

The segment routing architecture consists of two main components, the data-plane and the control-plane of segment routing. The data-plane part is concerned with the segment routing header, as seen in figure 5.12. It is an additional part of a IPv6 header and contains a pointer to the active segment of the packet and the sequence of segments (segment lists). The pointer is the instruction which the current node must execute. After execution, the next segment in the list becomes the active one. Nodes can support different data-plane operations: CONTINUE (forwarding action), PUSH (set active segment in header), and NEXT (mark next segment as the active one) [38]. As mentioned before a segment identifier (SID) identifies the segments and there are three main types:

- Node SID
- Adjacency SID
- Service SID

The control-plane part is responsible for the SIDs to be communicated among the nodes using protocols such as IGP. Routers can have a database with all current segments stored in it. Further, the control-plane tells the ingress node how to select a path for packets. For this there are different methods such as the distributed constrained SPF calculation (shortest path with some criteria) or the SDN controller based approach, which can make use of the PCEP protocol [38].

In SRv6 schemes, IPv6 is used as the data-plane technology while MPLS is the other used option (SR-MPLS). SRv6 uses the segment routing header (SRH) to define the explicit path by storing the IPv6 segment list information in it. In order to determine how a packet should be forwarded, the SRH field segments left and segment list are required in order to retrieve the IPv6 destination address. SRv6 comes with three programming capabilities, described by Lanjun (2021) as the three-dimensional programming space. It consists of path programming, where the paths can be defined by multiple segments with the segment lists. Secondly service programming is made possible, due to flexible programming capabilities with the SRv6 SIDs, which are 128-bit IPv6 addresses. These

addresses consist of a locator, a function, and optional arguments which all together provide the capability to program different services. Finally, the application programming can be achieved, through type-length values capturing irregular information of the network and forwarding plane. This is why SRv6 can interact with SDN networks and achieve service-driven networks [36].

The forwarding process of SRv6 works as follows. All the SIDs of the nodes located on the path from A to B are encapsulated in reverse order in the SRH. The segment left field in the SRH points to the segment list which is current. Then, the first node takes the value of the segment list and writes it to the destination address field in the IPv6 header. Then the packet is forwarded to the next node. It is important to mention, that devices which do not support SRv6 can still be involved in this process. In this case, the device would simply forward the packet based on the current destination address. On the other hand, SRv6 devices look up their local SID tables for matching SID (e.g., End.X SID). The final node decapsulates the packet and forwards the packet to the end host. Another way to look at the forwarding of SRv6 is the SRv6 working mode, which incorporates a (SDN) controller instance, which receives network topology information and delivers the path information to the ingress node. SRv6 works with both a traffic engineering policy or a best effort policy [36].

Many advantages come with the SRv6 approach. Besides offering extensibility and programmability, the number of required protocol types is reduced. Such protocols include for instance LDP or RSVP. Further, SRv6 meets the diversified requirements of novel services, provides high reliability, and offers potential for cloud applications. Further, native IPv6 devices without the capabilities to support SRv6 can still be used in the forwarding process of SRv6. This allows for SRv6 devices to work together with traditional IPv6 devices. When it comes to the costs, this is a important factor which speaks in favor of SRv6.[36]. Other use cases of segment routing in general are traffic engineering using SR tunnels, service function chaining, and segment routing based network resiliency [38].

One recent use case of SRv6 programmability was proposed by Gramaglia et al. (2020). The research was able to show advantages of segment routing in the context of 5G networks which use network slicing. The paradigm enables operators to manage different virtual instances of their respective mobile 5G network. This way, a diverse collection of services can be delivered, creating much more flexibility. At the same time, the overall resource utilization is increased [40]. The tunneling protocol which is mainly used is the GPRS tunneling protocol (GTP). It was shown that SRv6 has the potential to be a valid alternative approach. Two different modes could be possible, depending if GTP should be replaced completely or not: traditional and enhanced modes. In traditional mode, the network architecture remains the same and GTP is mapped to SRv6 encapsulations. More advanced functions are possible in enhanced mode, where packets can traverse mode nodes (more than one SID in headers). Services such as traffic engineering are made possible. Overall, SRv6 simplifies the transport of user data much simpler by adding tunnel endpoint identifiers (TEID) being encoded in the protocol stack. As shown in the research, SRv6 also uses much less computational resources because the path information is indicated upfront. Finally, SRv6 can replace underlay transport layers such as MPLS, making it the sole transport layer protocol in 5G networks [40].

5.4 Comparison Of Techniques And Discussion

In the following section, we will go deeper into the points raised above in order to provide a more complete overview draw a comparison of the various data plane programmability approaches. POF-FIS is adaptable in terms of implementing protocol rules and quickly deploying services, whether current or new ones. POF-FIS, being the primary southbound interface element in SDN, is independent of the high-level programming language, target platform or northbound interface. Therefore, Administrators can design the forwarding mechanism using C, Java, P4, or any high-level language. Administrators can also directly control the POF-FIS to build the entire forwarding process. The generic POF-FIS considerably enhances network element adaptability. POF-FIS enables the flexible network components' forwarding capabilities to be completely released, resulting in greater efficiency and more expressive forwarding behavior. Even if POF enables network devices to offer full programmability for any new feature based on existing or new protocols, POF has encountered several difficulties. Many manufacturers have created SDN forwarding devices and controllers that do not support POF. If POF is to be utilized globally, it must be standardized as a component of the SDN deployment. Beside POF, P4 and Domino, as high-level programming languages, also try to improve the OpenFlow architecture design. As the control plane cannot define how packets should be treated to best satisfy the demands of the control applications with OpenFlow, P4 and Domino suggest a step toward more adjustable switches. An administrator determines how the forwarding plane handles packets without regard for implementation concerns. A compiler converts the program into a table dependency graph that may be translated to a variety of specialized target switches, including efficient hardware implementations. Just like POF, P4 also has some challenges to overcome. In P4, there is no iteration concept. Only the state machine of the parser can produce loops. Recursive procedures are not supported. As a result, the work produced by a P4 program is linearly dependent solely on the header sizes. Due to the lack of loops, P4 is also ineffective for deep packet inspection as P4 applications, in general, cannot undertake any useful processing of the packet payload. A few years after P4, Domino introduced the ability to express data plane algorithms while maintaining performance of high-speed switches at line-rate. Like with P4, iterations and memory access are not supported by the language. However, Domino is suitable for small sets of modifiable packet headers and can use stateful algorithms, which P4 is not that suitable for [33]. As Domino targets solely a single switch device, there have been further developments of technologies, such as SNAP which addresses multiple hardware targets [41]. Another remaining challenge in recent developments, are the flexible and platform independent programming abstractions for stateful packet processing on data plane devices. Even though performance can be greatly reduced, SDN controllers still utilize this stateful processing. In order to tackle this, Domino introduced the abstraction of packet transactions, which helped to express stateful algorithms on the data plane without having to define match-action tables. As seen in the Domino sub-chapter, the atoms are responsible for encompassing the instruction sets, which are then configured by the compiler [17]. A similar concept like POF, is the SRv6 (segment routing over IPv6) technology, which came a few years after. Thus, it can be regarded as an extension of that technology. It shows promising use cases in fields such as 5G and cloud services, as it combines segment routing with IPv6 and MPLS. It also supports native IPv6 devices for forwarding pack-

ets, which can greatly reduce the costs of new hardware. It also introduced huge benefits such as network scalability and flexibility when it comes to data-plane forwarding. The figure 5.13 should highlight the key aspects of the various data plane programmability techniques which were discussed in this report.

	P4	Domino	POF	SRv6
<i>Year of introduction</i>	2013 (2014)	2016	2013	2017 (SR: 2015)
<i>Concept</i>	Defining how packets traverse programmable dataplane blocks with P4	Packet transactions with Domino language, Banzai model, Compiler	Packet parsing and flow matching based on a sequence of generic key assembly and table lookup instructions	Segment routing in combination with IPv6
<i>Primary Goals</i>	Protocol independence, Target independence, Reconfigurability in the field, (OpenFlow enhancement)	Data-plane algorithms while maintaining the performance of a line-rate switch	Enhancement to OpenFlow-based forwarding architecture, Support any new protocols without modifying any code	Simplified control plane (less protocols), traffic control, native IPv6 packets/devices supported
<i>Use cases</i>	Traffic engineering, Telemetry, 5G VNF offloading	Research, small sets of packet headers modifiable	All existing network application OpenFlow supports + user-defined forwarding protocols	Cloud services, traffic engineering, 5G networks / network slicing

Figure 5.13: Comparison of different data plane programmability approaches

5.5 Conclusion

This report presented many emerging technologies tackling the challenges of data plane programmability in recent years. Programmable data planes are the next step in allowing switches to conduct sophisticated packet processing. A data plane which is programmable goes beyond the abilities of a switch that solely employs OpenFlow. A configurable data plane can route packets based on factors other than those specified in Layer two or IP packet fields when referencing the OSI model. Programmable data planes simplify to perform tasks such as executing achieving load balancing by transmitting packets with the same characteristics out to various ports. The paper also demonstrated that programmable data planes are able to implement various sorts of policies. By incorporating programmable data planes policies based on packet rates can be easily implemented. Additionally, the flexibility to forward packets based on any format may be essential for supporting IoT devices. Ever since SDN introduced the notion of separating the data plane from the control plane, new opportunities arose, however accompanied by downsides as well. With all the promising data plane technologies such as the programming languages P4 and Domino, and the southbound APIs such as POF and SRv6, the overall concept of SDN still remains with some challenges to face. One of them is that a programming language like P4 does not yet have the flexibility that other programming languages have, making it challenging to define desired methods. Data plane programmability involves programming which is never straightforward. Programming the network is still comparatively new, with few mature tools to guide and support the development process. The tendency indicates that even more money will be invested in researching the field of data plane programmability. Even if P4 is currently the most used and best known

when talking about data plane programmability, it is not guaranteed which technology will prevail in the future as, depending on the scenario, they do not always compete and can be used combined.

Bibliography

- [1] Yu, J., Wang, X., Song, J., Zheng, Y., Song, H.: *Forwarding programming in protocol-oblivious instruction set*, IEEE 22nd International Conference on Network Protocols (pp. 577-582). IEEE, 2014.
- [2] Song, H., Gong, J., Wang, X., Zheng, Y.: *Principle and Implementation of Protocol Oblivious Forwarding*, 2012.
- [3] McKeown, Nick and Anderson, Tom and Balakrishnan, Hari and Parulkar, Guru and Peterson, Larry and Rexford, Jennifer and Shenker, Scott and Turner, Jonathan: *OpenFlow: enabling innovation in campus networks*, ACM SIGCOMM computer communication review, 38(2), 69-74, 2018. <https://dl.acm.org/doi/pdf/10.1145/1355734.1355746>.
- [4] Kreutz, D., Ramos, F. M., Verissimo, P. E., Rothenberg, C. E., Azodolmolky, S., Uhlig, S.: *Software-defined networking: A comprehensive survey*, Proceedings of the IEEE, 103(1), (pp. 14-76), 2014. <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6994333>
- [5] Bosshart, P., Daly, D., Gibb, G., Izzard, M., McKeown, N., Rexford, J., Schlesinger, C., Talayco, D., Vahdat, A., Varghese, G., Walker, D.: *P4: Programming protocol-independent packet processors*, ACM SIGCOMM Computer Communication Review, 44(3), (pp. 87-95), 2014. <https://dl.acm.org/doi/pdf/10.1145/2656877.2656890>
- [6] Filsfils, C., Camarillo, P., Leddy, J., Voyer, D., Matsushima, S., Li, Z.: *Segment routing over IPv6 (SRv6) network programming*, IETF RFC 8986, 2021. <https://www.rfc-editor.org/rfc/rfc8986.pdf>
- [7] Sivaraman, A., Cheung, A., Budiu, M., Kim, C., Alizadeh, M., Balakrishnan, H., Varghese, G., McKeown, N., Licking, S.: *Packet transactions: High-level programming for line-rate switches*, Proceedings of the 2016 ACM SIGCOMM Conference (pp. 15-28), August 2016. <https://dl.acm.org/doi/pdf/10.1145/2934872.2934900>
- [8] Lee, B. K., John, L. K.: *NpBench: A benchmark suite for control plane and data plane applications for network processors*, In Proceedings 21st International Conference on Computer Design (pp. 226-233), 2003. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1240899>

- [9] Celesova, B., Val'ko, J., Grezo, R., Helebrandt, P.: *Enhancing security of SDN focusing on control plane and data plane*, 7th International Symposium on Digital Forensics and Security (ISDFS) (pp. 1-6), 2019. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8757542>
- [10] Wickboldt, J. A., De Jesus, W. P., Isolani, P. H., Both, C. B., Rochol, J., Granville, L. Z.: *Software-defined networking: management requirements and challenges*, IEEE Communications Magazine, 53(1), 278-285, 2015. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7010546>
- [11] Isyaku, B., Mohd Zahid, M. S., Bte Kamat, M., Abu Bakar, K., Ghaleb, F. A.: *Software defined networking flow table management of openflow switches performance and security challenges: A survey*, Future Internet, 12(9), 147, 2020. <https://www.mdpi.com/1999-5903/12/9/147/pdf>
- [12] Li, Chung-Sheng, Liao, Wanjiun: *Software defined networks*, IEEE Communications Magazine, 51(2), 113-113, 2013. researchgate.net/profile/Chung-Sheng-Li-2/publication/260670412_Software_defined_networks_Guest_Editorial/links/552424630cf2caf11bfcc125/Software-defined-networks-Guest-Editorial.pdf.
- [13] Enns, R., Bjorklund, M., Schoenwaelder, J.: *NETCONF configuration protocol*, RFC 4741, 2006. <https://www.hjp.at/doc/rfc/rfc4741.html>
- [14] Enns, R., Bjorklund, M., Schoenwaelder, J., Bierman, A.: *Network configuration protocol (NETCONF)*, 2011. <https://www.hjp.at/doc/rfc/rfc6241.html>
- [15] Stancu, A., Avram, A., Skorupski, M., Vulpe, A., Halunga, S.: *Enabling SDN application development using a NETCONF mediator layer simulator*, 2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN) (pp. 658-663). IEEE, July 2017. <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7993873>
- [16] Yu, F., Katz, R. H., Lakshman, T. V.: *Gigabit rate packet pattern-matching using TCAM*, Proceedings of the 12th IEEE International Conference on Network Protocols (pp. 174-183), October 2004. <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1348108>
- [17] Michel, O., Bifulco, R., Retvari, G., Schmid, S.: *The programmable data plane: abstractions, architectures, algorithms, and applications*, ACM Computing Surveys (CSUR), 54(4), (pp. 1-36), 2021. <https://dl.acm.org/doi/pdf/10.1145/3447868>
- [18] Pfaff, B., Pettit, J., Koponen, T., Jackson, E., Zhou, A., Rajahalme, J., Gross, J., Wang, A., Stringer, J., Shelar, P., Amidon, K., Casado, M.: *The Design and Implementation of Open vSwitch*, 12th USENIX symposium on networked systems design and implementation (NSDI 15) (pp. 117-130), 2015. <https://www.usenix.org/system/files/conference/nsdi15/nsdi15-paper-pfaff.pdf>
- [19] Farrel, A., Vasseur, J.-P., and J. Ash: *A Path Computation Element (PCE)-Based Architecture*, RFC 4655, DOI 10.17487/RFC4655, 2006. <https://www.rfc-editor.org/info/rfc4655>

- [20] Otani, T., Ogaki, K., Caviglia, D., Zhang, F., and C. Margaria: *Requirements for GMPLS Applications of PCE*, RFC 7025, DOI 10.17487/RFC7025, 2013. <https://www.rfc-editor.org/info/rfc7025>
- [21] Crabbe, E., Minei, I., Medved, J., and R. Varga: *Path Computation Element Communication Protocol (PCEP) Extensions for Stateful PCE*, RFC 8231, DOI 10.17487/RFC8231, 2017. <https://www.rfc-editor.org/info/rfc8231>
- [22] King, D. and A. Farrel: *A PCE-Based Architecture for Application-Based Network Operations*, RFC 7491, DOI 10.17487/RFC7491, 2015. <https://www.rfc-editor.org/info/rfc7491>
- [23] Khorsandroo, S., Sanchez, A. G., Tosun, A. S., Arco, J. M., Doriguzzi-Corin, R.: *Hybrid SDN evolution: A comprehensive survey of the state-of-the-art*, Computer Networks, 192, 107981, 2021. <https://arxiv.org/pdf/2103.16444.pdf>
- [24] Li, S., Hu, D., Fang, W., Ma, S., Chen, C., Huang, H., Zhu, Z.: *Protocol oblivious forwarding (POF): Software-defined networking with enhanced programmability*, IEEE Network, 31(2), (pp. 58-66), 2017. <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7884951>
- [25] Pawel, P.: *P4 programming Language - Introduction to network programming with P4 Course*, CodiLime, 2020. <https://www.youtube.com/watch?v=UEMAvXXNwsY>
- [26] Da Costa Cordeiro, W. L., Marques, J. A., Gaspar, L. P.: *Data plane programmability beyond openflow: Opportunities and challenges for network and service operations and management*, Journal of Network and Systems Management, 25(4), (pp. 784-818), 2017. <https://link.springer.com/content/pdf/10.1007/s10922-017-9423-2.pdf>
- [27] Zanna, P., Radcliffe, P., Chavez, K. G.: *A Method for Comparing OpenFlow and P4*, In 2019 29th International Telecommunication Networks and Applications Conference (ITNAC) (pp. 1-3). IEEE, 2019. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9077951>
- [28] The P4.org Architecture Working Group: *P416 Portable Switch Architecture (PSA)*, working draft, 2021. <https://p4.org/p4-spec/docs/PSA.pdf>
- [29] Fattaholmanan, A., Baldi, M., Carzaniga, A., Soulé, R.: *P4 Weaver: Supporting Modular and Incremental Programming in P4*, In Proceedings of the ACM SIGCOMM Symposium on SDN Research (SOSR) (pp. 54-65), August 2021. <https://dl.acm.org/doi/pdf/10.1145/3482898.3483353>
- [30] Barefoot: The World's Fastest and Most Programmable Networks. https://barefootnetworks.com/media/white_papers/Barefoot-Worlds-Fastest-Most-Programmable-Networks.pdf.
- [31] XPliant Ethernet Switch Product Family. <http://www.cavium.com/XPliant-Ethernet-Switch-Product-Family.html>.

- [32] P. Bosshart, G. Gibb, H.-S. Kim, G. Varghese, N. McKeown, M. Izzard, F. Mujica, and M. Horowitz.: *Forwarding Metamorphosis: Fast Programmable Match-action Processing in Hardware for SDN*, In SIGCOMM, 2013. <https://dl.acm.org/doi/pdf/10.1145/2534169.2486011>
- [33] Sivaraman, A., Cheung, A., Budiu, M., Kim, C., Alizadeh, M., Balakrishnan, H., Varghese, G., McKeown, N., Licking, S.: *Packet Transactions: High-Level Programming for Line-Rate Switches*, MIT CSAIL, University of Washington, Barefoot Networks, Microsoft Research, and Stanford University (SIGCOMM 2016), 2016. <http://web.mit.edu/domino/>
- [34] S. Palkar, C. Lan, S. Han, K. Jang, A. Panda, S. Ratnasamy, L. Rizzo, and S. Shenker: *E2: A Framework for NFV Applications*, SOSP, 2015. <https://dl.acm.org/doi/pdf/10.1145/2815400.2815423>
- [35] Abdullah, Li, S., Hu, D., Fang, W., Zhu, Z.: *Source routing with protocol-oblivious forwarding (POF) to enable efficient e-health data transfers*, IEEE International Conference on Communications (ICC) (pp. 1-6). IEEE, 2016.
- [36] Lanjun, L.: *SRv6*, Huawei Technologies Co., Ltd. , 2021. <https://support.huawei.com/enterprise/en/doc/ED0C1100200080>
- [37] IETF: *Source Packet Routing in Networking*, 2013. <https://datatracker.ietf.org/wg/spring/charter/>
- [38] Filsfils, C., Nainar, N. K., Pignataro, C., Cardona, J. C., Francois, P.: *The segment routing architecture*, 2015 IEEE Global Communications Conference (GLOBECOM) (pp. 1-6). IEEE, December 2015. https://dSPACE.networks.imdea.org/bitstream/handle/20.500.12761/205/the_segment-routing_architecture_2015.pdf?sequence=1
- [39] Abdullah, Z. N., Ahmad, I., Hussain, I.: *Segment routing in software defined networks: A survey*, IEEE Communications Surveys Tutorials, 21(1), (pp. 464-486). IEEE, 2018. <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8462720>
- [40] Gramaglia, M., Sciancalepore, V., Fernandez-Maestro, F. J., Perez, R., Serrano, P., Banchs, A.: *Experimenting with SRv6: a Tunneling Protocol supporting Network Slicing in 5G and beyond*, 2020 IEEE 25th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD) (pp. 1-6). IEEE, September 2020. <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9209260>
- [41] Arashloo, M. T., Koral, Y., Greenberg, M., Rexford, J., Walker, D.: *SNAP: Stateful network-wide abstractions for packet processing*, In Proceedings of the 2016 ACM SIGCOMM Conference (pp. 29-43), August 2016. <https://dl.acm.org/doi/pdf/10.1145/2934872.2934892>

Chapter 6

Digital Sovereignty and Digital Colonization: A Technical Discussion Regarding Challenges and the Digital Future of Countries

Jiaming Tong

Contents

6.1	Introduction	133
6.2	Background	134
6.2.1	Sovereignty and Digital Sovereignty	134
6.2.2	Colonization and Digital Colonization	135
6.2.3	Technologies, Politics and Digital Sovereignty	135
6.3	Challenges for a Digital Sovereignty	137
6.3.1	Data Security and Data Privacy	137
6.3.2	Cyber Security	138
6.3.3	5G Infrastructure	139
6.3.4	Facebook and the Digital Colonialism	140
6.4	The Path for Digital Sovereignty	140
6.4.1	National Cybersecurity Strategy (NCCS)	140
6.4.2	General Data Protection Regulation (GDPR)	141
6.4.3	Investment in Data Security and Privacy	141
6.4.4	Fighting against Cybercrime	142
6.4.5	IT-Security Law 2.0	142
6.4.6	5G Security	142
6.5	Summary and Conclusions	142

6.1 Introduction

The Internet and digital systems have become a key not only for communication and innovation but for the economy, military, and sovereignty of countries worldwide. Research suggests that modern societies face three existential challenges: nuclear war, ecological collapse, and technological disruption. It is the last of the three that can affect the developmental trajectory of states and societies [1].

