NATIONAL AND KAPODISTRIAN UNIVERSITY OF ATHENS

DEPARTMENT OF PHYSICS

SECTION OF ELECTRONIC PHYSICS AND SYSTEMS

Comprehensive analysis of 5G NR: Overview of the Physical Layer, Network Security, and its vulnerabilities to Jamming Attacks

LOUKAS DROSOS

Bachelor's thesis

AM: 1110201900062

Supervisor Professor:

Anna Tzanakaki

Associate Professor

Department of Physics, N.K.U.A.

ATHENS 2024

Thankings

I would like to thank Mrs. Anna Tzanakaki, Associate Professor of the Department of Physics of the National and Kapodistrian University of Athens, who gave me the opportunity to work on the specific subject of Wireless Networks and with her help and full guidance helped me to prepare this thesis.

Also, I would like to thank Mr. Manolopoulos Alexandros-Ioannis, PhD Candidate of the Department of Physics of the National and Kapodistrian University of Athens, who helped me and guided me in all the procedures required for the completion of this thesis.

Table of Contents

List of Figures

1. 5G Overview

2. 5G NR

3. 5G Core Network

4. 5G NR Physical Layer

5. 5G System Security

6. Jamming Attacks In 5G NR

Abstract

5G is the fifth generation of wireless cellular networks and enables new use cases, such as high user mobility and connectivity of a huge number of devices to the network, while also promising higher data rates, lower latency, and enhanced reliability. At the same time, 5G made improvements in the security and privacy aspects over 4G and predecessors. However, the 5G standard is not without any security vulnerabilities. In fact, the security threat landscape of 5G is expanding due to the increased types of services 5G offers and the large number of devices the network supports.

The objective of this thesis is the study of security issues in 5G, emphasizing on the 5G NR physical (PHY) layer and its vulnerabilities to jamming attacks. The thesis includes an overview of the 5G network and its security weaknesses, 5G NR and the 5G core network and an analysis of the 5G NR PHY layer. Moreover, the thesis explores the different types of jamming attacks and examines mitigation techniques that can make the PHY layer more resilient to attacks in the next generation of wireless networks.

Περίληψη

Η επικοινωνία αποτελεί ένα πολύ σημαντικό κομμάτι της ανθρώπινης φύσης. Σήμερα, ένα μεγάλο μέρος της καθημερινής μας επικοινωνίας έχει ψηφιοποιηθεί και περιλαμβάνει την επικοινωνία ανθρώπου προς μηχανή αλλά και μηχανής προς μηχανή. Η ραγδαία αύξηση στη χρήση των δικτύων επικοινωνίας τα τελευταία χρόνια έχει θέσει ορισμένα πρότυπα για αυτά τα δίκτυα, όπως η χαμηλή χρονική καθυστέρηση στην επικοινωνία, οι υψηλές ταχύτητες μετάδοσης δεδομένων, η αξιοπιστία κάλυψης και η κινητικότητα κατά την ασύρματη επικοινωνία. Παρόλο που η προηγούμενη γενιά ασυρμάτων δικτύων κινητής επικοινωνίας, 4G LTE, έφερε σημαντικές βελτιώσεις στην τεχνολογία ασυρμάτων δικτύων, δεν μπόρεσε ακόμη να ικανοποιήσει όλες τις απαιτούμενες υπηρεσίες.

Το 5G είναι η πέμπτη γενιά ασύρματων δικτύων κινητής επικοινωνίας και σχεδιάστηκε για να ανταποκριθεί σε αυτές τις απαιτήσεις υπηρεσιών, επιτρέποντας νέα είδη χρήσεων, όπως είναι η υψηλή κινητικότητα χρηστών και η συνδεσιμότητα ενός μεγάλου αριθμού συσκευών στο δίκτυο, ενώ υπόσχεται επίσης υψηλότερες σε ταχύτητα μεταδώσεις δεδομένων, χαμηλότερη χρονική καθυστέρηση και βελτιωμένη αξιοπιστία. Ταυτόχρονα, το 5G έχει κάνει βελτιώσεις στους τομείς της ασφάλειας και της ιδιωτικότητας σε σχέση με το 4G LTE και τις προηγούμενες γενιές ασύρματων δικτύων. Η χρήση του 5G οδήγησε σε καινοτομίες που άλλαξαν την καθημερινότητα μας, όπως είναι η βιομηχανική αυτοματοποίηση, οι έξυπνοι οικισμοί και πόλεις, οι επικοινωνίες οχημάτων προς οχήματα, που συνδέουν την κοινωνία συνολικά και θα συνεχίσουν να το κάνουν και στο μέλλον. Ωστόσο, αυτές οι νέες τεχνολογίες και περιπτώσεις χρήσης του 5G δεν είναι χωρίς τις ανησυχίες τους σχετικά με την ασφάλεια. Πράγματι, ο κίνδυνος ασφαλείας του 5G επεκτείνεται λόγω της μεγάλου σε αριθμό είδους υπηρεσιών που προσφέρει, καθώς και εξαιτίας του μεγάλου αριθμού συσκευών που υποστηρίζει το δίκτυο, καθιστώντας απαραίτητες λειτουργίες ασφαλείας του δικτύου, όπως η πιστοποίηση, η ακεραιότητα και η εμπιστευτικότητα, δύσκολες να εφαρμοστούν.

Σκοπός της εργασίας είναι η μελέτη των θεμάτων ασφάλειας στο δίκτυο 5G, με έμφαση στο φυσικό επίπεδο (Physical Layer ή PHY) του 5G NR και τις αδυναμίες του σε επιθέσεις παρεμβολής. Συγκεκριμένα, το Κεφάλαιο 1 παρέχει με μια επισκόπηση του δικτύου 5G, η οποία περιλαμβάνει την εξέλιξη των δικτύων κινητής επικοινωνίας, το βασικό μοντέλου του δικτύου 5G, τις κύριες περιπτώσεις χρήσης του και τις αδυναμίες ασφάλειας του. Το Κεφάλαιο 2 συνεχίζει παρουσιάζοντας το πρότυπο 5G NR και την κατανομή του φάσματος του. Έπειτα, το Κεφάλαιο 3 εξετάζει συνοπτικά το δίκτυο 5G Core (5GC) και τις βασικές λειτουργίες του. Στο Κεφάλαιο 4 αναλύεται λεπτομερώς το φυσικό επίπεδο του 5G NR, εστιάζοντας στην διαμόρφωση συχνοτήτων, στην Ορθογώνια Πολυπλεξία Διαίρεσης Συχνοτήτων (OFDM), στις κυματομορφές πολλαπλών φορέων, στην αριθμολογία τους και στα φυσικά κανάλια και σήματα της δομής του φυσικού επιπέδου του 5G NR. Το Κεφάλαιο 4 αναλύει επίσης τεχνικές πολλαπλών κεραιών, αλλά και τεχνικές κωδικοποίησης καναλιού. Το Κεφάλαιο 5 εξερευνά τις

ασφαλείς πτυχές των συστημάτων 5G, μελετώντας αρχικά την εξέλιξη της ασφάλειας στις προηγούμενες γενιές ασυρμάτων δικτύων μέχρι το δίκτυο 5G, ενώ ταυτόχρονα επισημαίνει τις πιθανές απειλές των δικτύων 5G. Το Κεφάλαιο 6 διερευνά τους διάφορους τύπους παρεμβολέων και τεχνικών παρεμβολής, καθώς και τις αδυναμίες του φυσικού επιπέδου 5G NR σε αυτές τις τεχνικές. Επιπλέον, συγκρίνει την επίδραση των παρεμβολών στο 5G NR σε σχέση με τα δίκτυα 4G LTE, καθώς και στις λειτουργίες 5G SA έναντι NSA. Στο τέλος του Κεφαλαίου 6, εξετάζονται οι τεχνικές αντιμετώπισης για τον εντοπισμό και τον μετριασμό των επιθέσεων ανακοπής που μπορούν να κάνουν πιο ανθεκτικό σε επιθέσεις παρεμβολής το φυσικό επίπεδο του 5G NR αλλά και την επόμενη γενιά ασύρματων δικτύων.

Introduction

Communication has always played an important role in humanity's history. Nowadays, a big part of our everyday communication has been digitalized, including human-to-machine and machine-to-machine communication. The huge increase in usage of communication networks in recent years have set some standards for these networks, such as low latency communication, high data rate transfers, energy and spectral efficiency, and coverage reliability and mobility during wireless communication. Even thought 4G LTE made big improvements in the network technology to provide an enhanced mobile-broadband experience, it still couldn't satisfy all the required services. The fifth and current generation of mobile networks, 5G, is the successor to 4G LTE and has been designed to meet these service requirements while also introducing new features such as network slicing, software-defined networking (SDN), network function virtualization (NFV) and mobile edge computing (MEC). Furthermore, 5G New Radio (NR) is the radio access technology (RAT) that is developed for 5G networks, and provides the radio standard for the deployment of 5G [1], [2].

Alongside voice and data communications, 5G also introduced a multitude of new use cases and industry applications, such as vehicle-to-vehicle communications, health services, industrial automation, smart homes and cities, that connect society as a whole and will continue to in the future. However, these new technologies and use cases in the 5G architecture are not without their security concerns. The broadcast nature of 5G, which is limited when it comes to coverage bandwidth, make necessary security features like authentication, integrity and confidentiality challenging to implement. Also, wireless cellular networks face various problems when it comes to the security of the physical layer (PHY) and its vulnerabilities to possible attacks [3].

The thesis is organized as following: Chapter 1 of this thesis provides an overview of 5G, starting with the evolution of mobile communication, 5G's use cases and its security vulnerabilities. Then, Chapter 2 introduces the 5G NR standard and its spectrum allocation. Chapter 3 provides an overview of the 5G Core network and its main functions. In Chapter 4, the physical layer of 5G NR is examined in detail, focusing on modulation, OFDM, waveform, multiple antennas, and channel coding techniques. Chapter 5 explores the security aspects of 5G systems, tracing the evolution of security in wireless networks and detailing the security services and improvements in 5G over its predecessors, while at the same time analyzing the threat landscape of 5G networks. In Chapter 6 the different types of jammers and jamming techniques are investigated alongside the vulnerabilities of the 5G NR physical layer to jamming attacks. Additionally, it examines the impact of jamming in 5G NR compared to 4G LTE networks, as well as in 5G SA vs NSA modes. At the end Chapter 6, anti-jamming techniques for detection and mitigation of jamming attacks in 5G NR are explored.

1. 5G Overview

Wireless access technology has been an important part of mobile communication systems, allowing devices to connect with each other, as well as with radio base stations. The fifth generation of mobile communication introduced new technologies and use cases that have impacted the day to day lives of people and gave an opportunity to industries and corporations to evolve their business models based on these technological advancements in communication technologies. For these reasons, it is important to understand the changes the world has experienced when it comes to the evolution of communication [4].

1.1 Evolution of Mobile Communication

The first ever mobile telephony service was operated in 1947 in the USA. The equipment used for this service was installed in a vehicle due to its weight and its large power consumption. Today, after more than 30 years of evolution in mobile communication, technology has shifted from analog to digital communication, and has transitioned from voice to high-speed data transmission [5].

The first generation of mobile communication, 1G, emerged around 1980. It was based on analog communication and was limited to just voice transmission, it did, however, make mobile communication accessible to the masses. The main technologies of 1G were AMPS (Advanced Mobile Phone System) and NMT (Nordic Mobile Telephony), which were developed in the USA and Scandinavia respectively [4].

The second generation of mobile communication was introduced in the early 1990s and replaced the analog communication schemes of 1G with new digital ones that were transmitted on the radio link. Digital transmission in 2G allowed for limited data services alongside the usual voice services, providing the ability to users to send text messages via SMS (Short Message Service) and even e-mails. 2G networks introduced some innovative technologies for mobile communications, such as Code Division Multiple Access (CDMA) and, most importantly, Global System for Mobile Communications (GSM) was introduced in Europe and Digital-AMPS in the USA [4], [6].

In the early 2000s, the third generation of mobile communication, 3G, was first introduced and brought along innovative features such as multimedia message service (MMS), video streaming and HSPA (High-Speed Packet Access), which allowed fast wireless Internet access. At the same time, 3GPP introduced a universal mobile telecommunications system (UMTS) that was based on wideband CDMA (WCDMA) technology [4], [5].

In 2009, the fourth generation era of mobile communication, 4G and then LTE (Long Term Evolution) were introduced, while frequency schemes transitioned From Code Division Multiplexing to OFDM (Orthogonal Frequency Division Multiplexing). LTE introduces new features that ensured a higher level of mobile-broadband experience for users alongside higher data rates. At the same time, LTE introduced services such as

Multimedia Messaging Service (MMS), Digital Video Broadcasting (DVB), High-Definition TV content and mobile TV [4], [7].

Since its initial unveiling, LTE has progressed through various developmental phases, enhancing its performance, and expanding its capabilities. LTE's evolution has not only introduced new features, but it has also expanded the range of use cases applicable to LTE. Notably, important steps have been taken to enable cost-effective devices with extended battery life aligning with the requirements of massive Machine Type Communication (MTC) applications, while simultaneously reducing the LTE air-interface latency. Through these evolutionary steps, the development of LTE can accommodate a diverse array of 5G use cases. It is thus essential to view the evolution of LTE not merely as a distinct radio-access technology, but as an important part of the overall 5G radio-access solution [4].

The research and conceptualization of 5G wireless access technology, officially named 5G NR since 2016, started many years ago with innovative applications and business opportunities in mind. The research endeavors resulted in the establishment of 5G test-beds within academic institutions and industry settings. Similar to its predecessors in cellular technology, the evolution of 5G represents a globally synchronized initiative, dealing with the fresh spectrum allocations at both global and regional scales for the advancement of 5G and aligning with the global standardization of 5G NR in 3GPP based on the 5G requirements defined by International Telecommunications Union (ITU). To give a global picture of the 5G developments, we will discuss 5G's system model, its major use cases and the new technologies that it integrates [5].

1.2 5G Standardization

The operation of mobile devices and services on a global level is allowed thanks to certain agreements on multinational specifications and standards in the realm of cellular communication [4].

The Third-Generation Partnership Project (3GPP) is an organization that was started in 1998 and produces technical specifications that become standards for mobile communication. The members of 3GPP are ARIB, ATIS, CCSA, ETSI, TSDSI, TTA, and TTC, which are development organizations from Asia, Europe and North America and their original task was to develop globally applicable specifications for 3G networks. Nowadays, 3GPP is also responsible for the development and maintenance of specifications for 2G GSM, 3G WCDMA/HSPA, 4G LTE, and 5G NR/LTE evolution [5].

The International Telecommunication Union (ITU) is an agency of the United Nations that was established in 1865 and specializes in topics regarding information and communication technologies. The ITU Radiocommunication Sector (ITU-R) is a sector of ITU that manages the International Radio Frequency (RF) spectrum and creates global standards for mobile communication from given technical specifications made by organizations such as 3GPP, while also including countries that are not a part of the standardization bodies in 3GPP. The ITU-R is also responsible for spectrum allocation

for the International Mobile Telecommunications (IMT) systems. The IMT systems are serve as the requirements issued by the ITU-R for each generation of mobile networks, from 3G and beyond [5], [8].

Each new generation of mobile communication networks contains new features while also improving the ones from previous generations. The 3GPP specifications for each generation are split into releases, with each release having each own complete set of specifications. Every release has some new features, but also includes all the components required for a cellular network. Once the specifications for a release are completed, they are ready for implementation and cannot be altered except in case of some necessary adjustments. New features will have to be implemented in the next specification release. New releases are worked on before the completion of the current release for time efficiency. Each release needs to be compatible with previous releases so a user equipment (UE) that was developed during a specific release can work with cells implemented in previous releases [5].

1.3 5G System Model

A 5G system consists of various components that deliver the data and voice services utilized by billions of people worldwide on a daily basis. Users interact with their user equipment (UE), such as smartphones or broadband routers, which they use to connect to a base station (BS) over a radio interface named 5G New Radio (NR). These base stations connect to the 5G Core Network which is responsible for managing user subscriptions and exchanging data between users and external networks like the Internet. Figure 1.1 depicts a basic illustration of a standard 5G system [9].

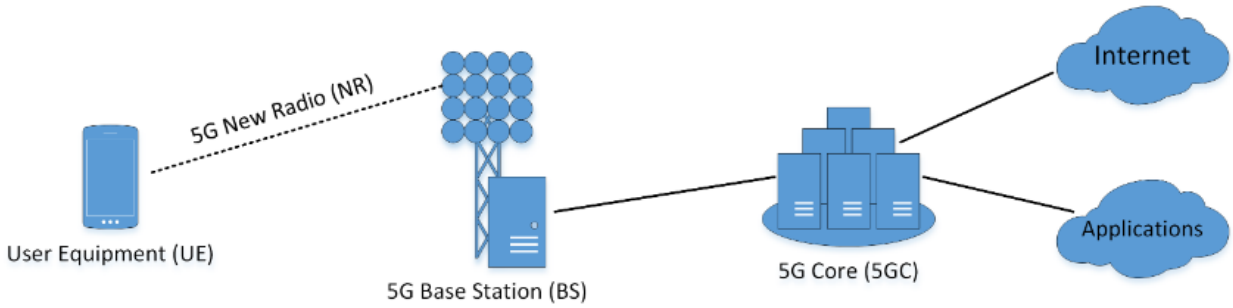


Figure 1.1 *The main components of a 5G system* [9].

5G systems can also include Multi-Access Edge Computing (MEC) and Network Slicing. Due to the extensive adoption of virtualization in 5G and the importance of secure network orchestration and management, these subsystems are included in the complete 5G system model depicted in Figure 1.2. The major components of the 5G system are depicted as either a single subsystem configuration or a combination of component configurations. Each 5G subsystem also has a set of attributes, such as the architecture

of the system (Stand Alone or Non-Stand Alone), information about the RAN and whether the UE can be authenticated or centrally managed [10].

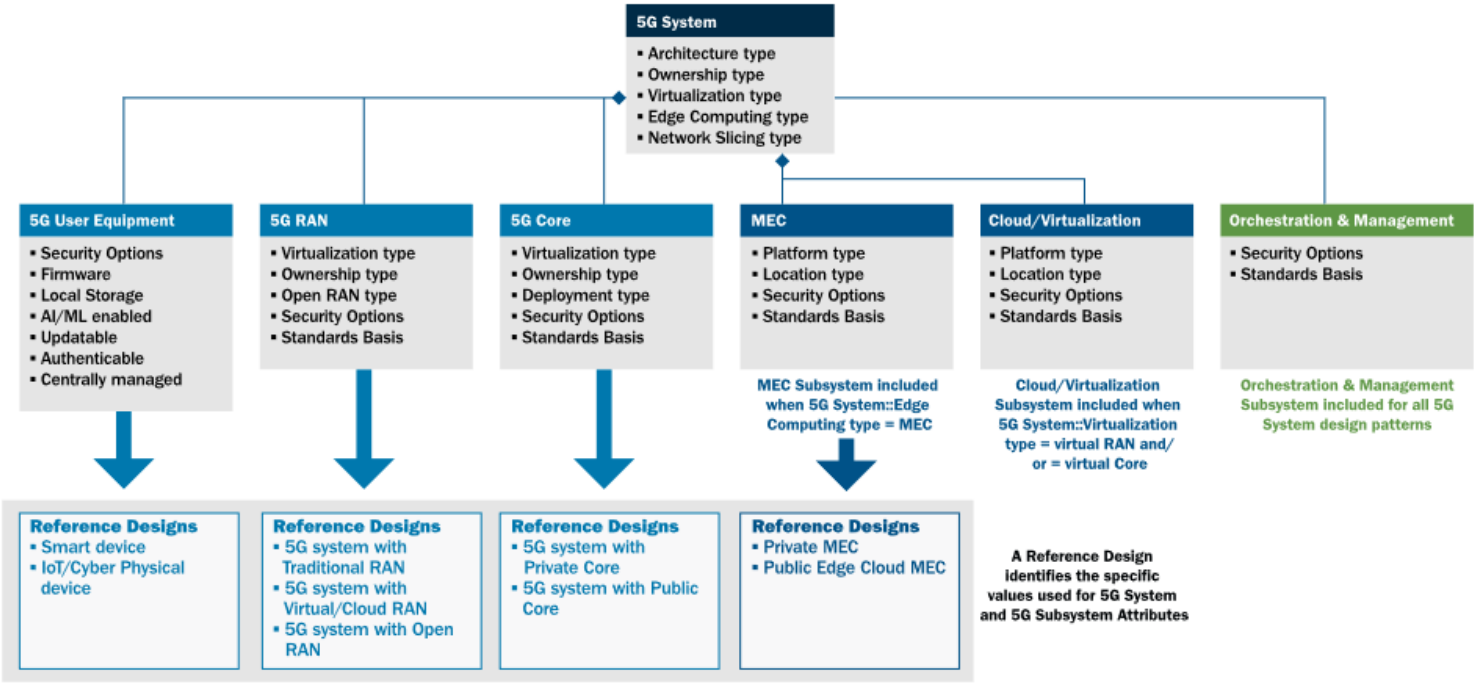


Figure 1.2 *Depiction of a complete 5G system* [10].

5G UEs give users access to the network services. 5G supports a plethora of device types, such as mobile devices, Internet of Things (IoT) and autonomous vehicles. These 5G UEs are composed of different hardware, software, interfaces such as air interfaces, local ports and sensors and can even include a subscriber identity module (SIM) [10].

The *gNB (Next Generation Node Base Station)* is a radio node that connects a 5G UE with a 5G NG core through the 5G NR air interface. It then splits the data into control and user plane segments and sends them to the various end points. The control plane contains the signaling traffic, while the user plane holds the data the user transmitted. Essentially, the gNB serves as the functional equivalent of a base station in a traditional cellular network and for simplicity reasons will be referred as base station (BS) [11], [12], [13].

The *5G RAN* facilitates all the logical access functions that connect a UE to the 5G system through the 5G NR air interface. The components of the 5G RAN are a central unit, a distributed unit, a radio unit, interfaces such as an air interface, midhaul and backhaul, management and possibly a RAN Intelligent Controller that controls the Open Radio Access Network configuration and optimization [10].

The *5G Core* is used to authenticate subscribers, to establish a secure connection to the network for end users and to provides access to the network's services. Apart from interfaces such as UE and RAN, the 5G Core encompasses a range of functions for the control plane, the user plane as well as network slicing [10].

MEC is a cloud service that handles specific tasks, such as routing, management and network capability exposure, in real time or near real time. Routing refers to packet forwarding within the MEC platform, as well as in the RAN and the core network. Network capability exposure refers to the exposure, in an authorized manner, of the radio network information service. Management in the MEC platform ensures the authentication and authorization of third-party applications. [4].

Cloud/Virtualization is an underlying infrastructure that supports cloud deployment virtual functions. It can be used either as a single virtualized platform or as multiple interconnected virtual platforms. This infrastructure consists of a virtualization layer and its hardware resources, resource allocation functions, a set of interfaces and other network functions [10].

Orchestration and Management includes all network management functions, such as configuration, performance and security functions, as well as any other functions that support deployment and management of virtualized infrastructure for the 5G RAN, 5G Core, MEC, and Cloud/Virtualization subsystems [10].

1.4 5G New Technologies

5G also introduced some new features and technologies. 5G utilizes *higher frequency bands*, including millimeter-wave spectrum up to 100 GHz, which allows for higher peak data rates and bigger network capacity. At the same time, 5G uses OFDM with adaptive numerology to configure the allocation of radio resources based on available spectrum and bandwidth [4].

Massive Multiple Input Multiple Output (*Massive MIMO*) is a new technology in 5G that consists of a large number of antennas at the base station which, combined with beamforming, increases the capacity of the network and provides lower latency during data transmission [4].

Another technology introduced in 5G is *Software Defined Networking* (*SDN*), which separates the control plane from the user plane. This separation allows for the dynamic control and management of the network via open interfaces [4].

Device-to-Device (*D2D*) communication is another 5G feature that enables the direct communication method between neighboring UEs to communicate with each other, without having to rely on intermediate network infrastructure. Because D2D communications don't use mobile networks communication they use less bandwidth resources and allow for more efficient spectrum allocation. Simultaneously, D2D reduces the traffic load of BSs, enhancing the stability of the network and giving cellular networks the ability to expand their services [14].

5G also introduced a new type of network called *heterogeneous network* (*HetNet*), which supports the integration and efficient operation of different types of radio nodes and technologies in a given area to counter high user density and poor signal reception for UEs in that area, while also providing seamless connectivity and service continuity.

These nodes may differ from each other in their spectrum usage, transmission power, coverage area and allocation and utilization of radio resources [15].

1.5 5G Use Cases

3GPP primarily concentrates on three major use cases for 5G NR: Ultra-Reliable and Low Latency Communication (uRLLC), Massive Machine Type Communication (mMTC), and Enhanced Mobile Broadband (eMBB) [16].

eMBB signifies a relatively straightforward progression from today's mobile broadband services, enabling even larger data volumes and further enhanced user experience. It addresses human-centric connectivity, including access to multimedia content and services, and helps those services control the increasing traffic volume that they generate by providing high data rates. The eMBB use case consist of small-area coverage scenarios, that require high user density, extremely high data rates and low mobility, but also extend to wide-area coverage scenarios, where the user density and data rates are lower, but the mobility is higher [4], [5].

mMTC corresponds to services that support massive device connectivity in IoT applications, for example, remote sensors, actuators, and monitoring of various equipment. Key requirements for such services include very low device cost and very low device energy consumption, allowing for very long device battery life of up to at least several years. Another distinctive feature of such devices is relatively small amount of transmitted data, so high rates are not required for this use cases. mMTC also combines random access and scheduling strategies when there are thousands of IoT devices waiting for access [5], [4], [17].

URLLC is primarily designed for machine-type communications (MTC). URLLC type-of-services are envisioned to require very low latency and have extremely high reliability. Examples that fall under this category include factory automation, traffic safety, Vehicle to Everything (V2X, communication between a vehicle and any entity that may affect or may be affected by the vehicle), robots' control and even remote medicine to name a few. To meet requirements set by this class of use cases, a specific set of 5G features has been specified. For instance, the support of mini-slot allows data transmission within a part of the slot, reducing transmission time on radio link between base station and UE (user equipment). Additionally, 5G imposes much stricter requirements on data processing time in both base station and UE, meaning the allowed time to process data is significantly shorter compared to LTE [5], [4], [18].

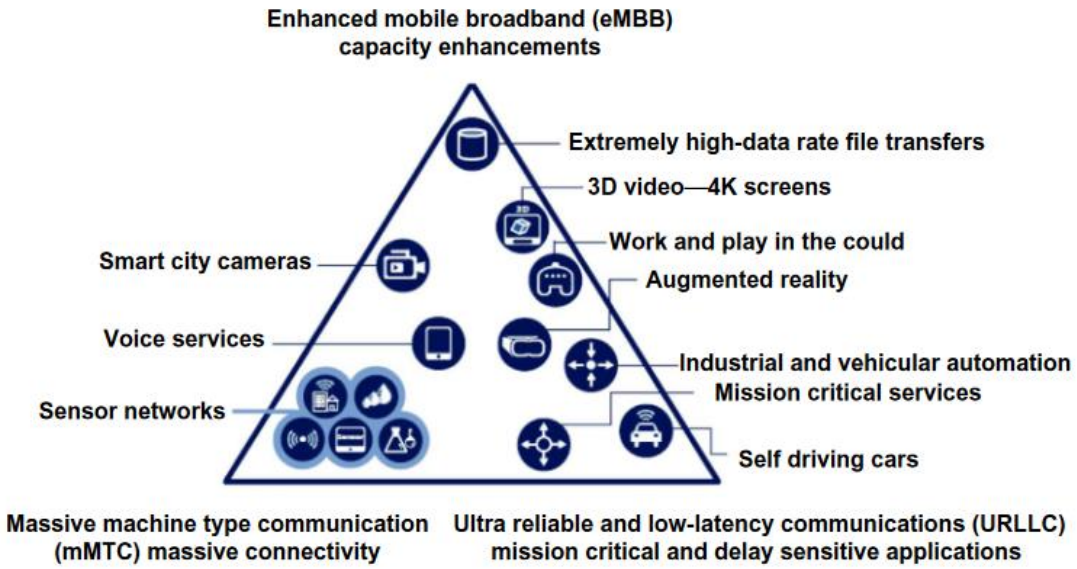


Figure 1.3 *The three major use cases for 5G* [4].

1.6 Vulnerabilities of 5G

5G's three main use cases, eMBB (Enhanced Mobile Broad Band), mMTC (Massive Machine Type Communication), and URLLC (Ultra Reliable Low Latency Communication), support applications such as Advanced IoT, industrial automatization and UHD video streaming, which require high bandwidth, low-latency communication, reliability, and low-cost implementation. These wider variety of applications alongside the vulnerabilities created in the 5G NR network due to the connection between devices are targets for malicious attacks that may have financial, political, or personal motives and can possibly interfere in military communications, damage the network's infrastructure, and even threaten personal safety [19], [7].

