Title: Vehicular Fog/Edge Computing to improve dependability and performance
Student: Mohamed Meskine

Problem description:

The number of objects connected to the Internet surpassed the world human population in 2010 and is expected to reach between 50 and 100 billion by 2020. Traditional centralized solutions are not suitable for this growth. Fog/Edge computing is a new technology that allows to move the computing, storage, and networking functions closer to the users. This will improve both network congestion and improve resource optimization, user experience and the overall performance of the network. This technology is important to meet the requirements of the 5G.

This thesis will focus on presenting a model for vehicular in the new communication generation. This model is based on the possibility of being able to exploit the characteristics of the new technology. This model will be applied in the vehicular field.

Vehicles will have within them a capacity to process local data and will be connected through two links with edge nodes. All data generated inside the vehicle will be processed. Data will be processed according to a hierarchy. We will have data with high and low priority. After this procedure of processing data, we will identify if those data needs information from the network or not. For example, the temperature inside the vehicle does not need to get information from the network. Contrary, knowing whether or not there are vehicles at a crossroads is essential, especially when the traffic lights are in night mode (yellow flashes).

The purpose of the thesis is to improve performance of this model based on the three aspects latency, mobility management and reliability.

In this model thanks to the presence of two links, it will be possible to increase the reliability. Migration process will be reduced thanks to the inter-connection between the nodes. For latency part it will be possible to decrease it thanks to the fact that the data processing will be done at the local level and also trying to make the two messages (from vehicle to node and vice versa) uncorrelated.

In this model we will provide results analytically or numerically. The simulation will be used only in case of necessity.

Date approved:	2020-01-15
Supervisor:	Norvald Stol, IIK

Abstract

We have more and more objects connected to the network. These objects need local information increasingly. This means that cloud computing is not suitable in some applications. This is also due to the high latency that it presents.

The thesis is focused on the vehicular field. In the presented application, we will use fog computing and 5G. These two technologies are important to guarantee a good performance. This is important because thanks to that we can avoid an accident.

The application is to provide that the user will be able to see the vehicles within a radius of 100 meters. The vehicles within this radius will be presented on the display inside the vehicle. Inside the vehicle there will be a processor that sends messages to the driver in case of emergency. In case that the warning remains, there will be automatic braking and/or steering.

The idea of the thesis is to improve three aspects: latency, mobility and reliability. Latency has been improved thanks to the idea of uncorrelation. Mobility, i.e. the time it takes to pass control from one small cell to another, has been improved thanks to the interconnection between the fog nodes and the information on the user's destination. Reliability is improved thanks to the interconnection between layers and besides thanks to the use of two links.

Preface

First of all, i would like to thanks Almighty ALLAH the most merciful and beneficent, who enabled and give guidance to me in completing this mammoth task. I would also like to acknowledge and express my sincerest gratitude to my supervisor Norvald Stol for providing help, scholarly advice, support, useful comments and continuous encouragement. I would also like to thank the Department of Information Security and Communication Technology. I dedicate this work to my family and my friends. Without their support and love, I could have not achieved any of the goals I had ever set in my life.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In the forthcoming fifth-generation (5G) cellular infrastructure, the Internet of Things (IoT) transmissions are deemed to reach 7 billion units by 2025 [RMZ+16]. Traditional centralized solutions are not viable for this growth. To meet the needs, new solutions must be chosen. Edge computing is identified to be compatible with the requirements. This technology allows to move the computing, storage, and networking functions from the central cloud to the edge of the network and enables the operators to manage the system efficiently. By pushing high data-intensive activities towards the edge of the network and performing data processing locally. Mobile network operators will benefit from this process by reducing traffic bottlenecks in core and backhaul networks. The user, on the other hand, will have a reduced energy consumption of the device. This is because the process is carried out by one of the nodes. Besides, since the process will be done close to users we will have a reduce latency. The process of moving cloud resources to the edge requires not only moving resources such as CPUs and storage but also stored content to provide faster data capture for end users.

This growth will give us two types of devices the low-end devices, such as wearable sensors and cameras, characterized by strict resource limitations (e.g., limited memory and processing capabilities), and the high-end devices equipped with high computational capabilities (e.g., autonomous cars, drones, robots, virtual reality-based systems) are increasingly populating the IoT [HBPT15], opening new market possibilities for highly integrated IoT applications [OSP+15]. In this thesis, we will focus on high-end devices. In particular, we will focus on the vehicular field. For the fact that we have an increase in the number of devices and sensors inside the vehicle. Some of these need to be connected to the network. They do it for two reasons, either to offer a service to passengers or to monitor our route efficiently. Edge computing is required to ensure real-time monitoring, especially during traffic peak times in a citywide area.

5G will play an important role in meeting vehicle requirements. This new telecommunication technology should guarantee the 4A (Anytime, Anywhere, Anyone, Anything). It should provide ultra-low latency. Besides, it should provide high throughput and five-nine coverage [Mat17]. For this new technology to guarantee these values, we need to use edge computing. This has been viewed as inseparable and an important ingredient of 5G networks.

This evolution brings along with its challenging requirements, especially in terms of increased bandwidth, mobility, and low latency, which need new solutions in both the radio access and the core network infrastructure.

1.1 Motivation

This technological advancement leads all objects to be smarter, autonomous in some cases even safer. Just like in our case in the vehicular environment. The motivation that led me to deal with this topic is because we could help in reducing the number of accidents and also improve traffic management. This choice is also dictated by the fact that I have a friend who was paralyzed after a car accident. This is because he could not see the vehicle coming the other way in a corner. Another reason that led me to choose this argument is the availability of two technologies that can guarantee almost instantaneous latency times. These two technologies are 5G and fog computing.

The choice of this topic was not that easy because a lot of research has been done in this area, so it is not easy to innovate. This factor pushed me even if it was hard to find new ideas in an area full of research. But despite this, the presence of different types of edge computing with different characteristics allows introducing new ideas and architectures. This prompted me to look for an idea that could help achieve certain goals in several aspects. Therefore, to present an application in the vehicular field with many advantages.

1.2 Methodology

The thesis does not include the use of software for scenario simulation. This means that no part deals with the implementation in software of the scenario, but there is an in-depth study that allows us to touch all the critical points concerning the vehicular environment.

The first step was to identify the topic and which user will be addressed by this service. This is because it is to make the appropriate choices. In our case, the system is aimed at all vehicles in circulation. This means that our idea will involve a few changes in the vehicle since it must be applied also to old vehicles.

We then have to identify the technologies that we want to use and make a study of them and see if they match the idea. Obviously, in the beginning, our idea was broadly based and with various. Through in-depth studies, we were able to find the idea that will allow us to get good performance. However, once we have the idea, we have to look in the related works to identify if it is already present. If there is such an idea, we try to innovate it. In our case, we did not find an idea that was close to our own. So we did not have to apply as many changes to the idea

Chapter 2

Edge Computing

The number of data produced at the edge of the network is increasingly increasing. Therefore a good solution is to allow data processing even close to the users. Edge computing allows processing and storage of data at the edge of the network [Sat17]. The word "edge" in this context refers to all computing and network resources between the user and the cloud. For example, if we consider the wearable objects we have that the mobile phone is the one that acts as edge computing. Another example could be in the area of smart homes. Where we have that the gateway acts as a link between the sensors and the cloud. We expect edge computing to have an impact on our society as big as cloud computing.

In the edge computing paradigm, the objects connected to it not only consume data, but they produce it. At the edge, things can not only require services and content from the cloud, but they can also perform computing operations from the cloud. Edge can perform computing offloading, data storage, caching and processing, as well as distribute the request and delivery service from the cloud to the user [SCZ$^+$16]. With these networked tasks, the Edge itself must be well designed to meet service requirements efficiently, such as reliability, latency, security, and privacy protection.

Edge Computing is a new solution that becomes essential to meet the requirements of the 5G. This new technology will favour integration with 5G. Likely the Edge Computing does not have the following limitations that the cloud has:

– The Responsiveness: the new 5G system should provide a latency of 1 ms. This vision, however, may be difficult or even impossible to achieve since the mobile user packages have to travel through the Radio Access Network (RAN), Core Network (CN) and the Internet before they reach the cloud server. Nowadays LTE networks reach 10 ms round-trip latency, 5 ms in RAN and CN and 5 ms on the Internet, this is possible only if the user and the server are located in the same country [Int14]. We have also the problem related to the bandwidth limitation and network uncertainty that makes the latency variable.

- The Backhaul Bottleneck: the 5G systems should guaranty the five nine and gigabit data rates [PDFF17]. However, even if these measures are taken to keep users connected, it may be difficult to achieve high data rates due to the backhaul bottleneck. The backhaul capacity should be at least comparable to that of the RAN, which is a challenging.

- Location-Aware Applications: in such applications, services are often requested by geographically adjacent users and data exchange takes place primarily at the local level. Pointlessly directing traffic to the cloud not only compromises responsiveness but also exacerbates backhaul tension. If this traffic can be managed in application contexts, users can be served with higher Quality of Experience (QoE) and cost-effectiveness.

In this section, we are going to discuss two different edge computing technologies. mobile edge computing (MEC) and fog computing. In the first part, we describe each technology in a settling way. In the second part, we will focus on the differences between the technologies. In the last part, we discuss which of the two technologies is suitable for the vehicular environment.

2.1 MEC

The acronym MEC at the beginning was referring to mobile edge computing [BS15], but then it has been changed to multi-access edge computing [HZXH15]. This is because we no longer just have mobiles, but we have a lot of things connected to the internet. MEC is a network architecture term established by ETSI [ETS] that enables cloud computing capabilities and an IT service environment at the edge of the cellular network and, more broadly, at the edge of any network [KAH+19]. MEC's underlying concept is that by executing applications and processing tasks nearer to the cellular customer, network congestion is minimized and applications function better. MEC is designed to be deployed in cellular base stations or other peripheral nodes, allowing flexible and fast deployment of new applications and services for customers. By combining IT and telecommunications network elements, MEC also enables mobile operators to extend their RAN to authorized third parties, such as application developers and content providers.

The MEC is characterized by high bandwidth, proximity, high latency and real-time knowledge of radio network and location information. This can be reflected in value and can generate opportunities for mobile operators, application and content providers, allowing them to perform complementary and profitable functions in their respective business models and enabling mobile broadband experience to be monetized.