Sovereignty and colonization are concepts that existed before human beings entered the information era. In the long past, there was neither electricity nor the Internet. Therefore, most of the people's assets were tangible (i.e. can be seen and touched), such as precious metals, food, lands, and so on. However, with the development of IT technology, human beings have entered the era of informatization and digitization, and assets are not only limited to those tangible assets mentioned above. Many emerging digital assets appeared, such as valuable data, cryptocurrencies, and Non-Tangible Tokens (NFT) [2].

The change of form of assets is affecting current politics. From Marxism, the concept that the base determines the superstructure is mentioned. The base refers to the mode of production which includes the forces and relations of production (e.g., employer-employee work conditions, the technical division of labor, and property relations) into which people enter to produce the necessities and amenities of life. The superstructure refers to society's other relationships and ideas not directly relating to production including its culture, institutions, political power structures, roles, rituals, religion, media, and state [3]. Since the economic foundations such as the form of property existence have changed with the development of science and technology, the superstructure in the political sense such as sovereignty and colonization also need to be redefined and discussed.

This article focuses on the topic of digital sovereignty and digital colonization. First, in this article, we introduce the concepts of digital sovereignty and digital colonization, including the context in which they emerged, such as how technologies gave birth to this new concept by comparing them with their traditional versions. i.e., how nascent technologies gave birth to this new concept by comparing them with their traditional versions. Then we will introduce some background about new digital technology and how they influenced our daily life and politics. After that, the challenges faced by the digital society are introduced, such as Europe's technological dependence on other countries, and the current issues of digital sovereignty, such as issues about data security, privacy, and overall cybersecurity. As shown in Figure 2.1, over half of the Europeans think that their country is too dependent on foreign digital technologies. Finally, this article gives some possible solutions and insights for some of the problems mentioned above, including an outlook on what to expect for the future of the digital world.

Figure 6.1: Share fo Europeans who think their country is too dependent on foreign digital technologies

6.2 Background

In this section, we introduce the main concepts required for a full understanding of this paper. This section also includes basic concepts about sovereignty and colonization as well as their digital versions.

6.2.1 Sovereignty and Digital Sovereignty

Different theories have different views on the concept of sovereignty space, such as Domestic sovereignty, Interdependence sovereignty, and so on, and this article will not delve into the definition of sovereignty. In the classic, post-Westphalian system, sovereignty is understood as the exclusive authority of the State over persons and things within a specified territory [4]. All three elements of this definition - the nature of power/authority, its exclusivity, and its territoriality - have been challenged by the invention of interconnected global communications networks, in short: cyberspace. Because cyberspace creates a space for storage of and access to information, as well as social interaction regardless of the user's location and irrespective of distances, it creates the perception of a space not restricted by - or even detached from - geography. In other words, cyberspace is perceived as a territorial [5].

The definition of digital sovereignty is still not clear enough, and the scope of the adaption of digital sovereignty is still controversial in the international and academic circles. While the 2013 and 2015 Reports of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security [GGE] have confirmed that sovereignty and the international norms and principles that flow from it apply to State conduct in cyberspace, they have left open the meaning and scope of sovereignty concerning the cyber domain [6].

The author believes that the scope of application of digital sovereignty includes but is not limited to personal data of the public or enterprises, digital assets of the public or enterprises, network security, and the government's control of public opinion in cyberspace, etc.

One commonly cited example of a digital sovereignty principle in play is the European General Data Protection Regulation (GDPR) which establishes key requirements for data handling related to European individuals or businesses. The GDPR extends that data sovereignty principle around the world and requires foreign dealers in European data to get in line [7]. In short, GDPR is also a code of conduct for companies or individuals just like other laws, but what is different from other laws is the scope of its use. The scope of application of traditional laws is the real world, while the GDPR has expanded to a new scope of application., that is, cyberspace or digital space, which can be regarded as a manifestation of digital sovereignty.

6.2.2 Colonization and Digital Colonization

The term colonization is derived from the Latin words *colere* ("to cultivate, to till") [8]. Colonization originally refers to large-scale population movements where the migrants maintain strong links with their or their ancestors' former country, gaining significant privileges over other inhabitants of the territory by such links.

In the Colonial Era, colonialism in this context refers mostly to Western European countries' colonization of lands mainly in the Americas, Africa, Asia, and Oceania. As mentioned above, this article also does not make too much discussion and research on the meaning of colonization itself. The author here summarizes the characteristics of traditional colonialism. The colonizers exploited the wealth such as resources and labor force of the colonies by establishing a dominant military position or other methods while dumping products to the colonies and exporting ideology and values to intervene in colonial politics. After this, new forms of colonialism such as economic colonization also emerged the difference is that this type of colonization is usually built based on capital and technology.

Digital colonialism refers to the use of digital technology to exercise political, economic, and social domination of another country or region [9]. Digital colonialism is to consolidate the unequal division of labor, and the dominant state uses its ownership of digital infrastructure, knowledge, and control over computing means to keep the countries of the South in a state of dependence for a long time. This unequal division of labor has evolved. Economically, manufacturing has moved down the value hierarchy, being replaced by an advanced high-tech economy firmly in the hands of tech giants.

In the information age, data is an important asset, and data can be used in various aspects, such as business analysis, machine learning, advertising recommendations, etc. Large Internet companies such as Google have 4.3 billion users worldwide. Based on 4.72 billion These companies can obtain huge amounts of data from their users and use it to enhance their services and reap huge benefits. The user attention brought by the number of users is also an important means for large Internet companies to make profits e.g., through advertisement. Figure 2.2 shows leading countries based on Facebook audience size. What's more, social media such as Facebook and Twitter also have an important influence on public opinion. Through content review and screening, these social media platforms can guide and control the public opinion of their users, and virtually shape the ideology of users.

Figure 6.2: Leading countries based on Facebook audience size as of January 2022

6.2.3 Technologies, Politics and Digital Sovereignty

In this subsection, we introduce some emerging technologies that give some examples of digital sovereignty and digital colonization.

6.2.3.1 Big Data

Big data can influence politics in many ways, such as public opinion analysis and aiding elections, etc. The government can obtain public information published by the public on the Internet through technologies such as web crawlers and conduct processing and analysis through natural language processing and other means to analyze public opinion and assist policy designation. In addition, the application of big data also has an impact on the election, like corporations, campaigns now know far more about their constituents than ever before what they read, which movies they stream, which shows they watch, where they shop, which products they buy [10]. If the elector can obtain these data about the voters, then we can say that whoever obtains the data has the initiative in the election. This shows that data is a very important property and should be considered as part of digital sovereignty.

6.2.3.2 Telecommunications

Founded in 1987, Huawei is a leading global provider of information and communications technology (ICT) infrastructure and smart devices [11]. Although two European alternatives are available in Ericsson and Nokia, the Chinese company has considerable market shares in some EU countries. In some cases, the entire mobile network is based on Chinese technology [12]. The author believes that telecom infrastructure, like other infrastructure such as water and electricity, plays a vital role in the security of digital sovereignty. Too much reliance on the technical support of a single country is like putting eggs in the same basket, which will bring certain risks to its security.

6.2.3.3 Digital Social Media

It is well known that during the 2020 US election, then-President Donald Trump's Twitter account was blocked for inciting violence. There is no doubt that Twitter's ban on Trump's account has had a certain impact on American politics. The CEO of a technology company can cut the president of the world's most powerful democracy from his supporters. Isn't it something to be wary of? This shows that the country's digital sovereignty is partly in the hands of some tech giants. After all, Twitter is an American company, and for the United States, the runaway of digital sovereignty is not particularly serious. But what if Twitter could also ban the accounts of leaders of other countries at will? "The right to freedom of opinion is of fundamental importance," Steffen Seibert, Merkel's chief spokesman, told reporters in Berlin on Monday, according to Reuters [13].

6.2.3.4 Digital Finance

Digital finance is the term used to describe the impact of new technologies on the financial services industry. It includes a variety of products, applications, processes, and business models that have transformed the traditional way of providing banking and financial services [14].

Even ten years ago, one could not have imagined that electronic or mobile payments would be as commonplace as they are today. Electronic payment has greatly changed the way people live and shop. You don't need to carry cash or credit cards, you only need to carry your mobile phone to make payments and shopping, which not only provides convenience for people but also avoids the risk of losing cash or credit cards. In China, electronic payment has a very high penetration rate, and over 80% of adults use electronic payment [15]. As shown in Figure 2.3, Popular mobile payments around the world include PayPal, Amazon Pay, eBay, Google Pay, Apple pay, Alipay, etc. Almost all of the above electronic payment methods are provided by companies in the United States.

Figure 6.3: Common electronic payments

6.3 Challenges for a Digital Sovereignty

In this section, we introduce the main challenges to digital sovereignty in this paper including the current situation of data security, data privacy, cyber security, and so on.

6.3.1 Data Security and Data Privacy

The Ponemon Institute's Cost of Data Breach Study found that on average, the damage caused by a data breach in the USA was \$8 million [16]. It is obvious that data security plays a vital role.

The information economy has proven to be a deteriorating issue in US-EU relations. In addition to concerns over data security, European politicians have also raised concerns about the market dominance of U.S. companies, especially in data storage and analysis. This week, Ms. Merkel warned of the risks of "reliance" on US technology [17].

However, the protection of data security and data privacy also faces many challenges. With the exponential growth of data, the protection of data security and data privacy is becoming more and more difficult. Threats to data security come from many aspects, including but not limited to, accidental leakage, malicious attacks, illegal collection, and utilization of data, etc.

Privacy is the ability of an individual or group to seclude themselves or information about themselves and thereby express themselves selectively. When something is private to a person, it usually means that something is inherently special or sensitive to them. The domain of privacy partially overlaps with security, which can include the concepts of appropriate use and protection of information [18]. In short, privacy can be understood as some information that is only relevant to an individual and does not want to be disclosed to the public.

The protection of private data also faces many challenges which are linked to technology moving at a quicker pace than the law. Even though many States have data protection legislation in place, some of it is not sufficiently updated to take technological advancements into account [19].

The opaque use of private data by businesses is also a challenge for data privacy protection. Companies often share their clients' information with third parties companies. This gives the third parties companies the ability to use the data as it is for their profit [20].

6.3.2 Cyber Security

All aspects of human life today are inseparable from information and communication technology. Almost everything is done on the Internet, such as shopping, transactions, and more. Cyberspace, like land, sea, and sky, is the fourth space in which people live. Therefore, cybersecurity is closely related to national security and sovereign security.

Just like in showing crime in life, there are all kinds of crimes on the web and these types of crimes are known as cybercrime. Cybercrime is any criminal activity involving computers, networked devices, or networks. Cybercrime consists of many kinds, including but not limited to cyber extortion, crypto-jacking, theft of personal information, pirated software, and more. Take cyber extortion, for example, where hackers encrypt corporate or personal data and demand a ransom to unlock the files. Cybercrime comes in many forms, such as cyberattacks, malware, phishing emails, and more. Cybercrime, like other crimes, can bring harm to people's property, reputation, and many other areas.

On May 12, 2017, WannaCry (Figure 2.4) launched a major cyberattack, infecting over 230,000 computers in 150 countries and demanding ransom payments on each. The attack was described by Europol as an unprecedented event in terms of its scale [21]. During the initial spread, WannaCry encrypted various files on the computer and asked users to pay bitcoins to a specified address. WannaCry threatens to delete the files if users fail to pay bitcoin within the deadline.

Figure 6.4: WannaCry, A major ransomware cyber attack in 2017

The major players operating in the Central & Eastern European Cybersecurity Market are Cisco Systems Inc., IBM Corporation, Fortinet Inc., Palo Alto Networks Inc., McAfee Corp, Sophos Group Plc, CrowdStrike Holdings Inc., FireEye Inc., Juniper Networks Inc., Check Point Software Technologies Inc, CyberArk Software Ltd., Trend Micro Inc., Zscaler Inc., Sonicwall Inc., Barracuda Networks Inc [22]. Table 2.1 shows the countries where these cybersecurity companies. From the table above, we can see that US-based cybersecurity companies dominate the cybersecurity market share in Central and Eastern Europe. Of the many Central & Eastern European cybersecurity providers mentioned above, only one is in Europe, but only geographically and not politically. From Figure 2.5 we can see that digital sovereignty in Europe is heavily dependent on the outside world, especially in the United States. This undoubtedly brings greater risk to their cyber security.

Table 6.1: The major players operating in the Central & Eastern European Cybersecurity Market and which country they belong to

Company	Country
Cisco	USA
IBM	USA
Fortinet	USA
Palo Alto	USA
McAfee	USA
Sophos	UK
CrowdStrike Holdings	USA
FireEye	USA
Juniper Networks	USA
Check Point	Israel
CyberArk	USA
Trend Micro	Japan
Zscaler	USA
Sonicwall	USA
Barracuda	USA

Figure 6.5: Location of the world largest tech companies

6.3.3 5G Infrastructure

As shown in Figure 2.6, Huawei has a great influence on the worldwide telecommunication field. Huawei's European activities began in 2000 with the opening of an R&D center in Stockholm. Since Huawei established its operations in Europe it has been focusing on customer-centric innovation, strong partnerships, and building close cooperation with nearly all main carriers in Europe. In 2015, Huawei ranked 4th on the European Patent Office's ranking of companies by the number of applications. Today Huawei has 2 regional offices in 33 countries. More than 10,000 people are employed in Europe, 1,570 in R&D, across the 18 R&D centers [23].

Figure 6.6: Worldwide Telecom Equipment Revenue

In recent years, Huawei has been seen as a security threat and has been jointly suppressed by the United States and its allies. Although Huawei is a private company, it is widely believed that Huawei has ties to the Chinese government. But while foreign-policy professionals have warned about the threats Huawei represents to national security and economic integrity, the exact nature and scope of such threats remain largely speculative [24].

5G is a critical infrastructure and will penetrate European society and its economy to an unprecedented extent. Proponents of a ban argue that Huawei is closely allied with the

authoritarian Chinese party-state, which could utilize Huawei equipment for espionage and sabotage [25].

The author of this article believes that although Huawei's security threat theory is speculative, it does not mean that it does not pose any digital sovereignty threat. The same is true for any company, not just Huawei. Too much reliance on a single enterprise on 5G infrastructure will inevitably increase the potential risks of digital sovereignty security.

6.3.4 Facebook and the Digital Colonialism

Many US tech companies have made a series of investments in the Global South, including but not limited to expanding their global reach, the most notorious of which is Facebook's Free Basics initiative. Free Basics is both an application and website that provides free of data charges access to a variety of basic services like news, weather, health information, job ads, and of course, Facebook [26].

The free internet service provided by Facebook only allows users to access limited resources, and if users want to read more, they need to pay. "Facebook is not introducing people to an open internet where you can learn, create and build things," said Ellery Biddle, advocacy director of Global Voices. Also, Ellery mentioned "It's building this little web that turns the user into a mostly passive consumer of mostly western corporate content. That's digital colonialism [27]."

The author believes that the main reason why Free Basics is accused of digital colonialism is that it does not meet the needs of users to connect to the Internet. It's just a good excuse for exploiting local markets and plundering user data under the guise of a charity. In addition, the content provided by Free basics also lacks neutrality, that is, only provides the content that Facebook wants to provide to users.

6.4 The Path for Digital Sovereignty

6.4.1 National Cybersecurity Strategy (NCCS)

The European Union Agency for Cybersecurity (ENISA) contributes to EU cyber policy, enhances the credibility of ICT products, services, and processes through cyber security certification schemes, and works with member states and EU institutions to help prepare Europe for future cyber challenges.

In a constantly changing cyber threats environment, EU Member States need to have flexible and dynamic cybersecurity strategies to meet new, global threats. A national cybersecurity strategy (NCCS) is a plan of action designed to improve the security and resilience of national infrastructures and services. It is a high-level top-down approach to cybersecurity that establishes a range of national objectives and priorities that should be achieved in a specific timeframe. Currently, all countries in the European Union have a

National Cybersecurity Strategy as a key policy feature, helping them to tackle risks that have the potential to undermine the achievement of economic and social benefits from cyberspace [28].

ENISA's work in support of these strategies focuses on the analysis of existing NCSSs; on the development and implementation of national information security services; outlining and raising awareness of the good practice, and providing guidance and practical tools for the Member States to assess their NCSSs.

6.4.2 General Data Protection Regulation (GDPR)

As technology progressed and the Internet was invented, the EU recognized the need for modern protections. So in 1995, it passed the European Data Protection Directive, establishing minimum data privacy and security standards, upon which each member state based its implementing law [29].

The European Union's General Data Protection Regulation (GDPR) in 2018, which requires websites to get users' consent before placing long-lived cookies, went into effect. This law applies to any company that collects data from people who live in the European Union, regardless of where the company is based. (When you visit a website and see a pop-up announcing, "this site uses cookies," you're seeing the effects of GDPR [30].)

Amazon has been fined 746 million euros by EU privacy regulators for violating EU General Data Protection Regulation regulations, the largest ever fine for a data privacy breach in the EU.

6.4.3 Investment in Data Security and Privacy

Most organizations are seeing positive returns on private investments, and more than 40% are seeing benefits at least twice that of their privacy spend, according to the Cisco Data Privacy Benchmark Study [31].

According to the Cisco Consumer Privacy Survey, the average annual privacy spending in 2019 was \$1.2 million. The average privacy spends of small businesses (250-499 employees) was \$800,000. Among large enterprises (10,000 or more employees), the average annual privacy spend was \$1.9 million, and 2% of these enterprises spent more than \$5 million [31].

There are many reasons why companies are willing to invest in protecting privacy, including but not limited to meeting customer expectations, increasing the company's transparency and trust among customers, demonstrating the company's values to the public, and avoiding fines and lawsuits that come with laws and regulations.

6.4.4 Fighting against Cybercrime

43 countries, including the United States, signed the Council of Europe's Convention on Cybercrime in November 2001. The U.S. Senate ratified the Convention on August 3, 2006. The Convention seeks to better combat cybercrime by harmonizing national laws, improving investigative abilities, and boosting international cooperation [32].

To combat cybercrime, the EU has enacted and enacted many laws such as the Council of Europe Convention on Cybercrime, Interim Regulation on the processing of personal and other data to combat child sexual abuse, Directive on non-cash payment, Regulation and Directive facilitating cross-border access to electronic evidence for criminal investigations, Directive on attacks against information systems, Directive on combating the sexual exploitation of children online and child pornography [33] and so on.

6.4.5 IT-Security Law 2.0

On April 23, 2021, the German Bundestag passed the T Security Act 2.0 [34]. The act aims to bring protection mechanisms and defense strategies up to date. The act reinforces the BSI's position as Germany's central cybersecurity agency, guaranteeing cybersecurity in mobile networks by banning the use of key components, while also proposing measures to strengthen consumer protection and provide more security for companies.