5G communication attacks can target the UE, the access networks and even the mobile operators core network. UE can be targeted for mobile malware attacks, which allow attackers to steal users' private data from their device. Attackers can also use the targeted devices to launch attacks against other users but can also target the cellular network itself. The access network can be targeted by attackers to gain access to valuable information such as packet scheduling and load balancing, by accessing the buffer status reports of the network's components, such as a BS. The attacker can then send false buffer status reports by pretending to be a legitimate UE to delay or even halt the access network's operations. Distributed Denial-of-Service (DDoS) attacks can be used against the core network to target and infect many users simultaneously, while attackers can also use SDN scanner attacks to passively collect network information such as the IP of the SDN controller and key network elements. Due to these threats, provision of an adequate security level in cellular networks is essential due to the continuously changing security threats in communication systems. Threats and attacks on prior cellular networks, such as 3G, 4G, LTE, etc., can still be applied in 5G systems [7].

2. 5G NR

The surging need for higher data rates and expanded bandwidth due to the increased number of users led to the deployment of the 5G network. 5G New Radio (5G NR), recognized as the global standardization of 5G, is in progress within the framework of the 3rd Generation Partnership Project (3GPP). 5G networks are the fundamental driver in terms of digitalization and advance communication for the industrial transformation [16].

2.1 5G NR Standardization

Despite LTE demonstrating considerable capability, there are requirements that LTE or its evolution cannot meet. LTE development started more than 10 years ago and since then the progress in technology has paved the way for more advanced technical solutions. 3GPP initiated the development of a new radio-access technology known as NR (New Radio) to try and meet these requirements and exploit the potential of new technologies. A workshop outlining the scope was held in the fall of 2015 and technical work started in the spring of 2016. The initial version of NR specifications was available by the end of 2017, while commercial requirements on early 5G deployments were already met in 2018. NR builds upon many structures and features of LTE, while at the same time being a distinct radio-access technology. That means that NR, unlike the LTE evolution, is not bound by the necessity for backward compatibility, and its requirements are also broader than those of LTE, resulting in a different set of technical solutions to be adopted [4].

The design principles for NR can be summarized as follows:

- NR design should be forward compatible, allowing for easy additions of features in coming releases.
- NR should offer ultra-lean transmission where transmissions are self-contained. NR should have a scalable numerology with a subcarrier spacing (SCS) in multiples of 15 kHz, consistent with LTE.
- NR should support dynamic TDD (Time Division Duplex), allowing any slot to be designated as a DL (Downlink) or UL (Uplink) slot on-the-fly.
- NR should support massive MIMO with hundreds of antenna elements, beamforming, MU-MIMO, and reciprocity-based operation.
- NR should be operable in both unlicensed spectrum and licensed spectrum [20].

Additionally, the requirements on NR release 18 include support for:

- *Further enhanced 5G performance,* which includes a study on network energy savings for NR, further NR coverage and mobility enhancements, NR MIMO evolution for downlink and uplink, enhancements of NR multicast and broadcast services and a study on expanded and improved NR positioning.

- *Flexible spectrum use,* which combines NR support for dedicated spectrum less than 5MHz for FR1, improvement of NR dynamic spectrum sharing (DSS), a study on the evolution of NR duplex operation and multi-carrier enhancements for NR.
- *Diverse 5G devices,* which consist of in-device co-existence (IDC) enhancements for NR, a study on low-power wake-up signal and receiver for NR, mobile terminated-small data transmission for NR, as well as NR side link evolution.
- *Evolved network topology,* which is made of a study on NR network-controlled repeaters, a study on enhancement for resiliency of BSs, and NR side link relay enhancements.
- *Data-driven and AI-powered 5G,* which involves artificial intelligence (AI)/machine learning (ML) for NG as well as a study on that topic, enhancements on NR QoE (Quality of Experience) management and optimization for diverse services [21].

2.2 Spectrum for 5G NR

Frequency bands for 5G New Radio (5G-NR) were established by the Third Generation Partnership Project (3GPP) based on the guidance from both the International Telecommunication Union (ITU) and the regional regulators, giving priority to the operators' commercial 5G plans. 5G's use cases require different frequency ranges, with each having different characteristics such as path loss and available spectrum bandwidth. For eMBB and mMTC services, the network needs to be accessible for as many users as possible, so good network coverage is crucial. Low frequencies (below 2 GHz) are widely used in 4G Long Term Evolution (LTE) and continue to be used in the 5G era, in fact they play an important factor when it comes to wide-area and indoor environments coverage. These services also require a substantial amount of spectrum to achieve high data rates, reliability with low latency, but the frequency ranges that have large channel bandwidth are located above 3 GHz and have increased path loss compared to lower frequencies. Therefore, 5G NR is designed by 3GPP to be flexible used over the full frequency range. Joint operation at both lower frequencies (e.g., below 6 GHz) and higher frequencies is supported, aiming to provide reliable coverage utilizing lower frequencies while also offering very high capacity and bitrates, when possible, especially in the millimeter-wave frequency range above 24 GHz [22], [5].

The frequency spectrum used in the 5G NR is divided into frequency range 1 (FR1), known as the sub-6 frequency range in the 5G context even though it spans from 410 to 7125 MHz, and frequency range 2 (FR2), which is often called the millimeter wave frequency range (mmWave) and covers a frequency interval of 24.25 to 71 GHz as per the latest published version (Rel. 18) of 5G NR. These frequency ranges may see extensions or additions with new ranges in future releases by 3GPP. 5G NR can be deployed in existing International Mobile Telecommunications (IMT) bands used by 3G UTRA (Universal Terrestrial Radio Access) and 4G LTE, but it also extends to the new bands defined for IMT at World Radiocommunication Conference 2019 (WRC-19) and may incorporate bands identified in future WRC or by regional bodies [9], [23], [4].

Frequency ranges are further sub-divided into distinct frequency bands with defined portions the frequency spectrum for the uplink and downlink, each one named by a unique identifier in the form "n" followed by a number. Table 2.1 below provides an overview of some NR frequency bands defined by 3GPP. The table gives a descriptive frequency, precise frequency ranges and duplex mode for each frequency band. The descriptive frequency acts as an additional name for the frequency band. The duplex mode dictates how the mobile network manages two-way communication in the uplink and downlink. 5G NR supports two duplex modes: frequency-division duplexing (FDD) and time-division duplexing (TDD). In FDD the uplink and downlink transmission occur simultaneously using different frequencies, requiring separate frequency ranges for each transmission. On the other hand, in TDD the uplink and downlink transmission occur at different times but on the same frequency, requiring only one specific frequency range for the frequency bands in TDD mode [9].

Frequency range	Frequency band	Descriptive frequency	UL frequency range [MHz]	DL frequency range [MHz]	Duplex
FR1 below 1 GHz	n5	850 MHz	824 - 849	869 - 894	FDD
	n8	900 MHz	880 - 915	925 - 960	FDD
	n20	800 MHz	832 - 862	791 - 821	FDD
	n28	700 MHz	703 - 748	758 - 803	FDD
	n71	600 MHz	663 - 698	617 - 652	FDD
FR1 1 -3 GHz	n1	2100 MHz	1920 - 1980	2110 - 2170	FDD
	n3	1800 MHz	1710 - 1785	1805 - 1880	FDD
	n71	2600 MHz	2500 - 2570	2620 - 2690	FDD
	n25	1900 MHz	1850 - 1915	1930 - 1995	FDD
	n40	2300 MHz	2300 - 2400		TDD
	n41	2500 MHz	2496 - 290		TDD
	n66	1700 MHz	1710 - 1780	2110 - 2200	FDD
FR1 3 - 6 GHz	n77	3.7 GHz	3300 - 4200		TDD
	n78	3.6 GHz	3300 - 3800		TDD
	n79	4.7 GHz	4400 - 5000		TDD
FR1 24 -30 GHz	n257	28 GHz	26500 - 29500		TDD
	n258	26 GHz	24250 - 27500		TDD
	n261	28 GHz	27500 - 28350		TDD

Table 2.1 *The most important 5G frequency bands for user equipment manufacturers. A 5G chipset must support the frequency bands given in FR1, but it's also recommended to support the bands given in FR2* [9].

The practical usability of all the frequency bands designated for 5G may vary, with some being more applicable than others based on the availability of the frequency spectrum at the national or regional level. For instance, the 3300 – 3600 MHz band was identified as a global IMT band by ITU Radiocommunication Sector (ITU-R) at WRC-15. In Europe early deployments focused on the 3600–3800 MHz band. Meanwhile, the United States, Japan, South Korea and China employed different allocations within the 3300–4200 MHz band for their initial 5G systems, while the 4400–5000 MHz band

was mainly used by China and Japan. Due to the divergent allocation of frequency bands to IMT across regions, there isn't a single band suitable for worldwide roaming. Nevertheless, extensive efforts have been put into defining a minimum set of bands that will enable truly global roaming, allowing multiband devices to efficiently support worldwide roaming. Many of these new bands have been identified at WRC-15 and WRC-19 and allow devices to use fewer bands to achieve global or near-global roaming, while at the same time promoting efficiency in equipment and deployment at a global scale [9], [5], [4].

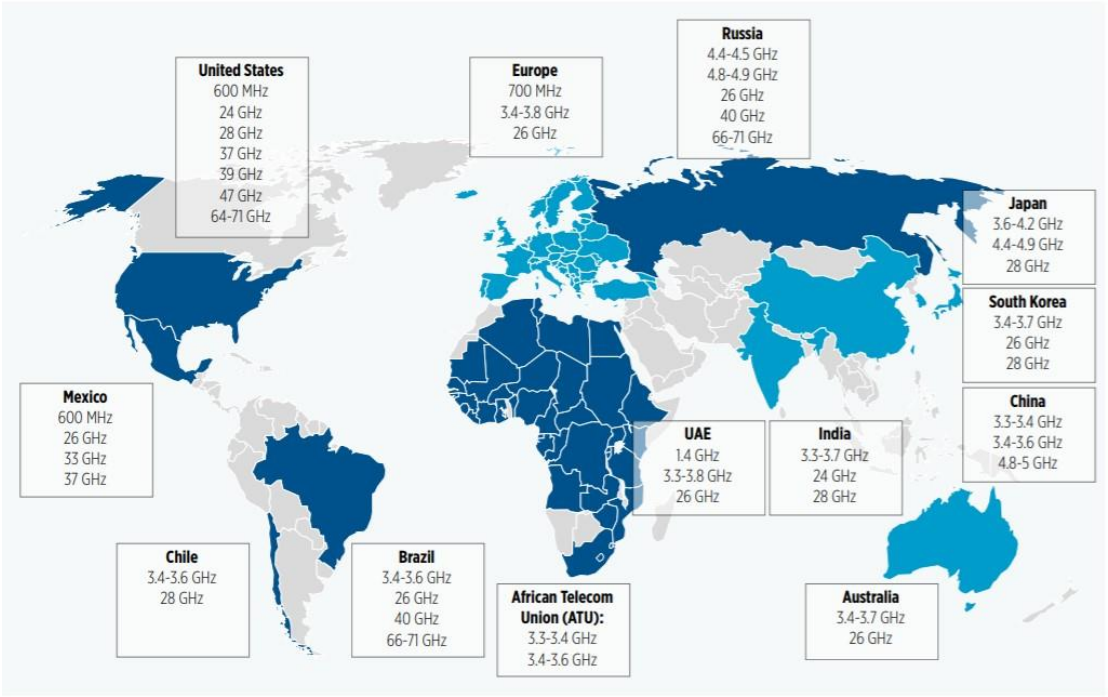
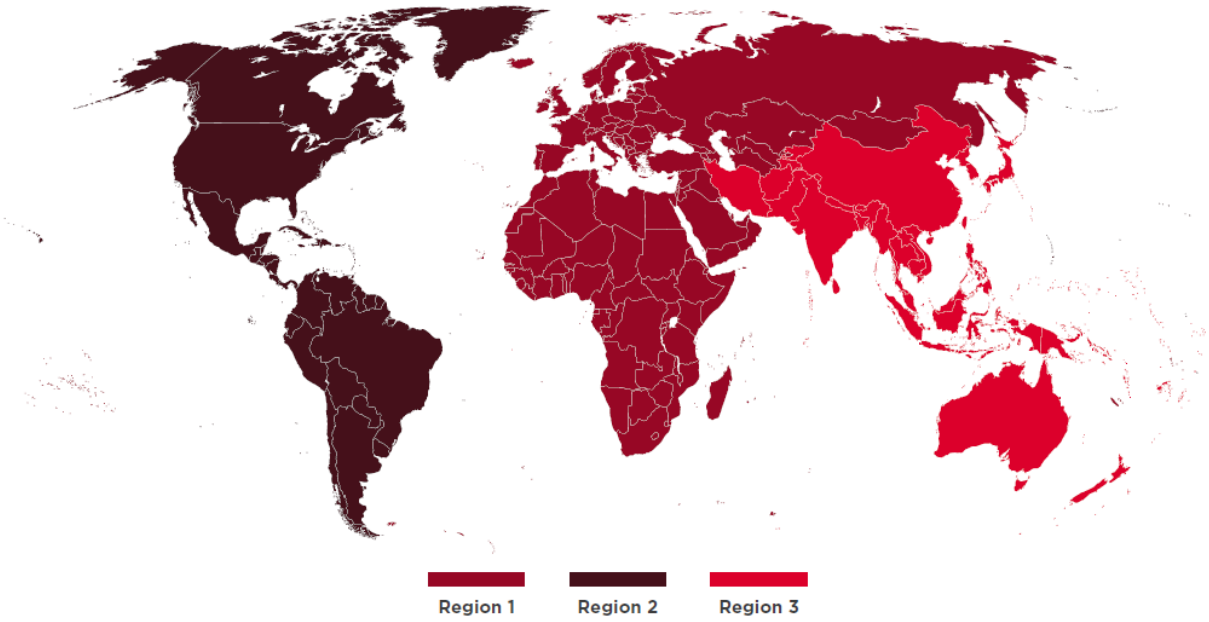


Figure 2.1 *5G Frequency Allocations Worldwide as of WRC-19* [24].

The latest ITU World Radiocommunication Conference took place from 20 November to 15 December 2023 in Dubai and focused on the expansion of 5G into all areas, aiming to provide affordable connectivity for people worldwide. WRC-23 specifically addressed both mid-band and sub-1 GHz frequencies for mobile applications. Among the WRC-23 resolutions, as depicted in Figure 2.2, new spectrum allocations were identified for International Mobile Telecommunications (IMT), including the 3300-3400 MHz, 3600-3800 MHz, 4800-4990 MHz and 6425-7125 MHz frequency bands across various countries and regions. This allocation aims to facilitate the expansion of broadband connectivity and support the ongoing development of 4G, 5G, and future 6G networks. Additionally, WRC-23 allocated the bands 15.41-15.7 GHz and 22-22.2 GHz in Radio Regulations Region 1 and certain Region 3 countries for the aeronautical mobile service, specifically for non-safety aeronautical applications. This allocation enables aircraft, helicopters, and drones to carry sophisticated aeronautical digital equipment for surveillance, mapping, and filming, while also providing the capacity to transfer large data from these applications using wideband radio links. Other important outcomes of WRC-23 were the progression and improvement of satellite services, including the allocation of additional frequencies for passive Earth exploration satellite

services, aiming to provide more accurate weather forecasting and climate monitoring. Regulatory actions were also adopted for the provision of inter-satellite links, enabling near-real time data availability that enhances the value of instrument data for low-latency applications such as weather forecasting and disaster risk reduction. Finally, WRC-23 approved the agenda items for the next World Radiocommunication Conference (WRC-27) and established the provisional agenda for WRC-31 [25], [26].



Bands	470-960 MHz	3300-3400 MHz	3600-3800 MHz	4800-4990 MHz	6425-7025 MHz	7025-7125 MHz	10-10.5 GHz
Region 1	AI 1.5 (IMT)	AI 1.2 (IMT)	AI 1.2 (IMT)	AI 1.1 (IMT)	AI 1.2 (IMT)	AI 1.2 (IMT)	
Region 2		AI 1.2 (IMT)	AI 1.2 (IMT)	AI 1.1 (IMT)		AI 1.2 (IMT)	AI 1.2 (IMT)
Region 3				AI 1.1 (IMT)		AI 1.2 (IMT)	

Figure 2.2 *WRC-23 IMT Agenda Items* [25].

2.3 5G NR Commercial Deployments

Certain 5G NR frequency bands overlap with current 4G LTE bands, meaning that the Mobile Network Operators (MNOs) may maintain the license to use these frequency blocks for 5G communication. Given that LTE operates below 6 GHz (low or mid frequency bands), NR can operate up to 100 GHz (low, mid, and high frequency bands), and that the existing LTE networks will remain in service for near future, it's crucial for 5G NR deployed in lower frequencies to coexist effectively with LTE in the same frequency bands. Extensive research in the wireless industry and standardization in 3GPP have been carried out in order to achieve efficient coexistence of both networks. The deployment of a 5G NR network started from an existing 4G LTE network, which by itself has good enough coverage. NR can coexist and interwork with LTE, which reduced the time to market for NR. Mobile Network Operators have two primary deployment options for 5G: Non-Standalone (NSA) and Standalone (SA) [5], [9], [22].

5G Standalone (SA) is a mobile network architecture that doesn't rely on existing 4G infrastructure for communication support. Instead, 5G SA networks are built with 5G infrastructure covering both the Radio Access Network (RAN) and the core network, complemented by cloud-native principles, such as virtualization and microservices. Consequently, 5G SA networks exhibit greater flexibility, scalability, and efficiency in their use of network resources, resulting in a better end user experience for consumers and lower costs for wireless carriers [27].

Non-Standalone (NSA) is a 5G Radio Access Network that operates on a legacy 4G LTE core, known as Evolved Packet Core (EPC), managing control plane functions. NSA includes both a 4G and 5G base station, with the 4G base station taking precedence. Because the NR control plane anchors to the EPC, radio frequency signals forward to the primary 4G base station. Since both radio frontends have to be powered up in parallel, NSA 5G is the most power efficient network and could be proven unsuitable for low-power applications in the IoT context. NSA 5G, also referred to as Release 15 by 3GPP, is considered the beginning stage of 5G [28], [11].

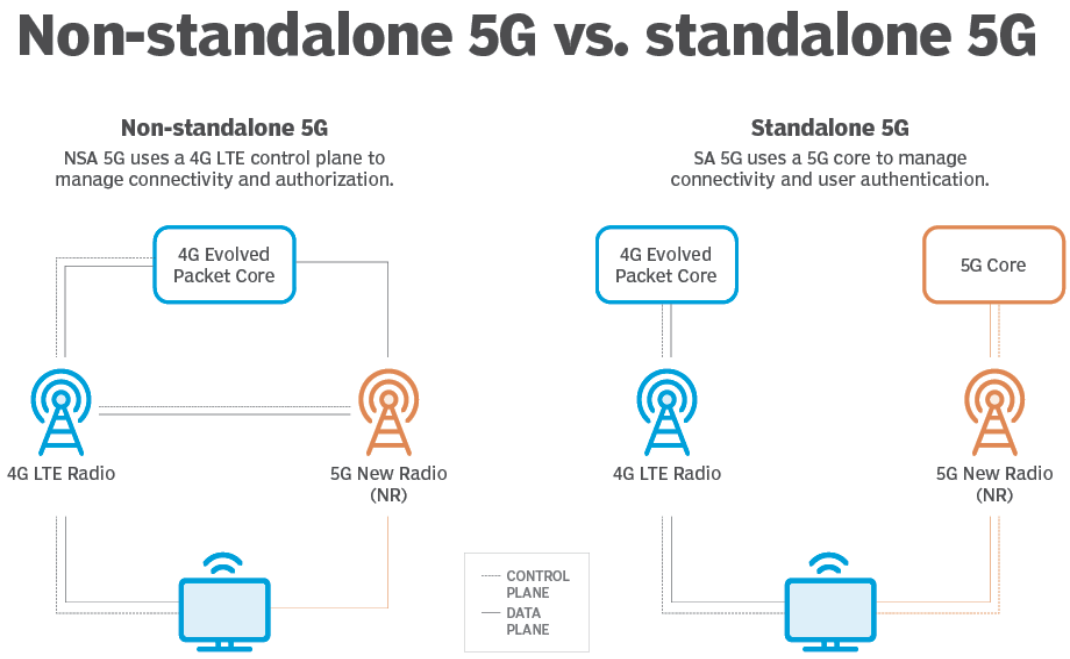


Figure 2.3 *Differences between NSA and SA 5G* [28].

NSA 5G shouldn't be confused with Dynamic Spectrum Sharing (DSS), another method of deploying 5G with 4G technology. DSS enables the sharing of a 4G frequency band between 4G and 5G services. An operator with a license for 4G use in a specific frequency range can use this 4G frequency for 5G radio communication through DSS. This adaptive approach allows spectrum usage to be adjusted based on demand [9], [28].

Even though 5G (NSA) is now globally adopted to provide enhanced smartphone experiences, 5G's key features require the use of 5G SA. However, the deployment of 5G SA by MNOs has been limited. In fact, as of March 2023, 83.6% of all announced 5G

devices support 5G SA, but only 22% of the 524 operators investing in 5G networks are investing in 5G SA. The primary reasons MNOs are in favor of NSA 5G over SA are its cost and deployment simplicity, since NSA utilize an existing 4G core network to connect to the 5G RAN. On the contrary, SA 5G utilizes a dedicated 5G core network, demanding significant investments in new infrastructure and equipment [29], [30].

Deploying NSA 5G serves as a strategy for some operators to gauge initial demand for 5G before committing substantial resources on infrastructure for an SA network. This, however, raises a particular concern for many operators in developing markets where 4G adoption is still limited. Deploying NSA 5G allows for backward compatibility, meaning users with existing 4G devices have access to the network. The lack of available 5G specifically dedicated spectrum also contributed to the decisions made by those that deployed NSA 5G, especially during the initial stages of the COVID-19 pandemic in 2020 which canceled most spectrum auctions and awards worldwide, comp some operators to launch 5G using 4G spectrum through DSS [29].

The 3GPP initially set an optimistic timeline for the deployment of SA 5G, expecting that "most operators" worldwide would have deployed SA 5G by 2023, however, that is not actually the case. The transition from NSA to SA 5G continues to be driven by individual operators, each one working on its own, with slower timelines and without a specific mandate from regulators [29].

3. 5G Core Network

The industry of wireless communication has experienced a significant transition from 4G LTE to 5G, with the focus points being a large increase in bandwidth, data speeds, and now not only end-users but also industries as customers. The architecture of 5G is split into the 5G Core Network (5GC) and the new Radio Network (NG-RAN), which supports the New Radio (NR) [31].

The RAN handles radio-related functions in the network, such as resource handling, scheduling, multi-antenna schemes, coding, and retransmissions. The 5G core network, on the other hand, is not related to the radio access. Instead, it handles functions such as authentication and management of end-to-end connections, which are necessary for the full network experience. The benefit of handling these tasks separately as opposed to leaving them to the RAN is that the same core network can serve multiple radio-access technologies at the same time [4].

Unlike the transition from 3G to 4G LTE, where the radio-access technologies where so different that LTE was not backwards compatible with 3G core networks, the radio-access network of 5G NR can connect with LTE's core network which is known as the Evolved Packet Core (EPC). This compatibility between them is necessary for the operation of 5G NR in non-standalone (NSA) mode, where the EPC handles tasks such as connection set-up and paging [4].

3.1 Core Network Architecture

The 5G system model is based on a service-based architecture and allows for many types of network services and deployable networks. These services can be updated or modified without losing performance and compatibility with already deployed networks. The 5G core network focuses on services such as the user and control plane functions and network slicing and is highly virtualized, with its main functions being runnable by even entry-level computer hardware [2], [4].

Fig. 3.1 illustrates a service-based representation of the 5G core network, focusing on its main services and functionalities. The *User Plane Function* (*UPF*) connects the RAN with the Internet and other external networks. This means that even if a device is constantly changing its location in the network, there will always be a route from the Internet to the UPF that provides a service to this device. The UPF performs tasks such as routing, processing and forwarding data, it handles QoS tasks (Quality-of-Service) and also analyzes traffic before data transmission [4].

The control plane has a number of functions, such as the *Session Management Function* (*SMF*) that allocates IP addresses to UEs and manages control of the user plane functions for that connectivity. Another one is the *Access and Mobility Management Function* (*AMF*) which is responsible for user authentication, security for user data and control signaling between the UE and the core network. *Non-Access Stratum* (*NAS*) is

the function that handles the operation between the AMF and the UE, while *Access Stratum* (*AS*) handles the operation between the UE and the RAN [4].

Other functions of the core network are the *Policy Control Function* (*PCF*), which is responsible for the enforcement of policy rules during user sessions, including rules for traffic routing and QoS enforcement, while the *Unified Data Management* (*UDM*) is responsible for access authorization and subscription management, and can work together with the AMF and AUSF. Other functions are the *Network Exposure Function* (*NEF*), which allows third parties access to certain network functions and network slices that provided as a service, and the *Network Repository Function* (*NRF*), which assists in the allocation of network functions. Specifically, each network function instance has its own status (activated or deactivated) that can be updated based on its availability, and NRF manages the information about these network function instances [4], [32].

The *Unified Data Repository* (*UDR*) stores data that define network or user policies, and user subscription data. The UDR can be used by other network functions such as the UDM, PCF and NEF. The *Authentication Server Function* (*AUSF*), although limited, its role is very important as it handles the authentication of UEs using the credentials created by the UDM. Finally, the *Network Slice Selection Function* (*NSSF*) that is responsible for the slice allocation to UEs, and the *Application Function* (*AF*) that manages traffic routing based on policy control [4], [31].

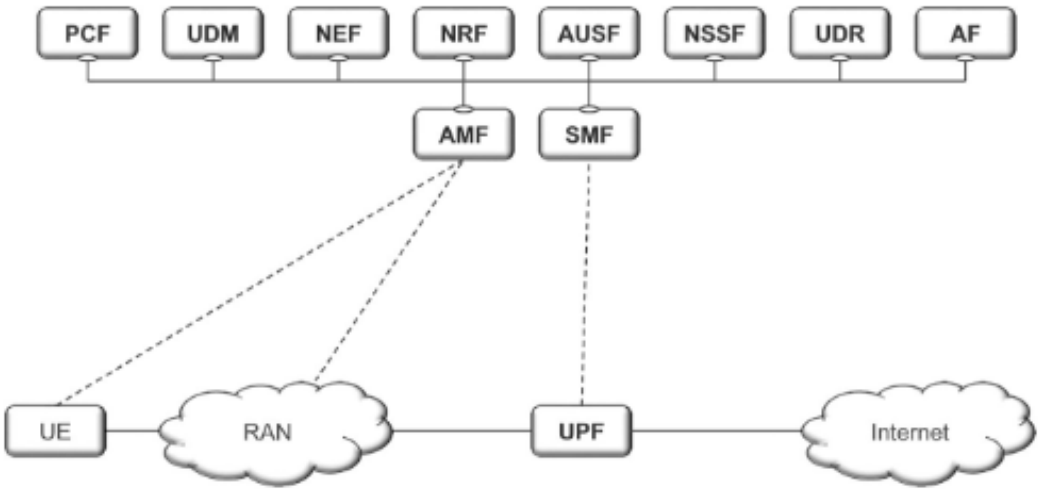


Figure 3.1 *5G core network service-based architecture* [4].

The functions of the core network can be either implemented in only one physical node, across a number of nodes, or even on a cloud platform. In standalone (SA) operation, the gNB, which is the base station for 5G NR, connects directly to the 5GC and handles both user plane and control plane functions [4], [31].

In non-standalone (NSA) operation, on the other hand, 5G NR is used simply for user data transmission when the UE is within coverage. Otherwise, the communication between devices and the network is handled by the LTE radio access on an EPC network that can support some features of 5G. This method allowed the early introduction of

5G NR in a non-disruptive way in already existing networks. The EPC core network connects to the base station of LTE, the eNB. In NSA operation, NR handles user plane the data while LTE is responsible for all the functions of the control plane. The eNB and gNB are connected with each other, allowing the transfer of user plane data from the EPC to the eNB and then to the gNB [4], [31].

In summary, LTE and NR can be implemented in four ways. They can either be used by themselves and handle all the tasks regarding signaling and data traffic, or they can be combined and can both handle tasks about data traffic, while the network with the larger coverage between the two will also handle signaling functions [31].