2.2 Fog Computing

Like the MEC technologies, fog computing [HSE+12] brings cloud computing capabilities closer to users. This technology was introduced by Cisco [JMJZ19], fog computing is promoted by the OpenFog Consortium [SCZ+16] founded by ARM, Cisco, DELL, Intel, Microsoft and Princeton University. Fog computing corresponds to a single-layer or multi-layer heterogeneous node architecture with storage, computational and connectivity capabilities. Such fog computing servers, distributed by manufacturers, are found at different layers between the cloud and the edge of the network: WiFi access points, bridges, routers, gateways or even end devices. Fog computing's primary idea is to utilize the available computing and storage capabilities of these legacy appliances to deliver real-time virtualized services close to users. Any fog node is composed of an abstraction layer and an orchestration layer that enables the deployment of various applications independently of the type of devices.

The fog network supports the concept of the Internet of Things (IoT), where most of the devices used daily by humans will be linked together. For example, phones, wearable health monitoring devices, connected vehicles and augmented reality using devices such as Google Glass are examples of things that will be connected to the network [Ant14].

2.3 Differences

These two technologies, despite their common goal, are supported by different types of companies. Such companies perform different roles with different skills. Besides, they act at various levels by different means. Hence, these edge computing solutions present different characteristic terms:

- Architecture: figure 2.1 shows the positioning of the nodes of the two edge calculation techniques. The MEC servers are located one hop away from the end-users by forming a three-layer architecture cloud, computing node, terminal equipment. Fog nodes can be located one hop or multiple hops away from the user and are interconnected composing an architecture with at least three layers.
- Communication Technology: MEC, integrated within RAN, exploits the mobile network. Fog computing uses different technologies such as mobile network, WiFi, Bluetooth, etc depending on the type of equipment and the manufacturer. Thus, it may be feasible to access fog computing servers through the use of non-IP-based communications, expanding the scope of this technology.
- Nodes Cooperation: fog computing technology support fully the inter-node communications and providing effective cooperation between the fog nodes

forming the network. Fog computing is supposed to be embedded in legacy devices (routers, gateways, etc.) on a virtual layer, without adding any computational or storage capacity. Therefore, the capacities of these devices are expected to be more limited than those of MEC servers.

- Deployment Cost: The use of fog computing in legacy devices seems to be the most economical and simple solution. It just needs a level of virtualization. Implementing MEC servers in the existing network architecture is more complex and more expensive.

- Coverage and User Proximity: using the cellular network, the coverage area of the MEC technology could be more important than the coverage area of fog computing. The fog computing nodes should always be located in the WLAN area while the MEC servers could be located at the base station level or the radio network controller level. Therefore the latency of the MEC applications could be affected. Moreover, fog computing using WiFi or short-range communication should be closer to the user.

- Flexibility: Fog computing is not built on a fixed distribution, so adding or removing servers for fog computing could be very simple. Besides, fog computing is independent of the network and could be easily managed. Therefore this solution is more flexible than MEC.

- Virtualization Layer: these two different architectures are based on a virtualization layer. However, depending on the implementation, different technologies are used. For example, fog abstraction layer, NFV or OpenStack.

2.4 Technology for our application

We presented two different technologies MEC and fog computing. Both of these technologies bring numerous advantages. As we have already presented, they present differences even if they have a common goal that is to increase the QoE.

For our application we will use fog computing for the following two reasons:

- Nodes cooperation: This is a fundamental fact for our application because many improvements would be obtained thanks to this cooperation. In the vehicular field we have high mobility and therefore it is important to accelerate the passage of information from one node to the other. This connection will help to to that. This connection between the various nodes will be done with optical fiber.

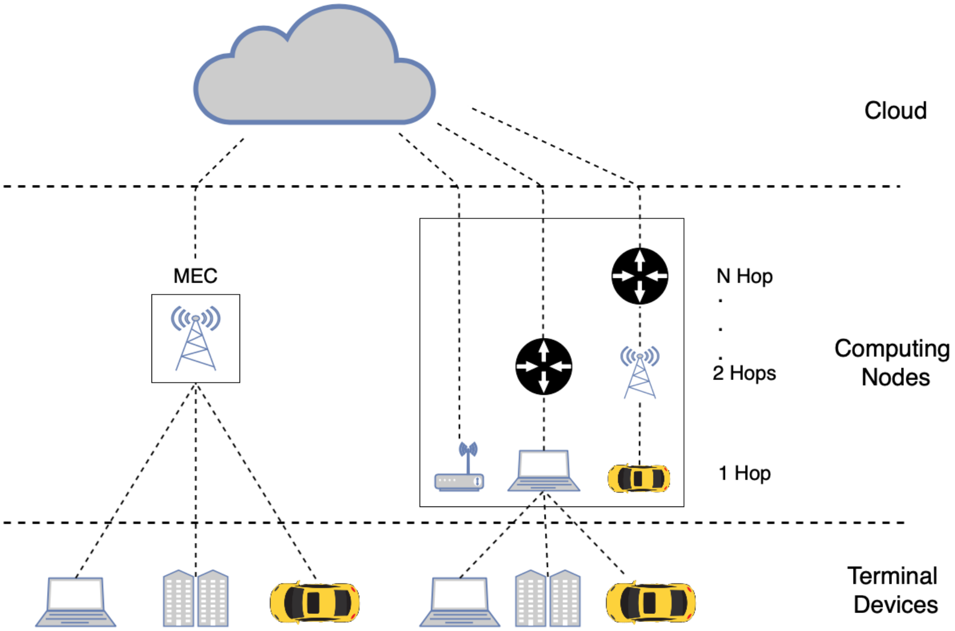


Figure 2.1: Edge computing technologies positioning. [Adapted from [MCK19]]

- Implementation cost: the cost of implementation fog computing is lower than that MEC implementation. This makes it easier to have more fog node and therefore decreases latency. In addition, it is easier to add or remove a fog node from the network.

We have disadvantages in using this technology, but for our application, it is fine that there are limitations in data processing and data storage. This is because our application does not require high computational capacity and neither required storage capacity.

Chapter 3

Architecture

This chapter is dedicated to representing an architecture that helps our application to achieve a good performance. But before we explain it we need some main concepts. One of those have already been covered in the previous chapter fog computing. So, it is good practice to describe the technologies present in the architecture. This description will not be detailed because in the following chapters some those it will be covered.

The chapter is structured as follows, we have as a first section a description of 5G. This description outlines the three fields where this new technology will be used. In the second part we treat two key technologies software defined networking (SDN) and network function virtualization (NFV) that allow to have a heterogeneous environment. In the final part we have a description of our architecture.

3.1 5G

In the coming years, global mobile data traffic is projected to increase at a 45% compound annual rate, representing a 10-times increase between 2016 and 2022 [Mat17]. Most of this growth is due to the widespread adoption of video mobile streaming. In addition, the IoT is moving from vision to reality, and out of the 30 billion connected devices it is expected to include by 2022, 18 billion will be IoT and machine-to-machine devices [Mat17]. The 5G networks will have to support these demanding new use cases in a cost and energy efficient way. However, 5G is designed to provide support for a variety of different use cases scenarios. Figure 3.1 shows the three main categories:

- Enhanced mobile broadband (eMBB): is better than what we currently have, but with enhanced performance and with a better QoE. It covers a variety of cases, including extensive coverage of areas and hotspots. In the wide area case, we want seamless coverage and high mobility, with greatly enhanced user data rates versus those currently offered. For hotspots, high user density and very high capacity is required, but the need for mobility is only at pedestrian velocity. Notice that the rate required for user data is significantly higher than that of a wide area coverage.

- Ultra-reliable and low latency communications (URLLC): it refers to the scenario when we have strict reliability, latency and availability criteria to be meet. For example tactile applications on the Internet, smart transport systems, vehicle-to-everything (V2X), transport security, remote medical surgery, smart grids, public protection and disaster relief, wireless industrial production control, etc.
- Massive machine type communications (mMTC): is a range of applications for which the traffic models are not even fully characterized. However, we know that an mMTC implementation could consist of a very large number of devices with a relatively low (or relatively high) volume of non delay-sensitive data. The devices must be low cost and have a long battery life.

Heterogeneous services in 5G can coexist in the same architecture of the network via network slicing [PFD11]. A network slicing allocates the network's computational, storage and communication resources between active services with the objective of ensuring isolation and certain levels of performance. Our interest is in the slicing of RAN communication assets for wireless access. Traditional approach to cut RAN is to assign radio resources orthogonally to eMBB, mMTC and URLLC devices in domains of time and/or frequency, consistently with the orthogonal assignment of cable communication resources. However, wireless resources are essentially different because of their shared nature.

The ability to manage applications that require different requirements absolutely needs network slicing. This technology is also favored by the use of SDN and NFV. In the following sections these two technologies are explained in more detail.

3.1.1 SDN

The SDN is making it easier for organizations to deliver flexible delivery and deploy applications, providing the ability to scale network resources according to application and data requirements and minimizing both CapEX and OpEX [NKW$^+$12]. We return back to the old type of designing, implementing and managing networks that divides network control (control plan) and forwarding process (data plan) to improve the QoE. Such network segmentation has several advantages in terms of network control and flexibility. It combines the benefits of system virtualization and cloud computing and enables a centralized deployment of intelligence that provides clarity on the network for easy network management and maintenance, as well as

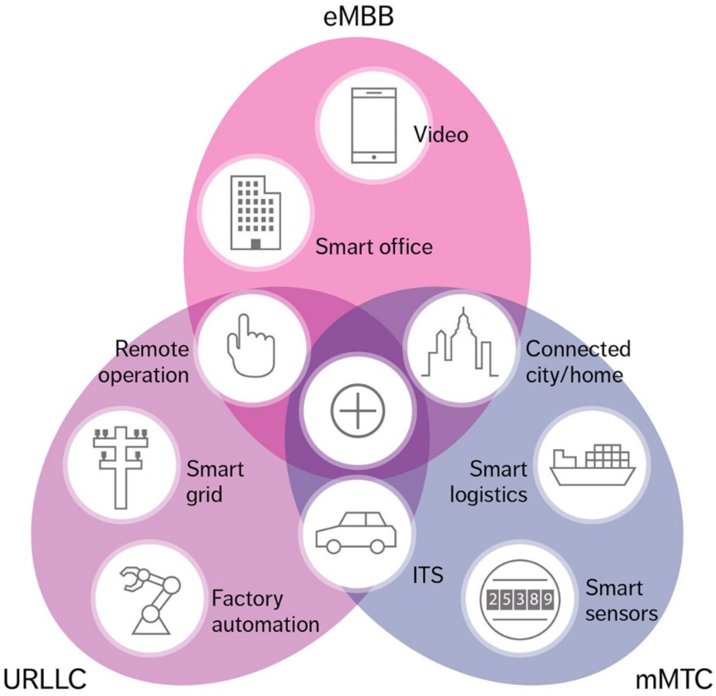


Figure 3.1: The three main 5G use cases and examples of associated applications [TWS+17].

enhanced network control and responsiveness. In traditional infrastructure, network deployment, configuration, and troubleshooting require high-level technical, network and system engineers and the operational costs required for large multivendor network provisioning and management. Indeed, due to the diversity and complexity of components on the network, their maintenance becomes very costly and the infrastructure underneath unreliable in the event of repeated network failures, particularly if no backup plans are foreseen within the infrastructure. One downside to the SDN is that since everything is centralized we have less reliability compared to distributed systems. The same applies to the dependability. In table 3.1 is represented a comparison with classic networking and the new SDN technology.