6.4.6 5G Security

In 2019 Member States of the EU, with the support of the Commission and the European Agency for Cybersecurity published a report on the EU coordinated risk assessment on cybersecurity in Fifth Generation (5G) networks to ensure a high level of cybersecurity of 5G networks across the EU [35].

Based on risk assessment demerits of EU member states, the report identifies key security threats, and sensitive assets identify some key security challenges and strategic risks, and forecasts the impact of the rollout of 5G networks.

6.5 Summary and Conclusions

In this article, the concepts of digital sovereignty and digital colonization are discussed, including a comparison of the concepts of non-digital sovereignty and colonization with real-world examples. That is, digital sovereignty is as important as land, sea, and sky, and is the fourth space of a country or group. Through the case of Facebook, we introduce and analyze the concept of digital colonization, that is, Facebook's Free Basic in India does not solve users' online needs but plunders a large amount of user data and opens up its potential digital market.

In addition to the concepts of digital sovereignty and digital colonization, this paper also analyzes the current challenges facing digital sovereignty security, such as issues faced in personal data and privacy protection, cybersecurity and cybercrime, reliance on other countries' 5G infrastructure, etc. Based on these issues, this article introduces the current efforts of the international community to realize the security of digital sovereignty, such as legislative activities made by different countries or organizations, and investments in related fields.

The author of this article advocates that with the gradual popularization of Internet applications, cyberspace, as the fourth space other than land, sea, and air, deserves equal attention to its sovereignty and security. In the future, better legal norms, as well as more related investments and service providers, will solve these existing problems. In addition to technical or legal means, international social organizations and groups should also move towards transparency, build trust, remove barriers, and follow the trend of economic globalization.

Bibliography

- [1] J. Rozpedowski: Digital Sovereignty in an Era of Global Surveillance, Disinformation, and Info-demics; May 2021. <https://www.geopoliticalmonitor.com/digital-sovereignty-in-an-era-of-global-surveillancedisinformation-and-info-demics/>
- [2] Non-fungible token — Wikipedia; Wikipedia contributors, https://en.wikipedia.org/w/index.php?title=Non-fungible_token&oldid=1090891003
- [3] Friedrich Engels: Engels's letter to J. Bloch; September 1890.
- [4] Tomas Gabris, Ondrej Hamulak: Pandemics in Cyberspace - Empire in Search of a Sovereign?; July 2021.
- [5] Tomas Gabris, Ondrej Hamulak: 5G and Digital Sovereignty of the EU: The Slovak Way; November 2021.
- [6] P. Roguski: Layered Sovereignty: Adjusting Traditional Notions of Sovereignty to a Digital Environment; 2019 11th International Conference on Cyber Conflict (CyCon), July 2019.
- [7] Digital Sovereignty; <https://www.techopedia.com/definition/33887/digital-sovereignty>.
- [8] Oxford English Dictionary Second Edition on CD-ROM(v. 4.0); Oxford University Press, 2009.
- [9] Digital Colonialism: The Evolution of Empires; July 2021, <https://iyouport.substack.com/p/7a4?s=r>.
- [10] Sean Illing: A political scientist explains how big data is transforming politics; March 2017, <https://www.vox.com/conversations/2017/3/16/14935336/big-data-politics-donald-trump-2016-elections-polarization>.
- [11] Our Company; <https://www.huawei.com/en/corporate-information>.
- [12] Kevin Liu: How Huawei Can be a Committed Partner in a Digital Europe; <https://www.huawei.com/en/voice-of-huawei-europe/how-huawei-can-be-a-committed-partner-in-a-digital-europe>.
- [13] Ryan Browne: Germany's Merkel hits out at Twitter over 'problematic' Trump ban; January 2021, <https://www.cnbc.com/2021/01/11/germanys-merkel-hits-out-at-Twitter-over-problematic-trump-ban.html>.

- [14] What is digital finance?; European Commission, https://ec.europa.eu/info/business-economy-euro/banking-and-finance/digital-finance_en.
- [15] Over 80% of adults use electronic payment; People's Daily Overseas Edition, December 2020, http://www.xinhuanet.com/fortune/2020-10/17/c_1126621847.htm.
- [16] Data Security; imperva, <https://www.imperva.com/learn/data-security/data-security/>.
- [17] Digital sovereignty does not need EU champions; Financial Times, <https://www.ft.com/content/2762d7dc-0607-11ea-a984-fbbacad9e7dd>.
- [18] Privacy; Wikipedia, <https://en.wikipedia.org/wiki/Privacy>.
- [19] Three challenges in data protection and privacy; January 2022, <https://weblog.iom.int/three-challenges-data-protection-and-privacy>.
- [20] Vernon Andrews: Analyzing Awareness on Data Privacy; ACM SE '19: Proceedings of the 2019 ACM Southeast Conference, April 2019, <https://dl.acm.org/doi/10.1145/3299815.3314458>.
- [21] Ransomware cyber-attack threat escalating - Europol; BBC News, May 2017, <https://www.bbc.com/news/technology-39913630>.
- [22] Central & Eastern Europe Cybersecurity Market Report: Historical Data 2016-2019 and Forecast Period 2022-2026; December 2021, <https://www.globenewswire.com/news-release/2021/12/13/2350480/28124/en/Central-Eastern-Europe-Cybersecurity-Market-Report-Historical-Data-2016-2019-an.html>.
- [23] Huawei in Europe; <https://www.huawei.com/ch-en/corporate-information/huawei-europe>.
- [24] Doowan Lee: Huawei Is Bad for Business; April 2021, <https://foreignpolicy.com/2021/04/30/huawei-china-business-risk/>.
- [25] Tim Ruhlig & Maja Bjork: What to Make of the Huawei Debate? 5G Network Security and Technology Dependency in Europe; January 2020, <https://www.ui.se/globalassets/ui.se-eng/publications/ui-publications/2020/ui-paper-no.-1-2020.pdf>.
- [26] Toussaint Nothias: The Rise and Fall... and Rise Again of Facebook's Free Basics: Civil Society and the Challenge of Resistance to Corporate Connectivity Projects; April 2020, <https://globalmedia.mit.edu/2020/04/21/the-rise-and-fall-and-rise-again-of-facebooks-free-basics-civil-and-the-challen>
- [27] 'It's digital colonialism': how Facebook's free internet service has failed its users; The guardian, <https://www.theguardian.com/technology/2017/jul/27/facebook-free-basics-developing-markets>.
- [28] ENISA <https://www.enisa.europa.eu/topics/national-cyber-security-strategies>.

- [29] What is GDPR, the EU's new data protection law?; GDPR.EU, <https://gdpr.eu/what-is-gdpr/>.
- [30] Data Privacy: 4 Common Issues and How to Solve Them; neeva, May 2021, <https://neeva.com/learn/data-privacy-4-common-issues-and-how-to-solve-them>.
- [31] Why are companies investing in privacy and GDPR compliance?; Data Privacy Manager, <https://dataprivacymanager.net/why-do-companies-invest-in-gdpr-compliance-what-are-benefits-of-gdpr-compliance/>.
- [32] Cybercrime: The Council of Europe Convention; September 2006, <https://www.everycrsreport.com/reports/RS21208.html>.
- [33] Cybercrime; European Commission, https://ec.europa.eu/home-affairs/cybercrime_en.
- [34] Bundesrat billigt IT-Sicherheitsgesetz 2.0; Federal Ministry of the Interior and Community, May 2021, <https://www.bmi.bund.de/SharedDocs/pressemitteilungen/DE/2021/05/it-sicherheitsgesetz.html>.
- [35] Member States publish a report on EU coordinated risk assessment of 5G networks security; European Commission, https://ec.europa.eu/commission/presscorner/detail/en/IP_19_6049.

Chapter 7

An Overview of Mitigation Techniques for Spoofing Attacks in IoT

Geyu Meng, Zi Ye

The Internet of Things is rapidly growing and IoT enabled devices are starting to appear in all aspects of our lives, since the enterprise IoT market grew 22.4% to \$157.9 billion in 2021, according to the March 2022 update of IoT Analytics' Global IoT Enterprise Spending Dashboard [1]. For something that has such wide use, IoT collects huge amounts of information about our daily life, including sensitive data, such as bio-data, home address, personal messages. While IoT devices still have security vulnerabilities and all attacks cannot be prevented, we cannot neglect the danger of sensitive data leaking or malicious data hijacking, which could have a devastating impact on society. As a result, the need for awareness of the security risks is increasing and it is essential for us to find mitigation techniques to protect private information from spoofing attacks. In this paper, we first introduce the architecture of IoT and the concept of spoofing techniques, before analyzing the security threats that can be encountered. Then we summarized some mitigation mechanisms and countermeasures for short-range communication protocols that can utilize these security techniques. Finally, we discuss their advantages and disadvantages within the context of IoT security.

Contents

7.1	Introduction	149
7.1.1	Internet of Things	149
7.1.2	Spoofing Attack	150
7.2	Security challenges of IoT	151
7.2.1	IoT Security Architecture	151
7.2.2	Security Requirements in IoT	153
7.2.3	Examples of Security Threats in IoT	154
7.3	Mitigation Technologies for spoofing attacks in IoT	154
7.3.1	Cryptography	155
7.3.2	Access Control	156
7.4	IoT Communication Protocols and their Security Measures	158
7.4.1	ZigBee	158
7.4.2	Bluetooth	165
7.5	Conclusion	169

7.1 Introduction

7.1.1 Internet of Things

7.1.1.1 What is IoT

The Internet of Things (IoT) refers to the combination of various information sensing devices, such as radio frequency identification (RFID) devices, infrared sensors, global positioning systems, laser scanners and other devices and the Internet to form a huge network [2].

The basic idea of IoT is to allow autonomous exchange of useful information between invisibly embedded different uniquely identifiable real world devices around us [3]. For example, it is capable of monitoring, sensing and collecting information of various environments or monitoring objects in real time through the collaboration of various integrated micro-sensors, and transmits the sensed information to the user terminal in a multi-hop relay mode through a random self-organized wireless communication network. The information is processed by the embedded system and transmitted to the user terminal in a multi-hop relay via a random self-organizing wireless communication network.

Overall, the Internet of Things is a network that extends and expands on the basis of the Internet, it is characterized by ubiquitous data sensing, wireless-based information transmission, intelligent information processing, and the user end can be extended and expanded to any object and object to exchange and communicate information.

Also, from the paper "Study on Security Problems and Key Technologies of The Internet of Things" by Xiaohui Xu [4], we could summarize the three characteristics of the IoT:

1. Comprehensive sensing: the use of RFID, sensors, or two-dimensional code to obtain information about objects at any time and anywhere.
2. Reliable delivery: the real-time information would be accurately transferred through the convergence of a variety of telecommunication networks and the Internet.
3. Intelligent processing: it makes use of cloud computing, fuzzy recognition and other intelligent computing technologies to analysis and process the massive data and information in order to implement intelligent control.

So how does IoT system work? Before revealing the secrets and providing a clear structure of this initiative, it's important to understand what the IoT architecture actually means. In essence, IoT architecture is the system of numerous elements: sensors, protocols, actuators, cloud services, and layers [6]. Given its complexity, there exist 4 stages of IoT architecture. The figure 7.1 below shows the detailed steps of collecting, processing and storing data in IoT.

Stage 1 consists mainly of networked things, such as wireless sensors and actuators. It is the basis for the development and application of IoT. Various sensing nodes sense the

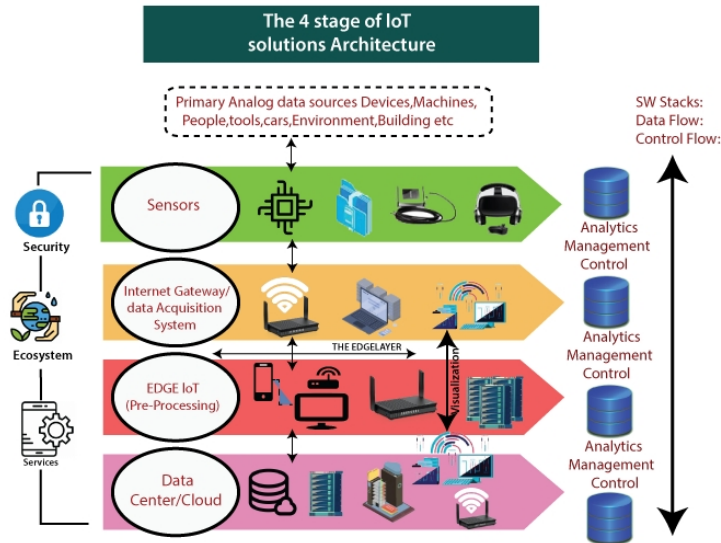


Figure 7.1: The 4 Stages of IoT Architecture

relevant information of the target environment and network themselves to the gateway access point, which submits the collected data to the back-end processing through the interconnection network.

Stage 2 covers sensor data aggregation systems and analog-to-digital data conversion, also known as the internet gateway/data acquisition system. It has a bonding effect. Data is transmitted using the Internet, wireless broadband networks. The importance of this phase is to process the large amount of information collected in the previous phase and to reduce it to the optimum size for further analysis. In addition, the necessary transformations in terms of time and structure also take place here.

Stage 3 deals with the appearance of edge IoT systems, also called the EDGE IoT. In this phase, between the various stages of the IoT architecture, the prepared data is transferred to the IT world. In particular, it is here that edge IT systems perform improved analysis and pre-processing. At the same time, other processing can be carried out here before the data center enters the phase.

Stage 4 controls the analysis, management, and storage of data (Data Center/Cloud). The main process of the final phase of the IoT architecture takes place in the data center or in the cloud. Precisely, it allows for in-depth processing, as well as a subsequent review of the comments.

7.1.2 Spoofing Attack

As defined by NIST [7], spoofing is the faking of the sending address of a transmission to gain illegal entry into a secure system. And the deliberate inducement of a user or resource to take incorrect action.

7.1.2.1 Types of spoofing attacks

What can hackers try to forge to make their spoofing attacks? Plenty things: an IP address, a phone number, a web page, GPS location... Some of these attacks take advantage of human gullibility, while others take advantage of hardware or software flaws [8]. Of all the nefarious scenarios that fit the pattern of spoofing attacks, we selected a few to explain in detail.

MAC spoofing In theory, every network adapter built into a connected device should have a unique Media Access Control (MAC) address that will not be encountered elsewhere. However, an attacker can exploit imperfections in some hardware drivers to modify or spoof MAC addresses. In this way, the criminal disguises his device as a device registered in the target network to bypass traditional access restriction mechanisms.

IP spoofing When packets are passed through the Hub to other network segments, the Hub simply copies the packets to other ports. Therefore, for networks using Hubs, there is no security and packets can easily be intercepted by users intercepted and analyzed by users and IP address spoofing.

DNS spoofing This is a type of spoofing in which an attacker impersonates a domain name server. the principle of DNS spoofing is that if you can impersonate a domain name server and set the IP address of the query to the IP address of the attacker, then the user will only see the attacker's homepage instead of the homepage of the website that the user wants to obtain.

ARP spoofing It is an attack technique that targets the Ethernet Address Resolution Protocol (ARP) by spoofing the gateway MAC address of the visitor's PC on the LAN, causing the visitor's PC to believe that the attacker's changed MAC address is the gateway's MAC, resulting in a network failure. This attack allows the attacker to access packets on the LAN and even tamper with them and can prevent specific computers or all computers on the network from connecting properly.

7.2 Security challenges of IoT

After having a basic concept of the IoT and spoofing attacks, we would focus on security risks in IoT in this part.

7.2.1 IoT Security Architecture

From the perspective of information and network security, the IoT is a heterogeneous converged network with multiple networks co-existing, which not only has the same security issues as sensor networks, mobile communication networks and the Internet, but also its

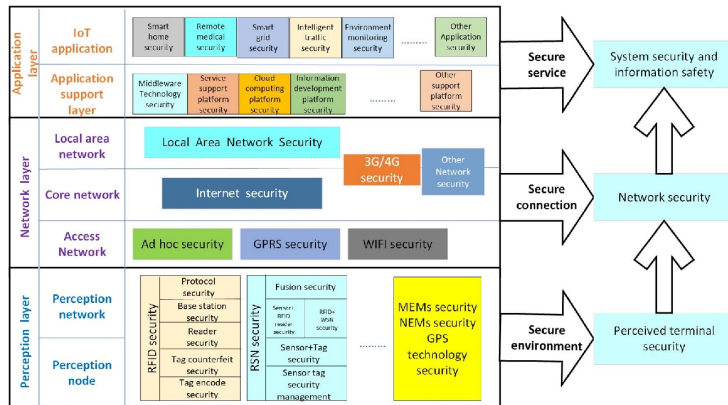


Figure 7.2: IoT Security Architecture

special security issues, such as privacy protection, heterogeneous network authentication and access control, information security storage and management [9].

In terms of the information processing process of the IoT, the whole process of information processing reflects the features and requirements of IoT security and reveals the security issues faced through the process of collection, aggregation, fusion, transmission, decision making and control.

As Figure 7.2 shows, from bottom to top, the architecture of the IoT is recognized as having three layers: the perception layer, the network layer and the application layer[10]. The IoT architecture layers are distinguished in order to track the consistency of the system, which should be taken into consideration before the IoT architecture process start. In addition, the fundamental features of sustainable IoT architecture include functionality, scalability, availability, and maintainability [5], without which the result of IoT architecture would be a failure. These requirements are addressed in the 4 stages of IoT architecture described above.

7.2.1.1 Security Threats in Different Layers of IoT Architecture

The following information comes from the paper "Overview of IoT Security Architecture" by J. Zhang, et al [10]., in which they introduced the security issues at the three different levels of IoT security architecture.

Sensory networks have information collection, transmission, and security issues. These include protocol security, base station security, tag forgery, and tag encryption security [37]. Issues related to wireless sensors include routing protocol security, cryptographic algorithms, and trusted management of nodes .

The core network has relatively complete security protection, but the large number of nodes in the IoT and their presence in clusters can lead to denial-of-service attacks when data is disseminated, as a large number of machines send data to congest the network.

The use of emerging technologies such as cloud computing will give attackers the opportunity to intercept and tamper with data and will also exploit vulnerabilities and flaws

in software systems and crack keys to achieve illegal access to database systems, causing significant damage.

7.2.2 Security Requirements in IoT

Once the analysis of the security threats at different layers of IoT is complete, we can examine the main challenges of designing and deploying security mechanisms at each layer.

7.2.2.1 Perception Layer Security Requirements

In the perception layer, it requires to use as little energy and bandwidth resources as possible to design streamlined and secure algorithms, key systems, and security protocols to solve the corresponding security problems. So the following aspects need to be considered:

To provide a secure peer-to-peer communication service, a corresponding key management scheme is required. It demands confidentiality and integrity, i.e., information is not tampered with. For wireless sensor networks, group communication is needed where source-side authentication is a necessity. Availability is also a fundamental requirement and goal of wireless sensor network security, meaning that security protocols are efficient and reliable and do not overload the nodes resulting in excessive power consumption. Security mechanisms must also provide the security mechanisms and algorithms to support this level of scalability to ensure that the sensor network remains up and running.

For the diversity of sensing networks, secure routing and connectivity, confidentiality and authentication are necessary. The symmetric cryptography scheme is efficient and computationally small, while the asymmetric cryptography provides higher security by using a pair of private key and public key, but at a higher cost. The security requirements of a sensing network should be based on the characteristics of the sensing network itself, the node characteristics of the service and the requirements of the user. In general, sensory networks are characterized by low power consumption, loose distribution, concise signaling, simple protocols, broadcast characteristics and little or no interaction, so security should be based on using as little energy and bandwidth resources as possible, designing algorithms, key systems and security protocols that are both concise and secure, and solving the corresponding security problems.

7.2.2.2 Network Layer Security Requirements

The network layer needs to establish corresponding authentication mechanisms, key negotiation mechanisms, key management mechanisms and algorithm selection mechanisms.

An encryption mechanism is required so that it can be used in conjunction with other security mechanisms; Additionally, a digital signature mechanism is also used to guarantee the non-repudiation of operations during communication and the recipient can verify the message. Access control mechanisms are also required to determine access rights to system resources based on the identity of the entity and information about its attributes.

7.2.2.3 Application Layer Security Requirements

The application layer of IoT is the core value of IoT. There are many typical applications in the application layer of IoT, such as: smart transportation, mobile phone payment, smart home, smart grid, smart city, smart water, food traceability and smart medical. These applications will generate a huge amount of data. Due to the large amount of data, cloud computing and cloud storage are needed to support the application layer. Diverse IoT applications face a variety of security issues, so the application layer needs a strong and unified security management platform, otherwise different applications require different security platforms, and these security requirements are not the same and can fragment the security trust relationship between platforms [11].