3.2 Network Function Virtualization

Network Function Virtualization (NFV) is a service deployment and management approach for cellular networks. In the 5G core network, the NFV involves virtualizing various 5G network functions, such as the Access and Mobility Management Function (AMF), Session Management Function (SMF), User Plane Function (UPF) and Authentication Server Function (AUSF), among others. These functions are called Virtual Network Functions (VNFs) because they can be run as software-based instances on a range of network server hardware and can be deployed or moved to other locations in the network when necessary, without requiring the installation of new equipment. The NFV can also be combined with an SDN in its implementation to enhance performance and make compatibility with already existing network deployments simpler [2].

NFV is used to separate the software implementations of network functions from their allocated network resources as well as from hardware infrastructure, allowing for better network management and more efficient resource allocation. The implementation of NFV is also cost-efficient, since it can be used to replace hardware appliances by virtualizing their functionality. This gives both service providers and corporations the ability to create and offer new services and improve the already existing ones, enhancing the experience of users [2].

3.3 Software Defined Networking

One of the most important implementations in the 5G core network architecture is the separation of the user plane and the control plane functions. The decoupling of these functions allows user and control plane resources and entities to be implemented to cloud-based deployments and different physical locations respectively [2].

Software Defined Networking (SDN) is an adaptable, cost-effective and efficient architecture which is used in modern applications that require high bandwidth. The main principle of SDN is the separation of the user and control plane functions, which provides enhanced flexibility when applied to the 5G core network. Specifically, it facilitates direct programmability of network control, enabling dynamic network configuration and efficient allocation of network resources, and allows the usage of the underlying infrastructure for network services [2].

A fundamental element for the SDN architecture is the OpenFlow protocol. OpenFlow is a standardized protocol that uses control plane Application Programming Interfaces (APIs) to facilitate user plane functions. The OpenFlow standard was unveiled in 2008 and it was the first SDN architecture that separated the control and user plane. Nowadays, there are many OpenFlow standards, which are also opensource specifications, and are all controlled by the Open Networking Foundation (ONF). At the same time, there are also other standards and organizations that create protocols for the SDN framework. However, OpenFlow is the standard responsible for the invention of Network Operating Systems. A Network Operating System is a software that manages the control and behavior of the network [2], [33].

The SDN architecture is illustrated in Fig. 3.2 and consists of three main parts, applications, controllers and network devices. SDN applications are software programs designed to interact with SDN controllers through APIs about their network resource requirements. These applications can also use information provided from the controller to construct a high-level representation of the network. The data provided by the controller can also be used for decision-making purposes. Some SDN applications include routing, load balancing, network management and analytics [2].

The SDN controller serves as a centralized entity within the core network architecture and is used to receive instructions and requirements from the SDN application layer. The controller communicates with the network components, such as switches and routers, via south-bound APIs, relaying them information it received while also receiving new information from them about the network and sending that information back to the SDN applications via north-bound APIs. Through these interactions, the controller creates an abstracted representation of the network, recording statistics and events. The control plane consists of one or multiple SDN controllers, which use south-bound interfaces to manage the network devices in the user plane, depending on requests from the application layer [2].

On the infrastructure layer, SDN networking devices are used data forwarding functions. The user plane consists of data forwarding elements, which handle the seamless forwarding and routing of data [2].

The APIs define the communication between the three parts of the SDN architecture and are categorized into north-bound and south-bound interfaces. A north-bound interface represents the connection between the SDN controller and SDN applications. On the other hand, the south-bound interface represents the connection between the SDN controller and the physical hardware of the network [2].

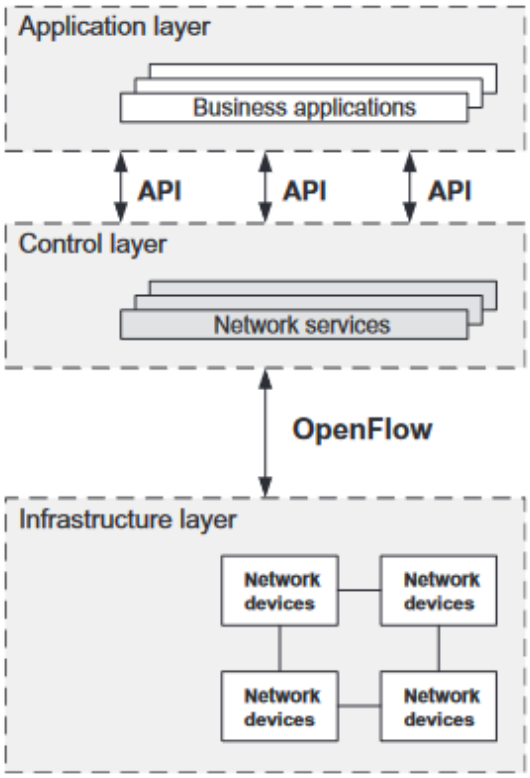


Figure 3.2 *SDN architecture* [2].

The control and user plane separation (CUPS) is an important innovation in 3GPP's 5G network and allows either distributed or centralized deployment of control plane and user plane functions, while also enabling the network programming via open interfaces. SDN compliments NFV when it comes to the network architecture, allowing functions to be run as software-based implementations on low-cost hardware or even cloud computing environments [2].

3.4 Network Slicing

One of the most important features of 5G, which is standardized by 3GPP, is network slicing. Network slicing deploys core network technologies, such as SDN and NFV, to dynamically allocate network functions, allowing 5G network operators to divide the physical infrastructure of a network into a number of isolated virtual networks, which are called network slices. Network slices are able to support different types of services dynamically based on requirements and allow for a more efficient use of communication channels. The creation of these slices aims to provide an enhanced level of QoS, such as improved functionality, guaranteed delay, throughput and reliability. From the perspective of a UE, network slicing is to categorize devices into slices based on their performance requirements, such as their transmission rate and their delay [2], [32].

The network slicing architecture is depicted in Fig. 3.3 and consists of three layers, the service instance layer, the network slice instance (NSI) layer and the resource layer. The service instance layer defines the services that a network can support. Each service is provided by a service instance, which is either a network operator or a third party.

Network slices are used to create NSIs based on the network characteristics, which are provided by a service instance, and can be shared with other service instances. The life cycle of NSIs and service instances are independent from each other. Specifically, the life cycle of an NSI. The NSI life cycle consists of several phases, including instantiation and configuration phases, an activation phase, a run-time phase, and finally a decommissioning phase. During its life cycle, NSI is controlled by a service instance, which may not be active through the entire duration of the run-time phase of the NSI it supports [2].

NSIs may consist of some sub-network instances, which can be shared with other NSIs. These sub-network instances are actually sets of network functions that consume the physical network resources. Network slices are complete networks that have different characteristics and functionalities based on the services they are required to support in a cellular network [2].

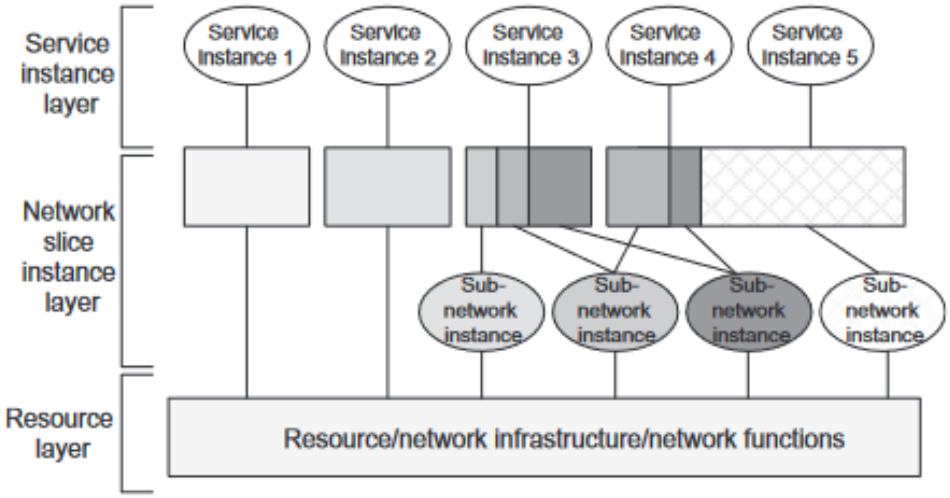


Figure 3.3 *Network Slicing architecture* [2].

4. 5G NR Physical Layer

Like every wireless technology, 5G NR's structure is built upon the physical layer, which is the lowest protocol layer in baseband signal processing and has the important role of transmitting and receiving wireless signals between devices. Essentially, as illustrated in Figure 4.1, between a base station and user equipment the physical layer is responsible not only for encoding the data bits into a radio signal at the transmitter and transmitting it over the air, but also receiving the radio signal at the receiver and decoding it back to digital bits. During this process, the transmitted data are organized into transport blocks, which are blocks of a fixed number of data bits. The modulation of digital bits on multiple carrier frequencies is accomplished through orthogonal frequency-division multiplexing (OFDM) [2], [9].

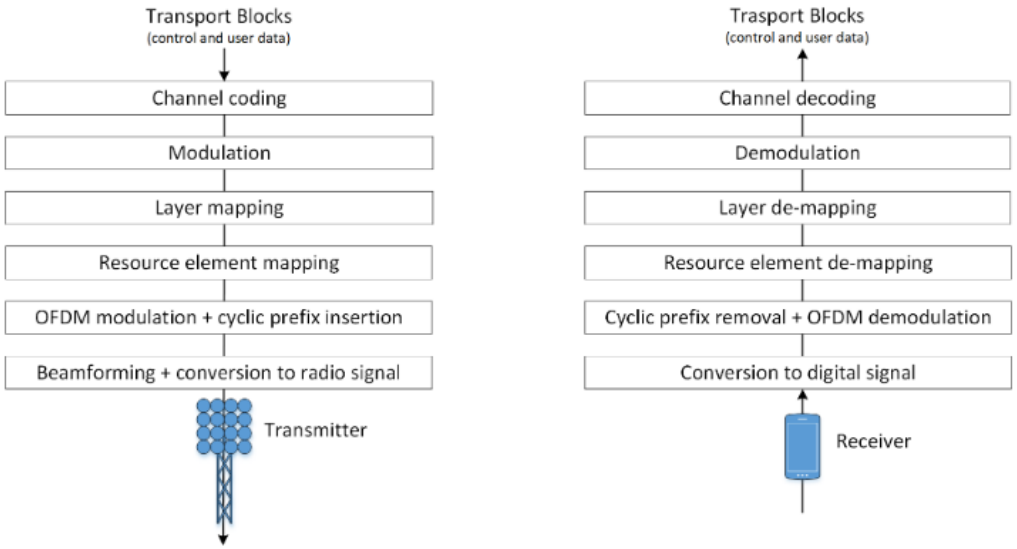


Figure 4.1 *5G NR's key components of the physical layer* [9].

A well-crafted and robust design of the physical layer is needed so the system can operate under any condition with minimum to no latency and without any signal delay, interference or fading [2]. That is why it is important to have a different set of protocols and functions that ensure the uninterrupted communication of devices between airwaves. The 5G NR protocol stack, as illustrated in Figure 4.2, is a layered radio protocol architecture that can be categorized into control plane architecture and user plane architecture. The user plane is responsible for the transmission of user data, while the control plane handles the connection establishment, mobility, and security of the connection. It also consists of three main layers: the physical layer, the data link layer, and the network layer [2].

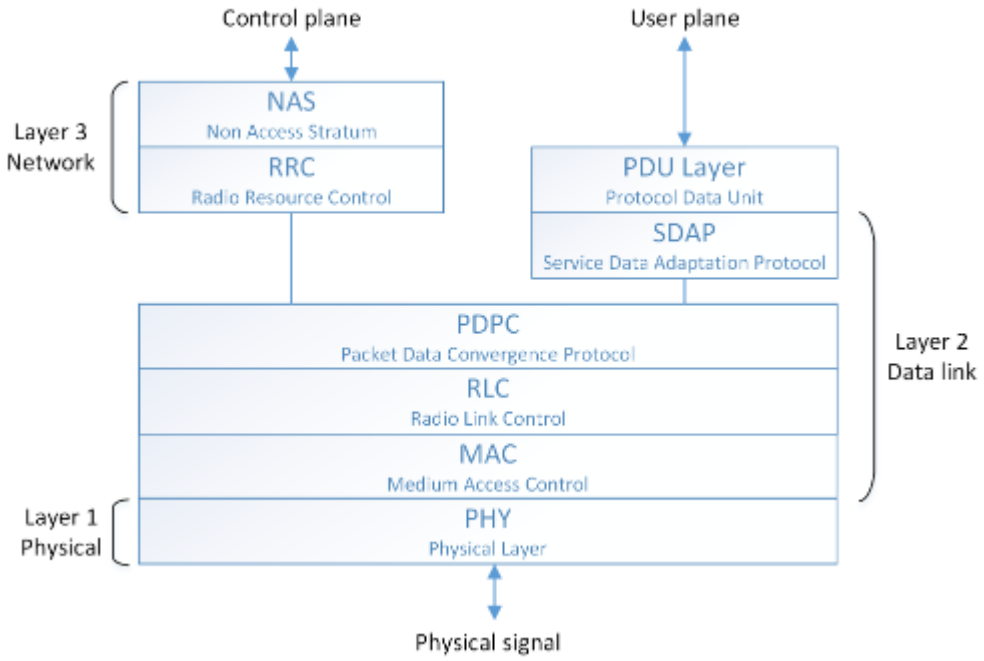


Figure 4.2 *Control plane and user plane radio protocol architecture for NR* [9].

Layer 1 of NR is the Physical (PHY) layer, which is the lowest layer of the protocol stack and deals with the coding and decoding of data, but also handles tasks like modulation and demodulation, transmission and reception of physical signals over the air interface, multiantenna processing, and mapping of signals to physical time-frequency resources [5].

Layer 2 of NR consists of the Service Data Adaptation Protocol (SDAP), Packet Data Convergence Protocol (PDCP), Radio Link Control (RLC) and Medium Access Control (MAC) sub layers. The SDAP Layer provides QoS (Quality of Service) management and mapping of data flows to specific radio bearers. It ensures that both DL (Downlink) and UL (Uplink) receive the necessary QoS parameters. The PDCP Layer handles the compression and decompression of data packets and their headers, the encryption and decryption of packets, but it is also responsible for reordering and detecting duplicate packets, ciphering, deciphering and ensuring the integrity of the transmitted packets. The header compression process reduces the number of bits to be transmitted over radio interface. The ciphering process protects from eavesdropping and ensures message integrity. The reordering and duplication detection processes ensure in-sequence delivery as well as removal of duplicate data units. The RLC layer primarily engages in error correction using an automatic repeat request (ARQ) mechanism, but also deals with the segmentation and resegmentation of (header compressed) data packets and ensures the in-sequence delivery of data units to higher layers. The MAC layer is also responsible reliable and error-free transmission, this time with the use of hybrid ARQ (HARQ) mechanisms. This layer also has the assignment of scheduling the transmission of data between multiple users, correctly allocating UL and DL physical time-frequency resources, traffic prioritization and control signaling [5], [34], [35].

Layer 3 of NR is the control plane and consists of the Radio Resource Control (RRC) and the Non-Access Stratum (NAS) sub layers. The RRC Layer handles the establishment, maintenance, and release of the radio connection between the UE and the BS, while the NAS layer is mainly responsible for control signaling, authentication, security, and mobility management. Control signaling originates either from a core network or from the RRC layer in the BS [5], [12] [13], [35].

4.1 Modulation

The key technology components of the NR physical layer are modulation, waveform, multiantenna transmission, and channel coding. Below, we provide a brief overview of these physical layer components, as well as explain the basic principle of Orthogonal Frequency Division Multiplexing (OFDM).

Modulation is the process of altering a carrier signal to transmit information, by converting digital bits into modulation symbols which are then encoded on frequency carriers. A modulation symbol is a complex number that has both real and imaginary parts and represents a certain number of bits. The symbol can be represented by a point in a two-dimensional complex plane, defined by its angle of rotation (phase) and its distance to the origin (amplitude). Different symbols have different phases and amplitudes. The carrier signal, often a high-frequency sine wave, is transmitted over a communication channel, with the transmitted information being a low-frequency signal, like an audio or video signal. Modulation ensures that the information can be transmitted over long distances without any fear of degradation by noise or interference. Additionally, having more complex modulation allows for significantly higher bandwidth [9], [36], [37].

The modulation order of a modulation scheme defines the number of bits per symbol. More bits that are transmitted over the same bandwidth (higher modulation order), means more efficient bandwidth and enhanced data rate. In practice however, the use of high modulation orders in a radio channel is constrained by two factors, noise and interference. The radio channel influences the phase and amplitude of the symbols, which may cause the receiver to confuse a symbol with a neighboring symbol of theirs during demodulation. Higher modulation orders reduce the distances between symbols, leading to more errors being detected. If the signal quality is not the best (low signal-to-noise ratio), a lower modulation order must be considered to prevent a multitude of bit errors [9].

Three common methods of modulation are Amplitude Modulation (AM) and Frequency Modulation (FM) and Phase Modulation (PM), all of them commonly used to "modulate" digital data onto a radio signal. AM works by adjusting the amplitude of the signal under constant frequency. This is different to FM which adjusts the frequency of the signal under constant amplitude. PM is similar to FM but adjusts the phase of the signal instead [37].

In cellular signals, the main modulation schemes are Quadrature Amplitude Modulation (QAM) and Phase Shift Keying (PSK). QAM is a high-bandwidth modulation method that is used in 4G and 5G cellular systems and performs amplitude and phase adjustments simultaneously at the signal. PSK is a phase-based modulation method that, unlike QAM, only adjusts the phase of the signal [37].

Like LTE, 5G NR supports modulation schemes where the symbols are given at different phases, such as quadrature phase-shift keying (QPSK), 16 quadrature amplitude modulation (QAM), 64 QAM and 256 QAM modulation formats for both uplink and downlink. Binary phase-shift keying (π/2-BPSK) is also supported in the uplink in order to achieve lower peak-to-average power ratio and better power efficiency at lower data rates, an important requirement for mMTC services. All these modulation schemes are illustrated in Figure 4.3. Given the diverse range of use cases for NR, it is likely that it could support an even bigger number of modulation schemes in the future. One of the modulation schemes that is considered is 1024 QAM, since fixed point-to-point backhaul already uses modulation orders higher than 256 QAM [5], [9].

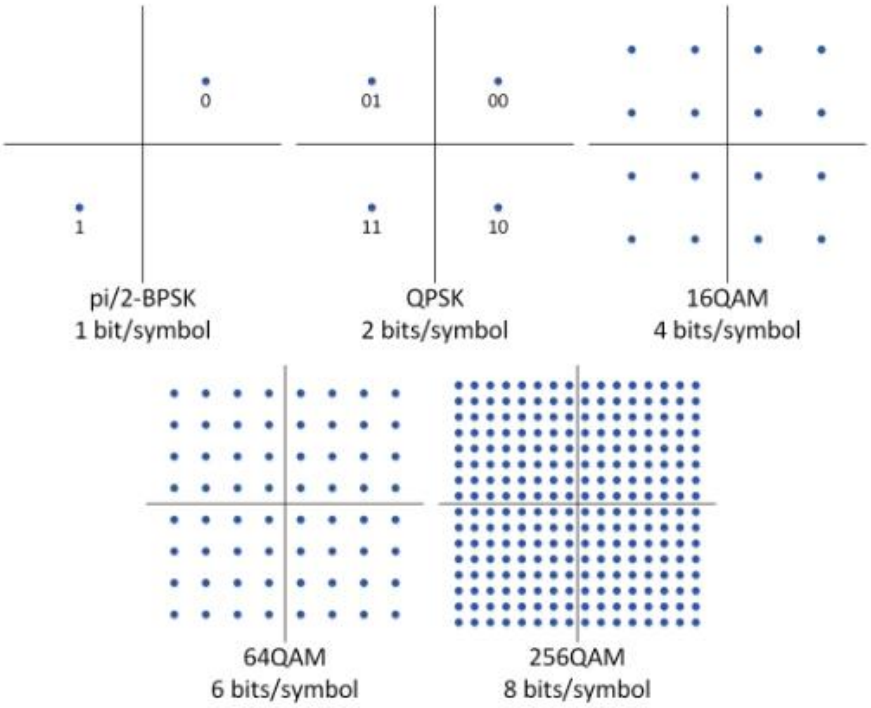


Figure 4.3 *Graphic representation of modulation schemes used in 5G NR. Modulation rates are given by the number of bits per symbol* [9].

4.2 Orthogonal Frequency Division Multiplexing (OFDM)

In a traditional parallel data system, the total signal bandwidth can be divided into N distinct and non-overlapping frequency subchannels. Each subchannel is modulated with a unique symbol and then all N subchannels are frequency multiplexed. Frequency Division Multiplexing (FDM) is a signal transmission technique that allows the transmission of multiple signals over a single channel. This is achieved by dividing the channel's available frequency spectrum into several smaller frequency channels, each one allocated to a different signal. The width of each frequency channel is determined by the bandwidth of its caring signal. The signals are then transmitted simultaneously in their respective frequency channels, and the signals are separated at the receiving end using a bandpass filter. A bandpass filter is a filter that permits only a specific frequency channel to pass through while blocking all other frequencies. In FDM, each signal occupies a distinct frequency channel, and the bandpass filter is used to isolate a frequency channel of our choice [38].

The general approach of preventing spectral overlap among subchannels was implemented to eliminate inter-carrier interference (ICI) and is illustrated in Fig 4.4 (a). The result of this method, however, was an insufficient utilization of the existing spectrum. A new solution to this problem was born in the mid-1960s, which would have the subchannels arranged in a way so that the sidebands of the individual carriers overlap without causing ICI and is illustrated in Fig 4.4 (b). To achieve this, the frequency channels must be mathematically orthogonal to the adjacent channels. It was from this constraint that the concept of Orthogonal Frequency Division Multiplexing (OFDM) was born [38].

Orthogonal Frequency Division Multiplexing (OFDM) is a modulation and multiplexing technique that divides the entire bandwidth into several orthogonal channels, referred to as subcarriers, each of which is narrower than the coherence bandwidth. The orthogonality of the subcarriers allows for high spectral efficiency, to the point that almost the entire frequency band is usable. The use of OFDM can also reduce the mutual interference between the subcarriers (ICI). By combining low data rates, OFDM systems can generate a large data rate with a long symbol duration, also reducing intersymbol interference (ISI). These systems can also lower the equalization complexity by converting the wideband signal into N narrowband flat fading signals with the use of the Inverse Fast Fourier Transform (IFFT) at the transmitter and the Fast Fourier Transform (FFT) at the receiver [39], [40].

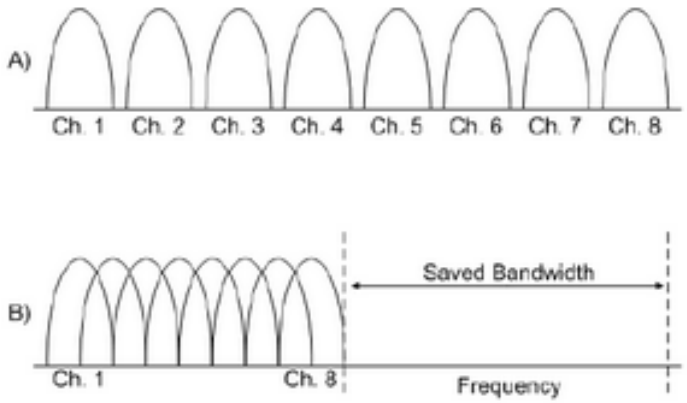


Figure 4.4 (*a)The frequency spectrum of eight channels using frequency division multiplexing, with guard bands placed between sub-carriers. (b) The frequency spectrum of eight channels using OFDM and the resulted saved spectrum. Adjacent channels are orthogonal to each other* [38].

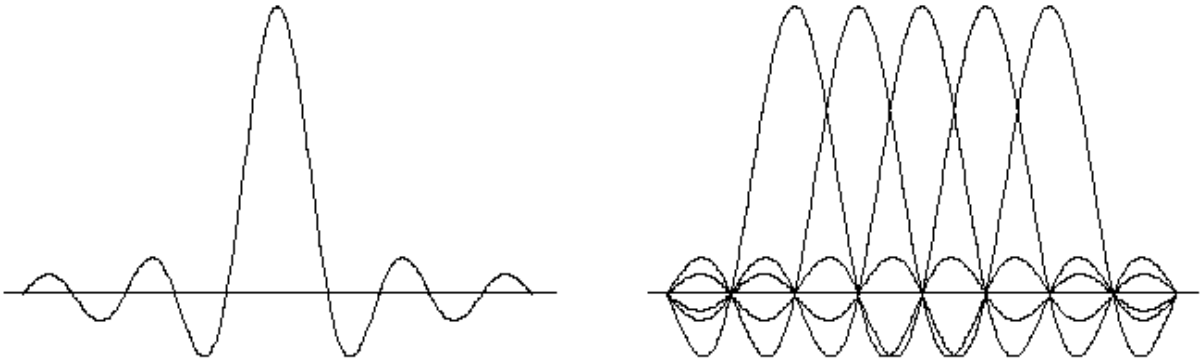


Figure 4.5 *OFDM spectrum for (a) a single subchannel, (b) 5 carriers* [41].

4.2.1 OFDM Block Diagram

OFDM transmits many closely spaced narrowband carriers in the frequency domain. To construct a complete OFDM circuit we must interconnect all of these carriers together, implementing the OFDM modulation block by block. A generic building block diagram of an OFDM system is depicted in Fig. 4.6. At the transmitter end, a serial-to-parallel (S/P) converter transforms the input data into many parallel data streams, each mapped onto corresponding information symbols for the subcarriers within one OFDM symbol. Then, training symbols (TSs) are inserted momentarily for channel estimation [42], [43].

The modulation of a waveform and its addition is mathematically equivalent to taking an IFFT. The time-domain OFDM signal is represented by a two-dimensional complex signal with real and imaginary parts, making IFFT is the correct choice to use in the transmitter, essentially converting frequency domain samples to time domain samples. In the OFDM block diagram, parallel data streams are modulated onto orthogonal subcarriers before being converted to the time-domain OFDM signal. In order to avoid channel dispersion a cyclic prefix is added into each OFDM symbol. The OFDM signal

undergoes digital-to-analog conversion and passes through a low-pass filter (LPF) in order to acquire the OFDM baseband signal. The baseband signal can be up-converted to an appropriate radio frequency (RF) passband with an in-phase/quadrature-phase (IQ) modulator and a band-pass filter (BPF) [42], [43].

At the receiver end, the OFDM signal is down-converted to a baseband signal with an IQ demodulator that has an analog-to-digital converter (ADC), and then the OFDM signal is demodulated by a Fast Fourier Transform (FFT) function. After being demodulated, the resulting signals go through a symbol decision process, where synchronization, channel estimation, and compensation are performed before a symbol decision is made. At the end of the OFDM system process, the sub-divided data streams are converted back to a single one by parallel-to-serial (P/S) operation [42], [43].

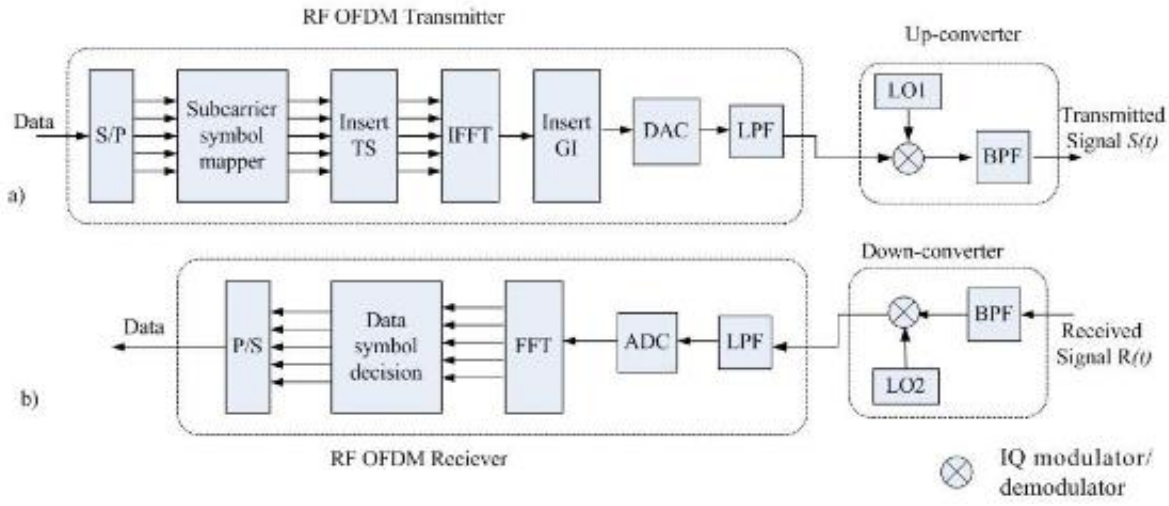


Figure 4.6 *Building blocks of an OFDM system* [42].