Since SDN decouples routing and forwarding decisions for network elements (access points, switches and routers) from the data plan, network management and administration is simple because the control plan only deals with information about the logical network topology, traffic routing, and so forth. In reverse, the data plan orchestrates network traffic in the configuration set out in the control plan. The control operations in SDN are centralized in a network policy controller.

Table 3.1: Software defined networking versus classical networking [NKW$^+$12].

Characteristics	SDN architecture	Classical architecture
Programmability	✓	
Centralized control	✓	
Error-prone configuration		✓
Complex network control		✓
Network flexibility	✓	
Improved performance	✓	
Easy implementation	✓	
Efficient configuration	✓	
Enhanced management	✓	

Several control platforms are available as open source, such as Beacon [Eri13]. Network management can be implemented on different levels (control, application and data plan). For example, providers can assign resources to customers through the application layer, configure and change network policies and logical entities on the control plane, and set data plan physical network elements.

3.1.2 NFV

The NFV comes out of the industry and offers to solve inconveniences, thus preventing the steady growth of the hardware devices. Furthermore, it makes it easier and more innovative in the network by exploiting virtualization technology to provide a new approach to network design [HGJL15]. The European Telecom Standards Institute (ETSI) [JAQ90] was chosen in 2012 by seven of the world's leading telecoms network operators as the location for the industry's NFV specification group. According to the NFV paradigm, classic middleboxes are handled as individual software modules, programmed to perform the role of a specific NFV, this enables the scalability and isolation of each function, so they can be independently handled. Furthermore, the NFV makes it easier to install and deploy NFVs on generic servers, thereby permitting NFVs to migrate dynamically from one server to another, in other words anywhere on the network [ETS13].

NFV provides solutions to most of today's network issues due to the extensive deployment of specific hardware equipment. It also offers network improvement and cost cutting opportunities. Furthermore, it allows the setup of hybrid environments in which functions running on virtualized resources coexist with functions running on physical resources [MSG+15]. These hybrid environments can be important in the transition to NFV. The traditional implementation of a network service demands that data traffic flows via a defined set of middleboxes in a given sequence, causing some processing depending on the function they perform. This is commonly known as middlebox orchestration [GKJ+13]. Actually, it is manually performed and set to the items in the router forwarding table. This task is a cumbersome and error-prone process. In addition, whatever positioning of these hardware middleboxes is going to become ineffective over time. This is because it is very expensive and unpractical to keep on changing the position of these hardware as network conditions change.

NFV implementation challenge is to achieve a scalable, fast and dynamic Network Function (NF) composition and allocation to run a network service. Nevertheless, as a network service requires a set of NFVs, the achievement of efficient NFV service coordination and management raises two issues:

- How to compose NFVs for a specific network service.
- How to allocate and program efficiently the NFVs of a network service on a substrate network.

Through its NFV technology group, ETSI is collaborating with network operators and equipment suppliers to foster NFV and is currently working on the first application mentioned above.

3.2 Combination between SDN and NFV

NFV is a concept that supplements SDN. Therefore, NFV is independent of SDN or SDN definitions. NFV disconnects software from hardware to enable flexible network distribution and dynamic functioning. NFV implementations usually use commodity servers to run software versions of network services that were formerly hardware-based. Such software-based services that are executed in a NFV environment are referred to as virtual network functions (VNFs) [Sta15]. The SDN-NFV hybrid combination is delivered for elastic, scalable, high efficiency NFV capabilities to speed service innovation and delivery using standard information technology virtualization [Sta15]. The SDN gives users the flexibility to easily monitor and control generic forwarding devices like routers and switches using SDN controllers. Additionally, NFV agility is delivered for network applications utilizing virtualized servers. It is fully feasible to deploy a virtualized network function (VNF) as a stand-alone entity utilizing already existing network and orchestration paradigms. Nevertheless, There are benefits inherent in leveraging SDN solutions to deploy and manage NFV infrastructure, particularly when it comes to NFV management and orchestration, and this is why multivendor platforms that incorporate SDN and NFV into concerted ecosystems are being defined.

3.3 Proposed Architecture

In this section we will describe the architecture of our system. In figure 7.3 we can see that it is not represented only the architecture of 5G technology, but we have fog computing integration. There are several architecture proposals that represent this integration, but do not describe the scenario that we would like to describe. For example in [KLL$^+$17] the authors focus on identifying the typology of the data (low priority or highly priority) and than deciding if it need local processing or not. In the same article we can read that they are built on much of the work done by the baseband unit (BBU). In addition, they are based on the use of SDN because architecture assumes a heterogeneous scenario. However, if we consider this article [VLG$^+$17] we have that presents an architecture very similar to the previous one. The innovation presented in this article is the use of NFV technology that we describe in the previous section. While our model will involve the use of these new technologies but in a way that we can achieve good performance. However, SDN and NFV they will not play a fundamental role in our vehicle application. This is because we will consider a homogeneous scenario.

In figure 7.3 we can identify four levels:

– Vehicular level: on this level that constitutes the lowest level of architecture, we have the vehicles. As can be seen we have that each vehicle relates to two links at two different levels. They are represented with two different dotted lines. This is to underline the fact that the two links use two different bandwidth frequencies.

– Fog node level: in this level we have mainly two components. The first is the small cell which is a component that allows the transmission and reception of signals. In addition, we have that it provides the data received to the fog node through the optical fiber and transmit the information received from the Fog node. The second component is the Fog node which can be identified as the intelligence closest to the user. The fog node has several functions. One of the most important functions for our application is to select the positions of the other vehicles to be transmitted to each user. Moreover, it has the task to help in the migration process between two small cells. In addition, has the task of transmitting the destination of each vehicle to the next level (BBU level). It also transmits useful information to the nearby Fog node.

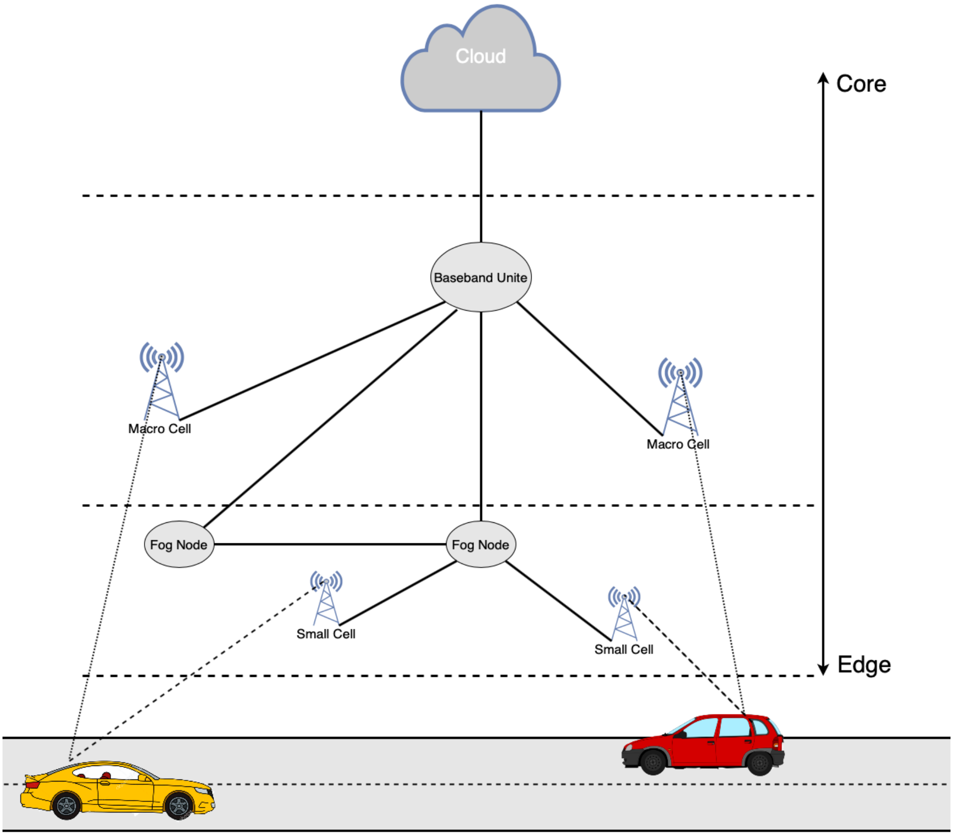


Figure 3.2: Proposed architecture in vehicular application.

– BBU level: at this level we have the BBU that does a mainly work of control and support to the fog nodes. Obvious if we consider the entertainment, we have that is to guarantee an excellent service to the vehicles. Now let's forget the entertainment part and focus on the monitoring vehicle part. The main task in case of normal operation is to provide information to the Fog nodes to which Fog node they must pass the control of the vehicle. Instead in case of problems with the fog nodes then the vehicle will communicate with the macro cell that transmit the information to the BBU that will provide the necessary information about the other vehicles position. This is possible for the simple reason that we have that the Fog node transmits data of the position of the other vehicles also to the BBU. So, with that information, it has the biggest picture of the scenario. This information can be used, for example, in an emergency to warn all vehicles in proximity of s trafficked area due to an accident to change roads. Another application could be to show the emergency vehicles the least crowded route and in the meantime warn the vehicles on the route of the emergency vehicle.