Therefore, the application layer requires authentication and access control for operation users, source encryption and integrity protection for industry sensitive information, certificate and PKI application to achieve identity identification, digital signature, anti-repudiation and security audit.

7.2.3 Examples of Security Threats in IoT

From this article written by Rudra Srinivas [12], there are several security incidents that make users feel less secure when using IoT devices.

For example, privacy Leaks. Hackers can cause considerable damage simply by identifying insecure IoT devices that leak (IP) addresses, which in turn can be used to pinpoint residential locations. What's more, home invasion is also on the list of the security threats in IoT. Insecure devices can spread IP addresses, and hackers can use these vulnerabilities to find users' residential addresses and sell this information to the wrong people. Further, hackers could also exploit the vulnerabilities of smart TV. There are several overlooked and neglected security issues with smart TVs. It says that security is an afterthought for several smart TV manufacturers, which makes them vulnerable to different types of threats. Not only can hackers take control of an insecure TV to change channels or volume controls, but they can also use the integrated camera and microphone to track your daily movements and conversations.

7.3 Mitigation Technologies for spoofing attacks in IoT

Since much of the information contained in IoT environments is potentially personal or sensitive data, there is a strong need to support anonymity and restrict access to the information. This puts security and privacy concerns at the forefront: the ability to manage the digital identities of millions of people and billions of devices is fundamental to success. In the following part we would put our focus on cryptography algorithms and access control as well as supporting the authentication mechanisms in constrained devices.

7.3.1 Cryptography

IoT devices and other interconnected designs face an increasing number of threats and therefore require more robust security methods based on multiple encryption algorithms.

7.3.1.1 Encrypted communication protocols

The biggest area of application for cryptography in the IoT is in securing communication channels. IoT-centric communication protocols (such as MQTT and AMQP) [13] allow developers to use the Secure Transport Layer Protocol (TLS) to ensure that all data sent over the network is unreadable by external parties. TLS ensures that data between two entities is not readable and not easily manipulated by third parties.

In addition to encrypting the primary data connection, it is also important to encrypt any secondary communication channels that are available [14]. For example, if an IoT device comes with a portal interface for consumer use, this should also be encrypted by default. For the same reason, insecure maintenance interfaces like remote login should be turned off to support secure methods like Secure Shell Protocol (SSH).

7.3.1.2 Asymmetric encryption algorithms

Asymmetric encryption algorithms [15] are provided with two keys, a public key and a private key. If data is encrypted with the private key, it can only be decrypted with the public key, and vice versa. This encryption is particularly suited to several aspects of the IoT infrastructure.

The first is the authentication of individual machines that join the IoT network. For example, an endpoint may need to connect to a central MQTT agent in order to publish data upstream. When joining the network, the use of a private key provides a secret and unique identifier for each machine and is virtually impossible to brute-force crack due to its length.

A second area where private key authentication can help in the IoT is message authentication between devices. A hash or other integrity checking algorithm will be calculated against the message and then encrypted and appended to the message using the private key. This check is then decrypted by the recipient of the message using the public key, proving that it can only be generated by the private key holder.

Finally, the result of the integrity check is verified to ensure that the message is not corrupted or altered during transmission. This method of electronic signature may be useful in situations where secure communication channels are not available.

7.3.1.3 Signed Firmware and Secure Boot

Electronic signature methods can be used to sign secure boot and firmware images [16]. Signing ensures that an authorized user or machine has placed an approval mark on the firmware before execution, which makes it difficult for a malicious individual to create malicious firmware and hijack a machine.

Secure boot takes advantage of this feature to ensure that any code running on the device is appropriate. The first bit of code that runs after the device boots up includes the ability to calculate and verify electronic signatures. In addition, the use of a private key infrastructure with secure boot gives maintainers a means of remediation in the event that the key used to sign the code is compromised.

7.3.2 Access Control

Access Control is the authentication and control of users' legitimate use of resources, controlling access to specific resources and thus preventing illegal access by some illegal users or improper use by legitimate users to ensure that the overall system resources can be used appropriately.

As IoT applications are multi-user, multi-task working environments, this opens the door to illegal use of system resources and therefore urgently requires IoT manufacturers to take effective security precautions for computers and their network systems to prevent unauthorized users from accessing the system and illegal use of system resources by legitimate users. This requires the use of an access control system.

Access control has three aspects [19].

1. Legitimacy: to prevent illegal access by unauthorized users and illegal access by illegal users.
2. Integrity: in a series of steps including data collection, information transmission and information storage, ensure that the data and information are intact and cannot be added, deleted, or altered at will.
3. Timeliness: within a certain time limit, ensure that the system resources cannot be tampered with by illegal users and guarantee the integrity of the system within the time limit.

7.3.2.1 Authentication

Authentication is the confirmation of the user's identity. Authentication must work in conjunction with an identifier. The authentication process first requires the user to enter an account name, a user logo or a registration mark to identify themselves. The account name should be kept secret and no other user should be allowed to have it. However,

in order to prevent the account name or user logo from being compromised and illegal user access, further authentication techniques are needed to confirm the user's legitimate identity. Passwords are a simple and easy means of authentication but are vulnerable to exploitation by attackers because they can be easily guessed and are weak. Biotechnology is a strict and promising authentication method, such as fingerprint recognition, retina recognition and iris recognition, but is not yet widely adopted due to its technical complexity.

7.3.2.2 Authorization

While authentication is the security practice of confirming that someone is who they claim to be, authorization is the process of determining which level of access each user is granted. In other words, authorization determines what a user is and is not permitted to do.

In practice, it is usually necessary to specify the user's access rights from 3 aspects: user type, application resources and access rules.

User type. For a user who has been identified and authenticated by the system, the system also imposes certain restrictions on his access operations. For a general-purpose computer system, there is a wide range of users and different levels and permissions. The types of users are generally system administrators, general users, audit users and illegal users. The system administrator has the highest authority and can access any resource in the system and has the right to all types of access operations. General users are subject to certain restrictions on access operations, and the system administrator assigns different access operation rights to such users as required. Audit users are responsible for auditing the security controls and resource usage of the entire system. Illegal users are those whose access rights have been removed or who have been denied access to the system.

Application resources. Application resources are the system resources that can be shared by every user in the system. It is the system resources that need to be protected within the system, so an Access Control Packet (ACP) needs to be defined for the protected resources. The Access Control Packet outlines an Access Control List (ACL) [20] for each resource or group of resources, which describes which user can use which resource and The ACL describes which user can use which resource and how.

Access rules. An access rule defines several conditions under which access to a resource can be granted. In general, rules allow a user to be paired with a resource and then specify which actions that user can perform on that resource, such as read-only, not allowed to execute, or not allowed to access. These rules are determined by the system administrator responsible for enforcing the security policy based on the principle of least privilege, i.e. when granting a user access to a resource, only the least privilege is granted to that resource.

7.3.2.3 File protection

File protection is the additional protection provided to a file so that it cannot be read by unauthorized users.

7.3.2.4 Auditing

Auditing is the process of recording all activities carried out on a user's system, i.e., recording the time and date when a user uses the system in breach of security regulations and the user's activities.

7.4 IoT Communication Protocols and their Security Measures

The smart IoT connected devices are vulnerable to threats and to minimize these security loopholes, the right protocols are required. The IoT communication protocols are modes of communication that ensure optimum security of the data being exchanged between IoT connected devices. Therefore, it is important to understand the IoT communication protocols and standards, which might reduce the security breaches at bay.

In this paper, we put our focus on two short-range communication protocols, specifically ZigBee and Bluetooth.

7.4.1 ZigBee

ZigBee is a short-range, low-power, low-speed wireless communication technology.

ZigBee is primarily a low-speed transmission network created for automated control data transmission and is low cost. Summarized in the paper "Study on ZigBee technology" by C. Muthu Ramya, et al [22]., ZigBee network has the following features:

Low power consumption In low-power standby mode, two No. 5 dry batteries can last from 6 to 24 months, thus eliminating the need for recharging or frequent battery replacement.

Low rate ZigBee operates at a lower rate of 20 to 250kbit/s, meeting the requirements of low rate data transmission.

Short latency ZigBee is highly responsive, typically taking only 15ms to switch from sleep to working and 30ms for nodes to access the network, resulting in further energy savings.

Proximity With an effective coverage range of 10 to 100m, it can basically cover an average home or office environment.

Large capacity ZigBee can be used in star, slice and mesh network structures to form large networks of up to 65,000 nodes.

Low cost ZigBee's simple and compact protocol greatly reduces its communication control requirements and ZigBee is royalty-free for the protocol.

High security ZigBee uses AES-128 encryption algorithms to provide data integrity checks and authentication capabilities.

License-free frequency bands Direct Sequence Spread Spectrum is used in the ISM(Industrial, Scientific and Medical) bands [26].

7.4.1.1 Architecture of ZigBee Protocol Stack

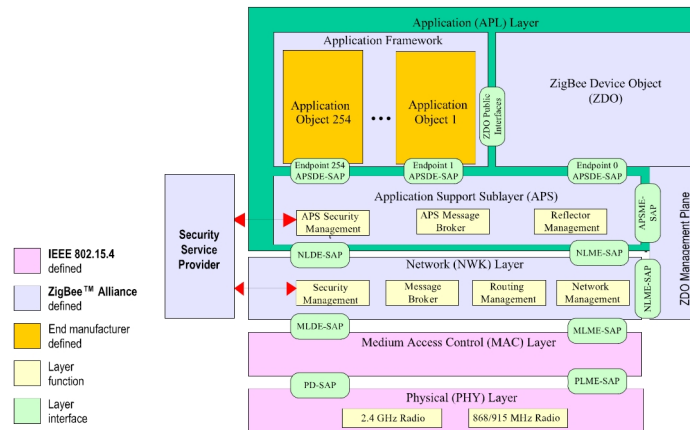


Figure 7.3: ZigBee Architecture

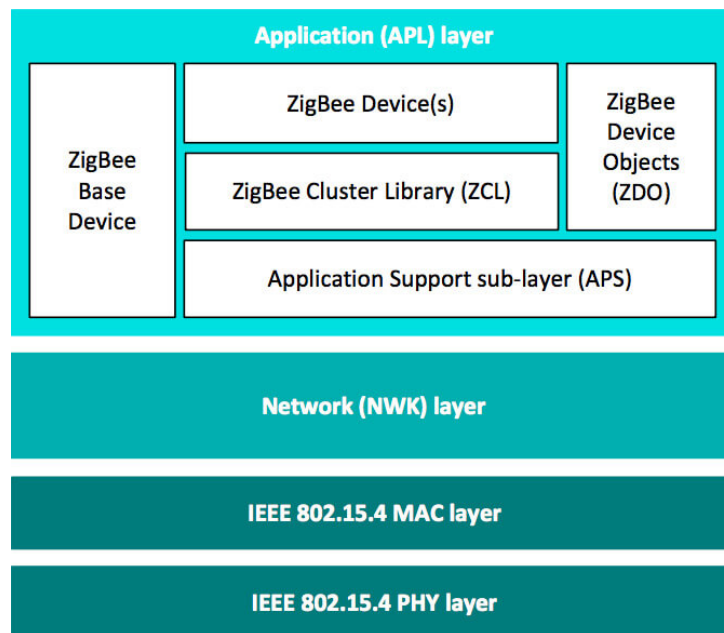


Figure 7.4: ZigBee Stacks

As shown in Figure 7.3 and Figure 7.4, ZigBee’s protocol stack consists of two parts, IEEE 802.15.4 [21] defines PHY (physical layer) and MAC (medium access layer) specifications; ZigBee Alliance [22] defines NWK (network layer), APS (application support sublayer), APL (application layer) specifications.

Physical layer The ZigBee physical layer provides an interface from the MAC layer to the physical layer radio channels through the RF firmware and RF hardware. The physical layer contains a Physical Layer Management Entity (PLME) that provides its interface to the physical layer management services by calling the physical layer management functions and is also responsible for maintaining a target database managed by the physical layer that contains basic information about the physical layer's individual domain networks. Physical layer functions: ZigBee activation; energy detection of the current channel; reception of link quality of service information; ZigBee channel access method; channel frequency selection and data transmission and reception.

Link layer Main functions of link layer: the coordinator generates and sends beacon frames, according to which ordinary devices synchronize with the coordinator; support for PAN networks; support for wireless channel communication security; support for time slot guarantee mechanisms and reliable transmission between the MAC layers of different devices.

MAC layer The MAC layer generates a network beacon for the coordinator; synchronizes with this beacon; supports device security; handles and maintains the GTS mechanism and provides a reliable link between two peer MAC entities.

Network layer Network layer functions: Configure a new device. For example, configure a new device as a ZigBee coordinator or try to join an already existing network, start a new network, join or leave a network; Performing network security, routing information frames to their destinations; discovering and maintaining routes between devices; discovering single-hop neighbors and storing their information; (single-hop devices do not require relay services); assigning addresses to devices that are part of the resulting network (coordinator or router).

Application layer ZigBee Application Layer consists of the Application Sublayer (APS), ZigBee Device Object, Framework for ZigBee (AF), ZigBee Device Template and Manufacturer Defined Hungry Application Object. It is for maintaining binding tables, passing messages between bound devices.

IEEE 802.15.4 supports three types of topologies [24]: Star, Mesh and Tree, as shown in figure 7.5, each consists of one coordinator, several routers and end-devices [25].

ZigBee Coordinator The ZigBee Coordinator is the device responsible for setting up, implementing, and managing the entire ZigBee network. It is responsible for configuring the security level of the network and configuring the address of the trust center (the default value for this address is the ZigBee coordinator's own address, otherwise the ZigBee coordinator can specify an alternate trust center). The ZigBee coordinator also maintains a list of currently associated devices and facilitates support for isolated scanning and rejoining processing to enable previously associated devices to rejoin the network. There is only one coordinator per network, so it can never go to sleep (there may be no coordinators in the network). The coordinator can also double as a router as required.

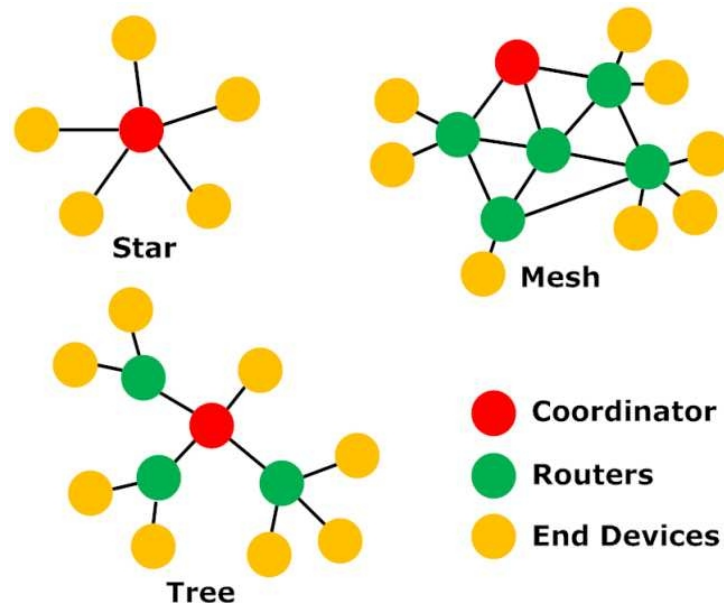


Figure 7.5: Topology of ZigBee

ZigBee Router A ZigBee router is an intermediate node device responsible for routing packets between end devices or between an end device and a coordinator. If security is enabled on the network, the router requires permission from the Trust Centre to join the network and can also double as an end device. In some cases, routers can allow other routers and end devices to join the network and will maintain a list of currently associated devices and facilitate support for isolated scanning and rejoin processing to allow previously associated devices to rejoin the network. As routers link multiple parts of the network, they cannot go to sleep.

ZigBee End-Device The terminal has the simplest function in a ZigBee network and can only join the network as the endmost child node device. It can only communicate with its parent node, and if two terminals need to communicate with each other, they must go through the parent node for multi-hop or single-hop communication. It is the largest number of nodes allowed to exist in a ZigBee network and is the only network device that allows low power consumption.

7.4.1.2 ZigBee in IoT

The conditions for the application of ZigBee technology are very low cost of node devices, small size of node devices, or having a large number of node devices in the network, but only for monitoring and control.

Therefore, the main areas of application for ZigBee technology include digital homes, industry, agriculture, medicine, home and building automation, consumer, and home automation markets and so on.

7.4.1.3 ZigBee Security Models

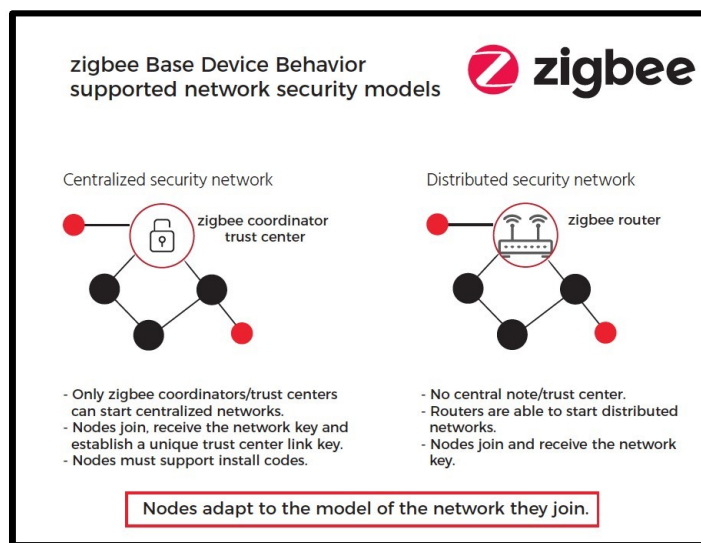


Figure 7.6: Security Models of ZigBee

ZigBee supports two types of security models, as shown in the above figure 7.6.

Centralized security model Complex but most secure model involving a third logical device: the trust center (network orchestrator, e.g. smart home gateway). The trust center is responsible for: configuring and authenticating the routers and end devices joining the network; generating the network keys used for encrypted communication over the network; switching to a new network key periodically or as required. Thus, if an attacker obtains a network key, it will have a limited lifetime before it expires; establishing a unique Trust Centre link key for each device as it joins the network to communicate securely with the Trust Centre; and maintaining the overall security of the network.

Distributed security model Simple, but less secure. This model only supports routers and end devices. Routers form a distributed network and are responsible for registering other routers and end devices. The router issues a network key (for encrypting messages) to newly joined routers and end devices. All nodes in the network use the same network key to encrypt messages. Similarly, all nodes are pre-configured with a link key (used to encrypt the network key) before they are registered into the network.

The ZigBee standard supports the following optional security services [29]:

Encryption/Decryption ZigBee frames can optionally be protected with the security suite AES-CCM to provide data confidentiality, data authentication and data integrity. AES-CCM is a smaller variant of AES (Advanced Encryption Standard) in modified CCM mode (counter with CBC-MAC).

Replay protection Each node in a ZigBee network contains a 32-bit frame counter that is incremented each time a packet is transmitted. Each node also keeps track of the previous 32-bit frame counter for each device (node) it is connected to. If a node receives a packet from a neighboring node with a frame counter value that is the same as or less than the previously received frame counter value, the packet is discarded. This mechanism enables replay protection by tracking packets and discarding them (if they have been received by the node). The frame counter can have a maximum value of 0xFFFFFFFF, but if it reaches the maximum value, no transmission can take place. The only time the frame counter is reset to 0 is when the network key is updated.

Device authentication The ZigBee standard supports both device authentication and data authentication. Device authentication is the act of confirming that a new device joining the network is a real device. A new device must be able to receive the network key and set the appropriate attributes within a given time period to be considered authenticated. Device authentication is performed by the Trust Centre. The authentication process differs in residential and commercial models.

Secure over-the-air (OTA) firmware updates OTA updates allow manufacturers to add new features, fix flaws in their products and apply security patches when new threats are identified. However, OTA updates can also pose a potential security vulnerability if the protocol does not provide sufficient protection or if the device manufacturer does not use all available safeguards. ZigBee devices and associated silicon platforms provide multiple layers of security to update the device in the field and ensure that the updated code image has not been maliciously modified.

Logical link-based encryption Another key security tool is the ability to create application-level secure links between pairs of devices in a network. This is managed by establishing a unique set of AES-128 encryption keys between a pair of devices. This allows a logical, secure link to be established between any two devices in the network, thus supporting a "virtual private link" between a pair of devices in the network and many other devices. This measure limits the ability of an attacker to obtain network keys by intercepting or injecting messages to be executed by other devices.