4.2.2 Orthogonality

Let $\{s_{n,k}\}_{k=0}^{N-1}$ with $E\left|s_{n,k}\right|^2 = \sigma_s^2$ be the complex symbols to be transmitted at the *n*th OFDM block, then the OFDM-modulated signal can be represented by,

$$s_n(t) = \sum_{k=0}^{N-1} s_{n,k} e^{j2\pi k\Delta f t}, \qquad 0 \le t \le T_s \quad (4.1)$$

where T_s, Δf, and N are the symbol duration, the subchannel space, and the number of subchannels of OFDM signals, respectively, and $f_k = k\Delta f$ is the frequency of the modulated subcarrier [44]. In order for these signals to be orthogonal, the integral of the products for their fundamental period has to be zero,

$$\frac{1}{T_s}\int_0^{T_s} e^{j2\pi f_k t} e^{-j2\pi f_i t} dt = \frac{1}{T_s}\int_0^{T_s} e^{j2\pi k\Delta f t} e^{-j2\pi i\Delta f t} dt$$

$$= \frac{1}{T_s}\int_0^{T_s} e^{j2\pi(k-i)\Delta f t} dt = \begin{cases} 1, & \forall\ integer\ k = i \\ 0, & otherwise \end{cases} \quad (4.2)$$

Taking the discrete samples with the sampling instances at $t = nT_s$, $n = 0, 1, 2, \ldots, N-1$, equation (4.2) can be written in the discrete time domain as [45],

$$\frac{1}{N}\sum_{n=0}^{N-1} e^{j2\pi k\Delta f n T_s} e^{-j2\pi i\Delta f n T_s} = \frac{1}{N}\sum_{n=0}^{N-1} e^{j2\pi(k-i)\Delta f n T_s}$$

$$= \begin{cases} 1, & \forall\ integer\ k = i \\ 0, & otherwise \end{cases} \quad (4.3)$$

To sum up, for the OFDM signal to be demodulated at the receiver the orthogonal condition $T_s\Delta f = 1$ must be in effect, since it makes $e^{j2\pi k\Delta f t}$ orthogonal to each other for different *k*. If the orthogonal condition is in effect, the receiver can spot the transmitted symbols $s_{n,k}$,

$$s_{n,k} = \frac{1}{T_s}\int_0^{T_s} s_n(t) e^{-j2\pi k\Delta f t} dt \quad (4.4)$$

if there is no channel distortion [44].

The above orthogonality is a crucial condition to ensure the OFDM signal is ICI-free. The advantage of OFDM diminishes if the orthogonal condition is not in effect between the subcarriers. When more subcarriers are required, the modulation, synchronization, and demodulation processes result in a not very cost-efficient OFDM circuit. This can turn the analog implementation of the Fourier Transforms impractical. OFDM modulation and demodulation can be implemented using Inverse Discrete Fourier transform (IDFT) or Discrete Fourier Transform (DFT) [38], [42].

4.2.3 The Fourier Transform

The Fourier transform is a widely used method for obtaining frequency spectrum of signals and relates to continuous signals that are not restricted to time or frequency domains. The process of signal processing, however, can be made easier with sampled signals. In the case of infinite spectrum while signal sampling the result may be affected by aliasing, and if there are no time restrictions it can cause problems in terms of storage space. To address these problems, signal processing uses the Discrete Fourier Transform (DFT). DFT is a mathematical operation and, to be precise, the discrete version of the Fourier Transform (FT), that takes a signal in its time domain representation and, as the name suggests, transforms it to its frequency domain representation. The Fast Fourier Transform (FFT) is another algorithm used for the same computation that can produce faster results with reduced complexity and works best with signals whose frequency components remain constant with time [46].

It is easy to get these two confused. Often, one may see a phrase like "take the FFT of this sequence", which really means to take the DFT of that sequence using the FFT algorithm to do it efficiently. Although a lot of modulators can be used to illustrate the basic principles of OFDM modulation and demodulation respectively, these are not the

most appropriate modulator/demodulator structures for actual implementation. Due to its specific structure and the selection of a subcarrier spacing *Δf* equal to the per-subcarrier symbol rate *1/Ts*, OFDM allows for low-complexity implementation by means of computationally efficient Fast Fourier Transform (FFT) processing. To confirm this, consider a time-discrete (sampled) OFDM signal where it is assumed that the sampling rate f_s is a multiple of the subcarrier spacing *Δf* – that is, $f_s = \frac{1}{T_s} = N\Delta f$. The number of subcarriers N_c, together with the subcarrier spacing Δf, $N_c\Delta f$, determines the overall transmission bandwidth of the OFDM signal. This implies that *N* should exceed N_c with a sufficient margin.

The FFT implementation of a sequence actually means to use the DFT of that sequence using the FFT algorithm to do it efficiently. Due to OFDM's structure and its subcarrier spacing *Δf* that is equal to the per-subcarrier symbol rate *1/Ts*, OFDM uses FFT processing in order to have low-complexity implementation for a sequence. This can be proven using a sampled time-discrete OFDM signal with a sampling rate f_s that is a multiple of the subcarrier spacing *Δf*, $f_s = \frac{1}{T_s} = N\Delta f$. The transmission bandwidth of the OFDM signal is determined by the number of subcarriers N_c and the subcarrier spacing Δf, $N_c\Delta f$. This implies that *N* should always surpass N_c [47].

The time-discrete mathematical expression of the baseband OFDM signal $s(t)$ is:

$$s_n = s(nT_s) = \sum_{k=0}^{N_c-1} s_{n,k} e^{j2\pi k\Delta f n T_s} = \sum_{k=0}^{N_c-1} s_{n,k} e^{\frac{j2\pi kn}{N}} = \sum_{k=0}^{N-1} s'_{n,k} e^{\frac{j2\pi kn}{N}} \quad (4.5)$$

where,

$$s'_{n,k} = \begin{cases} s_{n,k}, & 0 \le k < N_c \\ 0, & N_c \le k < N \end{cases} \quad (4.6)$$

The sampled OFDM signal which is depicted by the sequence s_n, is the size-*N* Inverse Discrete Fourier Transform (IDFT) of the block of modulation symbols $\{s_{n,k}\}_{k=0}^{N-1}$. This block can be extended with zeros to length *N* and can be calculated using the FFT. Implementing OFDM modulation includes the application of IDFT processing together with digital-to-analog conversion. OFDM demodulation can be similarly implemented using FFT processing with a sampling rate $f_s = \frac{1}{T_s}$, followed by a size-*N* DFT or FFT [47].

4.2.4 Cyclic-Prefix Insertion

In high-data-rate communication, the issue of Inter-Symbol Interference (ISI) can occur because of decreased time duration due to increased data rates and can lead to self-interference due to multipath delay spread. This interference is then decoded incorrectly at the receiver end. To mitigate the effect of ISI, time duration must be kept greater than the maximum delay T_{max} of the multipath channel. A guard interval is also used before the data period, allowing ISI to occur within the guard interval. The data are retrieved once the guard interval is removed [48].

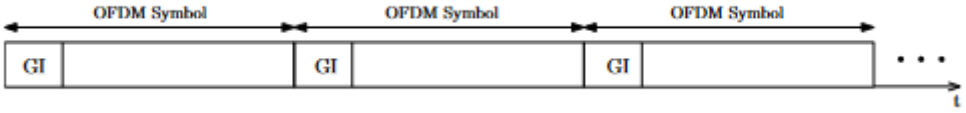


Figure 4.7 *OFDM symbol and its guard interval*[28].

In OFDM systems, a very effective timed-domain guard interval to use is a cyclic prefix (CP), which is the cyclic extension of the symbol. The CP is inserted between OFDM symbols to deal with ISI [49]. Fig 4.8 illustrates the use of a cyclic prefix. The cyclic extension of the symbol means that the last part of the OFDM symbol is copied and inserted at the beginning of the OFDM symbol prior to its transmission [47].

Let T_{CP} and T_s represent the length of cyclic prefix in terms of samples and symbol duration. If no cyclic prefix is inserted, the length of the OFDM symbol is T_s, as shown in (4.1). When the cyclic prefix is inserted, the transmitted signal's length is extended from $T = T_s$ to $T = T_{CP} + T_s$. This extension leads to a reduction in the OFDM symbol rate, and it can be mathematically expressed as:

$$\widetilde{s_n}(t) = \sum_{k=0}^{N-1} s_{n,k} e^{j2\pi k\Delta f t}, \qquad -T_{CP} \le t \le T_s \quad (4.7)$$

The term cyclic prefix comes from the fact that $\widetilde{s_n}(t) = \widetilde{s_n}(t + T_s)$ for $-T_{CP} \le t \le 0$ [44], [47], [49].

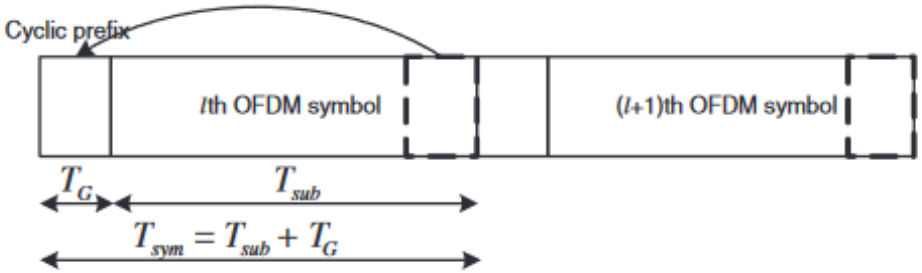


Figure 4.8 *Two consecutive OFDM symbols with cyclic prefix. Each symbol's cyclic prefix length is* $T_G = T_{CP}$*, and illustrates an OFDM symbol of length* $T = T_G + T_s$ [28].

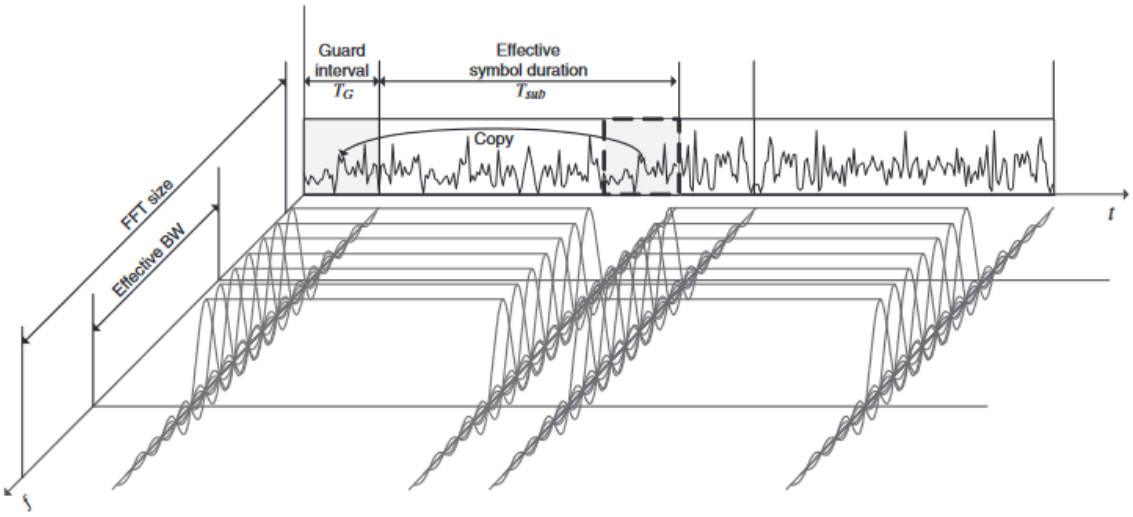


Figure 4.9 *Two consecutive OFDM symbols with a cyclic prefix in the time and frequency domain* [45].

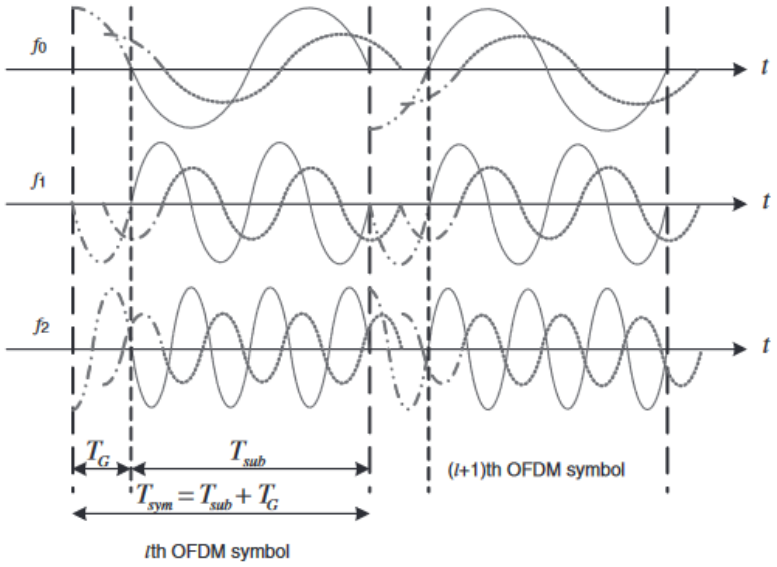


Figure 4.*10* *ISI effect on two consecutive OFDM symbols* [45].

The cyclic prefix is designed to be equal to or longer than the maximum delay T_{max} of a multipath channel. On the receiver side, sampling of an OFDM symbol starts somewhere in the interval (T_{max}, T_{CP}), ensuring that, for a duration of T_s. for two consecutive OFDM symbols the ISI effect of the first OFDM symbol on the next one is confined within the cyclic prefix. The cyclic prefix insertion not only prevents any change in the FFT of the second OFDM symbol and adds robustness against timing errors on the receiver side, but also plays an important role in maintaining the orthogonality among the subcarriers over the time duration T_s, such that,

$$\frac{1}{T_s}\int_0^{T_s} e^{j2\pi f_k(t-t_0)} e^{-j2\pi f_i(t-t_0)} dt = 0, \quad \forall\ integer\ k \neq i \qquad (4.8)$$

for the first OFDM signal that arrives with a delay of t_0, and

$$\frac{1}{T_s}\int_0^{T_s} e^{j2\pi f_k(t-t_0)} e^{-j2\pi f_i(t-t_0-T_s)} dt = 0, \quad \forall\, integer\ k \neq i \quad (4.9)$$

for the second OFDM signal that arrives with a delay of $t_0 + T_s$ [49], Η πηγή που καθορίστηκε δεν είναι έγκυρη..

In case of a cycle prefix shorter than the maximum delay of a multipath channel, for two consecutive OFDM symbols the tail part of the first OFDM symbol affects the head part of the second one, resulting ISI effect on these symbols, as illustrated in Fig. 4.11 [45].

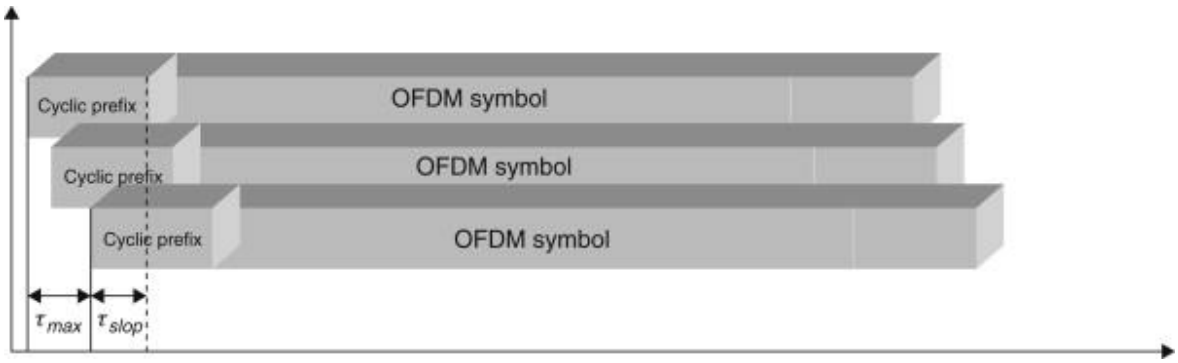


Figure 4.11 *ISI effect of a multipath channel on OFDM symbols* [45].

The primary advantage of cyclic-prefix insertion lies in its ability to make an OFDM signal insensitive to time dispersion if it does not exceed the length of the cyclic prefix. However, cyclic prefix insertion causes a power loss in the demodulation since the demodulator uses only a fraction $T_s/(T_s + T_{CP})$ of the received signal power. Additionally, it also creates bandwidth loss, since it reduces the OFDM symbol rate without reducing the overall signal bandwidth. A solution is the reduction of the subcarrier spacing *Δf*, which also implies an increase in the symbol time. It is crucial to note that the cyclic prefix does not necessarily need to cover the entire length of the channel time dispersion. In general, there exists a trade-off between the power loss resulting from the cyclic prefix and the one resulting from signal corruption because the cycle prefix does not cover residual time dispersion. There is a point where no matter how much the length of the cyclic prefix is increased to reduce signal corruption, it will not justify the corresponding additional power loss [47].

4.3 Waveform

Waveforms are an essential part of every communication technology. The main categories of waveforms are single-carrier and multicarrier waveforms. Single-carrier waveforms are generally characterized by how power efficient they are, since they have a low peak-to-average power ratio (PAPR), making them well-suited for coverage limited scenarios as well as improving battery life of user equipment. Multicarrier waveforms, on the other hand, not only provide high spectral efficiency and flexible

resource allocation in the frequency domain but can also be seamlessly integrated in multiantenna technology [5].

Waveforms are designed in ways that help accommodate 5G NR's use cases, service requirements and deployment scenarios in frequencies from below 1 GHz to 100 GHz. Several single carrier and multicarrier waveforms have been acknowledged by the 3GPP to be used in different scenarios for 5G NR, with multicarrier waveforms being the primary candidates to be used [5], [34].

Multicarrier waveforms split a high-rate data stream into multiple low-rate ones and transmit all of them over several subcarriers at the same time. For each subcarrier the symbol duration increases, making the amount of time dispersion caused by a multipath channel to decrease. Thus, a multicarrier waveform is robust to the ISI and can allocate various subcarriers to a number of users. Multicarrier waveforms are also resilient against narrowband interference, which only corrupts a few subcarriers, compared to single-carrier waveforms where it can make an entire link stop working completely [5].

However, the 3GPP also recognized the fact that using multiple waveforms will create a complex 5G NR system design. To satisfy the need for a single waveform and balance the performance of NR systems, the 3GPP decided to use cyclic prefix OFDM (CP-OFDM) for both uplink and downlink transmissions up to at least 52.6 GHz. This differs from LTE, where CP-OFDM is exclusively used for downlink transmissions, while uplink transmission uses DFT-Spread OFDM (DFTS-OFDM). This makes NR's design simpler, particularly when it comes to wireless backhauling and Device-to-Device (D2D) communications. Moreover, the single carrier DFTS-OFDM waveform is also an option for uplink transmission when coverage is limited. In practice, a BS can choose between CP-OFDM or DFTS-OFDM for the uplink waveform, and a UE should be able to support both OFDM and DFTS-OFDM [5].

4.3.1 Numerology

A numerology is a set of parameters that define an OFDM system, like Subcarrier Spacing (SCS), Cyclic Prefix (CP), Symbol Length, and Transmission Time Interval (TTI). The 5G NR numerology can serve diverse purposes and achieve various performance figures. 5G NR supports a number of scalable OFDM numerologies, each one offering different services and solutions to physical layer challenges at high frequencies. Numerology is denoted as a variable $\mu = \{0, 1, 2, 3, 4\}$ and is defined by SCS, which is the distance in frequency between two subcarriers. One significant change from LTE to 5G NR is the use of a single SCS of 15 kHz in LTE, whereas the NR standard defines several SCSs. The NR numerology for 15 kHz is based on the exponentially scalable SCS as defined by $f\,[\text{kHz}] = 15 \cdot 2^{\mu}$, so the defined SCSs are 15, 30, 60, 120 and 240 kHz [9].

Table 4.1 provides a summary of 5G NR's numerologies and their characteristics. Not every numerology is supported by all frequencies and channel bandwidths. FR1 and FR2 support different frequencies, with FR1 supporting 15, 30 and 60 kHz SCS and FR2 supporting 60, 120 and 240 kHz. It can also be seen that the 60 kHz numerology only

supports data transmission and does not support the transmission of signals, while the opposite holds for 240 kHz which can only be used for signal synchronization [9].

Numerology μ	Subcarrier spacing	RB bandwidth	Slots per subframe	OFDM symbol length	Frequency range	Channel Bandwidth	Adoption
0	15 kHz	180 kHz	1	66.67 μs	FR1	5 - 50 MHz	Data & Sync
1	30 kHz	360 kHz	2	33,33 μs	FR1	5 - 100 MHz	Data & Sync
2	60 kHz	720 kHz	4	16,67 μs	FR1 & FR2	10 - 200 MHz	Data
3	120 kHz	1440 kHz	8	8,33 μs	FR2	50-400 MHz	Data & Sync
4	240 kHz	2880 kHz	16	4,17 μs	FR2	-	Sync

Table 4.1 *Numerologies defined in 5G NR* [9].

4.3.2 Physical Channels and Signals

The radio frame structure in 5G NR is differently structured in the time domain and the frequency domain. In the time domain, it is divided into radio frames, each with a duration of 10ms. Each radio frame is further divided into ten subframes of 1ms, with each one being divided into time slots, each one consisting of 14 OFDM symbols. The number of time slots the subframes have depends on the numerology used. In the frequency domain, the radio frame structure, defined for transmission bandwidth, is divided into resource blocks. A physical resource block (PRB) is made of 12 orthogonal subcarriers in the frequency domain and a slot in the time domain. An NR carrier is constrained to a maximum of 3300 subcarriers or 275 PRB. The carrier bandwidth cannot be bigger than 100 MHz in FR1 and 400 MHz in FR2. This is a notable bandwidth difference compared to LTE's maximum carrier bandwidth of 20 MHz [9], [20].

The resource grid consists of resource elements (REs), modulation symbols that carry information on a radio frame and are modulated on a single subcarrier in the frequency domain and a single OFDM symbol in the time domain. The information these elements carry is either in user data form, control data form, reference or synchronization signal form. The user and control data of these REs can form physical channels, while the reference and synchronization signals can form physical signals. These channels and signals are summarized along with their functions for 5G NR in Tables 4.2 and 4.3 for downlink and uplink respectively [9].

Downlink physical channels and signals	Abbreviation	Function
Physical downlink shared channel	PDSCH	Carry user data in downlink
Demodulation reference signal for PDSCH	PDSCH DM-RS	Channel estimation to demodulate PDSCH
Phase-tracking reference signal for PDSCH	PDSCH PT-RS	Track and compensate for phase errors in PDSCH
Physical downlink control channel	PDCCH	Carry control data in downlink
Demodulation reference signal for PDDCH	PDCHH DM-RS	Channel estimation to demodulate PDCCH
Physical broadcast channel	PBCH	Carry information required for initial access
Demodulation reference signal for PBCH	PBCH DM-RS	Channel estimation to demodulate PBCH
Primary synchronisation signal	PSS	Synchronisation, physical cell ID
Secondary synchronisation signal	SSS	Synchronisation, physical cell ID, signal quality measurements
Channel state information reference signal	CSI-RS	Downlink channel estimation, signal quality measurements and beam management
Positioning reference signal	PRS	Positioning measurements

Table 4.2 *Downlink physical channels and signals in 5G NR* [9].

Uplink physical channels and signals	Abbreviation	Function
Physical uplink shared channel	PUSCH	Carry user data in uplink
Demodulation reference signal for PUSCH	PUSCH DM-RS	Channel estimation to demodulate PUSCH
Phase-tracking reference signal for PUSCH	PUSCH PT-RS	Track and compensate for phase errors in PUSCH
Physical uplink control channel	PUCCH	Carry control data in uplink
Demodulation reference signal for PUCCH	PUCCH DM-RS	Channel estimation to demodulate PUCCH
Sounding reference signal	SRS	Uplink channel estimation and beam management
Physical random-access channel	PRACH	Carry user's initial access message to the BS

Table 4.3 *Uplink physical channels and signals in 5G NR* [9].

4.3.2.1 Synchronization Signals

In 5G NR, a cell represents the geographic area a BS covers and is a fundamental unit of the basic building block that provides wireless connectivity. Each cell is identified by a distinct 36-bit identifier referred to as Physical Cell Identity (PCI). UEs use PCIs to distinguish different cells during operations such as cell search, cell selection and other reselection procedures. The first operation a terminal performs when it is activated is

cell search. Cell search can also be applied to find neighboring cells and dynamically allocate resources from one cell to another. Attackers aim to disrupt this operation in order to deny access to many terminals [50], [51].

5G NR contains two synchronization signals, the Primary Synchronization Signal (PSS) and Secondary Synchronization Signal (SSS). These signals are used for frame, slot and symbol timing, but can also convey the PCI of a cell. 5G NR has 1008 distinct PCIs, since the PSS has only three possible combinations while the SSS has 336 combinations. The PSS is formed using a correlation m-sequence of length 127 which is mapped to a consecutive set of 127 subcarriers within a single OFDM symbol. An m-sequence is a type of pseudorandom binary sequence characterized by spectral flatness except for the DC term. This sequence establishes time synchronization and exhibits low correlation with other sequences, enabling the UE to differentiate between nearby BSs operating on the same carrier frequency. The correlation-based nature of synchronization signals enhances their resilience to basic interference. The SSS is also a sequence of length 127, mapped to a different OFDM symbol in the same subcarriers as the PSS. The SSS uses a Gold sequence, which is a combination of two m-sequences and has very low cross-correlation allowing a UE to distinguish between several nearby BSs on the same carrier even at low Signal-to-Noise-plus-Interference ratio (SINR) [9], [52].

4.3.2.2 Reference Signals

5G NR uses reference signals as pilot signals for channel estimation and channel equalization. 5G NR introduced four main reference signals, Demodulation Reference Signal (DM-RS), Phase Tracking Reference Signal (PT-RS), Sounding Reference Signal (SRS) and Channel State Information Reference Signal (CSI-RS). Unlike in LTE where reference signals where constantly exchanged, in 5G NR these reference signals are separated by the physical channel because there are no cell specific reference signals in 5G NR, and are transmitted only when it is required [52], [53].

DM-RS is a reference signal used to ensure the correct demodulation of the received data. The system can beamform the DMRS and transmit it only when it is required in either the downlink (DL) or the uplink (UL). DM-RS has a unique design and mapping for each of the 5G DL and UL physical channels. The DL signals are the Physical Downlink Shared Channel DM-RS (PDSCH DM-RS), Physical Downlink Control Channel DM-RS (PDCCH DM-RS), Physical Broadcast Channel (PBCH DM-RS), while the UL signals are the Physical Uplink Shared Channel DM-RS (PUSCH DM-RS) and the Physical Uplink Control Channel (PUCCH DM-RS), and are used for channel estimation to demodulate their respective DL or UL channels [9], [52], [53].

As the frequency of operation increases, so does the phase noise of a transmitter. This can cause phase rotation of all the sub-carriers in an OFDM signal and is known as common phase error (CPE). The Phase-Tracking Reference Signal (PT-RS) is used to minimize this effect at the transmitter and the receiver, which is especially common at mmWave frequencies. PT-RS is present in the Physical Downlink Shared Channel (PDSCH PT-RS) in the DL and in the Physical Uplink Shared Channel (PUSCH PT-RS) in

the UL and is used to track and compensate for phase errors in its respective physical channels [9], [52], [53].

CSI-RS is a DL-only reference signal that is used by the UE for channel and radio signal quality estimation and can also be used to select an optimal modulation scheme. Even though it is configured to a specific UE, CSI-RS can share its resources with other UEs. It can also be used for CSI-Acquisition and can assist in beam management and in MIMO operations [9], [53].

SRS is a UL-only reference signal that has similar functionalities to CSI-RS, since it is also configured to a specific UE and it is used in beam management and in Massive MIMO operations, but can also be used for estimation of the UL channel [9], [53].

4.3.2.3 Physical Broadcast Channel (PBCH)

The Physical Broadcast Channel (PBCH) transmission takes place in the same slots as the PSS and SSS, however, it spans more subcarriers than the two synchronization signals. In fact, if the carrier's frequency is below 3 GHz the PBCH occupies 240 subcarriers across 12 OFDM symbols, while if the frequency is above 3 GHz it occupies the, across 24 symbols. The PBCH carries essential information, also known as the Master Information Block (MIB), such as the subcarrier spacing, the position of DL reference signals and the DL control channel, that is necessary for a UE to attach to a cell [9], [52].

4.3.2.4 Physical Random-Access Channel (PRACH)

In the 5G NR random access procedure for a UE to connect to a cell, after receiving the PSS, SSS, and PBCH and synchronizing to the cell in time and frequency, the UE transmits a preamble over the Physical Random-Access Channel (PRACH). When the BS receives the preamble, it estimates temporal synchronization parameters and allocates the resources necessary to continue the communication with the UE. These parameters and resources are communicated back to the UE that transmitted the preamble. The BS also broadcasts possible time and frequency locations of the PRACH that a UE can connect to [9], [52].