- cloud level: this level does not play an important role about vehicular application. But for the entertainment part is concerned, it has the task of providing the required data from the users in case it is not present in the BBU.

These four levels represent the main levels that make up our architecture. Of course, all levels are important for the proper functioning of the system. Thanks to the connections between the various levels in some cases is guaranteed a good service even in the presence of components that fail. This architecture will guarantee a fast and reliable service.

Chapter 4

Vehicular Application

This chapter is dedicated to describing our system, that present a model of a vehicular application. In this chapter, we are going to present a detailed description of the system. Besides, the functioning of the system itself is also covered. We also propose how our system will behaves in different situations.

The second part, which concerns the dimensioning of the system, is structured coherently until the result is reached. We start with a description of the element that plays an important role in our system. One of those is the presence of the small cells. We also illustrate the peculiarities involved in the use of high frequencies.

The final part shows how it is possible to calculate the channel capacity. Thanks to the previous information it is possible to estimate the required system capacity.

4.1 Relatives work

Many articles are dealing with the vehicular field in the new 5G technology. Some of these deal with the possibility of using Long-Term Evolution (LTE) and Dedicated Short Range Communication (DSRC). For example, article [DRC+16] presents cooperation between the two technologies. This helps to increase coverage, but still far to guarantee the required latency. Besides, we have a strong dependency on speed.

The recent papers focus on the use of edge computing. This is dictated by the high performance provided, especially in terms of latency. In this article [ZML+17] is represented as one of the edge computing technology that is the MEC. The idea of the article is to exploit the communication between vehicles for off-loading of data. The use of vehicular communication is essential to maintain latency requirements. This is because the MEC does not provide communication between MEC. Another reason is that if we want to communicate between MEC we must go through the cloud. This means an increase in latency. In article [LWJ+17] it is used the same edge computing technology as the present article, to ensure an efficient emergency service. The MEC's task is to inform the vehicles of the presence of an accident and to suggest the drivers divert the route. The time needed for rescue to reach the accident with this technology decreases considerably. In the following article [LWZ^{+}17] is presented on how to integrates different types of access technologies. This is possible using SDN. To ensure good performance the MEC is used. In this other article [NWH18], a new access technology non-orthogonal multiple access (NOMA) is presented. Here the MEC is used to guarantee low latency. This article deals with the collection of network traffic. The last article [NHW^{+}19b] using MEC deals with energy consumption. They act on scheduling to minimize energy consumption.

let's focus now on the articles that use our same technology fog computing. Analyzing this article [HLC^{+}16] we find that it presents a new architecture. In this architecture, we have that fog node are mobile. This means that the vehicles are fog nodes. This article [ZML^{+}17] is based on communication between vehicles like some articles already mentioned. Relying on this type of communication we must deal with the reliability factor. In this other article [NHW19a] they deal with fog node response time. In the same article, we have that the data is prioritized with a priority level. This selection then allows identifying where this data should be processed locally or not. In this article [XZ17] as in the previous one [HLC^{+}16] we have that it is the vehicles that act as fog node. However, the vehicles are not common ones, but public ones like buses. It is represented that also the taxis act as fog node.

Many articles are dealing with the vehicular field, but many of them do not have an application that can be used. In the sense that in many cases an idea is represented, but without explaining how it can be used in real life. The one that presents a solution usually have problems with latency and/or reliability. In this article, we present a vehicular model that allows reducing latency, mobility management and to increase the reliability. Besides, this model will decrease the number of accidents and it will help for better traffic monitoring. Our model is not based on vehicular communication, this allows us to have an excellent service even in the absence of other vehicles.

4.2 System Description

The presenting system is based on simplicity. This is since we want the system to be usable even by users who own an older car. This is possible thanks to the use of only essential elements. Thanks to the simplicity of the system it is possible for all vehicles to use it. The underlying idea for this system is to guarantee a service that can be used from today, but at the same time with future requirements. The system is about seeing what others see. Let's explain better, each vehicle will be able to see the vehicles present in a predefined radius on a screen. The screen represents the vehicles as in figure 4.1. The quality of the display depends on the user's choice. This choice does not affect the functionality of the system. The idea also came to us thinking about some video games we used to play as kids. In some of them, you can know what happens outside the radius, represented on the full screen. That was fundamental. For example in this case figure 4.2. knowing the position of the goalkeeper is essential to enable you to adopt the right moves to score. In the vehicular field is even more important, because it is not about victory or defeat, but about life or death. There are several situations where knowing what is going on the other side is essential. Here we present some examples:

Figure 4.1: Screen view.

- When the traffic light starts blinking after a certain time of the night.
- When you are about to face a road curve. Especially in the case where you cannot see the vehicles that are coming.
- At the entrance of a highway where there was the acceleration lane.

There are numerous cases where information on the location of other vehicles is useful. We have presented what we believe to be the most frequent and dangerous situations.

4.2.1 System Details

In this subsection we are going to describe the components or services that characterize our system:

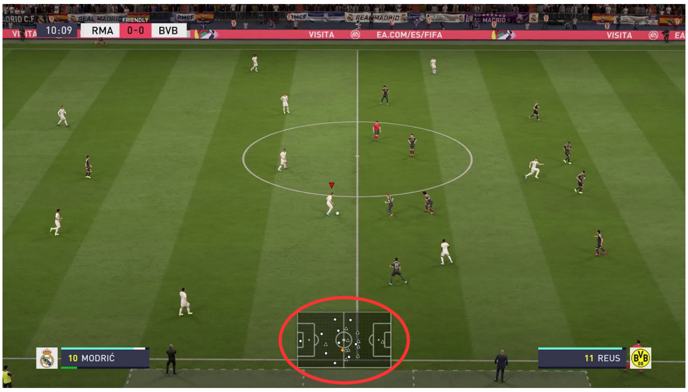


Figure 4.2: Representation of a radar in a video game [SPO].

- Localization and positioning system: these two information are fundamental for the proper functioning of our system. Therefore, these terms seem interchangeable, but they are not. We have that localization incorporates positioning. Because positioning is knowledge of coordinates, but the location in addiction it represents you in an environment. How we get these two values will be discussed later. In addition, several technologies will be presented that give different accuracy.

- Screen: represent the main objective for the driver. This is because it represents to the driver the vehicles present is a certain range. Its functionality is presented in the previous section.

- 5G: We already mentioned the characteristics of this technology in the previous chapter. Thanks to its numerous advantages it will be used in our system as a communication system. From the standard, we can read that two frequency bands will be used. The first involves frequencies ranging from 410 MHz to 7125 MHz [3GP19b]. The second range from 24250 MHz to 53600 MHz [3GP19b]. Both will be used in our system. The motivation will be explained later.

- Processor: This will play an important role in our system. It will map the positions received for the network on the screen present inside the car. But it will also translate the received warnings messages into actions. it will also make calculations for the distance needed to brake.

- Fog nodes: In the system, they will do a vital job to achieve the requirements proposed by 5G. We will have one for every three small radio base stations. This is in case we are in a trafficked environment. In the section where the system is dimensioned, we will also talk about this factor where we will explain what is meant by trafficked environment. The main task will be to transmit the position of the other vehicles to each vehicle. The cooperation between the Fog nodes will perform a key task to reduce the time needed for performing the handover procedure.

– WiFi: The system will also includes WiFi for passengers. It might seem a contradiction, with the fact that we have said that our system will contain only basic elements. It is introduced to take advantage of the connections that we are going to adopt. Since we are going to use two frequencies, why not use one to increase the Quality of Service (QoS).

4.2.2 System Functioning

This section presents how our system will work. We will provide two situations. In one is presented a normal functioning of the system. On the other hand, we will have the anomaly situation and discuss how it will be handled. Figure 4.3 represents our system where we have vehicles connected to the small cell using the 5G.

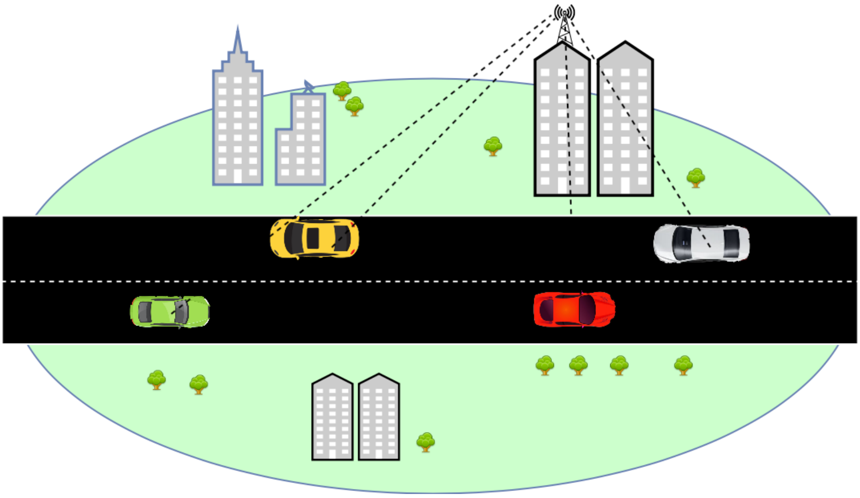


Figure 4.3: Scenario representation.

Normal system behave

Let's consider a user who wants to reach a destination such as work. When the driver enters the vehicle, he enters the destination that he wants to reach. This information will be sent to the small cell which in turn transmits it to the Fog node through the fiber optic connection. The Fog node will send this information to the baseband unite.

Our system, as already said, it will use two frequencies for communication. The high frequencies millimeter wave (mmWave) for information traffic with a high priority level. The second frequency will be used for the entertainment part. The information related to the destination of the passenger, is used so that the baseband unite know in advance which area he will be going to and to which fog node he will connect to. This mechanism, as we will see, will reduce the handover timescales. The vehicle must send there position, velocity and direction to the Fog node frequently. This information about each vehicle will then be forwarded to other vehicle.

On the screen, the user will see the route to the destination and the other vehicles on the road. This will be possible because the Fog node transmits to the vehicle the position of all the vehicles within a radius of 100 meters. The processor that is located inside the vehicle will have the function to map these positions on the map.