Runtime key updates Periodically, or when required, the Trust Centre will actively change the network key. The Trust Centre generates a new network key and encrypts it using the old network key, thus distributing it throughout the network. After the update, all devices will continue to retain the old network key for a short period of time until every device on the network has switched to the new network key. In addition, the device will initialize its frame counter to zero upon receipt of the new network key.

Network interference protection In low-cost ZigBee nodes, the use of band-selective filters to protect the network from interference may not be an option due to cost or node size constraints. However, IEEE 802.15.4 and the fundamental properties of ZigBee networks (such as low RF transmission power, low duty cycle and CSMA/CA channel access mechanisms) help to reduce the impact of the presence of a ZigBee wireless network on other nearby systems and vice versa.

7.4.1.4 ZigBee Security Vulnerabilities and Risk Mitigation

Vulnerabilities in a ZigBee network can be attributed to protocol issues or poor implementation of the protocol by the developers [28]. Below are some vulnerabilities examples summarized from the paper "Security Analysis of Zigbee" by Xueqi Fan, et al [30].

Passive key sniffing is the process of transmitting a key through a ZigBee device joining a secure network, so an attacker can listen to the traffic on the ZigBee network and wait for a new device to join the network, thus sniffing the packet that transmits the key.

Active key sniffing is used to obtain keys in response to the security weakness of trust center rejoin, which is when a device may not have a network key currently in use, thus requiring the rejoin command to be sent without network layer security and allowing the device to be able to obtain the current network key. Based on the above process, an attacker could easily forge this process to actively obtain a transmission key packet by exploiting the security weakness of the trust center rejoin.

A tampering attack modifies the frame counter to be greater than the current frame counter of the latest packet, tampers with the data, re-encrypts it using the link key and sends it out.

A number of mitigation mechanisms can be implemented to address the security vulnerabilities of ZigBee.

In ZigBee technology, asymmetric key security mechanism is used, where keys are generated by the network and application layers according to the actual application needs and are managed, stored, transmitted, and updated. Security mechanisms are provided by the security service delivery layer. However, the overall security of the system is defined at the template level, which means that the template should define what type of security should be implemented in a particular network.

The ZigBee protocol stack class defines security for the MAC, network and application layers. Its security services include methods for critical process creation and transport, device management and framework protection. Security at the MAC and network layers essentially serve the same purpose: securing single-hop transmissions. the MAC layer arbitrates access to shared media and controls single-hop transmissions between neighboring devices. The ZigBee Alliance has added a network layer security option to include features that are not possible at the MAC layer, including the ability to reject data frames that cannot be authenticated. Both security layers use a global key that is shared by all ZigBee devices on the network, and the MAC and network layer security is suitable for applications that need to prevent attacks on a specific infrastructure, such as preventing an illegal device from maliciously penetrating the network. If a developer needs to establish a route between two devices and the framework of that network layer is insecure, then an illegal device may intercept packets.

A security scheme may be used when a device is operating in ZigBee secure mode. A security scheme consists of a set of operations performed on frames at the MAC layer to provide security services. The name of the security scheme indicates the length of the symmetric encryption algorithm, pattern and integrity code. For all security schemes in

the ZigBee technology standard, the Advanced Encryption Standard(AES) [27] algorithm will be used.

7.4.2 Bluetooth

Bluetooth is a standard developed for short-range radio frequency communication and is already commonly used in various scenarios such as personal electronics, automotive electronics, medical devices. Despite its common uses, Bluetooth has become a household name in the Internet of Things community. It is a serious technology used for IoT applications.

Since its emergence, Bluetooth has gone through several versions from 1.0-5.0 and then Bluetooth mesh, each with different features. Generally, Bluetooth 4.0 and above is referred to as Low Power Bluetooth (LE) and those prior to Bluetooth 4.0 as Bluetooth BR/EDR/HS (Bluetooth 1.0, 2.0, 3.0).

One thing worth discussing is Bluetooth version 4.0, also known as Bluetooth Low Energy or BLE, which is ideal for IoT applications. Here we will compare the classic Bluetooth technology with Bluetooth low energy technology and their main application areas.

	Classic Bluetooth technology	Bluetooth low energy technology
Data payload throughput (net)	2 Mbps	~100 kbps
Robustness	Strong	Strong
Range	Up to 1000m	Up to 250m
Local system density	Strong	Strong
Large scale network	Weak	Good
Low latency	Strong	Strong
Connection set-up speed	Weak	Strong
Power consumption	Good	Very strong
Cost	Good	Strong

Figure 7.7: Comparison Between Classic Bluetooth and BLE

As Figure 7.7 shows, compared to traditional Bluetooth, BLE has the advantage of fast search, fast connection, ultra-low power consumption to maintain connection and transmit data, but the weakness is the low data transmission rate, the physical bandwidth is only 1M, the actual transmission speed is between 1-6KB.

7.4.2.1 Bluetooth in IoT

Bluetooth is a standard developed for short-range radio frequency (RF) communications and is primarily used to establish wireless personal area networks. Bluetooth has been integrated into many types of commercial and consumer devices including mobile phones, laptops, cars, printers, keyboards, mice, headsets, and more recently medical and personal devices (such as smart watches, music speakers, home appliances, fitness equipment and

trackers). This allows users to form self-organizing networks to transmit language and data before various devices.

As the demand for IoT continues to grow, so do Bluetooth applications, expanding from an initial focus on solving point-to-point interconnection problems to broadcast communications for indoor positioning and location-based services [31]. The status quo, driven by the Bluetooth mesh network, has created an emerging market for Bluetooth that requires a reliable wireless solution to build large-scale networks of devices.

7.4.2.2 Bluetooth Security Modes

Internationally, there are two guiding standards for Bluetooth security, NIST 800-121-R1 and IEEE 802.15.1. NIST 800-121-R1 details a complete set of Bluetooth security workflows, including identification and authentication of the sender's identity, confidentiality of the message in transit and the hierarchy of authorizations to access that message. The IEEE 802.15.1 focuses more on Bluetooth security standards.

There are five basic security services [32] specified in the Bluetooth standard.

1. **Authentication:** Based on the Bluetooth device address, the identity of the device being communicated with is verified. Bluetooth does not provide a native user authentication mechanism.
2. **Confidentiality:** ensuring that only authorized devices can access and view transmitted data to prevent information leakage due to eavesdropping.
3. **Authorization:** allowing control of resources by ensuring that a device is authorized before it is allowed to use a service.
4. **Message integrity:** verifying that messages sent between two Bluetooth devices have not been altered during transmission.
5. **Pairing/Binding:** creates one or more shared keys and stores these keys for subsequent connections to form trusted device pairs.

The discoverability mode of a Bluetooth device can also affect the security of the device. A device in discoverable mode is more vulnerable to attack. Device names, categories, service lists and technical information can all be accessed at a range of approximately 10 meters. In addition, each Bluetooth device has a 48-bit unique identifier called BD_ADDR. This address is similar to a MAC address, the BD_ADDR address is an address assigned to the hardware by the manufacturer and is used as a unique device identifier.

Bluetooth devices also have a total of four standardized access security modes. Security Mode 1 (no security); Security Mode 2 (service-level mandatory security); Security Mode 3 (link-level mandatory security); and Security Mode 4 (service-level mandatory security with key pairing policy). The different security modes determine the level of service security available.

Security mode 1 is considered insecure. In this security mode, the security features (authentication and encryption) are never activated and therefore the device and the connection are vulnerable to attacks. In practice, Bluetooth devices in this mode are not adversarial and do not employ any mechanism to prevent other Bluetooth devices from establishing a connection. If a remote device initiates a pairing, authentication or encryption request, the secure mode 1 device will accept the request without any authentication. Secure Mode 2 is a service-level enforced security mode which allows the security process to be initiated after the link has been established but before the logical channel has been established. In this security mode, the local security manager controls access to specific services. Access control and interfaces with other protocols and device users are maintained by a separate centralized security manager. This policy allows different security policies and trust levels to be defined for applications with different security requirements running in parallel to restrict access, and access to certain services can be granted without providing access to other services. Security Mode 3 provides the best security. It is a link-level enforced security mode in which the Bluetooth device initiates the security process before the link is fully established. A Bluetooth device operating in secure mode 3 authorizes authentication and encryption for all connections to the device. As a result, service discovery cannot even take place until authentication, encryption and authorization have taken place. Once a device has been authenticated, service-level authorization is not normally performed by a secure mode 3 device. When an authenticated remote device uses a Bluetooth service without the knowledge of the local device owner, service level authorization should be enforced to prevent authentication abuse. Secure Mode 4 uses Secure Simple Pairing (SSP), where the Elliptic Curve Diffie-Hellman (ECDH) key agreement replaces the obsolete key agreement during link key generation.

Also, when two devices try to connect for the first time, a trusted relationship needs to be established through authentication. Based on two sets of data, the BD_ADDR address and a link key, authentication is performed based on the Challenge/Response method. Once established, the link is saved for future pairings. Therefore, when connecting, it is first determined whether the device has previously been authorized as a trusted device. If the device database lists it as a trusted device, access to local services is granted. If the device is not listed as a trusted device, it must be re-authenticated before it can be authorized.

Bluetooth technology has a number of built-in security features [33]. They include:

Adaptive frequency hopping Bluetooth uses FHSS (frequency hopping spread spectrum) to ensure robustness against interference. It hops frequencies 1600 times per second across all 79.1 channels in the 2.4Hz band. If any interference is detected on a particular frequency, then, after 1,600th of a second, another frequency is immediately activated to send the message.

E0 encryption algorithm The key length of the cipher is typically 128 bits and uses stream encryption.

Invisibility This prevents the device from responding to scan attempts. The 48-bit BD_ADDR address of the device will also be hidden.

Pairing Only after pairing can the device communicate.

7.4.2.3 Bluetooth Security Vulnerabilities and Risk Mitigation

As Bluetooth Technology becomes widespread, vulnerabilities in its security protocols are increasing which can be potentially dangerous to the privacy of a user's personal information. Bluetooth attacks depend on exploiting the permission request/grant process that is the backbone of Bluetooth connectivity [34]. Here are a few examples of the mobile security threats in Bluetooth [35].

MAC spoofing attacks The MAC spoofing attack operates during the formation of the Piconet (micromesh) at the time of link key generation. Devices are able to authenticate themselves to each other by generating link keys. During the attack, the attacker can imitate the MAC information of other users and is also able to terminate others' connections or sniff and modify data using special tools.

PIN code cracking The attack occurs during the device pairing and authentication process. The attacker uses a channel tool (uber tooth) to collect the RAND and BD_ADDR of the target device and then uses the E22 algorithm to brute-force test all PIN alignments and compare them with the previously collected data until the correct PIN is found.

Man-in-the-middle attack A man-in-the-middle attack occurs when a device attempts to pair. During the attack, messages are maliciously forwarded between devices without using a shared key for authentication.

Bluetooth Hijacking Attacks In a Bluetooth hijacking attack, the attacker sends unauthenticated phishing messages, such as address book business cards, to the target device, tricking the user into revealing its Bluetooth authentication code. This allows the attacker to access files on the target device.

Local PIN code calculation During the attack, the attacker attempts to intercept and analyse the IN_RAND value, the LK_RAND value, the AU_RAND value and the SRES (signed response) value, which is an important matching variable required for authentication. The attacker uses brute force cracking to obtain a PIN that can be used to determine the correct SRES value.

Risk mitigation can be achieved in Bluetooth systems by applying countermeasures to address specific threats and vulnerabilities.

Organizations should mitigate the risks to their Bluetooth implementation by applying countermeasures to address specific threats and vulnerabilities. Each organization should assess the acceptable level of risk on the basis of the many factors that will influence the level of security implemented in that organization. To be effective, Bluetooth security should continue throughout the lifecycle of a Bluetooth solution.

Adequate levels of knowledge and understanding can also be provided to those who will be involved with Bluetooth-enabled devices [36]. Organizations using Bluetooth technology should establish and document security codes that address the use of Bluetooth devices and user responsibilities. These codes should include a list of approved uses of Bluetooth

and the types of information that can be transmitted over Bluetooth networks. The security code should also specify a scheme for proper password usage. Where feasible, a centralized approach to security policy management should be used in collaboration with the endpoint security products installed on the Bluetooth device to ensure that the policy is universally enforced locally.

7.5 Conclusion

The popularity of Internet of Things has grown in leaps and bounds over the years. In parallel, spoofing attacks on IoT devices has grown incrementally as well. This paper first introduces the basic ideas of IoT and spoofing attacks, a knowledge that serves as a prerequisite to understanding the security challenges of IoT and what mitigation technologies could be used against those spoofing attacks. For further information regarding short-range communication protocols and their security measures, we carefully surveyed the most important aspects of two technologies that are commonly used in IoT: ZigBee and Bluetooth. Both technologies have their own advantages and limitations.

While IoT vulnerabilities can be exploited and security risks cannot be easily avoided, the burden lies with the developer to address these issues and utilize the best of existing protocols, or research on new technologies that overcome the challenges.

Bibliography

- [1] Allan Tan. "What lies ahead for IoT in 2022", May 2022. <https://futureiot.tech/what-lies-ahead-for-iot-in-2022/>.
- [2] R. Porkodi and V. Bhuvaneshwari. "The Internet of Things (IoT) Applications and Communication Enabling Technology Standards: An Overview", 2014 International Conference on Intelligent Computing Applications, 2014, pp. 324-329. <http://dx.doi.org/10.1109/ICICA.2014.73>.
- [3] U.Farooq, M., Waseem, Muhammad, Mazhar, Sadia, Khairi, Anjum, Kamal, Talha. "A Review on Internet of Things (IoT)", International Journal of Computer Applications, March 2015.
- [4] Xiaohui, Xu. "Study on Security Problems and Key Technologies of the Internet of Things", 2013 International Conference on Computational and Information Sciences (2013): 407-410. <http://dx.doi.org/10.1109/ICCIS.2013.114>
- [5] Sobin, C. C. "A survey on architecture, protocols and challenges in IoT", Wireless Personal Communications 112.3 (2020): 1383-1429. <http://dx.doi.org/10.1007/s11277-020-07108-5>.
- [6] Paul Stokes. "4 Stages of IoT architecture explained in simple words", Dec 2018. <https://medium.datadriveninvestor.com/4-stages-of-iot-architecture-explained-in-simple-words-b2ea8b4f777f>.
- [7] CSRC Content Editor. Spoofing - nist glossary. <https://csrc.nist.gov/glossary/term/spoofing>.
- [8] David Balaban. "11 types of spoofing attacks every security professional should know about" , Mar 2020. <https://www.securitymagazine.com/articles/91980-types-of-spoofing-attacks-every-security-professional-should-know-about>.
- [9] Minoli, Daniel et al. "IoT security (IoTSec) considerations, requirements, and architectures", 2017 14th IEEE Annual Consumer Communications & Networking Conference (CCNC) (2017): 1006-1007. <https://doi.org/10.1109/CCNC.2017.7983271>
- [10] J. Zhang, H. Jin, L. Gong, J. Cao and Z. Gu. "Overview of IoT Security Architecture", 2019 IEEE Fourth International Conference on Data Science in Cyberspace (DSC), 2019, pp. 338-345. <http://dx.doi.org/10.1109/DSC.2019.00058>.

- [11] Liu, Shan, et al. "The Research on IOT Security Architecture and Its Key Technologies" 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC). IEEE, 2018. <http://dx.doi.org/10.1109/IAEAC.2018.8577778>
- [12] CISCOMAG. "10 IoT Security Incidents That Make You Feel Less Secure", Jan 2020. <https://cisomag.eccouncil.org/10-iot-security-incidents-that-make-you-feel-less-secure/>.
- [13] Uy, Nguyen Quoc, and Vu Hoai Nam. "A comparison of AMQP and MQTT protocols for Internet of Things." 2019 6th NAFOSTED Conference on Information and Computer Science (NICS). IEEE, 2019.
- [14] J. M. Carracedo et al., "Cryptography for Security in IoT", 2018 Fifth International Conference on Internet of Things: Systems, Management and Security, 2018, pp. 23-30. <http://dx.doi.org/10.1109/IoTSMS.2018.8554634>.
- [15] A. Yousefi and S. M. Jameii, "Improving the security of internet of things using encryption algorithms", 2017 International Conference on IoT and Application (ICIOT), 2017, pp. 1-5. <http://dx.doi.org/10.1109/ICIOTA.2017.8073627>.
- [16] L. Auer, C. Skubich and M. Hiller, "A Security Architecture for RISC-V based IoT Devices," 2019 Design, Automation & Test in Europe Conference & Exhibition (DATE), 2019, pp. 1154-1159. <http://dx.doi.org/10.23919/DATE.2019.8714822>.
- [17] He, Weijia, et al. "Rethinking Access Control and Authentication for the Home Internet of Things (IoT)." 27th USENIX Security Symposium (USENIX Security 18). 2018.
- [18] S. Khanji, F. Iqbal and P. Hung, "ZigBee Security Vulnerabilities: Exploration and Evaluating," 2019 10th International Conference on Information and Communication Systems (ICICS), 2019, pp. 52-57. <http://dx.doi.org/10.1109/IACS.2019.8809115>.
- [19] Andaloussi, Yasmina, et al. "Access control in IoT environments: Feasible scenarios." *Procedia computer science* 130 (2018): 1031-1036. <https://doi.org/10.1016/j.procs.2018.04.144>.
- [20] Ben Lutkevich. "access control list (ACL)". <https://www.techtarget.com/searchnetworking/definition/access-control-list-ACL>.
- [21] Ergen, Sinem Coleri. "ZigBee/IEEE 802.15. 4 Summary." UC Berkeley, September 10.17 (2004): 11.
- [22] C. M. Ramya, M. Shanmugaraj and R. Prabakaran, "Study on ZigBee technology", 2011 3rd International Conference on Electronics Computer Technology, 2011, pp. 297-301. <http://dx.doi.org/10.1109/ICECTECH.2011.5942102>.
- [23] Aju, Omojokun G. "A survey of zigbee wireless sensor network technology: Topology, applications and challenges." *International Journal of Computer Applications* 130.9 (2015): 47-55.

- [24] Salih, Mohammed A. Abdala, Alaa Mohammed. "Design and performance analysis of building monitoring system with wireless sensor networks." *Iraqi Journal of Science* 53.4 (2012): 1097-1102.
- [25] Somani, Nisha Ashok, and Yask Patel. "Zigbee: A low power wireless technology for industrial applications." *International Journal of Control Theory and Computer Modelling (IJCTCM)* 2.3 (2012): 27-33.
- [26] Wikipedia. ISM radio band. https://en.wikipedia.org/wiki/ISM_radio_band.
- [27] Vishruta Rudresh, "ZigBee Security: Basics", Nov 2017. <https://research.kudelskisecurity.com/2017/11/01/ZigBee-security-basics-part-1/>
- [28] Vishruta Rudresh, "ZigBee Security: Basics", Nov 2017. <https://research.kudelskisecurity.com/2017/11/21/zigbee-security-basics-part-3/>
- [29] L. Babun, H. Aksu, L. Ryan, K. Akkaya, E. S. Bentley and A. S. Uluagac, "Z-IoT: Passive Device-class Fingerprinting of ZigBee and Z-Wave IoT Devices," *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1-7. <http://dx.doi.org/10.1109/ICC40277.2020.9149285>.
- [30] Fan, Xueqi, et al. "Security analysis of ZigBee" *MWR InfoSecurity* (2017): 1-18.
- [31] MOKO BLUE, Why Bluetooth IoT?, Nov 2020. <https://www.mokoblue.com/why-bluetooth-iot/>
- [32] Elizabeth Montalbano, "Bluetooth Spoofing Bug Affects Billions of IoT Devices", Sep 2020. <https://threatpost.com/bluetooth-spoofing-bug-iot-devices/159291/>
- [33] L. Xing, "Reliability in Internet of Things: Current Status and Future Perspectives", in *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 6704-6721, Aug. 2020. <http://dx.doi.org/10.1109/JIOT.2020.2993216>.
- [34] Panse, Trishna, and Prashant Panse. "A survey on security threats and vulnerability attacks on bluetooth communication." *International Journal of Computer Science and Information Technologies* 4.5 (2013): 741-746.
- [35] Lonzetta, A.M.; Cope, P.; Campbell, J.; Mohd, B.J.; Hayajneh, T., "Security Vulnerabilities in Bluetooth Technology as Used in IoT", *J. Sens. Actuator Netw.* 2018, 7, 28. <https://doi.org/10.3390/jsan7030028>.
- [36] Minar, Nateq Be-Nazir Ibn, and Mohammed Tarique. "Bluetooth security threats and solutions: a survey." *International Journal of Distributed and Parallel Systems* 3.1 (2012): 127.
- [37] Jing, Q., Vasilakos, A.V., Wan, J. et al., "Security of the Internet of Things: perspectives and challenges", *Wireless Netw* 20, 2481-2501 (2014). <https://doi.org/10.1007/s11276-014-0761-7>.