4.3.2.5 Uplink and Downlink Physical Control Channels

In 5G NR there are two control channels, the Physical Downlink Control Channel (PDCCH) and the Physical Uplink Control Channel (PUCCH). The Physical Downlink Control Channel (PDCCH) is responsible for delivering control information to the UEs on a per-slot basis. It schedules downlink transmissions, uplink transmissions, modulation and hybrid-ARQ information. On the other hand, the Physical Uplink Control Channel (PUCCH) is utilized by the UE to transmit various control information to the BS. This information includes hybrid-ARQ acknowledgments, scheduling requests, and channel state information [52].

4.3.2.6 Uplink and Downlink Physical Data Channels

The 5G NR data channels that transmit and receive user data from the BS to the UE and vice versa are the Physical Downlink Shared Channel (PDSCH) and the Physical Uplink Shared Channel (PUSCH). The PUSCH carries UL user data and multiplexed control information. If the PUSCH wants to transmit multiple signals, it can use either 16 QAM, 64 QAM and 256 QAM to modulate the signals before it simultaneously maps them into resource blocks. The PDSCH is used to transmit DL user data and system information [9], [19].

4.4 Multiple Antennas

Multi-antenna techniques involve the use of multiple antennas at either the transmitter's side or the receiver's side, and even both sides, combined with signal processing. A communication system that uses multiple antennas at both sides is commonly referred to as a MIMO (Multiple-Input Multiple-Output) system. Multi-antenna techniques can be used to enhance system performance, improve system capacity, coverage, and provide higher per-user data rates. The use of multiple antennas in communication systems also improves link reliability, since it enhances resilience against fading effects in the radio channel, but can enable spatial processing as well, combining signals from or to multiple antennas to increase the Signal-to-Noise Ratio (SNR). Multiple antennas can also be combined, and if done correctly, their combination can suppress undesired signals, improving the Signal-to-Interference Ratio (SIR). Multiple antennas can transmit multiple data streams over several antennas utilizing the same time-frequency radio resource, thereby also enhancing spectral efficiency. These multiple data streams can either be transmitted towards a single UE, termed as single-user MIMO (SU-MIMO), or towards multiple UEs, termed as multi-user MIMO (MU-MIMO) [5].

Multi-antenna technologies have been in development for many decades to such a level that they were adopted by the mobile network industry, alongside hardware and software technologies that have greatly advanced as well. Before the introduction of advanced multi-antenna techniques, such as MIMO, LTE used simple antenna diversity techniques. Now, LTE contains an amplitude of multi-antenna techniques at its disposal, but it is 5G NR where multi-antenna techniques revealed their true potential supporting data-rates of minimum 100 Mb/s in wide areas, as well as a maximum of 20 Gb/s in hotspots and provide up to 3 times spectral efficiency or 100 times energy efficiency improvements compared to LTE-Advanced networks. However, the use of multiple antennas in 5G NR is different in low and high frequency bands [5], [20], [54].

In the low frequency bands, such as FR1, multiantenna techniques in NR are built upon the techniques used in LTE. Since the spectrum in these frequencies is congested, achieving higher data rates requires higher spectral efficiency. For this purpose, advanced techniques like spatial multiplexing are vital. Antenna technologies can also deal with the physical size limit on the number of antenna elements feasible in an array at lower frequencies by enhancing spatial domain utilization. NR also introduced a new

scalable CSI framework, which includes a high-resolution CSI reporting mode that improves MU-MIMO operation [5], [54].

In the high-frequency bands, such as FR2, where LTE was not originally designed for, multi-antenna techniques take on a more important role in system design. Here, NR introduced a set of procedures known as beam management, that ensure the alignment of the transmitter and receiver beams in order to establish highly directional transmission links between the UE and the BS, capitalizing on beamforming gains. In contrast to the lower frequencies, the main challenge at higher frequencies is not spectral efficiency but spectrum coverage. NR adopts a beam-centric design in the millimeter-wave spectrum, where data transmissions, control and broadcast signals are all beamformed. This differs from earlier cellular systems, where only data transmissions were beamformed. NR also introduced UE beamforming in millimeter-wave bands, since a bigger number of antenna elements can be fit within UE form factors. Analog beamforming is also prevalent in millimeter-wave bands for handheld devices due to their hardware limitations. Therefore, NR specifications also include support for analog beamforming procedures [2], [5], [2], [54].

4.4.1 Multiple-input multiple-output (MIMO)

MIMO operates on the principle of employing multiple antenna elements at both the transmitter and receiver ends. These antenna arrays are composed of smaller units known as antenna elements. Prior to the wireless transmission of a digital data stream, it is processed within a digital baseband unit, where it is encoded and modulated into an OFDM signal, among other physical layer tasks. That processed signal is converted into an analog form via digital-to-analog conversion on a radio frequency (RF) chain. The RF chain also handles tasks such as filtering and modulating the carrier frequency [9], [55].

Single-user MIMO (SU-MIMO) is illustrated in Fig. 4.12 and refers to the scenario where a single user device simultaneously transmits or receives multiple data streams. SU-MIMO techniques are best suitable for point-to-point communication and aim to enhance channel capacity and reliability by using space-time or space-frequency codes along with spatial multiplexing schemes. In an SU-MIMO transmission, the whole process of the transmitters and the receivers is coordinated, giving it the advantage of MIMO processing [2].

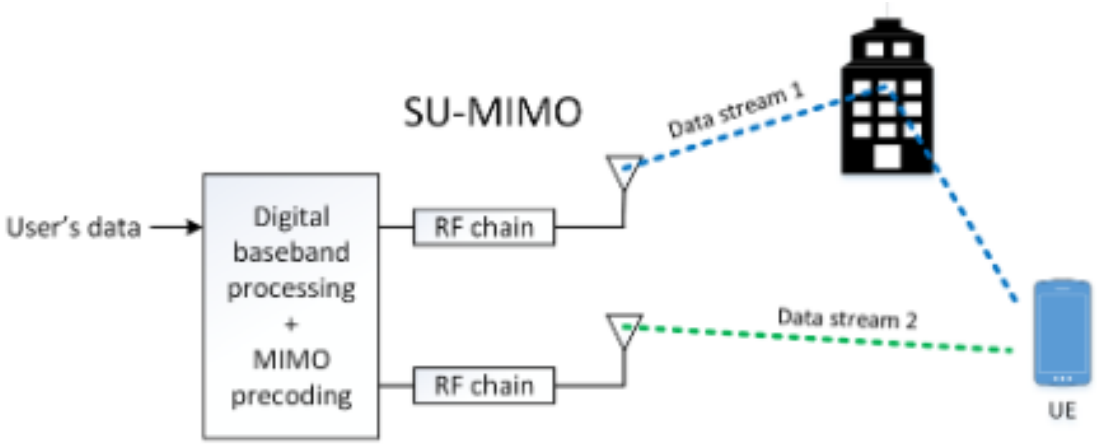


Figure 4.12 *Single-user MIMO (SU-MIMO) system* [9].

Within the physical layer's layer mapping procedure, MIMO recoding techniques are employed to distribute the data streams across the physical antennas. The number of distinct radio channels is limited by the number of antennas t at both the transmitter and receiver ends. For example, if there are two parallel data streams, the user equipment must be equipped with at least two receiving antennas. This configuration is referred to as 2x2 MIMO, because of the use of two antennas at both the transmitter and receiver, as depicted in Figure 4.12. In 5G NR, the maximum number of data streams per user is eight in the downlink (8x8 MIMO) and four in the uplink (4x4 MIMO) [9].

In multi-user MIMO (MU-MIMO), data streams are allocated among multiple users without any coordination between them. In MU-MIMO systems, the uplink and downlink channels are distinct. Downlink is the communication from the cellular BS to the UE, facilitating downloads, whereas uplink is the communication from the UE to the BS, facilitating uploads. In the uplink scenario, users transmit data to the BS over the same channel. The BS faces the challenge of distinguishing and separating the signals transmitted by individual users, which often requires multi-user detection methods to mitigate interference. In the downlink channel, the BS simultaneously transmits to a number of users over the same channel, causing inter-user interference by mixing signals intended for other users. While multi-user detection techniques can potentially help users overcome the multiple access interference, implementing such techniques at the receivers is often complex and challenging [9], [2], [55].

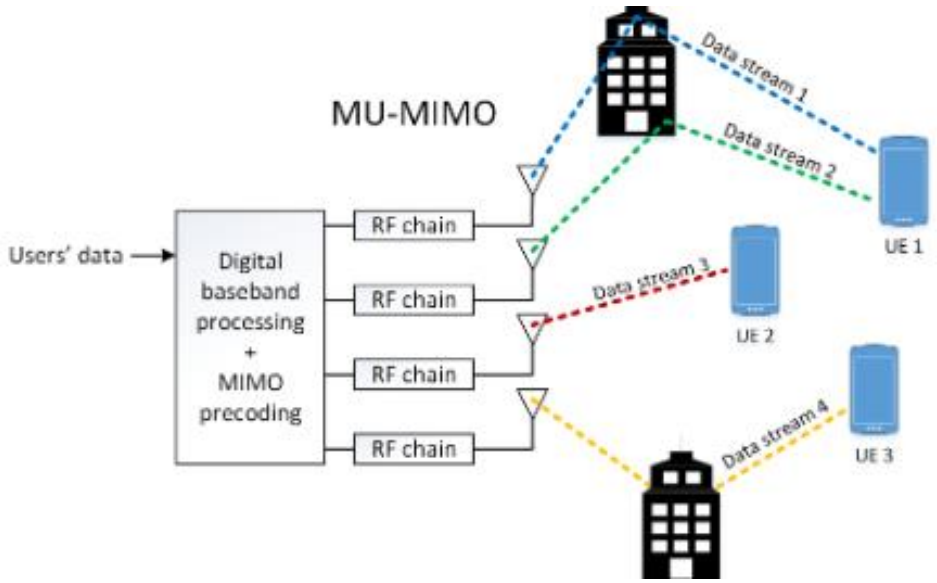


Figure 4.13 *Multi-user MIMO (MU-MIMO) system, with a 4x4 MIMO setup. The system has four transmitting antennas to transmit four data streams, as well as four receiving antennas* [9].

4.4.1.1 Mathematical description for MIMO communications

In a generic MIMO system, a MIMO transmitter consists of N_{TX} transmit antennas, a MIMO receiver consists of N_{RX} receive antennas, and there are $N_{RX} \times N_{TX}$ paths or channels between transmit and receive antennas [2].

The equation between transmitted and received signals for a specific subcarrier at a certain time point can be written as:

$$y = Hx + e, \qquad (4.10)$$

where *y* denotes the received signal, *x* denotes the transmitted signal, and *e* denotes the additive impairments. The impairments consist of thermal noise, distortions, and interference. The complex-valued matrix *H* represents the impact of the radio channel for the given subcarrier and time instance, and since multiple signals can be transmitted as well as received, this channel is a MIMO channel. The dimensions of the vectors and the matrix correspond to the number of transmit and receive antennas [20].

Bit Error Rate (BER) and Channel Capacity are two performance metrics usually used to evaluate the communication performance of MIMO systems [56].

4.4.1.2 Bit Error Rate

Bit Error Rate (BER) is a metric used to evaluate the reliability of digital communication systems by measuring the proportion of incorrectly transmitted bits to the total number of transmitted bits during the communication process. Its mathematical definition can be given as:

$$P_b = \frac{N_e}{N_b}, \qquad (4.11)$$

where N_e denotes the incorrectly transmitted bits, and N_b denotes the total transmitted bits [56].

4.4.1.3 Channel Capacity

The channel capacity depicts the maximum rate of information transmission that a communication system can achieve when it has a bit error rate of approximately zero. It is defined as the maximum mutual information between the input and output signals of the channel, indicating how much information is preserved from the transmitted signal to the received signal across the channel. The channel capacity is determined by optimizing the input distribution to maximize mutual information while considering the physical constraints of the channel and system power limitations. Channel capacity constraints the data transmission rate and is another metric for evaluating the performance and the effectiveness of communication systems. Mathematically, the channel capacity is expressed as:

The channel capacity is determined by optimizing the input distribution to maximize mutual information while considering the physical constraints of the channel and system power limitations.

$$C = max\ I\ (input; output), \qquad (4.12)$$

where the channel capacity is denoted by *C* and the mutual information between *x* and *y* is denoted by *I (x; y)*. If both the transmitter and receiver possess perfect CSI, a MIMO channel's capacity of an $N_{RX} \times N_{TX}$ can be captured precisely using the following equation:

$$C = log_2 \det\left(I_{N_{RX}} + \frac{\rho}{N_{RX}} HH^H\right), \qquad (4.13)$$

where ρ denotes the transmit SNR [56].

4.4.2 Beamforming

The 5G radio interface has a vast frequency spectrum, with different frequency bands being suitable for different use cases. The frequency bands above 6 GHz a high bandwidth and have a high absorption rate, restricting the geographical coverage a single cell can achieve. The implementation of omnidirectional antennas is challenging at these frequencies since the antenna to wavelength ratio results in more directive antennas than at lower frequencies. The solution to these challenges is the use of beamforming, which, due to the dimensions of the antenna elements, is quite easy to implement at these frequencies [11].

Beamforming is an analog, digital or hybrid technique used in communication systems to concentrate the radiated output power of an antenna in a specific direction. The result of this process is a beam and is facilitated using an array antenna composed of multiple antenna elements. The number of antenna elements in the array defines the width and the density of the created beam, with more antenna elements resulting in a narrower and more concentrated beam. Beams are used in a communication system

in order to improve signal power and signal quality and its beam's direction depends on the amplitude and phase of every antenna element in an array antenna [9].

4.4.2.1 Analog Beamforming

Analog beamforming involves the control of the antenna elements directly by the antenna's hardware. This setup typically includes a single radio frequency (RF) chain and multiple phase shifters that control the phase of each element within the antenna array. Analog beamforming operates at RF frequencies or an intermediate frequency and can use active beamforming antennas to steer the beam to virtually any angle, providing flexibility in beam control. Although in theory analog beamforming can support many antenna elements, in reality a system with only analog beamforming is limited to transmitting one data stream at a time. This limitation arises because all antenna elements are connected to a single RF chain, resulting in the same only phase-shifted signal being transmitted to each element [9], [2], [57].

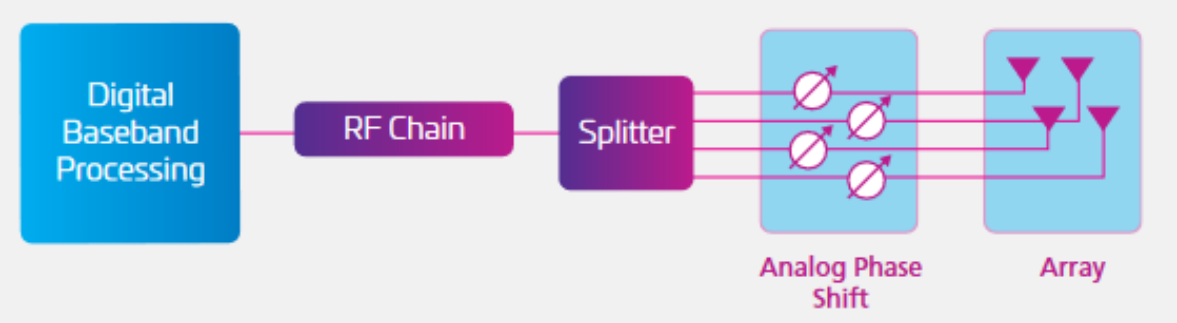


Figure 4.14 *Analog beamforming architecture* [57].

4.4.2.2 Digital Beamforming

In contrast to analog beamforming's limitation to one RF chain even with the use of large number of antenna arrays, digital beamforming can in theory overcome the scalability limitations of analog beamforming and support as many RF chains as there are antenna elements. Its antenna element has its own unique signal and therefore must have its own RF chain as well, where adjustments to the phase and amplitude of the signal take place. Through digital control of each antenna element, the system can form multiple beams, with each beam carrying a distinct data stream [9], [2], [58].

Digital beamforming is used in 5G NR to provide high flexibility in the direction settings of one or multiple beams, allowing for user-specific beams that can track users' movements. However, even though despite digital beamforming's theoretical performance superiority compared to other beamforming architectures, it may not always be the most practical solution for implementation. The large number of RF chains alongside high complexity in signal processing can increase costs, elevate energy consumption, and complicate integration in mobile devices. As a result, digital beamforming's best use is in base stations where performance needs outweigh mobility concerns [9], [2].

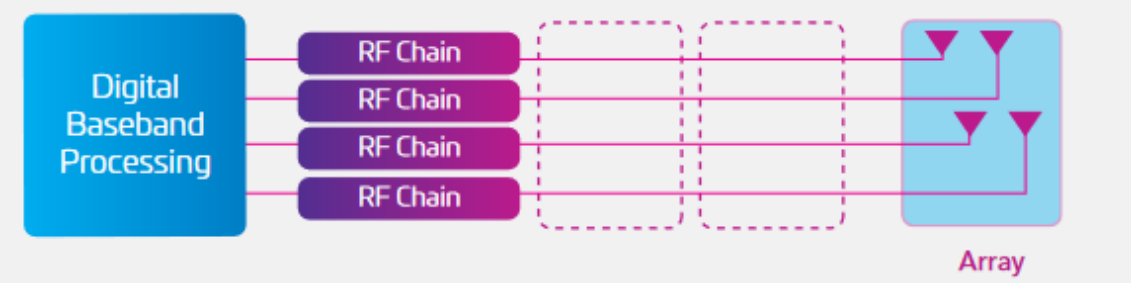


Figure 4.15 *Digital beamforming architecture* [57].

4.4.2.3 Hybrid Beamforming

Hybrid beamforming is a combination of both digital and analog beamforming, employing digitally controlled RF chains, splitters, and analog phase shifters. It uses a small number of antenna elements in the analog domain to narrow of the beams while also using less RF chains, resulting in lower power consumption and overall reduced system cost. At the same time, digital beamforming is used to provide flexible beam steering and to support multi-user operation. Hybrid beamforming is particularly well-suited for applications where digital beamforming is optimal, but cannot be fully implemented due to size and power constraints [9], [57], [58].

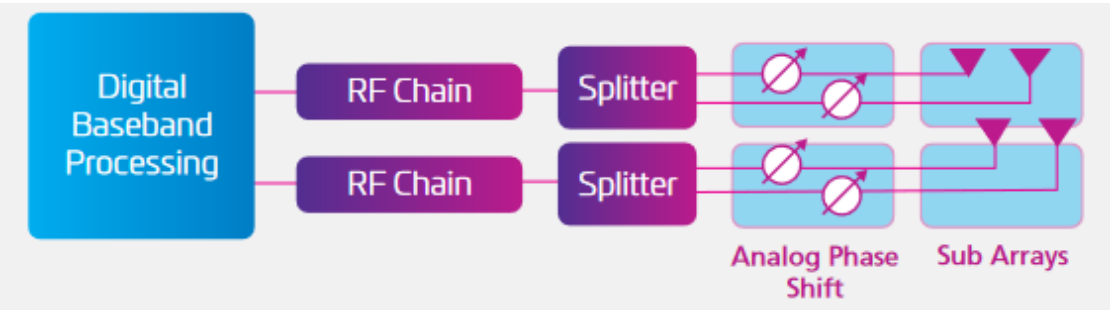


Figure 4.16 *Hybrid beamforming architecture* [57].

4.4.3 Spatial Multiplexing

A transmit beamformer operates by taking an information-carrying symbol stream and directing it to the antenna elements of an array through multiplication with a beamforming vector. This process results in transmission along a specific spatial direction, known as a spatial signature. Spatial multiplexing is the multiplexing of different spatial signatures on the same time or frequency resource in order to increase the total data rate of the system. Depending on if these symbol streams are directed on a single device or different devices, we are referring to either SU-MIMO or MU-MIMO respectively [20].

For MU-MIMO, where the direction of all UEs is different from each other in terms of angles as seen from the BS, spatial multiplexing is achieved by pointing a narrow beam pattern toward each UE. The UE of interest receives a strong signal while some interference incurs on the other UEs [20].

In SU-MIMO spatial multiplexing, all UEs are effectively aligned in the same direction as perceived from the BS and relies on the distinct propagation paths of the propagation channel. Each data stream with a corresponding spatial signature is directed towards a particular propagation path. These data streams are then separated on the receiver side by means of spatial filtering in case of different arrival angles [20].

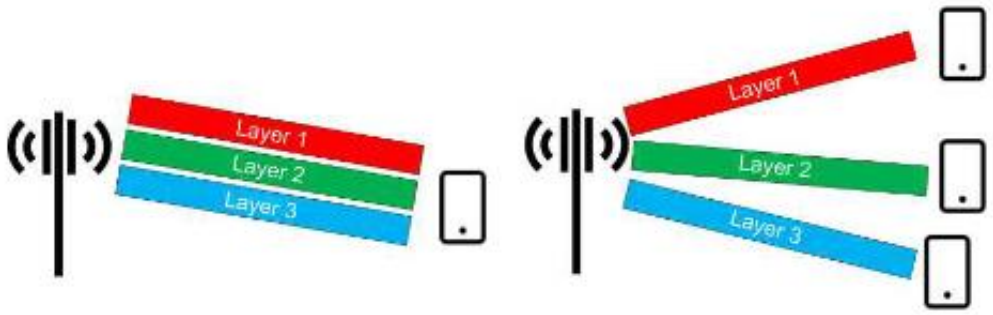


Figure 4.17 *Spatial multiplexing of three different data streams (layers) from a single UE (SU-MIMO) and from multiple UEs (MU-MIMO) on the same time or frequency resource* [20].

4.4.4 CSI Acquisition

The known channel properties between the transmitter and receiver in a wireless communication system are referred to as Channel State Information (CSI). These properties include details about signal propagation in a communication link, as well as factors such as scattering, fading, and power decay based on distance. In 3GPP in particular, CSI refers to the reports from the communication channel about these channel properties. Most of the channel properties are actually based on the downlink transmission rather than the entire channel and are measured based on some reference signals that are transmitted in the downlink. In NR, this downlink reference signal used for computing CSI is the CSI-RS. A CSI report that occurs at standard time instants is defined as periodic, while it is aperiodic if it only occurs when it is requested. A CSI report can also be semi-persistent if it is transmitted periodically until further notice. Since CSI can often be a limiting factor on the performance of multi-antenna systems, these advanced systems, such as MU-MIMO, require high quality CSI in order to provide reliability and high data rates in their communication. The performance benefits that CSI offers may diminish if a system cannot obtain the necessary CSI quality [5], [59].

CSI estimation typically occurs at the receiver's side by transmitting predefined reference signals known to the receiver and correlating them with the corresponding reference signal before it is sends them back to the transmitter. Therefore, CSI can be easily acquired at the receiver (CSIR) while it is more difficult at the transmitter (CSIT), resulting in them possibly having different values. Two alternatives for CSIT acquisition are feedback-based and reciprocity-based CSI acquisitions. Feedback-based CSI is measured based on the feedback received from a node that has estimated the CSI on the transmitted reference signals. Reciprocity-based CSI, on the other hand, depends on the measurements of the received reference signals [5], [59].

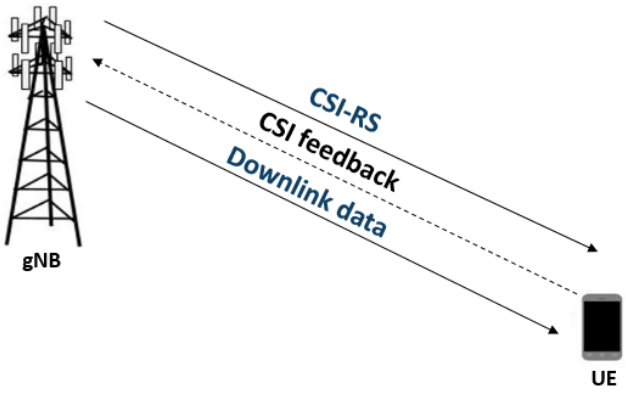


Figure 4.18 *An overview of CSI-RS transmission. The UE uses the CSI-RS to measure the CSI feedback and reports it to the BS. The BS receives that feedback and transmits the according downlink data to the UE* [60].

4.4.5 Massive MIMO

Massive MIMO technology is widely used in wireless communications, such as 5G NR and IoT, for the benefits it offers. One of them is channel hardening, which provides communication quality almost equal to that of non-fading channels, even with the increase in size of the antenna array. Another feature of Massive MIMO is favorable propagation, a phenomenon where the channels of different users become more and more orthogonal with each other in the spatial domain while the number of antennas at the BS increases, reducing inter-user interference and improving spectral efficiency. The latter, alongside data transmission speeds, can be further improved by the efficient utilization of bandwidth resources from the system. Massive MIMO can also process more data streams compared to traditional MIMO systems, improving network capacity. Massive MIMO can also use beamforming techniques for power concentration in desired directions and to minimize ISI, thereby enhancing signal quality and reliability [56].

Massive MIMO builds upon MU-MIMO, being able to use an even larger number of antennas for both transmitting and receiving signals. Unlike traditional MIMO systems, which are typically configurated with tens-of-antennas, Massive MIMO systems can operate using hundreds or even thousands of antennas and can also utilize beamforming to enable even more simultaneous data streams from multiple users. These data streams are then spatially separated at the BS due to the different angles they arrive in and their different propagation channels. Massive MIMO, however, is still required to operate in a multipath propagation environment in order to be able to provide multiple SU-MIMO data streams [9], [56].

4.4.6 Base Station Antennas

Because of 5G NR's large frequency spectrum from below 1 GHz up to 100 GHz, the antenna systems are designed differently for different frequency ranges. BSs can use different types of antennas depending on the frequency band and the services they are required to provide [5].

In the lower frequencies in 5G NR, where wavelengths are longer in size and extended coverage is necessary, antenna elements in a group antenna have bigger distances between them. A big enough size of wavelength, however, could cause antenna elements to have such distances between them an oversized physical antenna would be created at the base station. In practice, antennas operating at frequencies below 1 GHz support only up to 8x8 MIMO [9].

In the higher frequencies in 5G NR, mobile networks require hundreds of antenna elements in an array antenna. Massive MIMO antennas are well suited for these frequency ranges due to them being able to serve multiple users transmitting or receiving data streams simultaneously [9].

Base station antennas that operate at mmWave frequencies can be even smaller in size, despite the fact the increased number of antenna elements to thousands. Massive MIMO antennas are used by commercial mobile networks for these frequencies because they can support large bandwidths while simultaneously serve many UEs [9].

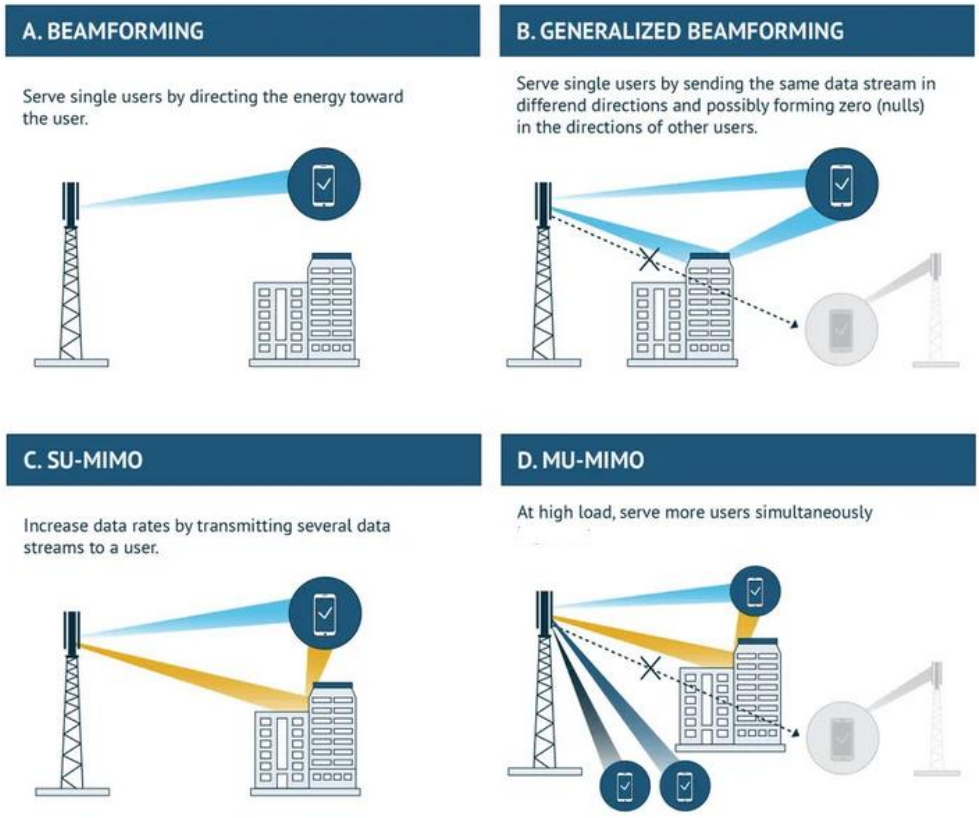


Figure 4.19 *Base stations using different 5G NR technologies* [61].