Abnormal system behave

One thing needs to be defined before we start describing what happens in abnormal situations. What we mean by the abnormal situation. With the term abnormal, we do not indicate when some components cease to function or starts giving errors. These problems will be dealt within the reliability section. So, we refer to when a road danger arises.

Let's consider a situation as we did in the privies section. This time our he is coming back from a party. he wants to go home. Once in the car he types the destination and the system starts.

Let's take a look at the situation when he come to an intersection with blinking traffic lights. As you know some traffic lights after midnight (the hour can vary depending on the state and area) starts only blinking. In this situation is a good habit to slow down/stop and then check before crossing an intersection even if you have priority. However, many cross without paying attention to others. This behaviour in some cases leads to road accidents

Now let's see how our system behaves in this situation figure 4.4. As a hypothesis, we have that both vehicles coming to the intersection can transmit their position, speed and direction. They must also be able to receive the position of the other vehicle. Our system will be considered as a level 2 figure 4.5, so, it will also be able to brake or steer automatically under indication from the processor. Before we get to that. Vehicle A (without precedence) receives on the screen that it must slow down because another vehicle is coming over at the intersection. Vehicle B is notified of the presence of vehicle A and to reduce speed even if it has the priority. If the driver of vehicle A does not decrease speed, the processor starts to calculate based on the speed of the vehicle the distance required to avoid hard braking and at the same time stop before the crossroad. Once this distance has been reached, automatically soft braking begins.

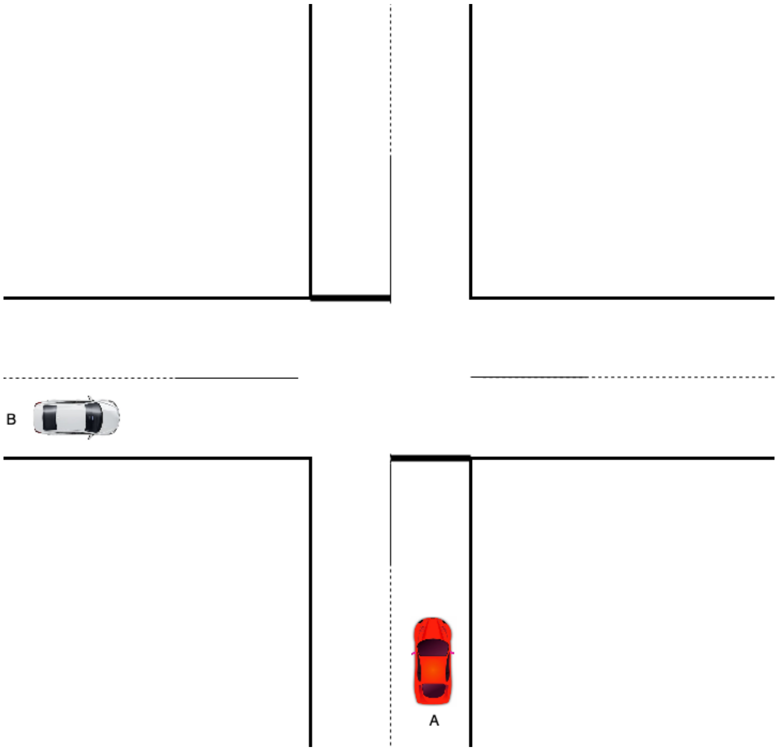


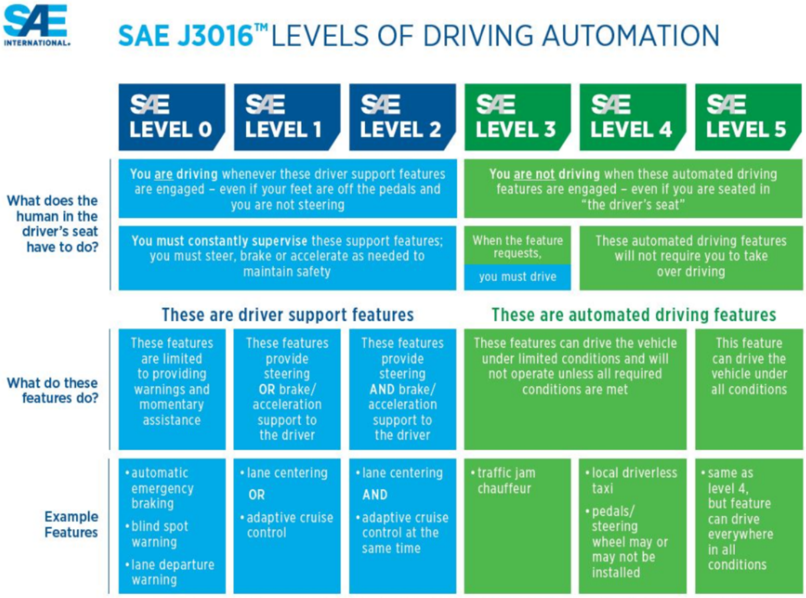
Figure 4.4: Two vehicles at the intersection.

With our system, both users will get home safely. This system also allows better monitoring of the roads as there is no need to stop in case there are no cars in the situation considered above.

4.3 Small Cells

Small cells are wireless senders and receivers that are intended to provide network coverage to smaller areas. So, while macro with an higher transmitting power towers keep the network signal alive over long distances, small cells can reach short distance. They are adapt to more densely developed areas such as cities. Small cells have lower coverage. This is also due to the transmitting power and frequency used to transmit the signal.

The 5G era brings with it the need for multi-gigabit speed and ultra-reliable, this can be accomplished through small cells. Small cells help bring new 5G connections



SAE J3016™ LEVELS OF DRIVING AUTOMATION

	SAE LEVEL 0	SAE LEVEL 1	SAE LEVEL 2	SAE LEVEL 3	SAE LEVEL 4	SAE LEVEL 5
What does the human in the driver's seat have to do?	**You are driving** whenever these driver support features are engaged – even if your feet are off the pedals and you are not steering			**You are not driving** when these automated driving features are engaged – even if you are seated in "the driver's seat"		
	You must constantly supervise these support features; you must steer, brake or accelerate as needed to maintain safety			When the feature requests, you must drive	These automated driving features will not require you to take over driving	
	These are driver support features			These are automated driving features		
What do these features do?	These features are limited to providing warnings and momentary assistance	These features provide steering **OR** brake/ acceleration support to the driver	These features provide steering **AND** brake/ acceleration support to the driver	These features can drive the vehicle under limited conditions and will not operate unless all required conditions are met		This feature can drive the vehicle under all conditions
Example Features	• automatic emergency braking • blind spot warning • lane departure warning	• lane centering **OR** • adaptive cruise control	• lane centering **AND** • adaptive cruise control at the same time	• traffic jam chauffeur	• local driverless taxi • pedals/ steering wheel may or may not be installed	• same as level 4, but feature can drive everywhere in all conditions

Figure 4.5: SAE J3016 "Autonomous Driving Levels" [BMLS17].

where some devices require specific requirements. They can be used to address even the densest environments such as arenas and train stations, improving economy for mobile operators and opening the door to more affordable data plans. Small cells are a long way from being a new concept since they started their implementations with 3G, and more so with 4G, they are now ready for significant growth with 5G.

The emphasis on including small cells as a critical component of the 5G deployment plan is highlighted by 5G mmWave. 5G and mmWave are used almost synonymously, but there are fundamental differences between the two. The mmWave technology is just a part of what the future 5G networks will use. The 5G is using low-band and sub-6 GHz frequencies [ARS16], both are part of the standard.

The term mmWave is used to refer to a particular part of the radio frequency spectrum ranging between 24 GHz and 53 GHz [3GP19b], which have a very short wavelength. This part of the spectrum is unused, so mmWave technology is designed to significantly increase the amount of bandwidth offered. The lower frequencies are occupied by TV and radio signals, like today's LTE networks, which generally lie between 800 and 3,000 MHz [3GP17]. One other positive aspect of this short wavelength is that it can transmit data more rapidly, even if its transfer distance is shorter.

The objective with mmWave is to increase the data bandwidth available over smaller, densely populated areas. It will be a key part of 5G in many cities, powering data in sports stadiums, malls, and convention centres, as well as anywhere data congestion, might be a problem. Out in rural towns and villages, sub-6 GHz and low bands below 2 GHz will probably play a more crucial role in ensuring consistent coverage.

4.3.1 Path Loss

The signal transmission over a wireless radio channel is affected by the path loss which depends mostly on the distance between the receiver antenna and the transmitter antenna, the antenna specifications and the operating frequencies. Furthermore, the behaviours of obstructing objects in the radio channel such as walls, terrain, buildings, vegetation, weather condition and other objects have an impact on the path loss.

Any computational procedure for the prediction of radio wave propagation can be defined as a propagation model (or also a field prediction model). In practice, a propagation model may consist of:

- A simple analytical, closed-form formula.
- An integral/differential expression which requires a numerical, computer-aided resolution.
- A software tool based on the general theory of propagation (Maxwell's equations).

The prediction accuracy depends on the capability to take into account all the aspects involved in the propagation phenomena (path loss, multipath, angle/time/frequency dispersion, etc.). Usually, the more reliable the model (i.e. the greater the accuracy), the higher its complexity and the correspondent computational burden. However, a prediction error greater than zero is unavoidable for every propagation model.

In [RXM$^+$17] we have formulas for the calculation of the path loss in different conditions and environments. This article unlike other older articles, where there are formulas for the calculation of path loss for frequencies below 6 GHz. This article deals with mmWave in both the two situations line-of-sight (LOS) and non-line-of-sight (NLOS). To be safe we are going to take the worst case scenario. So, we are going to consider the case when the transmitter and the receiver are NLOS.

Historically, standard bodies provide omnidirectional path loss models by assuming unitary gain antennas for the generality. Nevertheless, it is important to observe that omnidirectional path loss models will be unsuitable for use in the directional antenna system analysis unless the antenna models and true spatial and temporal multipath channel statistics are known or properly modelled [MR17], [RMSS15]. The model we are going to adopt will use this assumption.