Chapter 8

Threat Hunting: An Overview of Proactive Cybersecurity Methods

Heman Tanos, Karin Brunner

The risks from cyberattacks have steadily increased in recent years and no reversal of the trend is expected. Reactive measures such as response and forensic investigation alone are no longer sufficient to keep the dangers of cyberattacks at bay. Proactive measures are therefore recommended as a supplement. These are applied before vulnerabilities could be exploited. Examples of proactive methods are penetration testing, awareness training, honeypot/deception, and moving target defense (MTD). Another important proactive method is threat hunting. Data, hypotheses, tools, and expertise are the four prerequisites for successful hunting. Current approaches are relatively limited to ad-hoc monitoring of various malicious activities like disguised network connections or data exfiltration captured in an abundance of logs. With the help of machine learning (ML), people try to automate these unstructured processes. Various academic articles show that ML gives quite good results in intrusion detection. Another helpful technique to support threat hunting is Natural Language Processing (NLP). NLP can be useful in the threat hunting pipeline by extracting knowledge from threat intelligence and performing data management for subsequent processing.

Contents

8.1	Introduction	175
8.2	Background	176
8.2.1	Threats and Defence	176
8.2.2	Proactive Approaches	176
8.3	Threat Hunting	178
8.3.1	Requirements	178
8.3.2	Current Approaches	179
8.4	Case Study Proactive Cybersecurity	192
8.5	Summary	193

8.1 Introduction

According to an annual survey conducted by [2], companies worldwide rate the threat posed by cyber perils as the most threatening for the year 2022. The 2'650 experts from 89 countries surveyed believe that for example issues like a business interruption and natural disasters will cause less damage than cyber risks.

But is this threat real or is it just an assumption? Unfortunately, the companies assess the risks correctly. Cybersecurity Ventures estimates that the costs for global cybercrime will grow from 3 trillion USD in 2015 to 10.5 trillion USD in 2025. This would correspond to annual growth of 15 percent [3]. In view of the high expected costs, it is essential for companies to deal with the topic of cyber security.

Figure 8.1 shows the same trend, namely an increase in cybercrime incidents as well as an increase in the costs caused by them. This graph was published by the FBI's Internet Crime Complaint Center (IC3). IC3 collects and analyses data related to cybercrime and makes recommendations [11].

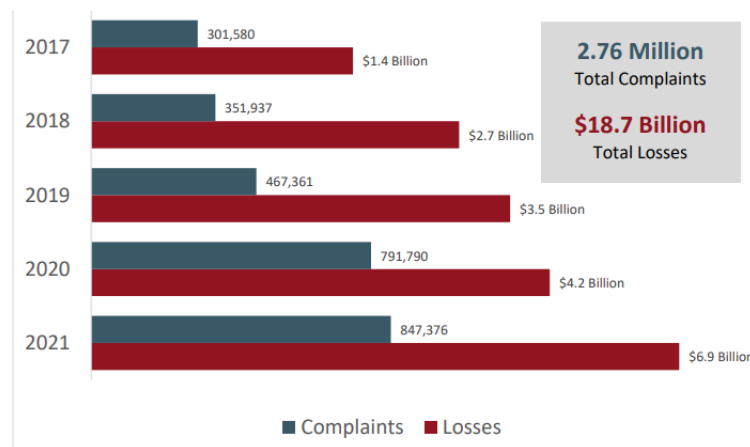


Figure 8.1: Complaints and Losses over the Last Five Years [11]

In order to counteract the increasing threats, the reactive cybersecurity methods have been constantly developed in recent years. However, proactive methods have also been added to further increase protection. This paper aims to give an overview of these proactive cybersecurity methods and to examine current approaches in the context of threat hunting. The first Section will shed light on current threats and possible defence approaches, focusing on the differentiation between reactive and proactive cybersecurity. In the next Section, some proactive methods will be explained. Chapter 8.3 then forms the main part of this work. The approach is explained in detail, possible advantages and disadvantages are described, useful tools are presented, and the approach's usefulness is discussed. The last part then covers a case study on proactive cybersecurity.

8.2 Background

This Chapter serves to provide the reader with background knowledge for the following Chapter 8.3. First, in Section 8.2.1 the difference between reactive and proactive approaches are described and it is explained why both are needed. In the following part 8.2.2 some important proactive approaches in the context of cybersecurity are presented.

8.2.1 Threats and Defence

According to The European Union Agency for Cybersecurity (ENISA), ransomware, malware, and cryptojacking are among the prime threats in 2021. In a ransomware attack, data is encrypted and the affected company has to pay a high amount of money to the attackers to decrypt the data again. The payment is usually made in a cryptocurrency. Ransomware is considered to be the top threat of 2021. Malware is a type of malicious software that executes unwanted processes to the detriment of the victim. It has been a prime threat for many years even though its importance decreased in 2021. One threat that has emerged only in recent years is cryptojacking. Cryptojacking, also known as hidden cryptomining, involves using computers to generate cryptocurrencies without their owners' knowledge [1].

There are many different strategies for avoiding cyberattacks or minimizing their damage. [4] roughly divides these strategies into two approaches: reactive and proactive. Reactive approaches are about reacting when an attack has been successful. These include for example response and forensic investigation. Reactive approaches have been widely used for years and help to reduce the damage caused by cyberattacks. They are also comparatively cheap to implement and easy for management to understand as there is a cause-effect relationship.

In proactive approaches, on the other hand, action is taken before vulnerabilities are exploited or data has been stolen. The biggest advantage of this is that both finances and reputation are unharmed. It also allows new threats to be identified and investigated before they can cause damage. The only disadvantage is the high costs since security experts are needed and there are not many off-the-shelf solutions available. In summary, it can be said that both approaches make an important contribution to protection against cyber risks and that a combination is therefore the best solution [5, 4].

8.2.2 Proactive Approaches

This Chapter looks at different proactive methods in detail. The methods are explained and possible advantages and disadvantages are pointed out.

8.2.2.1 Penetration Testing

In penetration testing, a company hires experts to bypass its own cybersecurity and hack the company. The company wants to find out where its security vulnerabilities are before they can be exploited by real hackers with malicious intentions [4].

In 2003, the Bundesamt fuer Sicherheit in der Informationstechnik published an implementation concept for penetration testing, which is still up-to-date. Among other things, this concept shows the benefits of penetration testing and explains which types there are. The concept distinguishes between two possibilities on which information basis a penetration test can be carried out. The two possibilities are called black-box and white-box testing. In black-box testing, only publicly accessible information is available to the tester. These are the conditions that a real hacker would find and have to overcome in order to hack the organisation. In white-box testing, the attacker has internal information, which could be the case, for example, in an attack by a (former) employee or an external service provider.

Thanks to penetration testing, the company has the opportunity to improve the security of its systems. However, this type of testing also poses challenges for the company. Penetration testing can only be carried out by experts, as it requires a lot of knowledge in different areas such as system administration, IT security products (e.g. firewalls) and hacking tools. Each test must be adapted to the respective organisation and is therefore difficult to standardise. This individuality makes penetration testing costly. In addition, the tests can affect system performance, and in the worst case, they can even lead to a system failure. The last point to consider is the short-lived nature of the results. Even if a penetration test today leads to the result that a company is well protected, a new dangerous vulnerability may be reported tomorrow [10]. An example of penetration testing will be shown in Chapter 8.4 "Case Study Proactive Cybersecurity".

8.2.2.2 Awareness Training

When explaining security awareness training, it is easiest to first state the main goal: Security awareness training helps to reduce the risk of phishing. Phishing is the attempt to obtain information by exploiting human inattention and then misusing it for malicious purposes. There are two ways to tackle phishing. Phishing can be combated by technical and non-technical means. Technical means include tools that prevent the system's vulnerabilities from being exploited. Non-technical means focus on human error/inaccuracy. Many studies show that a combination of non-technical and technical solutions offers the best protection and that only one method alone is not sufficient. Security awareness training belongs to the non-technical solutions. Awareness training increases user awareness and thus reduces the risk of phishing and other security risks. A well-designed training programme is characterised by two features. First, it should convey information in a way that will be remembered for a long time. Secondly, it should help users to apply what they have learned in other safety areas. To achieve the two points mentioned above, awareness training should be conducted regularly [6]. An example of security awareness training can be found in Chapter 8.4 "Case Study Proactive Cybersecurity".

8.2.2.3 Honeypot and Deception

The idea behind a honeypot is to lure attackers with a bait and spy on them. Thanks to this method, valuable information can be gained about the attackers' actions and goals. The bait is usually created from fake data and is maintained in isolation from the rest of the system. This ensures that the attackers cannot cause any real damage. Honeypots can be created for different purposes, such as malware honeypots, email/spam honeypots or database honeypots. However, there are two major disadvantages to the honeypot method. Firstly, there is a lot of manual work behind it, as the honeypots are individual. Secondly, they are often discovered relatively quickly because realistic updates of the activities are missing. Deception technology offers a solution for these two drawbacks. This is an automated solution to the honeypot method. With the help of this technology, a network of perfect decoys can be built and maintained and the security team is supported in tracking the attacks [8].

8.2.2.4 Moving Target Defense (MTD)

According to [12], the static nature of network architecture helps to ensure the smooth running of day-to-day operations. However, it also has a significant disadvantage: it makes networks easier to attack. Potential attackers have a long time to identify and observe a network and its vulnerabilities and can then attack at the right moment. The defenders are at a disadvantage, they have to guard the network around the clock and do not know in advance when there will be an attack. Nonetheless, there is a collection of Moving Target Defence Strategies that provide a remedy for this problem. With Moving Target Defense Strategies, the configurations of a possible attack target (e.g. a network) are constantly changed, so that the complexity for attackers increases. Moving target defence can be used, for example, to combat so-called computer worms. Computer worms are programs with malicious intentions. [13] shows that by changing the IP address of the host, the spread of the malicious program can be prevented.

8.3 Threat Hunting

Threat hunting refers to the proactive measures performed with the goal of detecting attackers, adversaries, malware and other cyber threats whose intrusion has been successful and who are lurking to carry out further damages.

8.3.1 Requirements

The requirements of threat hunting are data, hypothesis, tools and expertise [7].

Data

Most threat hunting measures need a lot of data, may it be system log, security log, network log, email log or application logs. The approaches based on natural language processing also need as input threat intelligence data from alerts, blogs, social media and the like.

Hypothesis

The threat hunter needs to know what he/she is looking for, in other words, he/she must form a hypothesis based on past observation.

Tools

Threat hunting relies on the availability of tools, either built-in like `nmap` and `sysmon` or third party.

`nmap` is a command often used for security auditing by scanning ports on local and remote network addresses to look for network information like open ports, running services, service type etc. It can detect some known Denial of Service (DoS) attacks and some common malware. It works by sending packets of data to ports and analyzing the response.

`sysmon` is a command for monitoring and logging process creation, network connection, file changes, registry modifications and further administrative events that is often used for discovering malicious activities. Apart from the process name and id, it also logs the name of the user who initiates the command, the directory it is started from, its hash, its parent process id and further information.

There are plenty of third party software available to automate some of the processes and routine tasks.

Expertise

Besides technical understanding (knowledge about tools, event ID, port numbers in case of `nmap`, `sysmon`; typical behaviors: e.g. what other machines a certain machine talks to / queries), a threat hunter must have deep knowledge about the business' IT environment and user space.

8.3.2 Current Approaches

The following subsections introduce current threat hunting approaches starting from the intuitive ones based on common sense and best practices. Then two framework-based, formalized approaches are described that try to bring in some structure. Finally, more recent approaches based on machine learning and natural language processing are presented.

8.3.2.1 Ad hoc techniques

The key activity related to threat hunting involves monitoring the various log files and watching out for *anomalous behavior* or *outliers*.

Some most common hypotheses in threat hunting are [20, 30] :

Masquerading

The threat hunter's *hypothesis* here is that an attacker who has succeeded to gain access to a machine will do its best to avoid detection. In practice it would try to achieve this by making its identity look legitimate or by concealing its location or other file metadata.

Since `sysmon` logs an event 4688 when a new process is created, one can monitor the log file, verify the process names with a list of known malwares and take appropriate actions in case a suspicious process is identified. One can also keep a baseline containing whitelisted processes and dll names and perform an incremental audit. Since process names are susceptible to masquerading (*flying underneath the radar*; tricking the operating system (OS) to run something in a privileged mode e.g. `lsass.exe`¹), a threat hunter can monitor the hash value of the binary instead of the process name. The hashes of two otherwise very similar looking strings (like `svchost` vs `srvcHost`) can be significantly different. The downside is that a hash changes at every patch, causing a burden of updating the whitelist very often.

A threat hunter can also benefit from an antivirus software when the latter found and removed a malware. It will show where the malware was removed from, from the `Downloads` folder or from `C:\Windows\Temp` or `C:\Windows\System32`. The location will show where the malware tries to hide. Furthermore, one should be highly suspicious of malware that was put in `C:\Windows\System32` because administrator right would be needed in order to write to that folder.

The "take appropriate action" starts from manual (or visual) inspection, compares the suspect with the baselined item and considers how long the latter has been on the baseline. The baseline can be made specific to a machine or a user. When a new server comes into service or a new user is onboarded, one allows an observation period of a few weeks.

When some suspicious process name or hash is identified, it should be googled and compared to entries in VirusTotal, a security community that analyzes suspicious artifacts and network site addresses and shares breaches with its community members.

Scripting abuse

The threat hunter's hypothesis here is that the intruder might try to not execute the malicious code as an EXE but as a script. Hence, monitoring the hash of EXE's might

¹Local Security Authentication Server, a process to validate user logons.

not be enough. On windows, scripts run under `cscript`, `wscript` or `powershell`. Scripts running under `cscript` include script name and user name, hence best practices call for monitoring unusual script names or unusual user names.

Powershell allows a number of monitoring options. The monitoring results are written into the powershell audit logs.

Best practices found that a powershell script is often started from some parent process. Hence a threat hunter should monitor closely parent processes that start a powershell script and become suspicious of unusual cases like an outlook / excel / word process that starts a powershell script. Baselining the parent processes that start a powershell script might be useful.

If `sysmon` finds a process or script changes registry keys, starts or ends external connection, or a finance chief who never uses powershell is found to suddenly start using powershell, the threat hunter should investigate in more details.

Lateral movement

The hypothesis here is : An attacker might gain access to a machine and start from there masquerading as a regular user. He might have gotten hold of admin access and now attempts to extend his lateral kill chain.

On windows every logon process is logged as an event having an ID equal to 4624 (*An account was successfully logged on*), providing a way to aid auditing. A threat hunter should monitor lateral movements, watch out for users who newly log on although they have been on a computer for along time, creating a suspicion for a credential theft. This monitoring process might want to exclude new computers.

New network connections must be monitored and checked for plausibility. For this purpose `sysmon 3` (network connection) or event 5156 (network connection has been allowed) can be monitored in the security log. Only new combinations of source and destination and unusual usage patterns need to be watched out for. New computers and new users should be excluded from this monitoring because they have not yet established a usage pattern.

Connection Proxy

The hypothesis here is : an adversary wants to connect to an outside server but does not want to be caught. Hence he does not connect directly but goes through a connection proxy.

A valid detection logic / rule could be : a process receives an inbound connection from an internal host and the same process initiates an outbound connection to the internet. Distinguishing an internal from an external host can be performed on the basis of whether the same subnet or an unknown one is used. smtp email server and other legitimate proxies should be excluded / filtered.

Persistence

An adversary who has gained access to a system tries to *persist* his access across restarts.

For this end, he would add an extra service to run in the background. Since a service is a series of registry entries and an action to start a service at system startup time [21] will be logged in the Windows Management Instrumentation (WMI) event repository, a threat hunter should watch out for event id 4697 (*A service was installed in the system*), bulk registry key readouts, registry key changes, whether administrator rights are needed, and make these events plausible. Furthermore, he could look out for attempts to modify services that do not correlate with patch cycles nor known software.

Domain Name Service (DNS) abuse

DNS is a vulnerable protocol that is often used in malicious activities [22, 23].

Best practices call for watching out for abnormally large DNS packets as a potential Indicator of Compromise (IoC) while baselining the normal size of DNS packets. Further, a threat hunter should monitor changes to `etc/hosts` in the security log or file system log.

A common attack hypothesis is DNS rebinding (<https://unit42.paloaltonetworks.com/dns-rebinding/>): after a malicious attacker persists itself on the victim's machine, it can assume the DNS resolver's job. It resolves an external IP address to look like an internal address, hence violating the same-origin principle applied to some web pages (javascript elements, images, css to render web pages etc).

Since DNS API is often used for address resolution, a threat hunter should watch out for API commands with a certain directionality with enhanced awareness.

Data exfiltration

An adversary steals data from a network by collecting, packaging and compressing / encrypting them before transferring them out of the network.

A sample threat scenario involving an insider applying the data exfiltration hypothesis would consist of the following steps:

- a contractor knows his contract is going to expire without being extended
- he would unsuccessfully attempt to login to unauthorized servers outside business hours
- a threat hunter would conduct file monitoring for add, delete, update, compress and transfer activities

- in case a suspicion is substantiated, an alert can be created, a dashboard updated and an administrator can disable the affected user account and alert security

8.3.2.2 The MITRE ATT&CK framework

In order to facilitate the sharing of past experience among the cybersecurity community, the not-for-profit organisation MITRE launched the project ATT&CK [32] in 2013.

ATT&CK [32, 33, 34], an acronym for Adversarial Tactics, Techniques, and Common Knowledge, is a public knowledge base of adversarial tactics ("what") and techniques ("how") used in real-world observations to carry out attacks.

Its major benefit is a common taxonomy based on real-world cases that has been used in many instances as basis for cybersecurity frameworks. It lets people learn from past experience to reduce the time to detect attacks.

The ATT&CK knowledge base is organized in a 2D-matrix with tactics as columns and rows as techniques in a tactic. For each technique one or more examples are given that include data source as well as prescriptive guidance on how to mitigate (prevent) and how to detect it. To ensure platform independence, it has a high level textual format.

Fig. 8.2 [54] shows an excerpt of the ATT&CK matrix (the right and the bottom parts shortened for space reason), fig. 8.3 [55] a sample tactic and fig. 8.4 [56] a sample technique.