4.4.7 User Equipment Antennas

5G UE includes a variety of categories like smartphones, tablets, laptops and mobile hotspot routers. Additionally, there are customer premises equipment (CPE) for 5G fixed wireless access (FWA) as well as different types of 5G modules for industrial applications, Internet of things (IoT), or vehicle- to-vehicle (V2V) applications [9].

Antennas in handheld devices are typically engineered for near omni-directional coverage to account for incident waves arriving from any direction. At conventional cellular frequencies, the size of a hand-held device usually aligns itself with the carrier wavelength. However, at millimeter-wave frequencies, handheld devices become relatively larger compared to the wavelength, making it more challenging to design omni-directional antennas for these devices. On the contrary, the physical size reduction of antennas paired with increasing frequency and the compact form factor of handheld devices make it possible to integrate millimeter-wave antenna arrays into devices. Omni-directional coverage at high frequencies can be acquired with the use of several directive antennas to cover different angular sectors. These antennas can either be discrete with a fixed beam pattern or antenna arrays that capable of dynamic beamforming. Each antenna requires its own transceiver in order to acquire instant omnidirectional coverage. Alternatively, a single transceiver can be switched to the antenna offering the best performance at any given time [5].

These devices must take measures against signal difficulties. In the case of signal blockage by the user putting its hand or finger over an antenna, the device must have antennas at different locations that can take the blocked antenna's place. Similarly, in the case of other nearby obstacles such as the user itself, other people or cars, the device has different angular coverages of the antennas so it can swiftly switch direction to find an alternative propagation path, while also enhancing spatial multiplexing. Signal loss from sudden movements and rotation of the device can be mitigated with the use of dynamic beamforming, only if the device is equipped with antenna arrays. The more suitable type of antenna arrays to be used in handheld devices are analog arrays because they comply to the strict regulations on cost and power consumption in handheld devices. The support of analog beamforming in handheld devices presents one of the new major challenges in 5G NR within the multiantenna area [5].

The support of the large frequency spectrum of 5G NR poses another difficulty in its implementation in handheld devices and the design of the antennas these devices will use. 5G is used alongside existing communication channels such as 4G, 3G, 2G, and Wi-Fi, increasing the number of antennas used in these devices. At 5G frequencies below 6 GHz UE antennas largely resemble those used in 4G LTE. On the other side of the spectrum, at mmWave frequencies above 24 GHz, devices need to integrate the antenna within their housing components, cover, or screen. The problem is that, in these frequencies, these components are no longer electrically thin and can significantly impact the radiating performance of the antenna [9].

4.5 Channel Coding

Forward error correction (FEC) schemes are essential parts of digital communication systems because they provide resilience against noise as well as other channel uncertainties, including imperfect channel-state information. The effectiveness of FEC designs is assessed based on factors like area, power consumption, and coding performance, making their design particularly difficult to cover all the different use cases of

5G NR. Another important part of wireless communications is channel coding. For 5G NR, channel codes are required to provide variable code rates and lengths for control information and for user data, while the latter can also benefit from hybrid automatic repeat request (HARQ). To meet the requirements of 5G communication, channel coding uses polar odes and low-density parity check (LDPC) codes [5], [62], [63].

Low-density parity-check (LDPC) code is a linear error correcting code used for the transmission of messages over a transmission channel impacted by noise. LDPC codes can have a noise threshold big enough so the probability of information lost during transmission can be very low. The decoding of LDPC codes can be achieved in time linear in their block length [64].

Polar code is another linear error correcting code that is used to transform the physical channel into virtual sub-channels. These virtual channels may or may not be reliable, with the data bits being allocated to the channels with the highest reliability. Polar codes can be easily encoded and decoded, which makes them power efficient [65].

NR employs LDPC codes for the data transmission for mobile broadband (MBB) services and polar codes for the control signaling. LDPC codes are attractive from an implementation perspective, especially at multigigabits-per-second data rates. Unlike the LDPC codes implemented in other wireless technologies, the LDPC codes considered for NR use a rate-compatible structure. This allows for transmission at different code rates and for HARQ operation using an incremental redundancy [5].

Figure 4.20 depicts an example of channel coding. Given an input signal *X*, when that signal passes through the channel encoder it becomes signal *X'* of longer length than signal *X*. The signal *X'* is then transmitted through a channel impacted by noise and has become signal X'_{no} when it reaches the channel decoder. The process ends with the decoder recovering the original input signal *X* from the corrupted message X'_{no} [66].

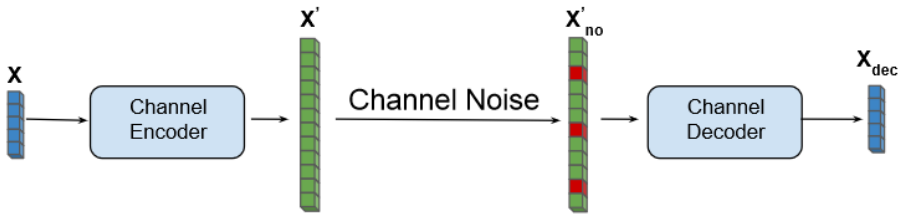


Figure 4.20 *Channel encoding and decoding of signal X* [66].

5. 5G System Security

The 5th generation of wireless communication systems is not simply an evolution of the legacy 4G cellular networks, but rather it is a system with many new service capabilities. 5G supports higher density of mobile broadband users than 4G LTE, Device-to-Device (D2D) Communications, Internet of Things (IoT) and massive Machine-Type Communications, while also providing lower latency and lower energy consumption. 5G is the backbone for a number of everyday network services and, like previous wireless networks, it is targeted for its security vulnerabilities. For this reason, 5G requires an adequate security level that can deal with the ever-expanding security threat landscape of 5G communication [7], [3].

5.1 Security Evolution in Wireless Networks

Wireless communication networks made their appearance in the last years of the 90th century, around 1980, with the first generation of wireless networks, 1G. Several generations of mobile networks have since been introduced, each one offering innovations and dealing with the weaknesses of the previous generations. These networks are still being developed today, with the latest generation, 5G, playing a critical part in everyday communication services. Alongside new features, every generation of cellular network has a more complex layer of security that keeps evolving to deal with the new security threats that appear [67].

5.1.1 Security in 1G

The 1G cellular system was first introduced in Japan in 1979 and was based on analog signal processing and was used purely for voice services. Due to it being the first iteration of a wireless network, however, 1G had many shortcomings. It was limited in coverage, with its range being restricted to single countries, but also offered really low speeds, with a maximum speed of 2.4 Kbps. Its service quality during calls was not the best, it couldn't handle a large number of users, it didn't make good use of its spectrum and also didn't include roaming services. Because it was based on analog communication, 1G networks also didn't provide any security parameters. 1G didn't allow for encryption, essentially giving attackers free access to the information of the devices and the information of the call. Scrambling was the first security measure introduced in 1G, and thus in wireless networks, and was developed to protect users against eavesdropping [67], [6].

5.1.2 Security in 2G

As the name suggests, 2G was the second generation of mobile networks and was introduced in 1991, a decade after 1G's appearance. 2G provided the ability to send text messages via SMS (Short Message Service) alongside voice communication. Four security measurements were also introduced in 2G, anonymity, authentication, signaling and user data protection. In order to provide anonymity to users, 2G communication systems use temporary identifiers that make it challenging for attackers to identify the

actual user of the system. The legit identifiers are used only when the device is powered on before the issuing of a temporary identifier. The network operators identify legit users using authentication mechanisms. One such mechanism used in cellular communications is cryptography, where secret keys are exchanged to provide a confidential communication between users. The authentication of legit users, alongside signaling and user data protection, was carried out through encryption with the use of Subscriber Identity Module (SIM) devices, physical devices that stored a cryptovariable used in the authentication process [67], [6].

However, 2G communication systems weren't without their security vulnerabilities. Only the network operators have the ability to authenticate users, allowing attackers to impersonate a legit operator to attack the user, since the users couldn't themselves authenticate operators. Furthermore, communications channels were falsely deemed as secure, so GSM encryption was only used on the radio interface, leaving channels vulnerable to DoS attacks and eavesdropping, while SMS were also very vulnerable to attacks. Finally, 2G standards were fixed on its release and 2G systems couldn't to upgrade their security functionality during this generation's lifespan [67], [6].

5.1.3 Security in 3G

Similar to 2G, the third generation of mobile networks, 3G, was introduced in 2001, a decade after 2G's first appearance. The main reasons for 3G's development was the need for higher data rates, but it also allowed for new services such as global roaming, highly improved voice quality, video calls and even video streaming via mobile networks [67], [6].

The 3G standard identified all the security vulnerabilities of 2G systems and included solutions for them, while also adding more security features. 3G's security architecture consists of five security features, network access security, network domain security, user domain security, application security and visibility and configurability security. These design implementations provided the UE a secure access to the network and could mitigate new threats that were detected by the system. However, 3G mobile communication systems were still targeted by attackers, even more that prior systems, due to the sheer number of devices that supported the 3G network. UEs were prone to denial of service, eavesdropping and impersonation attacks and gave attackers access to users' personal information. Specifically, attackers could impersonate either another user, a BS, or a network. Temporary identities or permanent encrypted identities were used to protect the users from impersonating attackers. User identification should be time restricted and data that contains the user's identity must be encrypted [67], [6].

5.1.4 Security in 4G

4G is the fourth generation of mobile networks and was introduced in 2009, following the same pattern as previous wireless networks of being introduced a decade after their predecessor. The entire 4G network's implementation was based on IP. In fact, all

mobile devices that supported 4G LTE used an E2E (End to End) architecture based on all-IP. 4G networks provided even higher data rates, up to 100 Mbps, allowing for fast data sharing between UEs [67].

4G also improved in the security department, introducing new features that could mitigate attacks detected in 2G and 3G networks. Such features were two new cryptographic algorithms EPS Encryption Algorithms (EEA) and EPS Integrity Algorithms (EIA), which used 256-bit keys, which have twice the length of the 128-bit keys used in 3G. Unlike previous generations of mobile networks, control and user planes traffic used different algorithm and key sizes. 4G networks use the Authentication and Key Agreement (AKA) protocol as a user and network authentication mechanism, whereas integrity and replay protection in 4G networks was guaranteed with the use of the NAS (Non-Access Stratum) and RRC (Radio Resource Control)-signaling protocol. IPsec protocols were used to encrypt the 4G backhaul traffic [7], [67].

However, due to the IP based 4G architecture that connects the 4G network to the internet network, 4G can be easily targeted by attackers that originate from the Internet. The 4G cellular network can be affected by attacks that target the functionalities of the IP protocol, such as IP address spoofing, User ID theft, Denial of Service (DoS), TCP SYN flood attack and intrusion attacks. Devices that supported the 4G networks required higher computing power, which proved to be ideal sources of attacks towards the cellular network. 4G networks also supported Wi-Fi and WIMAX and inherited all their security weaknesses due to their lower security levels [7], [67].

5.2 Security Services in 5G

The introduction of 5G brought forward new technologies, use cases and improvements in the wireless network structure. However, alongside those new features came certain security requirements in the form of security services such as authentication, confidentiality, etc., in order to protect users from the threat of attackers [3].

One such service is *authentication*, which is used to confirm the identity of the communicating entities, such as a UE, a BS or even a service provider, before the communication between these entities initiates. Authentication mechanisms are required to be both fast and effective in order to meet the 5G requirements of high data rates and low latency [3].

Another important security requirement in the 5G standard is *confidentiality*. In fact, there two types of this service, data confidentiality and privacy. The former is the requirement of ensuring user data are safe from attackers during transmission by preventing access to them from unauthorized users. Privacy, on the other hand, protects the information that could be obtained by monitoring a legitimate user's activity in the network. Traffic patterns could reveal to attackers information such as a user's location, health monitoring data, etc. Confidentiality in 5G applications can be achieved with the use of encryption algorithms, such as the symmetric key encryption algorithm that

uses a unique key known only to the transmitter and the receiver which encrypts and decrypts data [7], [6].

Integrity is used in the 5G standard to protect the authenticity of the data, ensuring the data are received in their correct state and are not modified, deleted or duplicated by unauthorized entities during transmission. Data integrity is important in the wide variety of use cases supported by 5G, especially when some of them have become a part of peoples' daily lives, such as certain time schedules or health information. It can also provide indications that a malicious user has tried to or managed to tamper with the user's data [3].

Network availability in 5G is the metric that shows if an authorized user can have access to the network's resources and services. The network should be accessible to legitimate users if it is requested, regardless of time and place. Availability also is used to measure the network's performance and its robustness against different types of attacks [7], [3].

Other security dimensions of 5G are *access control* and *non-repudiation*. The former is used to determine which users have access to the network's resources, services and applications, while blocking access to all other users. This way, no non-legit user can gain access to the network's infrastructure. The latter is a combination of authentication and integrity, authenticating the identity of a user that performed an action and then ensuring the integrity of that action [6].

5.3 Security Improvements in 5G over 4G

The fifth generation of mobile communication networks, 5G, was first deployed worldwide in 2019, a decade after the previous generation was introduced, and offers new features and services such as spectrum sharing, increased spectrum and beamforming, reducing interference while also serving a larger number of users at the same time. 5G can transfer high volumes of data with low latency. Another feature of 5G is network slicing, which improves the service quality over the physical channels and signals by slicing the network in multiple virtual sub-networks. Dynamic scaling of network functions can be achieved with the virtualization of the RAN and the 5G Core [10].

The biggest difference between 4G and 5G networks is the larger frequency spectrum used in 5G with the inclusion of mm-wave frequency bands. This extension of frequencies, however, doesn't by itself make 5G any safer for users, except from the fact that it is now more challenging for an attacker to jam the entire spectrum a 5G UE covers. Instead, 5G introduced other security features that were developed and implemented during its creation [11].

A lot of the new security features of 5G have been implemented in 5G SA, while security measures from previous generations networks are also present. 5G SA architecture took measures to protect *the privacy of users*, by encrypting the unique identifiers of the devices to defend against non-legit BSs. Privacy of UEs is also ensured by authenticating both the users and the network, while also protecting the control and user

traffic data and signals. Finally, 5G SA can restrict previous generations of radio technologies that a device may use, such as 2G and 3G, in order to mitigate their weaknesses [10].

5G SA architecture also includes measures to ensure the *security and privacy of RAN*. 5G SA architecture uses a large number of antennas combined with beamforming in order to reduce interference and mitigate eavesdropping attacks. RAN is also separated into distributed units (DUs) that are located near the antenna, and centralized units (CUs) that placed in a secure location and contain restricted information [10].

Network Slicing and Virtualization are two of the most important features introduced in 5G. Network slicing can isolate the control and user plane traffics but can also do the same for security attributes for some user classes. At the same time, network virtualization allows for quick reconfiguration of the network in order to defend against attackers [10].

Another improvement of 5G SA over 4G is the inclusion of *roaming security*. Control plane security is ensured with the use of security gateways during roaming interconnections. A connection between two networks also contains mechanisms that protect the user plane traffic [10].

Authentication in 5G SA is open to the RAN. In fact, the same authentication mechanisms are used by both 3GPP and non-3GPP access networks. Home networks also have the ability to authenticate UEs and BSs and are able to detect and protect users from non-legit BSs. For the *security of the Core Network*, 5G SA can use authentication and encryption mechanisms of other layers, such as the transport layer of the network [10].

5.4 5G Threat Landscape

Currently, 5G plays an important role in all aspects of society's daily lives, from vehicles and home appliances to businesses and health care. Due to its importance, however, successful attacks on 5G can have a big impact on the network and its infrastructure. This is the major challenge both present and future networks face, as new threats and security vulnerabilities keep on being introduced [6].

5.4.1 Attacks and Threats in 5G Wireless Networks

There are different types of attacks that can target a communication system. In an *active attack*, the attacker is in direct communication with the target system or network with the aim of altering transmitted information and disrupting the network's functions. Active attacks can compromise the integrity and the availability of the targeted system or network. A *passive attack*, on the other hand, monitors the communication of the target system, since the attacker is not in direct contact with its target, and aims to intercept private users' data, threatening the confidentiality of the network [68].

5.4.1.1 Eavesdropping

Eavesdropping is a passive attack method used by a malicious user to intercept messages from legitimate users without disrupting the normal communication. Because of its passive nature and the fact that eavesdropping can be performed without any signal transmission from the attacker's side, it is challenging for legitimate users to detect the origin of the attack [3], [69].

Countermeasures against eavesdropping include the encryption of signals when transmitted over the radio link, which makes the direct interception of signals more challenging, while channel coding with LDPC codes can also be used [3], [69].

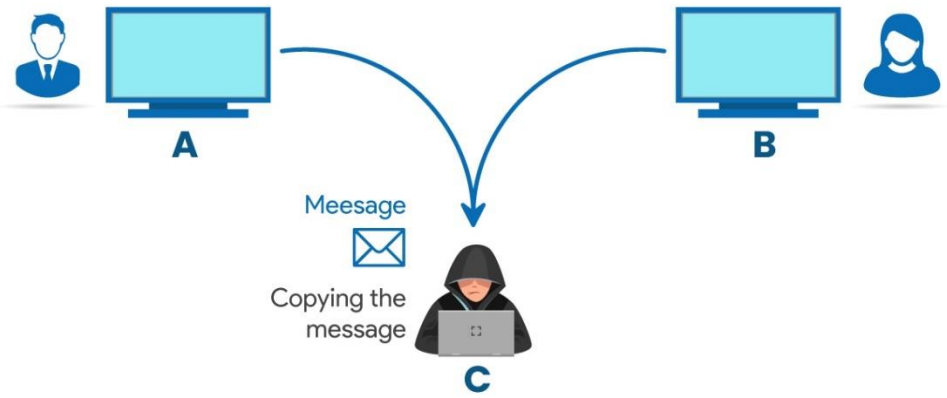


Figure 5.1 *Passive eavesdropping attack* [70].

5.4.1.2 Traffic Analysis

Another passive attack in wireless networks is traffic analysis. Even though this attack cannot by itself disrupt the legitimate communication, it can be used to monitor the users' activity in the network and discover their location as well as other communication patterns. Such information about the users can prove to be useful for other types of attacks [3].

5.4.1.3 Jamming

Jamming attacks are different from eavesdropping and traffic analysis attacks because they can fully disrupt a legitimate communication. Jammers can cause interference in the communication channel in the form of noise, decreasing its Signal-to-Noise Ratio (SNR). Other forms of jamming attacks are spoofing and sybil attacks. Spoofing attacks transmit deceiving signals that mimic the identity of a legit user, while sybil attacks use a single malicious node to operate a number of non-legitimate identities in the network simultaneously. Jamming attacks, their effects and their mitigation will be analyzed more in Chapter 4 [3], [69].

5.4.1.4 Man-In-The-Middle

Man-In-The-Middle (MITM) is an active attack that aims to compromise 5G network services such as confidentiality, availability and integrity. MITM attacks don't disrupt

the communication channel, instead they intercept transmitted messages, modify them and replace them with the legit ones. MITM attacks can be mitigated with the use of a mutual authentication system between the two endpoints of the communication channel [3].

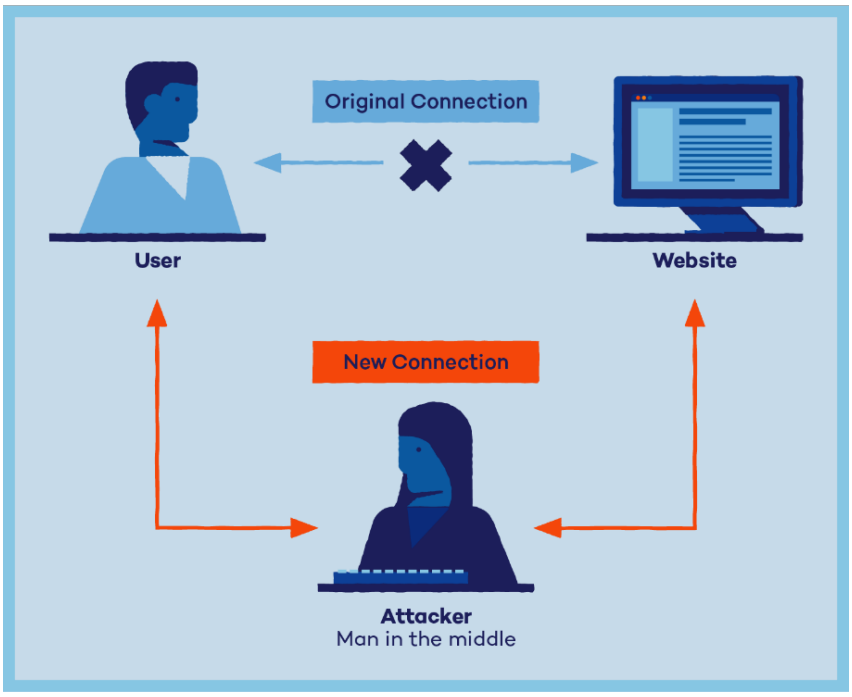


Figure 5.2 *Man-In-The-Middle attack* [71].

5.4.1.5 DoS and DDoS

Denial-of-Service (DoS) attacks are active attacks that are used to compromise the availability of the 5G network by overloading it with a large number of devices and a large number of data transmissions simultaneously, exhausting the network's resources. Distributed Denial-of-Service (DDoS) attacks are DoS attacks that use more than one adversary. DoS attacks can also fake overload the network to trick the legitimate users and disrupt communication. Attackers can use the large number of devices 5G wireless network support to target UEs or even the network itself. DoS attacks can be prevented with the use of authentication mechanisms to ensure the network is used by legitimate users, as well as some mechanisms that will prevent the network overload and limit services for the problematic devices [7], [3].

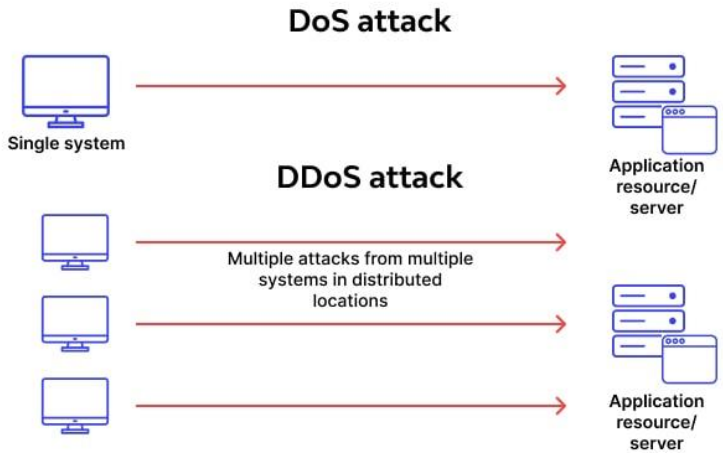


Figure 5.3 *DoS and DDoS attacks* [72].

5.4.1.6 Network Slicing

Attackers can also target specific network components and functionalities, such as network slicing and the MEC. Attacks that target the network slicing architecture of 5GC network can cause denial of service to some slices and exhaust their resources. At the same time, communication between slices in the network is no longer secure, since attackers can launch impersonator attacks against one or more slices. UEs that use the services of the affected slices put their own security in risk. For this reason, security requirement for 5G is the secure communication between all slices and the interfaces between them and the mutual authentication between all slices in the network [7].

5.4.1.7 MEC

Attackers can use the MEC as a security tool for them, making it provide security services to unauthorized users rather than authorized ones, thus allowing these users with malicious intensions access to the network and influence on it. The MEC environment can also be targeted with user plane attacks. This threats can be mitigated with the use of authentication mechanisms, so only legitimate users can have access to the MEC security services. Network operators must also limit to a degree the distortion of the network [7].

5.4.2 Security for Technologies of 5G Networks

The 5G infrastructure introduced four major new core technologies, Heterogeneous Networks, massive MIMO, Device-to-Device Communication and Software Defined Networking. These technologies have brought along significant improvements, but they are not without any security vulnerabilities.

5.4.2.1 Heterogeneous Networks

A heterogeneous network, or HetNet, combines variously powered radio nodes to cover an area efficiently, even if it has high user density or poor reception, ensuring high throughput for numerous devices. Nodes may have different transmission power, coverage area and radio access technologies from each other. HetNets allow a large

number of devices to connect to the network and transfer big amounts of data. Because endpoints connect to the node that has the highest Signal-to-Interference-plus-Noise Ratio (SINR), UE are vulnerable to eavesdropping and the location of the UE could be discovered. For this reason, randomness was added to the SINR to prevent attackers from discovering the location of the UE or even the network [15], [3].

A HetNet contains both high powered nodes and low powered nodes, which could overlap in high-density areas and cause load balancing problems. Specifically, the high density of small cells in HetNet can cause significant performance issues because of the many handovers that occur between different cells. It is important to monitor and balance the loads of these two types of nodes in a HetNet in order to prevent performance drops [15], [3].

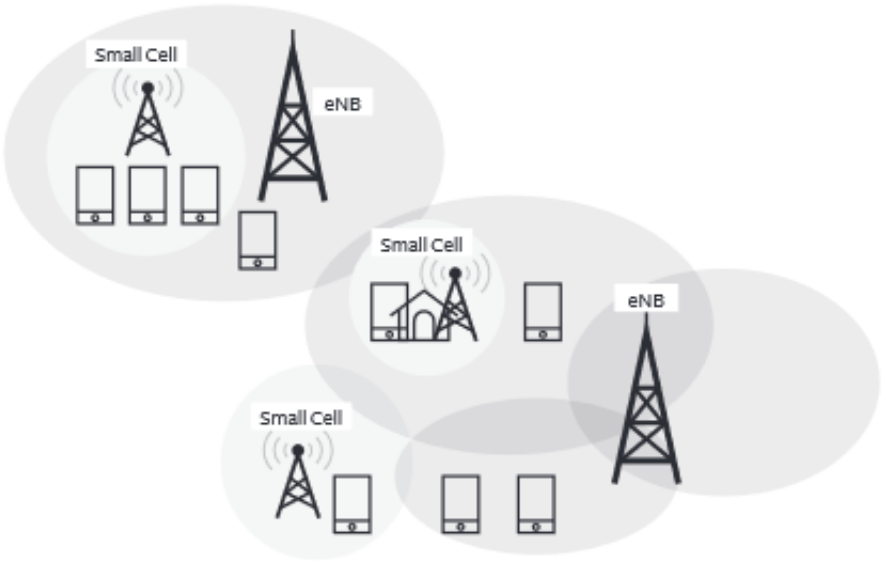


Figure 5.4 *Illustration of a heterogeneous network* [15].

5.4.2.2 Massive MIMO

Multiple Input Multiple Output (MIMO) architecture consists of several antennas in order to transmit multiple streams of data in wireless communications. Massive MIMO and its use of beamforming provides low latency over the air interface, increased capacity of frequency resources and more resiliency to interference, intentional or not. Due to the large number of antennas, massive MIMO provides security in the communication, since the signal can be transmitted without being decoded by an attacker, protecting users from eavesdropping [15].

Jamming massive MIMO systems can disrupt the channel estimation process, lowering the performance of the target system and allowing jammers to launch attacks against other unsuspecting users. These attacks can be assessed through analysis, simulations, and real-world experimentation and can be very dangerous if the attacked MIMO system doesn't have techniques for accurate channel estimation while under jamming [73].

5.4.2.3 Device-to-Device

Another feature of 5G is Device-to-Device (D2D) communications, which allows users to communicate with each other while not involving BSs in the process. D2D in 5G allows for more spectrum efficiency and also removes some traffic load from BSs. Direct communication of UEs can cause concerns when it comes to security, since it could be easier for attacks such as jamming and eavesdropping to target UEs [15], [3].

D2D communications can use parameters like distance to ensure a safe communication between endpoints. Devices use the distance between them to check for improvements in the security of communication. Distance reveals if the devices can cooperate jointly, if there is cooperation only from one side or no cooperation at all. Another security parameter used in D2D communications is authentication of UEs in order to detect and precent interference from non-legitimate users [15], [3].

5.4.2.4 Software Defined Networking

Software Defined Networking (SDN) in 5G communication systems separates the control plane from the user plane. This separation offers increased control capabilities and makes the management of the network easier with the use of flow-based forwarding schemes. The centralization of the control plane makes it vulnerable to DoS attacks, which can be mitigated with the dynamic assignment of the controller. The lack of authentication also makes the forwarding-control link vulnerable to MITM attacks, which can be simply prevented with signal encryption [3].