We are considering the close-in (CI) free space reference distance model (with a 1 m reference distance) [RMSS15], [SRT^{+}16]. The CI path loss model accounts for the frequency dependency of path loss by using a CI reference distance based on Friis's law as given by [Rap18], [MRSD15].

$$PL^{CI}(f_c, d)[dB] = FSPL(f_c, 1\,m) + 10n\log_{10}(d) + \chi_{\sigma}^{CI} \tag{4.1}$$

Where $\chi_{\sigma}^{\mathbf{CI}}$ is the shadow fading that is modeled as a zero mean Gaussian random variable with a standard deviation in dB, n is the path loss exponent (PLE) found by minimizing the error of the measured data to 4.1, d > 1 m, FSPL(f, 1 m) is the free space path loss (FSPL) at frequency fc in GHz at 1 m and is calculated by [Rap18], [Fri46]:

$$FSPL(f_c, 1\,\mathrm{m})[dB] = 20\log 10(\frac{4\pi f_c \times 10^9}{c}) = 32.2 + 20\log_{10}(f_c)[dB] \tag{4.2}$$

where c is the speed of light. Using 4.2 it is clear that 4.1 can be represented as given in the following equation:

$$PL[dB] = 32.2 + 31.7\log_{10}(d) + 20\log_{10}(f_c) \tag{4.3}$$

We have that d in the distance between the transmitter and the receiver and it is measured in metes and the frequency in GHz. This is valid for NLOS. For urban microcell. With a shadowing fading standard deviation 8.09 dB.

Shadowing

Let's consider a transmitter and a receiver separated by a certain distance and in the first Fresnel ellipsoid there is an obstacle (human, object,vehicles and etc) something that has a size in the same order of λ or larger. In this case we have the so-called shadowing.

The ellipsoid between the receiving and transmitting antennas is the Fresnel Zone figure 4.6. The First Fresnel Zone (FFZ) is the difference between the direct path ($\overline{XY}$) and an indirect path that touches a single point on the edge of the Fresnel zone ($\overline{XZY}$) is half the λ [GB13]. In formulas we have:

$$\overline{XZ} + \overline{ZY} - d = \frac{\lambda}{2} \tag{4.4}$$

Rewrite the term with the Z-point coordinates, assuming that d=d_1+d_2:

$$\sqrt{(d_1)^2 + (r)^2} + \sqrt{(d_2)^2 + (r)^2} - (d_1 + d_2) = \frac{\lambda}{2} \tag{4.5}$$

By assuming that the antenna distances to the Z point are significantly larger than the radius, by expanding the roots in series and preserving the first two terms, the expression is simplified:

$$\frac{(r)^2}{2}(\frac{1}{d_1} + \frac{1}{d_2}) = \frac{\lambda}{2} \tag{4.6}$$

Which can be solved for r:

$$r = \sqrt{\frac{\lambda \times d_1 \times d_2}{d_1 + d_2}} \tag{4.7}$$

The transverse radius of each Fresnel area is the longest in the middle of the LOS. We can calculate the maximum FFZ radius by making some assumptions:

- d_1=d_2=$\frac{d}{2}$.
- $\frac{d}{2}$»λ. This assumption is used in the previous equation 4.7 as well.
- λ=$\frac{c}{f}$

$$r_{max} = \frac{1}{2}\sqrt{\lambda d} = \frac{1}{2}\sqrt{\frac{cd}{f}} = 17.32\sqrt{\frac{d}{4f}} \tag{4.8}$$

The distance d is expressed in kilometres and the frequency f in GHz. The maximum radius is expressed in meters.

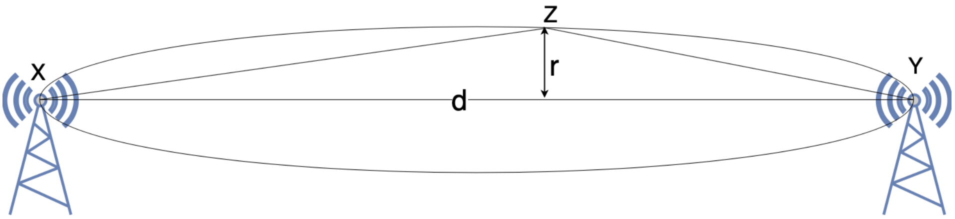


Figure 4.6: First Fresnel ellipse.

After explaining what first Fresnel ellipse is and finding out how to calculate it, now we can comeback to explain shadowing by considering a scenario. In figure 4.7 we are considering omnidirectional antenna, and we assume a LOS condition (when the transmitter and receiver see each other). Hence, with this type of scenario, we would expect the receiving signal level to be the same at the same distance regardless of the position of the receiver around the transmitter. This is true if there are no obstacles and if we have a plain surface figure 4.7.

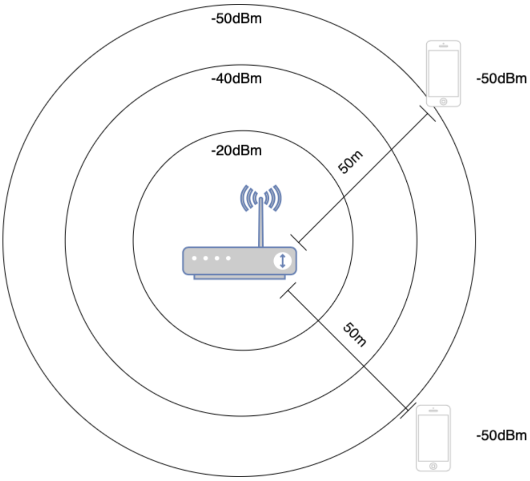


Figure 4.7: Scenario in presence of no obstacles.

The reality is not like the described situation, described previously few lines. In real life scenarios we have obstacles figure 4.8. In this case we have a building that cover the LOS. We are evaluating the same situation as before, but with the addition of an obstacle. We have that the receivers are at the same distance from the transmitter, but in a different location (different angle). The expected received signal in presence of an obstacle is much weaker figure 4.8. This phenomenon happens due to the fact that every time the signal goes through an object it weakens. In our case, we have a building. In this kind of situation it is even weaker because the signal has to pass through several walls. This is known as shadowing, or slow fading.

Fading

In the small scale phenomena, we consider position of the receiver in an area that size is in order of few wavelength. In this case, we observe a phenomenon that is the fading, completely separated with shadowing. Fading is considered a small scale phenomenon that means, when the position changes by a few wavelengths we have a change in the received signal. In shadowing we considered large scale phenomenon.

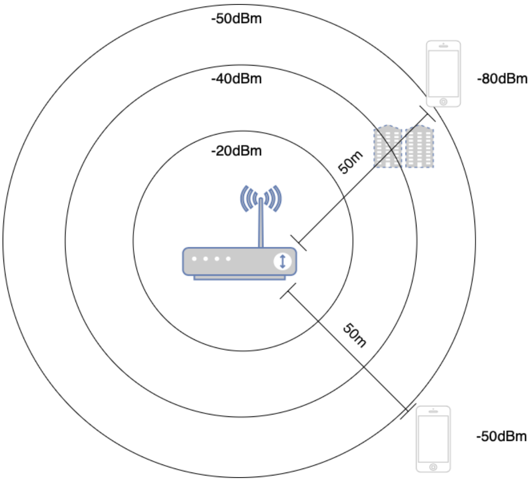


Figure 4.8: Scenario in presence of obstacles.

The cause of this small scale phenomenon is the presence of multiple path figure 4.9. The receiver might receive the signal from different paths due to reflection and diffraction. Each of those two phenomena is a potential source of a new path.

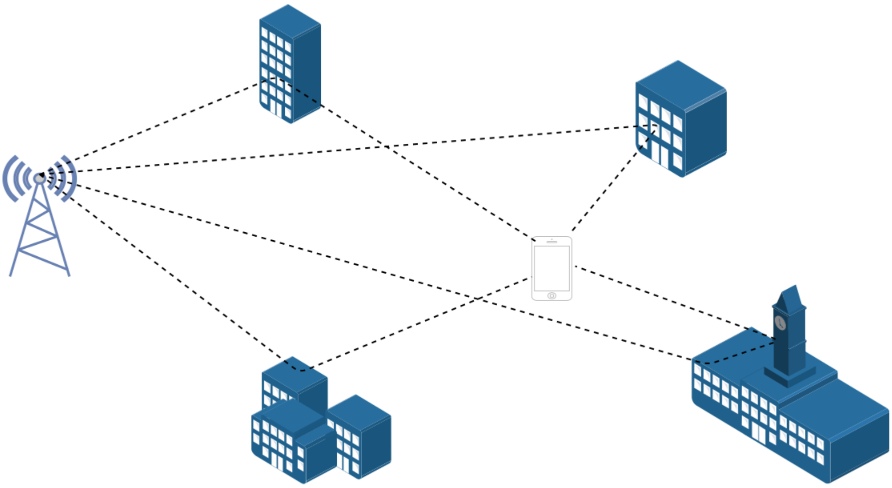


Figure 4.9: Multiple path scenario.

Fast fading happens when the channel coherence time is short compared to the application delay requirement. Here, the amplitude and phase change set by the channel changes significantly over the usage period. The coherence time is the duration of the time the channel's impulse response is considered non-variable. This channel change is much more significant in wireless communication systems due to Doppler effects. The Doppler effect is the frequency variation of a wave concerning an observer moving with respect to the source of the wave.

4.4 Channel Capacity

Within a communication system, we have the three main components: the source, the destination (user), and the media between them figure 4.10. We call the media the (communication) channel.

Figure 4.10: Diagram.

A channel could be in any form. It could be physical wires, cables, open environment in wireless communication, antennas and certain combination of them. In our case we are considering a wireless propagation.