Reconnaissance	Resource Development	Initial Access	Execution	Persistence	Privilege Escalation	Defense Evasion	Credential Access	Discovery	La...
10 techniques	7 techniques	9 techniques	12 techniques	19 techniques	13 techniques	40 techniques	15 techniques	29 techniques	9 tech
Active Scanning (2)	Acquire Infrastructure (6)	Drive-by Compromise	Command and Scripting Interpreter (8)	Account Manipulation (4)	Abuse Elevation Control Mechanism (4)	Abuse Elevation Control Mechanism (4)	Adversary-in-the-Middle (2)	Account Discovery (4)	Exploits Remote
Gather Victim Host Information (4)	Compromise Accounts (2)	Exploit Public-Facing Application	Container Administration Command	BITS Jobs	Access Token Manipulation (5)	Access Token Manipulation (5)	Brute Force (4)	Application Window Discovery	Internal Spearph
Gather Victim Identity Information (3)	Compromise Infrastructure (6)	External Remote Services	Deploy Container	Boot or Logon Autostart Execution (15)	Boot or Logon Autostart Execution (15)	Build Image on Host	Credentials from Password Stores (5)	Browser Bookmark Discovery	Lateral Transfe
Gather Victim Network Information (6)	Develop Capabilities (4)	Hardware Additions	Exploitation for Client Execution	Boot or Logon Initialization Scripts (5)	Boot or Logon Initialization Scripts (5)	Deobfuscate/Decode Files or Information	Exploitation for Credential Access	Cloud Infrastructure Discovery	Remo
Gather Victim Org Information (4)	Establish Accounts (2)	Phishing (3)	Inter-Process Communication (2)	Browser Extensions	Boot or Logon Initialization Scripts (5)	Deploy Container	Forced Authentication	Cloud Service Dashboard	Remo
Phishing for Information (3)	Obtain Capabilities (6)	Replication Through Removable Media	Native API	Compromise Client Software Binary	Create or Modify System Process (4)	Direct Volume Access	Forge Web Credentials (2)	Cloud Service Discovery	Remo
Search Closed Sources (2)	Stage Capabilities (5)	Supply Chain Compromise (3)	Scheduled Task/Job (6)	Create Account (3)	Domain Policy Modification (2)	Execution Guardrails (1)	Input Capture (4)	Cloud Storage Object Discovery	Replica
Search Open Technical Databases (5)	Trusted Relationship	Software Deployment Tools	Shared Modules	Create or Modify System Process (4)	Domain Policy Modification (2)	Exploitation for Defense Evasion	Modify Authentication Process (4)	Container and Resource Discovery	Remo
Search Open Websites/Domains (2)	Valid Accounts (4)	System Services (2)	User Execution (3)	Event Triggered Execution (15)	Event Triggered Execution (15)	File and Directory Permissions Modification (2)	Network Sniffing	Domain Trust Discovery	Softwa
Search Victim-Owned Websites	Windows Management Instrumentation	External Remote Services	Implant Internal Image	Exploitation for Privilege Escalation	Exploitation for Privilege Escalation	Hide Artifacts (9)	OS Credential Dumping (8)	File and Directory Discovery	Softwa
		Hijack Execution Flow (11)	Modify Authentication Process (4)	Hijack Execution Flow (11)	Hijack Execution Flow (11)	Impair Defenses (9)	Steal Application Access Token	Group Policy Discovery	Taint Sh
		Process Injection (11)	Scheduled Task/Job (6)	Process Injection (11)	Process Injection (11)	Indicator Removal on Host (6)	Steal or Forge Kerberos Tickets (4)	Network Service Scanning	Use A
		Valid Accounts (4)	Office	Scheduled Task/Job (6)	Scheduled Task/Job (6)	Indirect Command Execution	Steal Web Session Cookie	Network Share Discovery	Auth
				Valid Accounts (4)	Valid Accounts (4)	Masquerading (7)	Two-Factor Authentication Interception	Network Sniffing	Auth
								Network Service Scanning	Auth
								OS Credential Dumping (8)	Auth
								Steal Application Access Token	Auth
								Steal or Forge Kerberos Tickets (4)	Auth
								Steal Web Session Cookie	Auth
								Two-Factor Authentication Interception	Auth
								Masquerading (7)	Auth
								Indirect Command Execution	Auth
								Scheduled Task/Job (6)	Auth
								Valid Accounts (4)	Auth
								Process Injection (11)	Auth
								Hijack Execution Flow (11)	Auth
								Event Triggered Execution (15)	Auth
								External Remote Services	Auth
								System Services (2)	Auth
								User Execution (3)	Auth
								Windows Management Instrumentation	Auth
								Valid Accounts (4)	Auth
								Trusted Relationship	Auth
								Supply Chain Compromise (3)	Auth
								Stage Capabilities (5)	Auth
								Obtain Capabilities (6)	Auth
								Establish Accounts (2)	Auth
								Develop Capabilities (4)	Auth
								Hardware Additions	Auth
								External Remote Services	Auth
								Exploit Public-Facing Application	Auth
								Compromise Accounts (2)	Auth
								Acquire Infrastructure (6)	Auth
								Active Scanning (2)	Auth

Figure 8.2: ATT&CK Matrix [54]

In the **Valid Accounts** technique case, the *procedure examples* section listed some 3 dozens cases where this technique has been used to carry out the **Persistence** tactic, some of them are [56]:

Home > Tactics > Enterprise > Persistence

Persistence

The adversary is trying to maintain their foothold.

Persistence consists of techniques that adversaries use to keep access to systems across restarts, changed credentials, and other interruptions that could cut off their access. Techniques used for persistence include any access, action, or configuration changes that let them maintain their foothold on systems, such as replacing or hijacking legitimate code or adding startup code.

ID: TA0003

Created: 17 October 2018

Last Modified: 19 July 2019

[Version Permalink](#)

Techniques

Techniques: 19

ID	Name	Description
T1098	Account Manipulation	Adversaries may manipulate accounts to maintain access to victim systems. Account manipulation may consist of any action that preserves adversary access to a compromised account, such as modifying credentials or permission groups. These actions could also include account activity designed to subvert security policies, such as performing iterative password updates to bypass password duration policies and preserve the life of compromised credentials.

Figure 8.3: Persistence tactic [55]

- The financially motivated threat group FIN7 harvested valid administrative credentials for lateral movement via a zipped file downloaded from a malicious https site, documented in details in [36].
- The financially motivated threat group FIN5 attacked the hospitality industries via VPN, RDP ², Citrix connections, documented in [37].
- The Linux Rabbit malware acquired valid SSH accounts to login to Linux servers and IoT devices and installed cryptocurrency miners there, documented in [38].

The mitigation section in the **Valid Accounts** technique page [56] recommends the following policies:

1. Development practices

Developers should not hard code credentials into source code, config file and / or place the same knowingly or unknowingly in a publicly accessible repository.

2. Password policies

Default username and password should be changed upon deployment / installation. Password in the production environment should not be the same as in the development and test environment.

3. Privileged accounts

Powerusers, users with elevated rights and administrator accounts should be strongly protected and audited. Shared accounts should be used prudently.

4. User training

In some cases, users can be asked to identify/verify their peers as part of authentication. Users should be trained to be wary in such cases.

²Remote Desktop Protocol

Home > Techniques > Enterprise > Valid Accounts

Valid Accounts

Sub-techniques (4) ▼

Adversaries may obtain and abuse credentials of existing accounts as a means of gaining Initial Access, Persistence, Privilege Escalation, or Defense Evasion. Compromised credentials may be used to bypass access controls placed on various resources on systems within the network and may even be used for persistent access to remote systems and externally available services, such as VPNs, Outlook Web Access and remote desktop. Compromised credentials may also grant an adversary increased privilege to specific systems or access to restricted areas of the network. Adversaries may choose not to use malware or tools in conjunction with the legitimate access those credentials provide to make it harder to detect their presence.

The overlap of permissions for local, domain, and cloud accounts across a network of systems is of concern because the adversary may be able to pivot across accounts and systems to reach a high level of access (i.e., domain or enterprise administrator) to bypass access controls set within the enterprise. ^[1]

ID: T1078

Sub-techniques: T1078.001, T1078.002, T1078.003, T1078.004

- ⓘ **Tactics:** Defense Evasion, Persistence, Privilege Escalation, Initial Access
- ⓘ **Platforms:** Azure AD, Containers, Google Workspace, IaaS, Linux, Office 365, SaaS, Windows, macOS
- ⓘ **Permissions Required:** Administrator, User
- ⓘ **Effective Permissions:** Administrator, User
- ⓘ **Defense Bypassed:** Anti-virus, Application control, Firewall, Host intrusion prevention systems, Network intrusion detection system, System access controls

Figure 8.4: Valid Accounts technique [56]

The detection section in the **Valid Accounts** technique page [56] recommends the following actions:

1. Logon sessions

Threat hunters should watch out for single user logged in multiple machine or multiple users logged into a single machine simultaneously.

Threat hunters should align user login data with further administrative data, for example absence days.

2. User authentication

Threat hunters should perform periodic access reviews, especially for users with enhanced access rights.

The MITRE ATT&CK framework has enjoyed a certain level of acceptance in the cybersecurity community. It serves as basis of the experiments in [18] and [16]. It is used in microfocus's product ArcSight [39, 35], in Azure Sentinel [40], in an industrial control system [41], among others.

A number of practical tools are available for working with ATT&CK. Among others,

- STIX ³: a json compatible format for expressing threat information
- TAXII ⁴: an application protocol running over https. It enables people to consume (subscribe to) threat intelligence data in STIX format over an API.
- Powershell / Python connectors to access MITRE's json repository.

³stix.mitre.org

⁴taxii.mitre.org

8.3.2.3 Targeted Hunting Integrating Threat Intelligence

TaHiTI [9] is a threat hunting methodology created in a joint effort by the financial sector in the Netherlands, but also available to other organizations in other sectors. The idea is to make threat intelligence the focus of the threat hunting process and make the methodology a common approach for threat hunting in the financial sector, and to share best practices amongst each other. In TaHiTI threat intelligence is a major source of threat hunting hypothesis, might however at the same time generate new threat intelligence.

Like other threat hunting approaches, it aims at reducing the threat detection gap between the time of the breach and the time the attacker is detected. In its effort trying to uncover unknown attackers via known TTPs, it might even discover new TTPs, related TTPs or additional information on existing TTPs.

A TaHiTI process consists of 3 steps, namely trigger, hunt and finalization.

The major trigger source is threat intelligence, besides other processes like previous or current incident reports. The ATT&CK framework can serve as a guidance to look for TTPs. The TaHiTI trigger creates a ticket containing the trigger and the hypothesis.

The second step of a TaHiTI process picks up the ticket and enriches it with various information, e.g. from ATT&CK and refines the hypothesis. The actual hunting process starts with retrieving, analyzing data and confirming or rejecting the hypothesis. The analysis is carried out using a mixture of tools, manually or automated. In an exfiltration case, for example, one can analyze the number of bytes sent and received per source and destination, the ratio of request to response, mean and standard deviation of request and response size as well as various clustering and grouping operations can be carried out.

The third step of the TaHiTI process then makes use of the standardized reporting tool *MagMa* to document the process and update the ticket.

8.3.2.4 Machine-Learning (ML)

Unlike the adhoc, unstructured threat hunting procedures described above, machine learning provides a slightly more automatic way for threat detection and hence is a valuable tool for threat hunting.

Anomaly based NIDS ⁵ that are based on statistical evaluation relies on a stationary assumption of the traffic behavior. Ruled - based systems suffer from the need to always create new rules. A machine learning based IDS can adapt to dynamic network traffic patterns and learn from new network traffic behavior to detect anomalies [24].

The following paragraphs report some classical machine learning techniques, while the subsequent subsection will present some more recent results.

A number of academic articles [26, 27, 28, 29] have reported results applying machine learning techniques to detect intrusions, with quite good results. Basically, they study a data set and perform a classification on a number of features.

⁵Network Intrusion Detection System

UNSW-NB15 study

As an example, [25, 19] use the UNSW-NB15 data set with 47 features that contain realistic network data (packets) including attack behaviors simulated in a research lab in Australia. Some of the features are :

1. Source / Destination IP address
2. Source / Destination Port number
3. number of bytes
4. time to live, the maximum number of hops a data packet can travel between routers before it is discarded
5. packet loss, the fraction of network packets that fail to reach their expected destination
6. protocol
7. start time, various connection setup times, arrival time between packets, round trip time
8. = 1 if ftp login by user and password , 0 otherwise

There are 9 types of attacks : DoS, Fuzzers (random data to let system crash), worm etc.

The machine learning algorithm *classifies* a data record as being malicious or not dependent on its features. The classifier can be based on a decision tree, a logistic regression (which is a deterministic algorithm with a closed form solution), a neural network or one of the many clustering algorithms.

Figure 8.5 [25] shows classification results on UNSW-NB15 and another, simpler dataset, the KDD99 [24], (<http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html>), with 4.9 million connection records and 42 features.

Techniques	KDD99 data set			UNSW-NB15 data set	
	Reference	Accuracy (%)	FAR (%)	Accuracy (%)	FAR (%)
DT	(Bro-IDS Tool, 2014)	92.30	11.71	85.56	15.78
LR	(Witten & Mining, 2005)	92.75	-	83.15	18.48
NB	(Shyu et al., 2005)	95	5	82.07	18.56
ANN	(Witten & Mining, 2005)	97.04	1.48	81.34	21.13
EM clustering	(Salem & Buehler, 2012)	78.06	10.37	78.47	23.79

Figure 8.5: Classification results on network connection data sets [25](DT: decision tree, LR: logistic regression, NB : Naïve Bayes, ANN: artificial neural network, EM: expectation maximization clustering, FAR: false alarm rate)

NATO Communication & Information Agency (NCIA) study

Article [16] presents results using machine learning techniques on the open source data analytics platform KNIME [17] to detect/predict attacks in an isolated network with simulated activities. There were 93 millions windows event log entries. Noise reduction first discarded 51 % of the data. Limiting the windows log entries to (3243) time windows further makes the data manageable. Dimensional reduction further reduced the data by 35 %. A subset selection algorithm (a DNN ⁶ with 2 hidden layers) selected 8042 out of 30069 features.

The following experiments were conducted:

- Supervised Machine Learning to detect known cyber attacks according to MITRE ATT&CK. The threat log contains timestamps, tactics and techniques and serves as a source of the labels. A DNN in the form of a multilayer perceptron was trained to detect security threat from regression on 8042 features to predict 32 ATT&CK techniques. An F1-score over 0.99 can be achieved.
- Unsupervised Machine Learning to detect unknown cyber attack and APT ⁷. Applying the DBSCAN clustering algorithm can achieve an accuracy slightly over 80 %.

Further studies

Further practical work are as follows:

[42] that explains how to perform threat detection using supervised ML on file parameters (it states : "structure of file, content in file, use of different features in file formats, behavior of file when executed or opened in application") and unsupervised ML to detect anomaly.

In a machine learning evaluation study, [43] looked at techniques for consistent detection of Windows ransomware network traffic. This article reported achieving a True Positive Rate (TPR) of 97.1 % when applying the decision tree classifier on a dataset created from conversation-based network traffic features.

[44] proposed a neural network and tested it on Windows, Ransomware, Internet of Things (IoT) and a mix of different malware sample datasets. For example, an evaluation to detect IoT malware achieved an accuracy of 99.65 % , an AUC ⁸ of 0.99 and a MCC ⁹ of 0.992.

[48] reported 4 supervised machine learning algorithms (KNN, SVM ¹⁰, DT ¹¹ and RF ¹²) for detecting cyber attacks on a water treatment facility. KNN is a clustering algorithm

⁶Deep Neural Network

⁷Advanced Persistent Threat, an intruder who persists itself on a host and steals information over an extended period of time

⁸Area Under ROC Curve

⁹Matthew Correlation Coefficient

¹⁰support vector machine

¹¹decision tree

¹²random forest

based on the k nearest neighbors in the feature space. SVM is a classifier that finds the best hyperplanes separating the clusters. A random forest works by taking the majority votes from a set of decision trees. DT is found to beat the other three with an overall accuracy of 99.9 % and other improved metrics.

Many threat hunters are biased towards a single view, e.g. opcodes. [49] showed that combining multiple views improves the performance. The authors performed various multiview experiments on various OS and platforms and reported better accuracies and a low false positive rate than single view approaches.

In general, signature based techniques can achieve around 90 % accuracy. More advanced AI techniques can achieve 95 % but the false positive rate also rises [46].

Although ML is a promising tool, adversaries can also use it to improve their attacks [45]. Phishing emails can be trained to look like legitimate emails. Attackers can generate a huge number of false alarms to disturb machine learning. The more advanced techniques, like those in the next section, are often based on deep learning, requiring large amount of computing power to perform training.

8.3.2.5 Natural Language Processing

More recent work goes beyond static processing of log data to detect attackers. Among others, those employing NLP deal with how to make sense of the huge amount of textual data from multiple, disparate sources to assist the expert [31, 14, 18] in threat hunting.

Lexical similarity

According to [47], attackers tend to observe a certain fixed lexical style when naming malicious domain names. Hence one can apply techniques from NLP that look for lexical styles similar to those of known malicious sites to look for further / future malicious sites.

[47] reported successful detection of phishing attacks on companies such as Wells Fargo, Facebook, Dropbox and others using lexical similarities (minimum edit distance) to identify malicious websites and phishing domains.

securityKG

This work [15] has the goal to meet a number of challenges :

- data comes in myriads of formats, structured or unstructured
- threat intelligence consists not only of IoCs but also relationships between them
- existing approaches are static, ignoring sequence of steps

- labels needed for training in a machine learning setting are hard to come by

This work, co-authored by Microsoft, claims to be the first work in this space.

It builds a knowledge graph from data web crawled from 40+ major sources and constantly updates it.

Although its size is only 9 K lines of python code, it performs data collection, parsing and knowledge extracting, creating a knowledge graph, persisting it into a DB (Neo4j or RDBMS) and updating it by continuously ingesting new data (scalability). The extraction, based on CRF ¹³, had to overcome a number of challenges like nuances (special characters like dots and underscores) and labelling.

Labelling is trained using a training function that is noisy and whose accuracy is learned. This approach is known as data programming [51] and allows the creation of large training sets quickly.

The extracted knowledge persisted in a DB can later be used by various frameworks (e.g. [53]).

A front end is built to explore the knowledge graph [14].

ThreatRaptor presented next picks up where securityKG leaves off.

ThreatRaptor

According to [57]: "Threat intelligence is data that is collected, processed, and analyzed to understand a threat actor's motives, targets, and attack behaviors." It is predicted to gain increasing importance for organizations in cyber security decision making in the coming years [52].

Although there are plenty of data, they are often unstructured, unrelated or scattered all over the place, hence of limited use and need a better organization [50]. Furthermore, diverse infrastructure and OS call for OS independent procedures. ThreatRaptor [14] presents a framework aiming to conquer such challenges.

While IoCs cover malware's name, IP address, domain names and file hashes only, they ignore affected software names, higher level concepts and TTPs ¹⁴. In contrast to IP addresses that can change easily, TTPs are tied to attackers goals and hence harder to change.

ThreatRaptor continues the work done in SecurityKG and adds a query language created to allow the synthesis of OS independent queries. Fig. 8.6 [14] shows the architecture of ThreatRaptor.

¹³conditional random field

¹⁴Tactic, Technique and Procedure

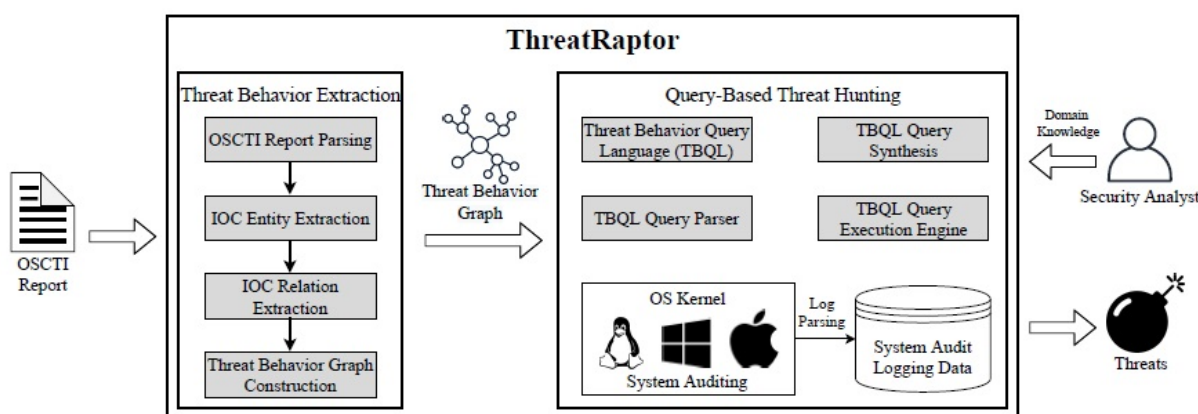


Figure 8.6: ThreatRaptor Architecture [14]

For evaluation, ThreatRaptor was applied to a number of attack cases including three multi-step attack cases where they have the ground truth system events. With knowledge extracted from OSCTI¹⁵ sources, they synthesized queries that were later applied to system audit logs of a test machine on which they constructed the 3 multi-step attack cases. The queries were able to find most of the malicious activities (precision¹⁶ 100%, recall¹⁷ 96.74 %, F1 score¹⁸ 98.34 %), proving the high performance of the (*automated*) threat behavior extraction pipeline. See fig. 8.7 [14] and fig. 8.8 [14].