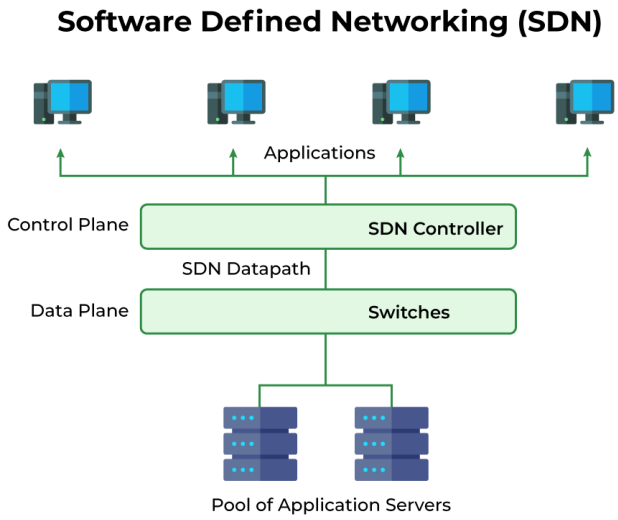


Figure 5.5 *Separation of user plane and control plane in SDN* [74].

5.4.3 Security Threats in 5G Subsystems

The previously described threats can severely affect the security services 5G networks offer and can be categorized into threat groups based on the effects they have on 5G systems and subsystems. Figure 5.6 illustrates the threats 5G subsystems face and can help 5G development agencies identify their systems' security capabilities and make any necessary adjustments to their security models [10].

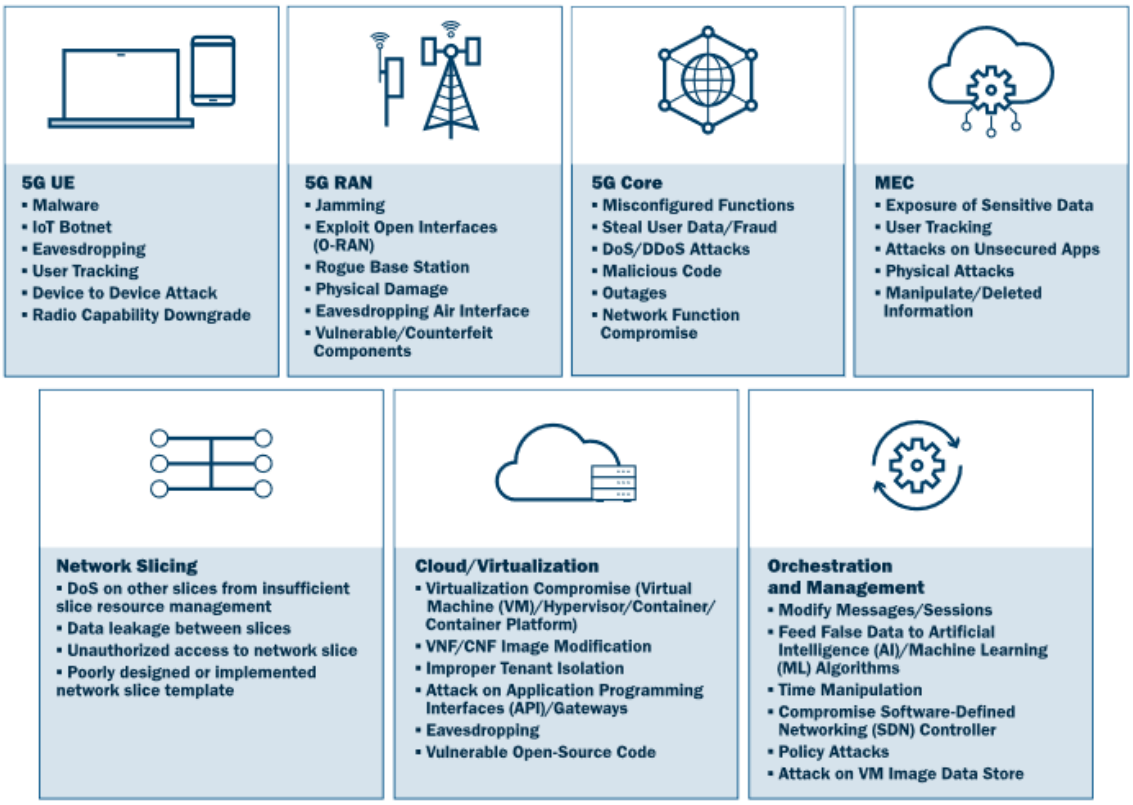


Figure 5.6 *Threats to 5G subsystems* [10].

5.4.3.1 General Cybersecurity Threats

General Cybersecurity Threats, such as network misconfigurations of the network or of a network's component, information leakage, human errors, hardware and software security vulnerabilities, unauthorized access to the network and unauthorized access attacks, impact all 5G subsystems. This weakness maybe exploited by attackers to reconfigure 5G components, redirect traffic to an attacker or gain access to legitimate users' data [10].

5.4.3.2 Virtualization Threats

Attacks on virtual machine (VM) software and container as a service (CaaS) platform, which are used for system operation, data storage, network connection and cloud-based organization and management of software, can have a huge effect on the 5G infrastructure. DoS attacks, side-channel attacks and cloud service usage misconfigurations can threaten the 5G Core, RAN, MEC, Network Slicing, Virtualization and Orchestration and Management. In a multi-user virtualization environment, excessive resource consumption by a single user can trigger a DoS event for other neighboring users, disrupting their operations. Side-channel attacks can also disrupt the operations of other users in the network and compromise 5G network aspects such as data confidentiality, integrity, and availability of the system. Side-channel attacks can also target 5G RAN and Core functions and give network access to malicious users, bypassing virtualization boundaries and revealing private users' information [10].

5.4.3.3 Network and Management Interface Threats

Attacks such as DoS, eavesdropping, jamming, address spoofing, message modification and deletion and access control attacks can severely damage the network, management, and over-the-air interfaces across all 5G subsystems. Over-the-air interface attacks occur between the UE and the RAN with the use of jamming techniques that can deny access to the network to legitimate UEs and even compromise 5G network services. The shared cloud infrastructure disguises some core network functions as legitimate users in the network to mitigate network interference by attackers, which could give them unauthorized access to the network's services and set the confidentiality of the network at risk [10].

5.4.3.4 Application and Service Threats

The deployment of 5G applications and services can be targeted by DoS and DDoS attacks, access control attacks, manipulation of Application Programming Interfaces (APIs) and malicious code injection, which can have deleterious effect on all 5G subsystems. Exploitation of applications and the deployment of malicious codes threatens the privacy of UEs. Meanwhile, vulnerabilities in the APIs at the MEC can give unauthorized access to malicious users and cascade into further attacks that originate from within the network [10].

5.4.3.5 Rogue Elements

The integrity of the 5G NR system can be threatened by unauthorized UEs, BSs or Radio Units in the RAN, as well as spoofed components in the MEC. These rogue elements can employ jamming techniques to detect a legitimate user's location and compromise MEC applications in order to access, modify or delete private users' data [10].

5.4.3.6 Privacy Threats

Attacks such as eavesdropping, location tracking, UE identification and spoofing that target UEs and the 5G Core and RAN can threaten data management and transmission operations in the 5G network. Attackers can intercept communications over the air interface between the RAN and UE and gain unauthorized access to the 5G Core, endangering private users' data, such as their identity and location [10].

5.4.3.7 Artificial Intelligence/Machine Language (AI/ML) Threats

Attacks that compromise the integrity, confidentiality, and availability of data within UEs, the RAN, and the Orchestration and Management subsystems can impact the reliability of, as well as other elements and services of the network, such as dynamic allocation of network functions, that rely on the accurate operation of AI/ML software and systems [10].

6. Jamming Attacks In 5G NR

5G NR, like all wireless cellular networks, operates in an open sharing environment, rendering it susceptible to interference. Interference hinders the performance of wireless networks and if it reaches high enough levels, it can become an obstacle to the receivers' ability to decode transmitted signals. This vulnerability can be exploited by malicious nodes to deliberately disrupt user communication over specific wireless channels and is commonly referred to as jamming attacks [75].

Security, privacy, and the threat of jamming attacks are always topics of discussion in wireless communication. New features introduced in every generation of cellular networks can be vulnerable to jamming attacks, especially when the implementation of sophisticated jamming attacks only requires some low-cost devices and some basic programming skills. For this reason, every generation of cellular networks has developed countermeasures against these attacks to protect users' privacy and prevent the disruption of all types of services, such as public and military communication services, emergency services and even national security services [75], [76].

Therefore, ensuring a high level of security and resilience to jamming attacks is a central requirement for 5G. The huge number of devices served by the 5G network is a weak point that can be targeted by attackers, since even the disruption of a small fraction of devices would cause functionality issues in the network. For this reason, new 5G NR standards released by 3GPP must discover and improve the networks' weak points and choose the proper anti-jamming techniques that are suitable for 5G's use cases [75], [76].

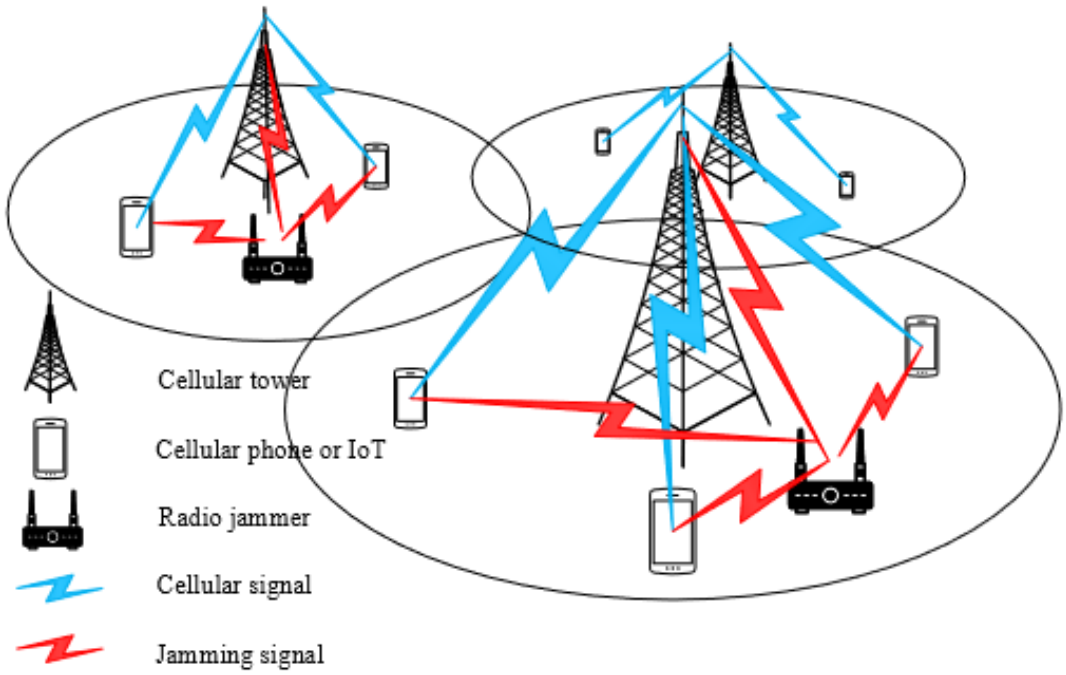


Figure 6.1 *Jamming attack in a cellular network* [77].

6.1 Classification of Jammers

Jammers are defined as malicious wireless nodes with the purpose of causing interference in wireless cellular networks. There are different types of jammers based on the methods they use to cause interference and are summarized in Table 6.1 [75].

Regular jammers continuously emit radio frequency signals without gaps in between that interfere with legitimate transmitted signals of the network. These signals are either legitimate or random bit sequences that congest the transmission channel of the legitimate network, making it appear to be busy and therefore hindering transmission from legitimate nodes. Because of their non-stop transmission, regular jammers do not need to monitor the actions of legitimate users. On the other hand, regular jammers require a substantial amount of power to perform their attacks, since their continuous transmission of radio signals consumes significant power and quickly drain the battery life of the malicious nodes they use [75], [78].

Deceptive jammers are similar to regular jammers that they continuously emit only legitimate sequence of bits into the communication channel by impersonating a legitimate node, delaying the receiver in the listening states. Deceptive jammers are more effective than regular jammers and are quite difficult to detect thanks to the similarities between the fake signal they transmit and the legitimate signal [75], [78].

Random jammers differ from both regular and deceptive jammers in terms of power consumption, since they don't continuously transmit frequency signals but instead they alternate between active and idle state, conserving their energy and reducing power consumption. During their active states, random jammers exhibit either regular or deceptive jamming behavior and transmit signals only for a predetermined period before switching back to an idle state. After some time, the jamming process starts again until the jammers enters an idle state and the same pattern continues [75].

Reactive jammers are different from the prior mentioned jamming strategies, which attempt to interfere with the communication channel without considering the activity pattern of the legitimate nodes and are defined as active jammers. Instead of continuously emitting signals, reactive jammers monitor the communication channel and are quick to respond when the transmitter is active, transmitting signals only during legitimate signal transmission. Because continuously monitoring the channel requires far less power than transmitting non-stop frequency signals, reactive jammers are more power-efficient than active jammers [75], [78].

Go-next jammers target only one frequency channel each time, but if detected by the transmitter, they have the ability to switch frequency channels and follow the transmitter to the next frequency it goes. Although go-next jammers are more power efficient compared to active jammers, they can still consume a lot of power if they follow a transmitter that performs frequency hopping at a really fast rate, potentially wasting energy [75].

Control channel jammers target the control channel to not allow the communication between the transmitter and the receiver but can also deny nodes access to the network [75].

Smart jammers possess the capability to pinpoint and target specific physical channels and signals within the communication system, provided they have knowledge of the exact time-frequency allocation of the intended channels and signals, while also being time-synchronized to the target system. These jammers come with a high level of complexity in order to be effective and efficient. However, open-source libraries and affordable software-defined radio (SDR) exist and can make the implementation of these attacks much more feasible. Smart jammers can use specially designed destructive signals that inflict significant damage to the targeted physical channels and signals while at the same time being power efficient [9].

Jammer type	Attack feature	Effect on 5G NR
Regural jammer - continuously transmits legitimate or random signals on a channel	Simple jammer, knowledge of the target system not required	Jamming signals congest the legitimate transmission channel
Deceptive jammer - continuously transmits legitimate signals on a channel	Simple jammer, more effective than regular jammers, knowledge of the target system not required	Jamming signals congest the legitimate transmission channel
Random jammer – switches between active and idle state of legitimate or random signal transmission on a channel	Power efficient jammers thanks to the limited in time transmissions	Jamming signals congest the legitimate transmission channel only during active states
Reactive jammer - transmits fake signals on a channel only during legitimate signal transmission	Power efficient, monitors the activity of the target system	Jamming signals congest the legitimate transmission channel only during legitimate signal transmission
Go-next jammer – transmits fake signals to only one frequency channel at a time, can switch to another frequency channel to transmit signals	Power efficient, monitors the activity of the target system and the frequency channel it uses	Jamming signals congest only the legitimate transmission channel that is currently used
Control channel – targets the control channel	Requires knowledge of the target system	Denies signal transmission and network access
Smart jammer- targets specific physical channels and signals	Power efficient but high complexity, requires time-synchronization to the target system and knowledge of the time and frequency allocation of the targeted channels and signals	Targeted physical channels and signals contest with specially designed destructive signals

Table 6.1 *Types of jammers, features and effect on 5G NR* [9], [75], [78].

6.2 Classification of Jamming Techniques

Jamming techniques vary based on their power efficiency, their complexity, the frequency bands they target, how detectable they are and the impact they have on the 5G NR communication system and are summarized in Table 6.2 [9].

Spot Jamming is a very popular jamming method, where a jammer focuses all of its transmitting power on a single frequency used by the target. This technique uses the same modulation as the original signal and requires enough power to override it. Although spot jamming, if successful, can be very powerful, since it only jams a single frequency each time it can be avoided simply by switching to another frequency [79].

Sweep Jamming is similar to spot jamming where the jammer only jams only one frequency each time, but here the jammer can rapidly shift his power from one frequency to another, thus being able to jam multiple frequencies in quick succession. This method is not very efficient since the jammer cannot affect all frequencies at the same time [79].

In *barrage jamming* the entire frequency bandwidth is jammed simultaneously by a single jammer. It is the optimal strategy for an attacker to use if he doesn't have any information about the target signal and aims to decrease the signal-to-noise ratio (SNR) at the receiver's side while also making all physical channels and signals contest with the jamming signal. The main drawback of barrage jamming is that the jammer's power is reduced proportionally to the number of frequencies it jams at a time, making it less powerful when used at a single frequency. For this reason, barrage jamming is used as a baseline of evaluation of other jamming methods [9], [79].

Partial-band jamming is a jamming method where, unlike barrage jamming, only a part of the bandwidth is targeted for attacks, meaning that only selected parts of physical channels and signals have to contest with the jamming signal. This method can cause denial of service while being power efficient at the same time and is especially effective when the attacker has knowledge of the bandwidth's vulnerabilities and can target them to damage the functionality of the communication system [9].

Smart jamming relies on the exploitation of the communication system's vulnerabilities known the jammer and attack these vulnerable spots of the target communication system. Smart jamming is a very efficient jamming method that causes a lot of damage without using a lot of power and can also be used without being detected by the target system, causing it even more problems [9].

Spoofing is a jamming method that transmits fake signals intended to mimic legitimate signals. For example, instead of the jammer injecting noise on top of the resource elements carrying the primary and secondary synchronization signals (PSS and SSS), it may be more efficient for a jammer to transmit fake PSS and SSS. This fake transmission of PSS and SSS can cause a UE a denial of service if it attempts to use these signals to establish a radio connection. Moreover, spoofing can be used to broadcast fake control information on the Physical Broadcast Channel (PBCH) and Physical Downlink Control

Channel (PDCCH) in order to deceive UEs into connecting to non-existent cells. Spoofing can also disrupt BS operations by flooding the Physical Random-Access Channel (PRACH) with numerous random-access preambles, since the BS would not be able to handle all these fake connection requests [9].

Jamming type	Attack feature	Effect on 5G NR
Spot jamming – targets only a single frequency	Simple method of jamming, requires knowledge of the frequencies used by the target	Targeted physical channel and signal contest with the jamming signal
Sweep jamming - targets only a single frequency, can rapidly switch to attack another frequency	Not power efficient, requires knowledge of the frequencies used by the target	Targeted physical channel and signal contest with the jamming signal
Barrage jamming – targets the entire bandwidth	Simple method of jamming, no knowledge of the frequencies used by the target is required	All physical channels and signals contest with the jamming signal
Partial-band jamming – targets only a part of the bandwidth	Power efficient, more effective if it has knowledge of the target system	Selected parts of physical channels and signals contest with the jamming signal, denial of service
Smart jamming - targets specific time and frequency resources	Power efficient but high complexity, requires time-synchronization to the target system and knowledge of the time and frequency allocation of the targeted channels and signals	Targeted physical channels and signals contest with specially designed destructive signals
Spoofing – transmits fake signals	Power efficient but high complexity, requires knowledge of the time and frequency allocation of the targeted channels and signals and possibly time-synchronization to the target system	Fake signals disrupt the legitimate communication

Table 6.2 *Types of jamming, features and effect on 5G NR* [9], [78].

6.3 Physical Layer Vulnerabilities of 5G NR

The physical layer of 5G NR consists of different types of physical channels and signals that are used on the downlink and uplink transmissions, with each channel or signal having each own vulnerabilities. Jammers can attack the protocol level and the physical layer channels and signals of 5G NR by exploiting their target's vulnerabilities, which depend on factors such as the percentage of resource elements the channel or signal occupies, the sparsity of these channels or signals in the uplink and downlink frame with respect to the entire time-frequency resource grid, the needed jamming power in order to distort the signal or channel and the level of complexity of the jammer, and are summarized in Table 6.3 [19].

Disrupted physical channel or signal	Impact	Damage
Synchronisation signals PSS and SSS	Interrupts cell search, disrupts synchronisation to the cell	Active connections are lost, new connections cannot be established
Reference signal DM-RS	Corrupts the demodulation of the received data, denial of service	Reduced or unavailable service
Reference signals CSI-RS and SRS	Poor MIMO and beamforming performance	Reduced service quality
Broadcast Channel PBCH	Prevents connection to a cell	New connections cannot be established
Random-Access Channel PRACH	Prevents connection and reconnection to a cell	Active connections are eventually lost, new connections cannot be established
Control channels PDCCH and PUCCH	Control information neccesary for the communication cannot be transmitted	Reduced performance, active connections are lost
Data channels PDSCH and PUSCH	Increased error rate in modulation symbols detection, corrupting data demodulation	Denial of service, gives access to crucial information about the communication
Massive MIMO	Disrupted channel estimation process	Poor performance, leaves other users vulnerable

Table 6.3 *Physical channels and signals and the effects caused by jamming* [9], [52], [73].

6.3.1 Jamming Vulnerability of Synchronization Signals

Primary Synchronization Signal (PSS) and Secondary Synchronization Signal (SSS) are the two synchronization signals of 5G NR and are used for frame, slot and symbol timing as well as to convey the Physical Cell ID of a cell. The PSS is formed using a correlation m-sequence, while the SSS uses a Gold sequence, which is a combination of two m-sequences. Due to their low cross-correlation, Gold sequences are resilient to jamming. A jammer aiming to disrupt the PSS and SSS requires either substantial jamming power or a specifically crafted signal that can disrupt the correlation. Time selective jamming of the PSS and SSS also requires time synchronization to the cell, as well as knowledge of the subcarrier spacing [9], [52].

A more effective and power efficient jamming method to use would be spoofing since, unlike noise injection on top of the existing PSS and SSS, fake transmission of PSS and SSS does not require time synchronization to the target's 5G NR frames and at the same time uses less power. The fake PSS and SSS can either interfere with the correlation of the legit synchronization signals or mislead the UE into connecting to a non-existent cell, causing denial of service during the initial cell search. In fact, depending on the chipset and the PSS and SSS blacklisting mechanism used, jammers may only need to spoof the PSS. As 5G NR specifications do not define the response of the UE upon detecting a valid PSS without an associated SSS, the impact of PSS spoofing varies depending on the implementation. The result of this attack could cause the UE to lose its active connections and be prevented from establishing new ones. Mitigation using

blacklisting mechanisms becomes harder and more sophisticated with the transmission of more fake PSSs [9], [52].

6.3.2 Jamming Vulnerability of Reference Signals

The reference signals used in 5G NR are the Demodulation Reference Signal (DM-RS), Phase Tracking Reference Signal (PT-RS), Sounding Reference Signal (SRS) and Channel State Information Reference Signal (CSI-RS), which provide channel estimation and channel equalization [52], [53].

The DM-RS is known to both the transmitter and receiver and is used to estimate the distortion of a received signal during transmission by comparing it to a predefined reference signal. This estimate is crucial for a communication channel to reduce any noise and interference effects that may appear and if disrupted can corrupt the demodulation of the received data and even cause denial of service. For this reason, jamming the radio resources (REs) carrying DM-RS can cause more impact than jamming the data channels separately. The best reference signals for jammers to target are those that are a vital part of the communication link between the transmitter and receiver and require the least power usage to jam, meaning that they have the least number of REs per frame. The jamming of the DM-RS requires not only knowledge of the DM-RS allocation in REs, but also synchronization to the target system, drastically increasing the jamming complexity [9], [19], [52].

The DM-RS for the PBCH is the most efficient to jam, since it is in the same spot every frame and the jammer can have knowledge of the cell ID and the PBCH's location if he is time-synchronized to the frame. The DM-RS for the PBCH occupies 1⁄4 of the REs that are allocated to PBCH and can also be jammed without any time-synchronization to the cell, by jamming the correct 60 subcarriers [19], [52].

Other types of reference signals targeted for jamming attacks are the Channel State Information Reference Signal (CSI-RS) and the Sounding Reference Signal (SRS). These signals provide channel estimation and can also be used by the UE to assist in beam management, resource allocation and in Massive MIMO operations. Jammers can exploit this knowledge to target these reference signals, causing inefficient MIMO and beamforming performance and reduce the service quality [9].

The mapping of the Phase-Tracking Reference Signal (PT-RS) for the PDSCH depends on parameters such as time density and frequency density. Therefore, the effectiveness of jamming the PDSCH PT-RS is unclear and requires information such as the frequency of PT-RS's activation as well as the previously mentioned densities [52].

6.3.3 Jamming Vulnerability of PBCH

The Physical Broadcast Channel (PBCH) is transmitted together with the PSS and SSS on the SSB and carries vital information to a UE in order to attach to a cell. The number of slots between the symbols assigned to the PBCH region depend on the carrier frequency. If the carrier frequency is below 3 GHz then the assigned symbols are within

two slots of each other, while they are within four slots of each other if the carrier frequency is above 3 GHz. Even though at higher subcarrier spacings the duration of one slot is lower, in both of these cases the symbols are actually close to each other. This is a vulnerability in the design of the PBCH and allows a jammer that selectively targets the PBCH to have a very low duty cycle but requires the jammer to have a small distance from the target. The jammer can be prevented from attacking with the use of localization-based detection techniques that discover the source of the attack. If the jammer is mobile, however, it is also necessary to use techniques that monitor the attacker's position [9], [52].

Jammers can also attack the PBCH in a time-selective manner by synchronizing with the target cell, denying UEs access to information necessary to connect to a cell and thus preventing new connections from being established. Otherwise, the jammer could attack the PBCH by jamming the subcarriers the PBCH is on using 100% duty cycle. Spoofing could also be used as a jamming method to lure UEs to connect to a non-existing cell causing denial of access to an existing cell [9], [52].

6.3.4 Jamming Vulnerability of PRACH

The Physical Random-Access Channel (PRACH) is used by a UE to establish or re-establish random access procedures. These procedures can be disrupted if the jammer has access to information about the resource allocation of the physical channel on the radio frame. Jamming the PRACH can cause denial of service to the new users attaching to the cell, since new connections cannot be established. It can also prevent an existing user to switch from an idle to an active state. A UE is in an idle state when it doesn't use any mobile services. To use these services and acquire the necessary radio resources, a UE switches to an active state. If all users are prevented from switching to their active states, all active connections will be lost. Jamming the PRACH is a difficult task due to its large number of possible time and frequency locations as well as the decoding of these locations in real time. Spoofing can be used if the jammer fails to determine these locations, flooding the PRACH with a large number of invalid preambles, as the 5G NR specifications do not specify what it should be done in this scenario [9], [75], [52].

6.3.5 Jamming Vulnerability of the UL and DL Physical Control Channels

The two control channels of 5G NR are the Physical Downlink Control Channel (PDCCH) and the Physical Uplink Control Channel (PUCCH). The channels exchange information between the UE and the BS such as HARQ acknowledgements, resource allocation, modulation, slot format and channel quality reports and their successful operations are vital for the communication system, since their disruption can reduce performance and even cause connections to fail [9].

The PDCCH is carried by the CORESET, which is a set of parameters and physical resources. Since the PDCCH can appear on any subcarrier, the jammer can use the information of the CORESET to attack the PDCCH by jamming specific sub-carrier using a

small duty cycle. If the jammer doesn't have knowledge of the information of the CORESET, it can instead target all the possible locations in which the PDCCH resides by jamming every subcarrier [52].

The PUCCH can choose between five different formats of transmitting subcarriers and symbols on each message. The PUCCH also has an option for intra-slot hopping, which depending on the hopping rate can defend the PUCCH against selective jammers. Because the 5G NR standard is public, jammers have knowledge of the intra-slot hopping and can jam the PUCCH without major costs. Moreover, the PUCCH utilizes MPSK modulation (with m=2 or 4) and employs either polar codes or repetition codes as error coding schemes, depending on the number of bits being transmitted. Polar codes are recognized for their relatively low resilience against jamming attacks [75].

6.3.6 Jamming Vulnerability of the UL and DL Physical Data Channels

In 5G NR, the Physical Downlink Shared Channel (PDSCH) and the Physical Uplink Shared Channel (PUSCH) are used to transmit and receive user data. Jamming the PUSCH can cause denial of service in an entire cell, while jamming the PDSCH can give jammers access to UEs' resources and thus allow them to target more than one UEs simultaneously [9], [19].

Even though it is possible to jam only the REs that carry user data, it is more effective for the jammer to use barrage jamming to attack the entire UL and DL transmission bandwidth, since the biggest part of these resources are data channels, leading to an increased error rate in modulation symbols detection, corrupting data demodulation and causing reduced or unavailable services [9], [52].