The channel capacity, C, is defined to be the maximum rate at which information can be transmitted through a channel [Sha48].

$$C = B_c \log_2(1 + SNR) \tag{4.9}$$

The capacity formula represent a theoretical value that can not be reached. In the formula we have the bandwidth B_c that is the difference between the upper and lower frequencies in a continuous band of frequencies figure 4.11. It is expressed in Hertz and, according to the situation, may refer either specifically to the pass-band bandwidth or to the base band bandwidth.In our case we are referring to the pass band bandwidth since that we are dealing with mmWave. A significant aspect of bandwidth is that any given bandwidth can transport the same information, independent of where in the frequency spectrum that bandwidth is situated. The

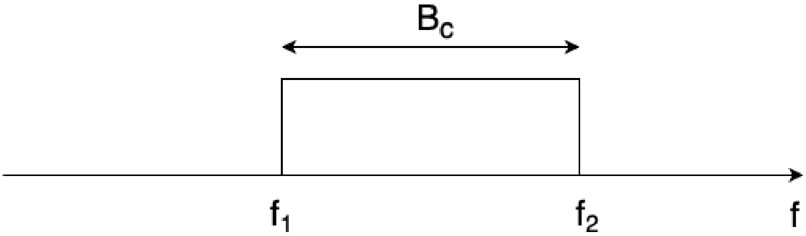


Figure 4.11: Bandwidth.

bandwidth can be calculated with the following formula:

$$B_c = (1+\alpha)\frac{R_b}{2\log_2 L} \tag{4.10}$$

The raised-cosine filter is a filter frequently used for pulse-shaping in digital modulation due to its ability to minimise intersymbol interference (ISI). The ISI is a form of distortion of a signal in which one symbol interferes with subsequent symbols. However, this is undesirable, as the previous symbols have a noise similar effect, thereby causing communication to become less reliable. The roll-off factor, α, is a measure of the excess bandwidth of the filter. The value of this factor goes from zero to one both values are included.

The fraction in 4.10 is composed of two elements. In the numerator we have the bit rate R, as the name suggests, representing the speed at which bits are transferred from one point to another. In other terms, it measures the quantity of data transmitted during a given period of time. The bit rate is most commonly measured in bits per second (bps). On the other hand we have L representing the modulation level.

Now back to the equation 4.9 where we have not yet explained one of the fundamental terms. Inside the logarithm we have the Signal to Noise Ratio (SNR) that compares the level of a desired signal to the level of background noise. SNR is defined as the ratio of signal power to the noise power, often expressed in decibels. SNR is calculated with the following formula:

$$SNR = \frac{P_{rec}(d)}{N_0 R_b} \tag{4.11}$$

The power received P_{rec} is the difference between the transmitted power and the path loss. This equation takes on meaning if the values are in decibels. It is possible to write it in a linear way:

$$P_{rec}(d) = \frac{P_t}{PL(d)} \tag{4.12}$$

In this case we will have the result in Watt. The bit rate can be achieved from the equation 4.10. The last term to be described in the equalization 4.11 is the thermal noise.

Thermal noise

Thermal noise is the electronic noise produced by the thermal shaking of the charge carrying elements (usually electrons) inside an electrical conductor at equilibrium, which occurs independently of any applied voltage [DR$^+$58]. Thermal noise is present in every circuit, and in sensitive electronic equipment such as radio boxes can submerge weak signals, and can be the limiting factor in the sensitivity of an electrical measuring device. Thermal noise grows with increasing temperature. Certain sensitive electronic equipment such as radio receivers are cooled to cryogenic temperatures to reduce thermal noise in their circuits. Thermal noise in an ideal resistor is roughly white, which means that the spectral power density is almost constant over the entire frequency spectrum. When limited to a finite bandwidth, thermal noise has an amplitude distribution that is almost Gaussian. The spectral density of a thermal noise is given by:

$$N_0 = kT \tag{4.13}$$

Where k is Boltzmann's constant in joules per kelvin, and T is the receiver system noise temperature in kelvins.

4.4.1 Network Capacity

Interference happens when there is at least one data blocks transmitting partially or simultaneously and it is partially or totally overlapping in the frequency domain. As represented in figure 4.12 we have a interference situation. One important thing to note is that both conditions must be fulfilled. If the two signals overlap only on the time axis we do not have interference. This also holds true in case they overlap in frequency, but not in time. There are several technologies that use this type of process to eliminate interference in an efficient way.

Figure 4.12: Interference situation between two data blocks.

All radio transmitters in a wireless network have a range of coverage. Other than this, the signal level drops to a limit under which it is not usable, and where it will not cause meaningful interference to mobile devices that are connected to a separate radio transmitter. This means that a channel can be recycled once out of range of the radio transmitter. Likewise, the reverse direction of the receiver, where it will only be capable of receiving signals with a certain power.

Typically, operators assign different frequency bands or channels to adjacent cells to reduce interference even when coverage areas overlap lightly when a cellular network is planned. Therefore, cells can be arranged in a so-called cluster. Clusters typically include seven cells, but further setups are possible. Different conflicting requirements must be reconciled when selecting the cell number in a cluster. This includes limiting interference levels and the number of frequency bands or channels that can be assigned to every cell site.

Interference between cells using the same frequency bands must be limited. The configuration topology of the cells has a big impact on this. The higher the number of cells in the cluster, the greater the necessary distance between cells using the same frequencies.

In our case we are assuming the case of no interference. This is because we will use a cluster with 7 cells using different frequencies bandwidth. In figure 4.13 it is shown three clusters and in each one as 7 different frequency bandwidth. The numbers represent different frequency bandwidth.

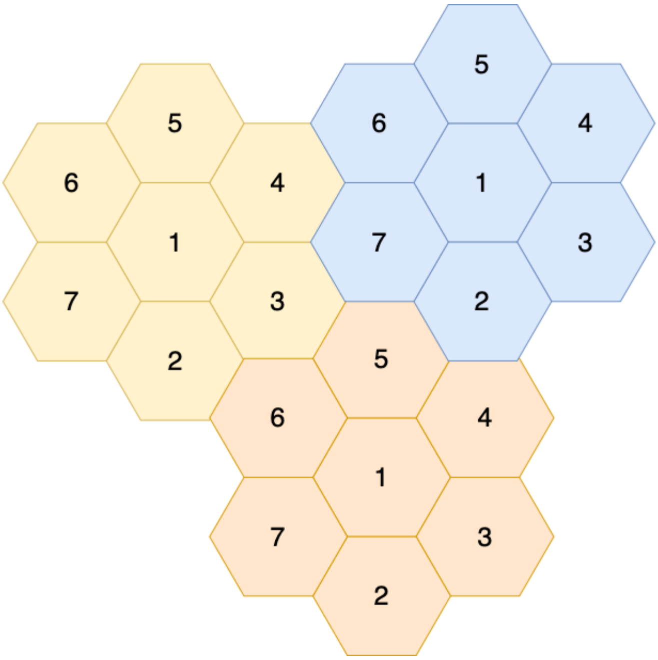


Figure 4.13: Three clusters.

This is a valid assumption, because we are using mmWave that do not travel for long distances. The second reason is the fact that we have no frequency bandwidth problems. The portion where the mmWave operate there are a lot of unoccupied frequency bandwidth.

Therefore, the estimation of the capacity of a single Fog node can be calculated as the sum of several contributions. In the figure we have a Fog node every three cells figure 4.14. The connection between the small cells and the Fog node is made with optical fiber. The following formula can be used for safety estimation:

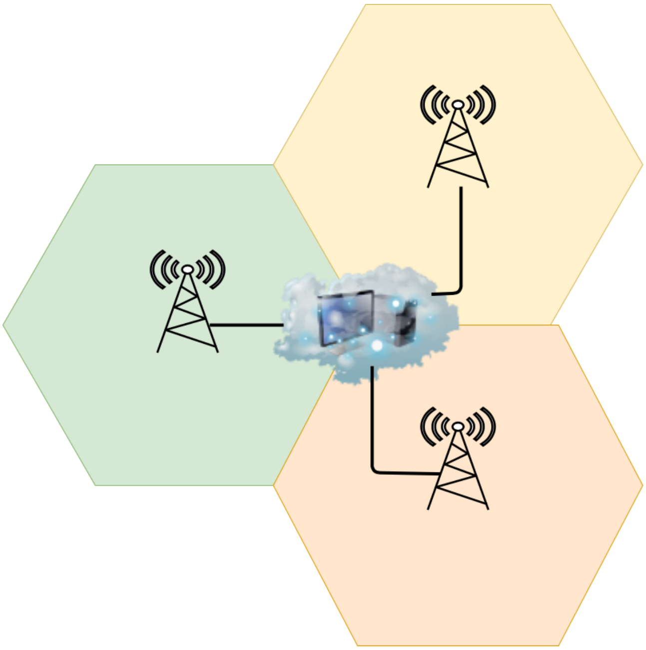


Figure 4.14: Three remote radio heads with a Fog node in the center.

$$C_{tot} = \sum_{i=1}^{N} C_i \tag{4.14}$$

We have that now the capacity has the subscript, for the sample reason that we have a dependence on distance. It is the distance between a small cell and a vehicle.

4.4.2 Multiple Access Techniques

Mobile standards use a diversity of multiple access (MA) techniques, which we point out in Table 4.1. Such techniques include frequency division multiple access (FDMA), time division multiple access (TDMA), code division multiple access (CDMA), and orthogonal frequency division multiple access (OFDMA). The duplex method used for two-way communication and the effective available physical resources to be allocated to each user are also described. Duplex methods are time division duplex (TDD) and frequency division duplex (FDD).

These multiple access techniques can be viewed as a kind of orthogonal multiple access, where users' access ideally does not interfere with each other because they

Table 4.1: Multiple access in different generations of cellular networks [VDP19].

Cellular generation	MA technique	Duplex method	Physical resources	Notable examples
1G	FDMA	FDD	Frequency	AMPS, NMT
2G	TDMA	FDD	Time slots	GSM, IS-54
3G	CDMA	FDD/TDD	Time slots/PN Codes	WCDMA
4G	OFDMA	FDD/TDD	Time/Frequency	LTE, LTE-A
5G	OFDMA	FDD/TDD	Time/Frequency	5G-NR

share the wireless medium. However, they are restricted to the number of available resources that make them orthogonal to each other. The exception is CDMA, in which transmission from the wireless device to the base station is inherently non-orthogonal [VDP19].

The frequency in the FDMA is split into channels that can be used by different users. In TDMA, time is divided into time slots to enable users to access the cellular system. In CDMA, users are separated by PN codes and broadcast on the entire frequency channel, all at the same time. In OFDMA, users are allocated to different frequency channels in various time slots. OFDMA is still used for the new generation 5G, where sub-carrier spacing and time slot duration are flexible and scalable to support highly variable needs and use cases. NOMA is also planned to be used in 5G.

Non-Orthogonal Multiple Access

NOMA [SKB$^+$13] is one of the most promising radio access techniques in next-generation wireless communications. NOMA offers a number of desirable potential advantages over OFDMA, which is the technique currently used, such as increased spectrum efficiency, reduced latency with high reliability and massive connectivity. NOMA's basic idea is to serve several users using the identical resource in terms of time, frequency and space.