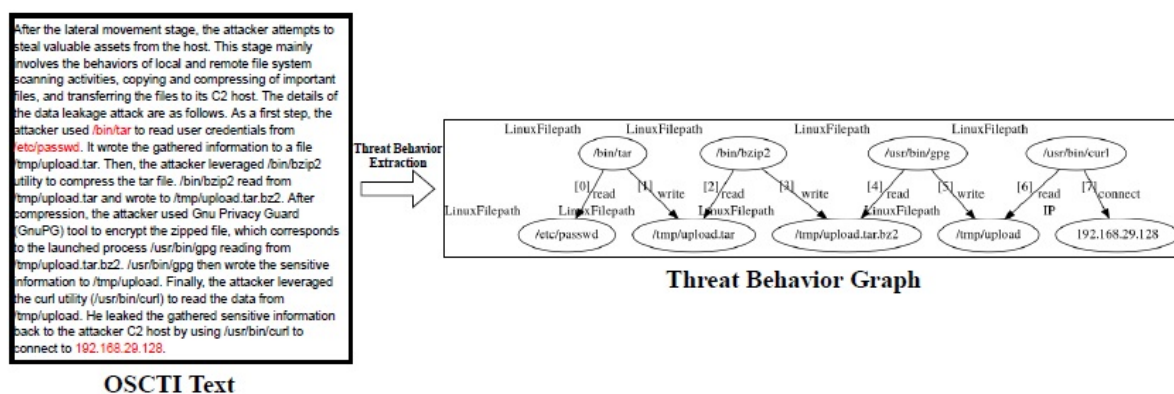


Figure 8.7: Dataleak Attack Case [14]

¹⁵Open Source Cyber Threat Intelligence

¹⁶precision = true positive / (true positive + false positive)

¹⁷recall = true positive / (true positive + false negative)

¹⁸F1 score = harmonic sum of precision and recall

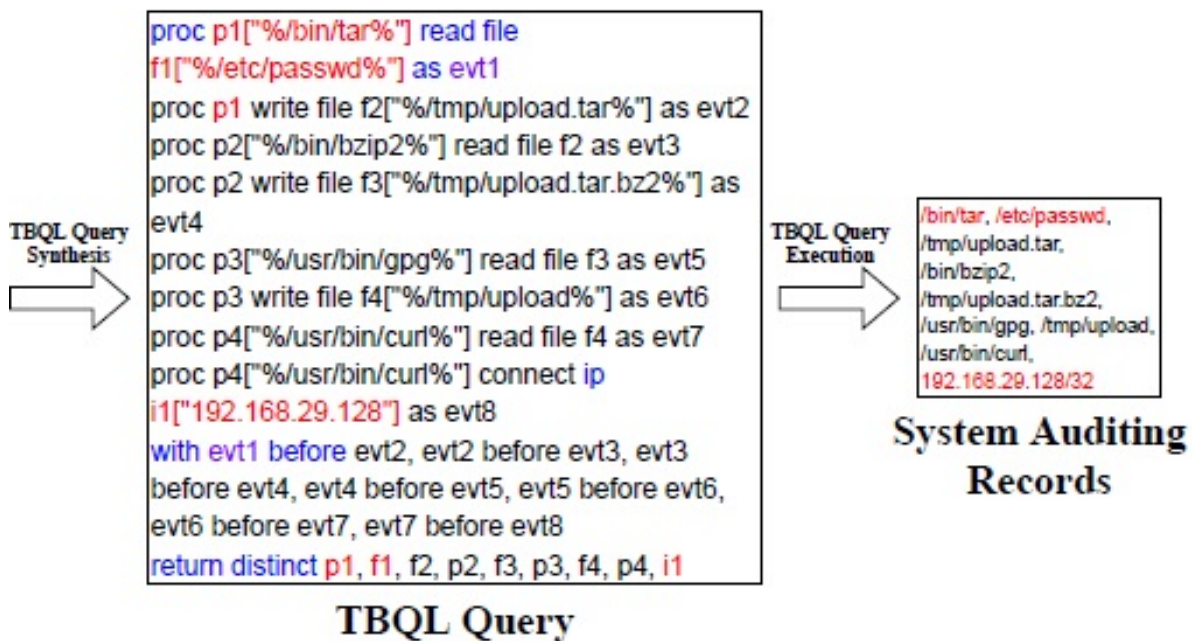


Figure 8.8: Dataleak Attack Case (cont.) [14]

8.4 Case Study Proactive Cybersecurity

This case study aims to find out whether a large Swiss bank uses the proactive methods presented in this paper. The bank involved will be referred to as the "bank" in the following, as it wishes to remain anonymous for confidential and private reasons. In order to obtain this information, two departments were interviewed. The "Security Operations" department is responsible for cyber defence and IT service continuity management. They know the current threat situation and the relevant cyber actors and their techniques. They also detect cyberattacks on the bank, analyse them and initiate measures. Attack simulations are another part of their field of activity. This department was chosen for the case study because its field of activity is suitable for achieving the objective of the case study. The second department involved is simply called "Security". It coordinates security-related topics and functions, as well as the subordinate specialised units. One of these specialised units is the Security Awareness Unit. This department was chosen because of its expertise in awareness training. The two departments mentioned, "Security Operations" and "Security", work together in day-to-day business.

Awareness training is compulsory for all employees of the bank. This takes the form of an online learning module followed by a test. Furthermore, various forms of phishing simulations are carried out during the year, in which the entire workforce is made aware of realistic phishing attacks. In these phishing simulations, it can be evaluated afterwards what percentage of the employees would have reacted correctly. In addition to the classic phishing e-mails, there are also e-mails without harmful content in circulation. These want to trick the recipient into making a value transaction, for example. Such e-mails are difficult for a security infrastructure to recognise, which demands even greater attention from the recipient.

Each application of the bank has its own **penetration testing** team. Some tests are also carried out by external people. However, some people work independently of these teams. Such a "Red Teamer" works on about 3 projects per year. His goal is to penetrate as deeply as possible into the bank through several stations. In one project he proceeded as follows: he sent e-mails to several employees, then he wanted to manifest himself on their laptops, so he wanted to get to the server and then to the Active Directory. In the Active Directory, the access rights of all employees are managed. If he could create a new profile there with as many rights as possible, he would have achieved his goal.

The Security Operation Team is responsible for **monitoring**. In their office, large screens are mounted on the wall so that all team members always have an overview of the current threat situation. The team works with various tools, including a SOAR tool (Security Orchestration, Automation and Response). However, this is used for reactive cybersecurity. A proactive method that is carried out is the monitoring described in Chapter 8.3.2.1 "Ad hoc techniques". They use VirusTotal to analyse process names and hashes. For example, if a vulnerability becomes known, they check the hashes of the relevant period to see if the bank is affected.

8.5 Summary

In recent years, an increase in cyberattacks has been observed. As this trend is expected to continue, companies and individuals should address the issue of cybersecurity. An optimal security strategy consists of reactive and proactive methods. Reactive methods are about reacting properly once a cyberattack has occurred. Examples are response and forensic investigation. Proactive methods, on the other hand, are used before vulnerabilities can be exploited by attackers. Penetration testing, awareness training, honeypot/decryption and moving target defence (MTD) are all methods that are applied proactively.

Penetration testing is an attempt to recreate a real cyberattack. The aim of this is to improve the security of a system and identify possible vulnerabilities. Awareness training starts at a different point. Through targeted training of users, their handling of cyber risks is to be steered in the desired direction. Awareness training is a useful method to reduce or even prevent the success of phishing attacks. The honeypot method is used to spy on attackers and gather useful information about their actions. The attackers are lured by a bait of false data. In the fourth proactive strategy, Moving Target Defence, the configurations of possible targets of attacks are constantly changed so that an attack is made more difficult.

This work has shown how threat hunting can be a necessary and useful proactive process in every organization. It needs data, expertise and tools to be fruitful. Currently it is in many circumstances still largely an adhoc, unstructured, manual process that should be automated. MITRE's ATT&CK is a framework to help organizations understand various types of attack tactics and techniques. It also provides recommendations for mitigation and detection. Several efforts and results in the machine learning and NLP areas for automating threat hunting have also been shown, each with its own strengths and weaknesses.

Bibliography

- [1] European Union Agency for Cybersecurity (ENISA): *ENISA Threat Landscape*, October 2021. <https://www.enisa.europa.eu/publications/enisa-threat-landscape-2021>, last visit March 4, 2022
- [2] Allianz: *Allianz Risk Barometer 2022*, Pressemitteilung, January 2022. <https://www.agcs.allianz.com/news-and-insights/news/allianz-risk-barometer-2022-press-de.html>, last visit March 4, 2022
- [3] Steve Morgan: *Cybercrime To Cost The World USD 10.5 Trillion Annually By 2025*, Cybersecurity Ventures, Special Report: Cyberwarfare In The C-Suite, November 2020. <https://cybersecurityventures.com/cybercrime-damage-costs-10-trillion-by-2025/>, last visit March 4, 2022
- [4] Dr. Erdal Ozkaya: *Cybersecurity : the beginner's guide : a comprehensive guide to getting started in cybersecurity*, Packt Publishing, May 2019.
- [5] Imunify360: *Proactive vs. Reactive Security: 5 Tips for Proactive Cyber Security*, August 2021. <https://blog.imunify360.com/proactive-vs.-reactive-security-5-tips-for-proactive-cyber-security>, last visit March 4, 2022
- [6] Melad Mohamed Al-Daeef, Nurlida Basir, Madihah Mohd Saudi: *Security Awareness Training: A Review*, Proceedings of the World Congress on Engineering 2017, Vol I , July 2017. <https://oarep.usim.edu.my/jspui/handle/123456789/1880>, last visit May 28, 2022
- [7] mcafee: *What Is Cyber Threat Hunting?* <https://www.mcafee.com/enterprise/de-de/security-awareness/operations/what-is-cyber-threat-hunting.html>, last visit March 4, 2022
- [8] Fidelis Cybersecurity: *What Is a Honeypot* <https://fidelissecurity.com/resources/edu/network-security/honeypots/#definition>, last visit March 4, 2022
- [9] Rob van Os, Marcus Bakker, Ruben Bouman, Martijn Docters van Leeuwen, Marco van der Kraan, Wesley Mentges, Armand Piers: *Threat Hunting Methodology*, A joint threat hunting methodology from the Dutch financial sector, December 2018. <https://www.betaalvereniging.nl/wp-content/uploads/DEF-TaHiTI-Threat-Hunting-Methodology.pdf>, last visit March 17, 2022.

- [10] Bundesamt für Sicherheit in der Informationstechnik: *Studie Durchführungskonzept für Penetrationstests*, November 2003. <https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/Studien/Penetrationstest/penetrationstest.html>, last visit April 14, 2022.
- [11] Federal Bureau of Investigation: *Internet Crime Report 2021*, 2021. https://www.ic3.gov/Media/PDF/AnnualReport/2021_IC3Report.pdf, last visit April 20, 2022
- [12] Jianjun Zheng, Akbar Siami Namin: *A Survey on the Moving Target Defense Strategies: An Architectural Perspective*, JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY 34 1 : 207–233, January 2019. <https://link.springer.com/content/pdf/10.1007/s11390-019-1906-z.pdf>, last visit April 22, 2022
- [13] Antonatos S, Akritidis P, Markatos E P, Anagnostakis K G: *Defending against hitlist worms using network address space randomization*, Proc. the 2005 ACM Workshop on Rapid Malcode pp.30–40, November 2005.
- [14] Peng Gao, Fei Shao, Xiaoyuan Liu, Xusheng Xiao, Zheng Qin, Fengyuan Xu, Prateek Mittal, Sanjeev Kulkarni, Dawn Song : *Enabling Efficient Cyber Threat Hunting With Cyber Threat Intelligence*, 2021 IEEE 37th International Conference on Data Engineering (ICDE).
- [15] Peng Gao, Xiaoyuan Liu, Edward Choi, Bhavna Soman, Chinmaya Mishra, Kate Farr, Dawn Song: *A System for Automated Open- Source Threat Intelligence Gathering and Management*. In Proceedings of the 2021 International Conference on Management of Data SIGMOD 21.
- [16] Arvind Kok, Ivana Ilic Mestric, Giavid Valiyev, Michael Street: *Cyber Threat Prediction with Machine Learning*. Information & Security: An International Journal 47, no. 2 2020: 203-220. <https://doi.org/10.11610/isij.4714>, last visit 10 April 2022.
- [17] Michael Berthold et al.,: *KNIME: The Konstanz Information Miner*. in Studies in Classification, Data Analysis, and Knowledge Organization (Springer, 2007), 319-326.
- [18] Prakruthi Karuna, Erik Hemberg, Una-May O'Reilly, Nick Rutar : *Automating Cyber Threat Hunting Using NLP, Automated Query Generation, and Genetic Perturbation*. CoRR 2021. arxiv.2104.11576v1.
- [19] Salvatore J. Stolfo, Wei Fan, Wenke Lee, Andreas Prodromidis, Philip K. Chan : *Cost-based Modeling for Fraud and Intrusion Detection: Results from the JAM Project*. Proceedings DARPA Information Survivability Conference and Exposition. DISCEX'00. 25-27 Jan. 2000.
- [20] Hunt Evil : *Your Practical Guide to Threat Hunting*. no date. <https://www.threathunting.net/files/huntpedia.pdf>, last visit 10 April 2022.
- [21] Randy Franklin Smith: *Catch Malware Hiding in WMI with Sysmon*. no date. <https://www.netsurion.com/articles/catch-malware-hiding-in-wmi-with-sysmon>, last visit 22 April 2022.

- [22] Srikrupa Srivatsan : *DNS over HTTPS misuse or abuse: How to stay secure*. March 11, 2020. <https://www.helpnetsecurity.com/2020/03/11/dns-over-http-abuse/>, last visit 22 April 2022.
- [23] Logrhythm : *Detecting DNS Tunneling*. December 17, 2014. <https://logrhythm.com/blog/detecting-dns-tunneling/>, last visit 22 April 2022.
- [24] Abhisek Divekar, Meet Parekh, Vaibhav Savla, Rudra Mishra, Mahesh Shirole : *Benchmarking datasets for Anomaly-based Network Intrusion Detection*. 3rd IEEE International Conference on Computing, Communication and Security (ICCCS), 2018. arxiv.1811.05372.
- [25] Nour Moustafa, Jill Slay : *The evaluation of Network Anomaly Detection Systems: Statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set*. Information Security Journal: A Global Perspective Volume 25, 2016. Issue 1-3.
- [26] Y. Liao and V.R. Vemuri : *Use of k-nearest neighbor classifier for intrusion detection*. In Computers & Security, Elsevier, vol. 21, no. 5, 2002.
- [27] N.B. Amor, S. Benferhat and Z. Elouedi, : *Naïve Bayes vs. decision trees in intrusion detection systems*, Proc ACM Symp. Appl. Comput, pp. 420-424, 2004.
- [28] Md. Al Hasan, M. Nasser, B. Pal and S. Ahmad, *Support Vector Machine and Random Forest Modeling for Intrusion Detection System (IDS)*, Journal of Intelligent Learning Systems and Applications, 2014.
- [29] J. Zhang, M. Zulkernine and A. Haque, *Random-forests-based network intrusion detection systems*, IEEE Transactions on Systems Man and Cybernetics Part C (Applications and Reviews), vol. 38, no. 5, pp. 649-659, Sep. 2008.
- [30] *An Overview to Threat Hunting: 7 Common Hunts to Get Started*. November 20, 2018. <https://logrhythm.com/webcasts/an-overview-to-threat-hunting-7-common-hunts-to-help-get-started/>, last visit 10 April 2022.
- [31] Jian-Hua Li : *Cyber Security meets Artificial Intelligence : a survey*. Frontiers of Information Technology & Electronic Engineering volume 19, pages 1462-1474 (2018).
- [32] The MITRE Corporation, <https://attack.mitre.org>, last visit 10 April 2022.
- [33] Katie Nickels : *Getting Started with ATT&CK: Threat Intelligence*. Jun 10, 2019. <https://medium.com/mitre-attack-getting-started-with-attack-cti-4eb205be4b2f>, last visit 10 April 2022.
- [34] *Using ATT&CK to Advance Cyber Threat Intelligence - Part 1*. May 24, 2018. <https://medium.com/mitre-attack/using-att-ck-to-advance-cyber-threat-intelligence-part-1-c5ad14d59724>, last visit 10 April 2022.

- [35] *UEBA and the Mitre Att&ck Framework: Detect, Investigate, Respond*. 2019-7-20. <https://community.microfocus.com/cyberres/b/sws-22/posts/ueba-and-the-mitre-att-ck-framework-detect-investigate-respond>, last visit 10 April 2022.
- [36] Loui, E. and Reynolds, J. August 30, 2021. <https://www.crowdstrike.com/blog/carbon-spider-embraces-big-game-hunting-part-1/>, last visit 22 April 2022
- [37] Higgins, K. October 13, 2017. <https://www.darkreading.com/analytics/prolific-cybercrime-gang-favors-legit-login-credentials>, last visit 22 April 2022.
- [38] Anomali Labs. December 6, 2018. <https://www.anomali.com/blog/pulling-linux-rabbit-rabbit-malware-out-of-a-hat>, last visit 22 April 2022.
- [39] CyberRes : ArcSight Intelligence and MITRE ATT&CK. Flyer. 2021. <https://www.microfocus.com/media/flyer/intersec-ueba-and-mitre-attack-flyer.pdf>, last visit 23 April 2022.
- [40] Jonathan Trull : *Threat hunting: Part 1 Why your SOC needs a proactive hunting team*. March 10, 2020. <https://www.microsoft.com/security/blog/2020/03/10/threat-hunting-part-1-why-your-soc-needs-a-proactive-hunting-team/>, last visit 12 April 2022.
- [41] Masumi Arafune, Sidharth Rajalakshmi, Luigi Jaldon, Zahra Jadidi, Shantanu Pal, Ernest Foox, Nagarajan Venkatachalam : *Design and Development of Automated Threat Hunting in Industrial Control Systems*. 2022. CoRR. arxiv 2202.01543.
- [42] Abubhav Arora : *Using Machine Learning for Threat Detection*. June 10, 2020. <https://fidelissecurity.com/threatgeek/network-security/using-machine-learning-for-threat-detection/>, last visit 12 April 2022.
- [43] Omar M. K. Alhawi, James Baldwin, Ali Dehghantanha : *Leveraging Machine Learning Techniques for Windows Ransomware Network Traffic Detection*. 2018. pp 93-106. Springer International Publishing. https://rd.springer.com/chapter/10.1007/978-3-319-73951-9_5 last visit 12 April 2022.
- [44] Amir Namavar Jahromi, Sattar Hashemi, Ali Dehghantanha, Kim-Kwang Raymond-Choo, Hadis Karimipour, David Ellis Newton, Reza M.Parizi : *An improved two-hidden-layer extreme learning machine for malware hunting*. Computers & Security (89), 2020.
- [45] Suphannee Sivakorn, Jason Polakis, and Angelos D. Keromytis: *I 'm not a human: Breaking the Google reCAPTCHA*. Black Hat ASIA 2016. <https://www.blackhat.com/docs/asia-16/materials/asia-16-Sivakorn-Im-Not-a-Human-Breaking-the-Google-reCAPTCHA-wp.pdf>, last visit 17 April 2022.

- [46] Eddie Segal : *The Impact of AI on Cybersecurity*. no date. <https://www.computer.org/publications/tech-news/trends/the-impact-of-ai-on-cybersecurity>, last visit 12 April 2022.
- [47] Jaikumar Vijayan: *Using Natural Language Processing to Identify Malicious Domains*. March 16, 2015 <https://securityintelligence.com/news/using-natural-language-processing-identify-malicious-domains/>, last visit 11 April 2022.
- [48] Prabhat Semwal, Akansha Handa : *Cyber-Attack Detection in Cyber-Physical Systems Using Supervised Machine Learning*. Handbook of Big Data Analytics and Forensics pages 131–140. First online 01 January 2022.
- [49] Hamid Darabian, Ali Dehghantanha, Sattar Hashemi, Mohammad Taheri, Amin Azmoodeh, Sajad Homayoun, Kim-Kwang Raymond Choo, Reza M. Parizi : *A multiview learning method for malware threat hunting: windows, IoT and android as case studies*. World Wide Web volume 23, pages 1241–1260 (2020).
- [50] Vector Guo Li, Matthew Dunn, Paul Pearce, Damon McCoy, Geoffrey M Voelker, and Stefan Savage. *Reading the tea leaves: A comparative analysis of threat intelligence*. Proceedings of the 28th USENIX Security Symposium, August 14–19, 2019. pages 851-867.
- [51] Alexander J Ratner, Christopher De Sa, SenWu, Daniel Selsam, and Christopher Ré. *Data programming: Creating large training sets, quickly*. <https://doi.org/10.48550/arxiv.1605.07723>, 2016. last visit 12 April 2022.
- [52] *Threat intelligence market analysis by solution, by services, by deployment, by application and segment forecast, 2018–2025*. no date. <https://www.grandviewresearch.com/industry-analysis/threat-intelligence-market>, last visit 12 April 2022.
- [53] *The Linux Audit Framework*, <https://github.com/linux-audit/>, last visit 12 April 2022.
- [54] <https://attack.mitre.org/matrices/enterprise/>, last visit 25 May 2022.
- [55] <https://attack.mitre.org/tactics/TA0003/>, last visit 25 May 2022.
- [56] <https://attack.mitre.org/techniques/T1078/>, last visit 25 May 2022.
- [57] Kurt Baker: *What is Cyber Threat Intelligence?*. March 17, 2022. <https://www.crowdstrike.com/cybersecurity-101/threat-intelligence/>, last visit 31 May 2022.