6.4 Jamming of NSA vs SA 5G

Jamming of 5G NR can have different results on the total system performance based on the use of the Stand Alone (SA) or Non-Stand Alone (NSA) network in the communication system. In SA, only 5G NR is used for control signaling and data transmission, whereas in NSA a 4G network is used alongside 5G NR as a supplementary data carrier. The 4G network continuously monitors the connection between the UE and 5G BS and assigns 5G radio resources to the UE. In case of connection loss, the 4G network can take over for the disrupted 5G NR network. Therefore, when the 5G channel is jammed in NSA the connection will not be terminated but instead the UE will be downgraded to operate on an uninterrupted 4G channel. However, in the case of SA the UE could either operate on a 5G channel of poor quality or just terminate the connection completely and go into "no service" mode [9].

6.5 Jamming of 5G NR vs 4G

5G NR is a commercial mobile technology that operates on higher frequencies than 4G LTE. Due to the increased signal path loss the cells operating in these low frequencies, especially mmWave frequencies, have small network coverage, reducing unwanted radio emissions outside of their coverage area. The use of massive MIMO enables

beamforming at the BS, concentrating the signal energy towards the user and thus further reducing the emissions in unwanted directions. This makes it challenging for the jammer to identify, disrupt and exploit the radio communication while at the same time strengthens the received signal in the downlink, providing further assistance against interference. Furthermore, 5G NR broadcasts in general less information than 4G LTE, reducing the chances of disrupting a 5G NR communication [9].

The physical layer of 5G NR is another important factor in its NR's resiliency against jamming attacks. Compared to 4G LTE where the physical channels and signals are statically allocated on the radio frame and can be easily targeted and attacked by jammers, in 5G NR their allocation is dynamic and highly configurable. For example, although in 4G LTE the physical uplink control channel (PUCCH) was vulnerable to jamming, in 5G NR it can be mapped dynamically to avoid being so easily jammed [9].

6.6 Anti-Jamming in 5G NR

The rise of 5G networks alongside the increasing number of the supported smart devices put a heavy load on existing cellular networks' performance. Due to the low frequency bands these networks use and the ultra-dense positioning of BSs, jammers can pose a challenging threat to the system's functionality and to frequency coordination among its users. For this reason, it is important to understand not only how to detect jamming attacks in a 5G system, but also how to mitigate them. Jamming detection and mitigation methods are summarized in Tables 6.4 and 6.5 respectively [80].

6.6.1 Detection of Jamming Attacks

Jammer detection mechanisms are required for the secure operation of a wireless communication system. However, locating a jamming scenario from a legitimate activity is quite challenging due to the different types of jammers and jamming attacks. Network congestion can make this situation more difficult because in the case of excess traffic loads the performance of the networks drops and the communication van be interrupted, causing confusion about the presence of the jammer. Smart-jamming attacks can be detected by monitoring the amount of energy that a physical channel uses or any performance changes of the communication over this channel. Jamming detection uses a threshold with some performance metrics such as the Packet Delivery Ratio (PDR), Packet Drop Ratio (PDR), Bit Error Rate (BER), and Signal-to-Noise Ratio (SNR), the Energy Consumption Amount (ECA) and Channel Utilization, while Received Signal Strength (RSS) and Signal Collision Ratio-Based Detection can also be used [19], [75].

Packet Delivery Ratio (PDR) cannot accurately determine jamming attacks by itself, since PDR packets could also be dropped due to bad connection, defected nodes or packet collisions. In order to improve the detection of jamming, Received Signal Strength is used with PDR (PDR with RSS). The thresholds of PDR and signal strength are pre-assigned to the network. If low PDR and high signal strength are detected, then

a jammer is present. If the PDR detected is higher than the threshold and the signal strength detected is lower than the threshold, then the channel is not interfered [19].

Bad Packet Ratio (BPR) represents the number of corrupted packets received by the receiver compared to the total number of packets received at a specific time. BPR identifies corrupted packets using cyclic redundancy check (CRC), which discards the damaged packets and approves the intact ones packets for transmission. This method is very for known for its simplicity in computation and can be applied in channels where acknowledgments are unnecessary [19].

Similar to BPR, *Bit Error Rate (BER)* calculates the ratio of corrupted bits to the total bits received by a node during transmission and is used to detect reactive jammers [19].

Energy Consumption Amount (ECA) is the power consumed by a node within a specified timeframe. It is calculated by squaring the squared value of voltage drop in the node's battery, multiplying it by the duration, and then dividing the result by the average resistance of the node. Jammers disrupt the node's normal operation by keeping it in an active state instead of an idle state, leading to higher power consumption than usual [19].

Signal-to-Noise Ratio (SNR) is the ratio of the received signal power to the received noise power at a particular node. SNR is a very accurate jamming detection technique at the physical layer, since when a system is jammed the SNR value drops. The jamming index can also be computed via SNR and packet dropped per terminal (PDPT) values. A BS uses the detection algorithm to learn how many packets the node received, how many packets it dropped and what is the power of the signal, and uses this information to calculate the PDPT and SNR and examine if a jammer is present [19].

Jamming detection at the physical layer can also be done *using ant system mechanisms*. These mechanisms assess whether interference has occurred at the system or not. In this approach, an agent is deployed to traverse the nodes iteratively, gathering information along the way. This data is then used to compile a set of routes leading to the destination, which is stored for future reference and redirection purposes. The presence or absence of jamming in the system is determined by metrics such as PDR, BER, SNR, energy levels, packet loss or coverage [19].

Jammed-area mapping protocol (JAM) is both a jamming detection and a jamming mitigation technique that is specifically applied to wireless sensor networks. This protocol works by identifying jammed regions within the network and rerouting packets to bypass these areas. JAM detects a jamming threat when a node's utility drops below a predefined threshold value. The system shares the jammed or unjammed status to neighboring nodes, notifying them of the situation. Upon receiving a jammed message, nodes devise countermeasures to effectively overcome the jamming attack [19].

Detection techniques	Detection parameters
Packet Delivery Ratio (PDR) with Received Signal Strength (RSS)	Low PDR, high signal strength
Bad Packet Ratio (BPR)	Cyclic redundancy check
Bit Errot Rate (BER)	Ratio of corrupted bits to the total bits
Energy Consumption Amount (ECA)	Consumed power in a set time
Signal-to-Noise Ratio (SNR)	Ratio of received signal power to received noise power
Ant System Mechanisms	Rerouting of the signal
Jammed Area Mapping Protocol (JAM)	Utility percentage of nodes

Table 6.4 *Jamming detection techniques and their parameters* [19].

6.6.2 Mitigation of Jamming Attacks

Alongside jamming detection mechanisms, wireless communication systems also require the use of some methods that will prevent the jamming attacks from occurring. 5G NR uses techniques such as Direct Sequence Spread Spectrum (DSSS), Frequency Hopping Spread Spectrum (FHSS), Game Theory, Unmanned Aerial Vehicle (UAV), massive MIMO, Scheduling and Deep Learning that can act against jamming at the 5G NR physical layer [75].

Direct Sequence Spread Spectrum (DSSS) is a mitigation mechanism against jamming attacks with limited power usage. DSSS changes the bandwidth that the signal carrying information is transmitted to one that is larger than the least required for transmission bandwidth. This transition of bandwidths transmits the signal through the channel without it being detected by eavesdroppers by multiplying the narrow band signal with a wideband pseudo-noise (PN) sequence. This results in a signal that has a similar spectrum to the PN and is the spreading factor that determines DSSS's resiliency against jamming attacks. However, DSSS is not without its flaws and limitations. DSSS cannot be easily implemented in all wireless devices due the larger bandwidth it requires alongside its high complexity. At the same time, the PN code of DSSS used in TDMA is a vulnerability that can be targeted in real-time by jammers allowing them to launch a follow-on jamming attack to disrupt the communication at low cost [75].

If the bandwidth increase is not enough to overcome the effect of some jammers, *Frequency Hopping Spread Spectrum (FHSS)* can be used. FHSS switches the transmission of the modulated signal carrier to a frequency different than the one its currently using. This is a different type of spread spectrum where the spectrum of the transmitted signal is now spread sequentially instead of instantaneously. There are two types of frequency hopping based on the hopping rate. In slow-rate hopping, a number of modulated symbols are conveyed within one frequency hop before hopping to another frequency, while in fast-rate frequency only one symbol rate is transmitted during several frequency hops. However, both of these methods have their weaknesses. Slow-rate hopping is weak against a smart jammer that can track the signals movements and

accurately predict the next hop before the transmitter switches to the next frequency. Fast-rate hopping, on the other hand, lowers the communication channel's performance since it requires the synchronization between the transmitter and the receiver. Both of these methods also require the use of a key known to both the transmitter and the receiver that contains the hopping pattern, which can be intercepted by a jammer and used to attack the communication channel [75].

Game theory is another jamming mitigation technique where the users can avoid any interference from jammers by switching to another accessible channel, minimizing the performance loss during transmission. Its name originates from the fact that channel hopping in this manner can be seen as a game between the legitimate user and the jammer. In fact, research has shown that game theory may be able to find the ideal strategy that deals with jamming attacks [75].

Timing channels are a recovering type of mechanism against jamming attacks, meaning that they restore the communication channel that is jammed rather than switching to a different channel or frequency. Essentially, the timing channel uses the timing patterns of the jammer's attacks and transmits information only when the jammer is in an idle state. Thus, a detection step is required before the timing channel is created [75].

An additional strategy to combat jamming attacks in 5G NR wireless communication systems is an *unmanned aerial vehicle (UAV)* which acts as a relay node when the BS is under heavy jamming. Deep reinforcement learning techniques are also used in UAVs to identify the ideal relay policy for mobile users within the 5G cellular network. Even though this framework can provide flexible solutions to the network to evade jammers, UAVs themselves are still susceptible to jamming attacks and their performance and operational capabilities are constrained by their limited power supply [75].

Massive MIMO communication systems also have built-in jammer suppression capabilities without changing the 5G NR specifications. Due to the high power and wide area required to accommodate the large number of antennas massive MIMO uses, massive MIMO techniques are well suited to deal with jamming attacks in the cellular uplink transmissions by using robust channel coding schemes that can recover the corrupted packets to their original state, exhausting the jammer. One such scheme is the computation of the legitimate channel and the jamming channel estimates, if the presence of jamming signal is detected from the BS. These estimates are used to construct linear receiver filters that are not affected by the impact of the jamming signal and can recover the legitimate signal [75], [77].

5G NR wireless communication networks can also use functions like *Software Defined Networking (SDN)* and *Network Functions Virtualization (NFV)* alongside *Deep Learning* to defend against jamming attacks. SDN separates the control plane from the user plane and make networks agile and flexible by enabling a dynamic and efficient network configuration which also improves performance. NFV separates the network's functions from the hardware, delivering equivalent network functionality without the

need for specialized hardware. Together they can drastically reduce the threat of jammers by dynamically allocating radio resources and scheduling them appropriately. The performance of these method can increase if combined with deep learning to discover the jammer's strategy [75], [81].

Mitigation techniques	Methods used
Direct Sequence Spread Spectrum (DSSS)	Changes the bandwidth used for transmission to a larger one
Frequency Hopping Spread Spectrum (FHSS)	Changes the frequency used for transmission
Game Theory	Changes the channel used for transmission
Timing Channels	Restoration of the communication channel after jamming
Unmanned Aerial Vehicle (UAV)	Relay node when the BS is under heavy jamming
Massive MIMO	Channel coding schemes that recover corrupted packets
Software Defined Networking (SDN) and Network Functions Virtualization (NFV) alongside Deep Learning	Dynamic allocation of radio resources, scheduling

Table 6.5 *Jamming mitigation techniques and the methods they use* [75].

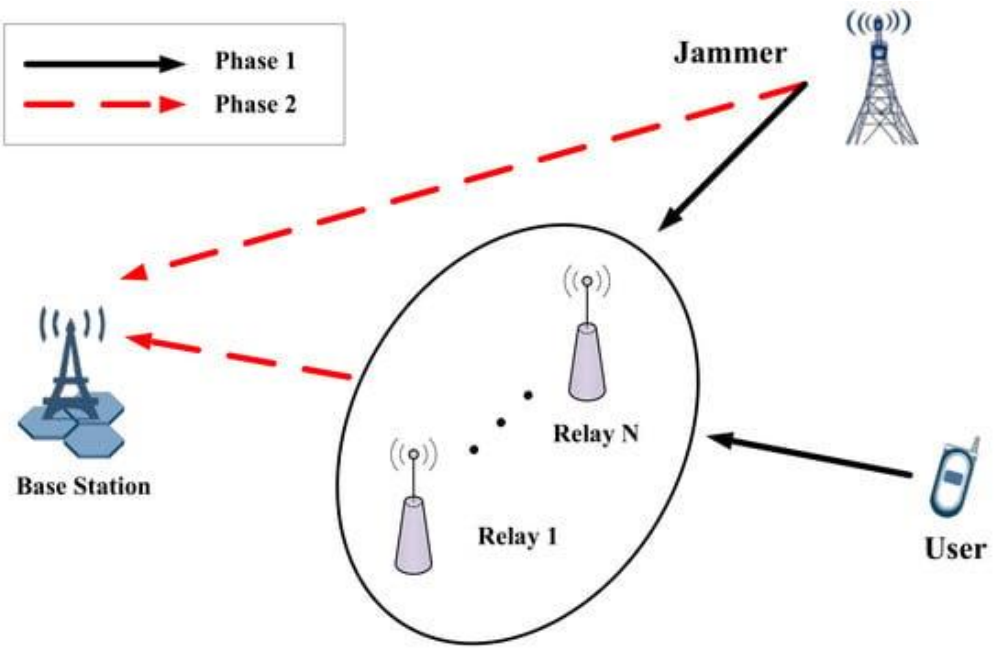


Figure 6.2 *Jamming mitigation using relay nodes* [82].

Conclusion and Future Works

In conclusion, this thesis has provided a comprehensive exploration of 5G, covering its evolution from LTE, introducing 5G NR and the 5G Core network, analyzing the aspects of the 5G NR physical layer, and overviewing the security of the 5G systems and their vulnerabilities to jamming attacks. The overview of 5G gave us an insight into its background, specifically about its use cases, its spectrum allocation, and its vulnerabilities, while the detailed examination of the physical layer has enhanced our understanding of modulation, waveform, multiple antennas, and channel coding techniques employed in 5G NR. Furthermore, the evolution of security in wireless networks is important so we can understand the milestones in technological advancements we have made thus far and will continue to make, but also reminded us of the ever-lasting threat landscape that we must deal with and the reasons why security and privacy of users is so important in the first place. Finally, an investigation into jamming attacks in the 5G NR physical layer has highlighted the classification of jammers and the techniques they use, as well as the impact of jamming attacks on different aspects of 5G NR.

Overall, this thesis has contributed to the understanding of 5G NR technology, its security challenges, and the potential threats posed by jamming attacks. As the deployment of 5G continues to expand, addressing these security concerns will be crucial in ensuring the reliability and integrity of next-generation wireless communications. We anticipate that this thesis could serve as a valuable resource for both industry corporations and academic researchers, offering insights and potential research directions for enhancing security in 5G wireless networks in the foreseeable future.

References

[1] P. B. R. B. R. B. H. E. E. F. F. K. P. K. N. M. N. P. O. J. P. J. S. M. S. J.-P. W. A. Z. Ghada Arfaoui, "A Security Architecture for 5G Networks," 2018.

[2] Ahmad Sassan, 5G NR Architecture, Technology, Implementation, and Operation of 3GPP New Radio Standards, Elsevier BV.

[3] Y. Q. R. Q. H. Donfeng Fang, "Security for 5G Mobile Wireless Networks," 2017.

[4] S. P. J. S. Erik Dahlman, 5G NR The Next Generation Wireless Access Technology, Academic Press, 2018.

[5] F. A. J. M. G. D. X. C. Ali Zaidi, 5G Physical Layer Principles, Models and Technology Components, Academic Press, 2018.

[6] S. S. T. K. J. O. A. G. M. Y. Ijaz Ahmad, "Security for 5G and Beyond".

[7] P. K. D. N. K. J. M. L. Rabia Khan, "A Survey on Security and Privacy of 5G Technologies: Potential Solutions, Recent Advancements and Future Directions," 2020.

[8] "Wikipedia, The Free Encyclopedia," [Online]. Available: https://en.wikipedia.org/wiki/International_Telecommunication_Union. [Accessed 2 4 2024].

[9] A. M. T. U. Ø. D. B. J. K. Agnius Birutis, A study of 5G New Radio and its vulnerability to jamming, Norwegian Defence Research Establishment (FFI) , 2022.

[10] D. D. M. B. T. D. D. M. Vincent Sritapan, "5G Security Evaluation Process Investigation Version 1," 2022.

[11] W. L. D. P. D. A. M. G. B. V. L. Gerrit Holtrup, "5G System Security Analysis," 2021.

[12] Frederic Launay, "NG-RAN Network – Functional Architecture," Wiley Telecom.

[13] "RF Wireless World," [Online]. Available: https://www.rfwireless-world.com/5G/Difference-between-ng-eNB-and-gNB-in-5G-NG-RAN.html. [Accessed 02 10 2023].

[14] "IPLOOK," [Online]. Available: https://www.iplook.com/info/what-is-device-to-device-communication-i00359i1.html. [Accessed 2 5 2024].

[15] Shane Fonyi, "Overview of 5G Security and Vulnerabilities," 2020.

[16] R. K. J. S. J. Anutusha Dogra, "A Survey on Beyond 5G Network With the Advent of 6G: Architecture and Emerging Technologies," 2020.

[17] "Cadence Design Systems," [Online]. Available: https://resources.pcb.cadence.com/blog/2023-5g-embb-urllc-and-mmtc-service-categories-for-a-smarter-tomorrow. [Accessed 03 08 2023].

[18] "Wikipedia," [Online]. Available: https://en.wikipedia.org/wiki/Vehicle-to-everything. [Accessed 03 08 2023].

[19] Sarada Adhikari, "PHYSICAL LAYER JAMMING DETECTION FOR 5G NR UPLINK," 2022.

[20] D. A. P. v. B. T. C. M. F. F. G. M. H. B. H. G. J. J. K. F. K. E. L. Henrik Asplund, Advanced Antenna Systems for 5G Network Deployments Bridging the Gap Between Theory and PracticeHenrik Asplund, Elsevier, 2020.

[21] Xingqin Lin, "An Overview of 5G Advanced Evolution in 3GPP Release 18," IEEE, 2022.

[22] Z. G. X. C. Lei Wan, "Enabling Efficient 5G NR and 4G LTE Coexistence," *IEEE Wireless Communications,* February 2019.

[23] "Wikipedia, the free encyclopedia," [Online]. Available: https://en.wikipedia.org/wiki/5G_NR_frequency_bands. [Accessed 01 08 2023].

[24] Nitin Dahad, "EE Times Asia," [Online]. Available: https://www.eetasia.com/gsma-warns-against-blocking-mmwave-spectrum-for-5g/. [Accessed 02 08 2023].

[25] "GSMA," [Online]. Available: https://www.gsma.com/spectrum/wrc-series/. [Accessed 02 08 2023].

[26] "ITU International," [Online]. Available: https://www.itu.int/en/mediacentre/Pages/PR-2023-12-15-WRC23-closing-ceremony.aspx. [Accessed 16 01 2024].

[27] Simmons Adam, "Dgtl Infra," 23 02 2023. [Online]. Available: https://dgtlinfra.com/5g-standalone-sa/. [Accessed 02 08 2023].

[28] Deanna Darah, "TechTarget," [Online]. Available: https://www.techtarget.com/searchnetworking/feature/5G-NSA-vs-SA-How-does-each-deployment-mode-differ. [Accessed 02 08 2023].

[29] "S&P Global Market Intelligence," [Online]. Available: https://www.spglobal.com/marketintelligence/en/news-insights/research/5g-tracker-94-markets-worldwide-have-commercial-5g-services. [Accessed 02 08 2023].

[30] "Qualcomm," 06 07 2023. [Online]. Available: https://www.qualcomm.com/news/onq/2023/07/its-time-for-5g-to-standalone. [Accessed 02 08 2023].

[31] P. H. M. O. L. F. S. S. C. M. Steffan Rommer, 5G Core Networks: Powering Digitalization, Academic Press, 2020.

[32] Dr. William Stallings, "5G Wireless: A Comprehesive Introduction," 2021.

[33] B. H. M. G. T. Z. Sherif Abdelwahab, "Network Function Virtualization in 5G".

[34] Dimitris Tsolkas, "5G Networks: The new era in mobile communications," [Online]. Available: https://eclass.uoa.gr/modules/document/file.php/D211/Winter%20semester%202020-2021/Kinita-Metaptyxiako-2018-2019-11-5G.pdf. [Accessed 02 10 2023].

[35] "Telcoma Global," [Online]. Available: https://telcomaglobal.com/p/5g-nr-new-radio-protocol-stack-layers-5g-communication. [Accessed 19 01 2024].

[36] Z. Q. F. C. Y. L. J. A. M. Yunlong Cai, "Modulation and Multiple Access for 5G Networks," *IEEE COMMUNICATIONS SURVEYS & TUTORIALS, VOL. 20,* 2018.

[37] Kyle Reis, "WAVEFORM," [Online]. Available: https://www.waveform.com/a/b/guides/modulation-coding-speeds. [Accessed 03 10 2023].

[38] W. J. B. H. H. R. Nick LaSorte, "The History of Orthogonal Frequency Division Multiplexing," *IEEE GLOBECOM 2008 - 2008 IEEE Global Telecommunications Conference,* 2008.

[39] Y. Y. W. L. Qiaoping Liu, "Application of OFDM technology in 4G mobile network," *Applied Mechanics and Materials (Volumes 631-632) ,* 2014.

[40] S. Y. A. S. R. M. Z. Zainab Sh. Hammed, "Massive MIMO-OFDM Performance Enhancement on 5G," IEEE , 2021.

[41] D. Matic, "JPL's Wireless Communication Reference Website," [Online]. Available: http://www.wirelesscommunication.nl/reference/chaptr05/ofdm/ofdmmath.htm. [Accessed 13 10 2023].

[42] M. D. L. A. M. M. B. M. Guoying Zhang, "A Survey on OFDM-Based Elastic Core Optical Networking," IEEE , 2013.

[43] S. D. P. P. B. Nilesh Chide, "Implementation of OFDM System using IFFT and FFT," ISSN, 2013.

[44] G. W. G. L. C. Y. S. L. Taewon Hwang, "OFDM and its Wireless applications: A survey," IEEE, 2009.

[45] J. K. W. Y. Y. C.-G. K. Yong Soo Cho, MIMO-OFDM Wireless Communications with MATLAB, ISBN, 2010.

[46] G. U. M. E. V. S. P. D. M. N. R. R. G. Gowri, "Performance Analysis Of DWT-OFDM and FFT-OFDM Systems," IJET, 2013.

[47] S. P. a. J. S. Erik Dahlman, 4G LTE/LTE-Advanced for Mobile Broadband, 2011.

[48] B. R. S. N.-h. D.C. Shah, "Effects of Cyclic prefix on OFDM system," 2010.

[49] H. J. A. A. G. A. Q. Yazen S. Almashhadani, "Performance Analysis of OFDM with Different Cyclic Prefix Length," 2017.

[50] "Telecomtrainer," [Online]. Available: https://www.telecomtrainer.com/nr-cell/. [Accessed 20 2 2024].

[51] "Converged wireless access for reliable 5G MTC for actories of the future," 2019.

[52] R. R. V. M. J. R. R. P. J. Marc Lichtman, "5G NR Jamming, Spoofing, and Sniffing: Threat Assessment and Mitigation," 2018.

[53] "Techplayon," [Online]. Available: https://www.techplayon.com/5g-nr-reference-signals-dmrs-ptrssrs-and-csi-rs/. [Accessed 20 02 2024].

[54] G. G. F. V. Donatella Darsena, "Beamforming and precoding techniques," 2020.

[55] Adnan Ghayas, "Commsbrief," 31 10 2021. [Online]. Available: https://commsbrief.com/mimo-massive-mimo-spatial-multiplexing-and-beamforming/. [Accessed 8 10 2023].

[56] A. L. Haonan Wang, MIMO Communications - Fundamental Theory, Propagation Channels, and Antenna Systems, 2023.

[57] Dr. Mohamed Nadder Hamdy, "Beamformers Explained," COMMSCOPE, 2023.

[58] Cadence System Analysis, "The Basics of Digital and Analog Beamforming with Phased Arrays," CADENCE SYSTEM ANALYSIS, [Online]. Available: https://resources.system-analysis.cadence.com/blog/the-basics-of-digital-and-analog-beamforming-with-phased-arrays. [Accessed 12 10 2023].

[59] "Wikipedia," [Online]. Available: https://en.wikipedia.org/wiki/Channel_state_information. [Accessed 14 02 2024].

[60] "MathWorks," [Online]. Available: https://www.mathworks.com/help/5g/ug/5g-nr-downlink-csi-reporting.html. [Accessed 13 10 2023].

[61] "Wevolver," [Online]. Available: https://www.wevolver.com/article/5g-antenna-design. [Accessed 6 3 2024].

[62] A. A. H.-P. L. K.-B. S. J. L. Jung Hyun Hae, "An overview of channel coding for 5G NR cellular communications," Industrial Technology Advances, 2019.

[63] S. P. Z. S. Mrinmayi V Patil, "Coding Techniques for 5G Networks: A Review," IEEE, 2020.

[64] "Wikipedia," [Online]. Available: https://en.wikipedia.org/wiki/Low-density_parity-check_code. [Accessed 6 3 2024].

[65] "Wikipedia," [Online]. Available: https://en.wikipedia.org/wiki/Polar_code_(coding_theory). [Accessed 6 3 2024].

[66] R. Z. H. C. F. Y. P. M. Xiyang Luo, "Distortion Agnostic Deep Watermarking," 2020.

[67] A. B. S. K. S. Sullivan, "5G Security Challenges and Solutions: A Review by OSI Layers," 2021.

[68] "Infosectrain," 11 05 2023. [Online]. Available: https://www.infosectrain.com/blog/active-attack-vs-passive-attack/. [Accessed 6 03 2024].

[69] P. W. A. A.-F. L. J. K. Z. Ning Wang, "Physical Layer Security of 5G Wireless Networks for IoT: Challenges and Opportunities," 2018.

[70] "Shiksa Online," [Online]. Available: https://www.shiksha.com/online-courses/articles/eavesdropping-how-to-prevent-it/. [Accessed 6 3 2024].

[71] "Panda Security," [Online]. Available: https://www.pandasecurity.com/en/mediacenter/man-in-the-middle-attack/. [Accessed 6 3 2024].

[72] [Online]. Available: https://www.formasup.fr/?k=what-is-ddos-attack-and-how-to-prevent-it-ii-Qd11zzHk. [Accessed 6 3 2024].

[73] J. M. K. K. M. K. M. A. M. A. V. S. S. W. L. A. C. J. M. P. R. S. A. Paweł Skokowski, "Jamming and jamming mitigation for selected 5G military scenarios," 2022.

[74] "GeeksForGeeks," [Online]. Available: https://www.geeksforgeeks.org/software-defined-networking/. [Accessed 6 3 2024].

[75] S. F. Youness Arjoune, "Smart Jamming Attacks in 5G New Radio: A Review," IEEE, 2020.

[76] Raghunandan M. Rao, "Perspectives of Jamming, Mitigation and Pattern Adaptation of OFDM Pilot Signals for the Evolution of Wireless Networks," 2016.

[77] H. Z. Hossein Pirayesh, "Jamming Attacks and Anti-Jamming Strategies in Wireless Networks: A Comprehensive Survey," 2021.

[78] B. E. H. M. A. N. Satish Vadlamani, "Jamming attacks on wireless networks: A taxonomic survey," 2016.

[79] D. G. C. K. G. P. Aristides Mpitziopoulos, "A Survey on Jamming Attacks and Countermeasures in WSNs," 2009.

[80] L. J. N. Q. Y. X. M. W. Yunpeng Zhang, "Anti-jamming channel access in 5G ultra-dense networks: a game-theoretic learning approach," 2022.

[81] "Equinix Interconections," [Online]. Available: https://blog.equinix.com/blog/2020/03/10/sdn-vs-nfv-understanding-their-differences-similarities-and-benefits/. [Accessed 22 02 2024].

[82] G. R. J. C. C. C. X. Y. Y. L. K. X. Zhibin Feng, "An Anti-Jamming Hierarchical Optimization Approach in Relay Communication System via Stackelberg Game," 2019.

[83] "EITC," [Online]. Available: http://www.eitc.org/research-opportunities/5g-and-beyond-mobile-wireless-technology/5g-and-beyond-technology-roadmap/5g-deployments-and-use-cases/5g-new-radio-use-cases%20. [Accessed 02 08 2023].

[84] G. H. L. A. L. E. P. N. S. Adriano Baratè, "5G TECHNOLOGY AND ITS APPLICATIONS TO MUSIC EDUCATION," *International Conference e-Learning 2019,* 2019.

[85] "ShareTechNote," [Online]. Available: https://www.sharetechnote.com/html/5G/5G_CSI_RS_Codebook.html. [Accessed 13 10 2023].