NOMA available techniques can be broadly divided into two main categories, power domain NOMA and code domain NOMA. Code Domain NOMA can be additionally classified into several multiple access techniques that are based on low-density diffusion and sparse code multiple access. Other closely related multiple access schemes in this context are grid multiple access, multi-user shared access and pattern division multiple access.

New studies prove that NOMA has the capacity to be applied in various 5G communication scenarios. Moreover, there is some existing proof of performance enhancement when NOMA is integrated with various effective wireless communication techniques, such as cooperative communications, MIMO, beamforming, space-time encoding, network encoding, full-duplex, etc. [ZYD+17].

Since the principle of NOMA allows multiple users to be superimposed on the same resource, this leads to interference for such systems. Consequently, existing resource management and interference mitigation techniques, especially for ultra-dense networks, need to be revisited due to the incorporation of additional interference this new technology brings. For the similar reason, beamforming and the resultant other problems (e.g., precoding) in massive-MIMO systems introduce additional challenges and need to be solved in order to achieve full utilization of these technologies.

4.4.3 Example

In this section we have a scenario dimensioning example. The idea that this example can be extended to other scenarios. We have a roundabout with several roads leading to it. In the Figure4.15 we can see the scenario.

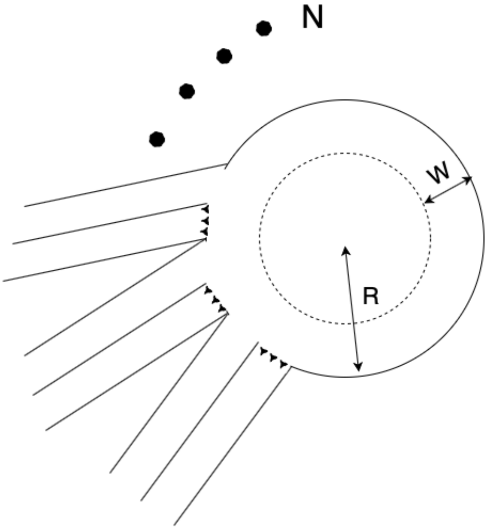


Figure 4.15: Example scenario.

We have the following parameters:

- N = Number of rows.

- R = Roundabout radius.
- W = Lane width.
- D = Circumference.

The circumference can be calculated:

$$D = 2\pi R \tag{4.15}$$

From this formula we can found the maximum number of rows:

$$N = \frac{D}{2\,W} \tag{4.16}$$

Now that we have defined the parameters it is possible to start to dimension the system. Let's give a numerical example. First, we need to do some assumptions:

- W = Lane width 2.50 m.
- A = Average length of the vehicles 4 m.
- R = Two lanes so it's 5 m.

the first thing to calculate is the maximum number of roads that can be had in this scenario. This is obtained by calculating the value of the circumference 4.15 and then dividing it by twice the width of the lane. This is because we consider the roads to be two-way. substituting number values:

$$N = \frac{2\,\pi\,2\,W}{2\,W} = \; 2\pi \approx 6 \tag{4.17}$$

now we know we can have a maximum of six roads with two lanes. Now let's calculate how many vehicles physically can be on the six roads. First, we calculate for one road, then it is easier to find for the remaining roads. Since we have what a small cell covers a 100 m radius. The figure 4.16 shows that we have 95 m left because 5m are occupied by the roundabout. So, a lane can have a maximum number of vehicles of 23. This value is obtained with the following formula:

$$N_v = \frac{95}{4} \approx 23 \tag{4.18}$$

the maximum number of vehicles on all the roads is now easy to calculate. Since we have 12 lanes, we can make the number of cars in a lane by the total number of lanes. This brings us to the maximum number of vehicles that can be in this scenario. The result is 276 vehicles. This is not the total number of vehicles because we did not consider the cars in the roundabout. let's calculate it with the following formulas:

$$N_{v1} = \frac{2\,\pi\,1.25}{A} \approx 2 \tag{4.19}$$

$$N_{v2} = \frac{2\,\pi\,3.75}{A} \approx 6 \tag{4.20}$$

since the cars are in the middle of the lane, we considered the radius from the middle of the roundabout to the middle of the lane. The total now stands at 284 vehicles.

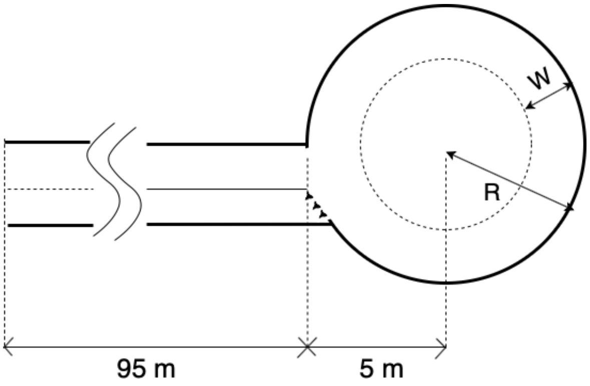


Figure 4.16: Numerical example.

Chapter 5

Localization and Positioning

This chapter describes the main positioning approaches, that are utilized to implement various localization systems, i.e., cellular, satellite. The reasonable classification can be made as suggested in [WAR06] by dividing the positioning techniques according to two groups:

- Mobile-based, whereby the location calculation responsibility lies on the user device side, through the use of its measurements and network data support.
- Network-based, which in reverse of the mobile phone approaches, a user device transmits its measurements to the network location server that calculates the final location.

As regards the particular techniques applied for localization, there are four main techniques. However, they are described below as a concept. Their implementations and accuracy in current systems are covered in the final parts.

5.1 Trilateration

The position of the user can be determined based on the reference stations of known locations [ME06]. Given that the radio wave has the well-known propagation speed value, the easiest idea is to quantify the time the radio wave travels from the reference station to the user's receiver and to compute the distance (this concept is know as the Time of Arrival (TOA) method [ME06]). The position of the user corresponds to the intersection of the different circles, which are formed by the different references base stations (see figure 5.1). At least three distances from the reference station are required to obtain the 2D position. In the numerical form, the trilateration can be defined as solving the whole set of the equation 5.1 for k = 1,-,k (where k is the number of reference stations):

$$r_k = \sqrt{(x - x_k)^2 + (y - y_k)^2} \tag{5.1}$$

where r_k correspond to the distance from the user to the k-th reference station, x, y are the 2-D user position coordinates and x_k, y_k are the k-th reference station coordinates. In [WAR06], the authors also cite the calculation of the position according to the difference in TOA.

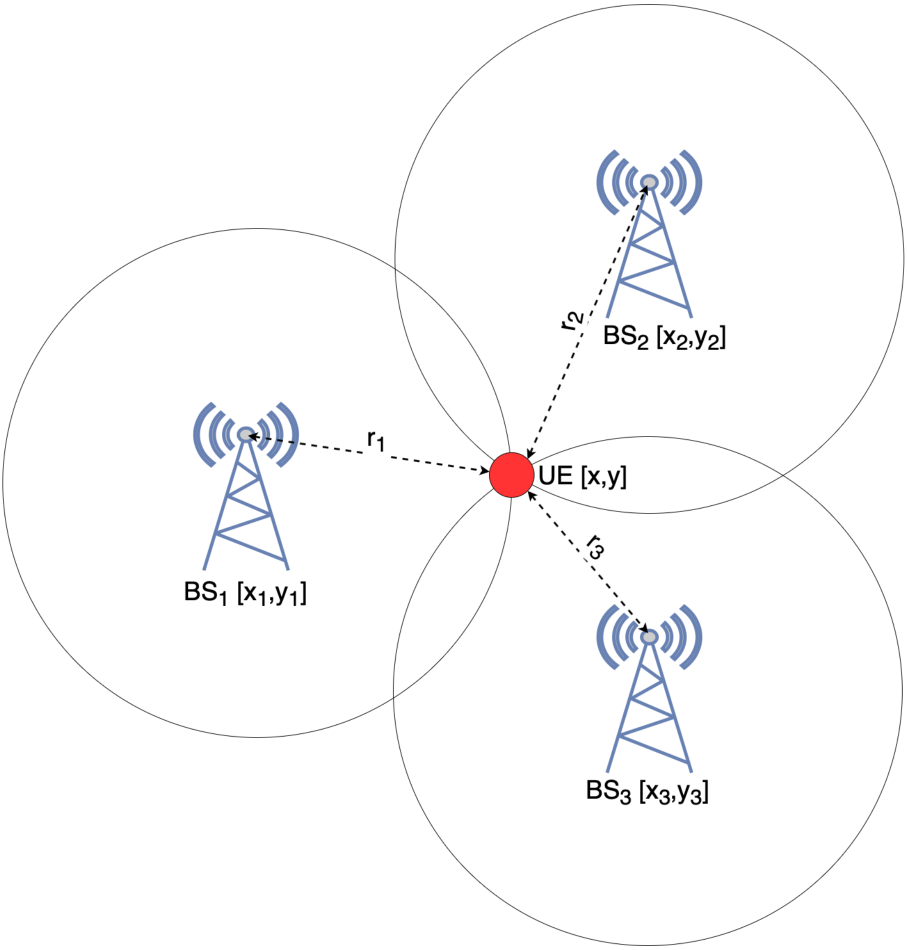


Figure 5.1: The concept of the trilateration positioning technique.

Positioning based on time requires precise measurements of the time interval, i.e., to reach the level of centimetre accuracy, time errors must be of the order of magnitude of 1 ns [PF18]. The synchronization of the receiver clocks and reference stations is essential to achieve a good accuracy. Hence, In [PF18], it can be observed that a user receiver has an inner clock with an accuracy much worse than that of a reference station. An extra measurement is needed to estimate the offset of the receiver's clock, but the reference station's transmitters still need to be synchronized [ME06].

5.2 Triangulation

The second localization method (triangulation), is base on the angles measured. To obtain the position of a user, we need at least two reference stations with their already known position. These requirements are to obtain 2-D positioning. However, increasing the number of reference station increases the accuracy of the position. This sort of measurement is named the Angle of Arrival (AoA). The position of the user is determined where the directions of the signals received intersect, as it is shown in figure 5.2.

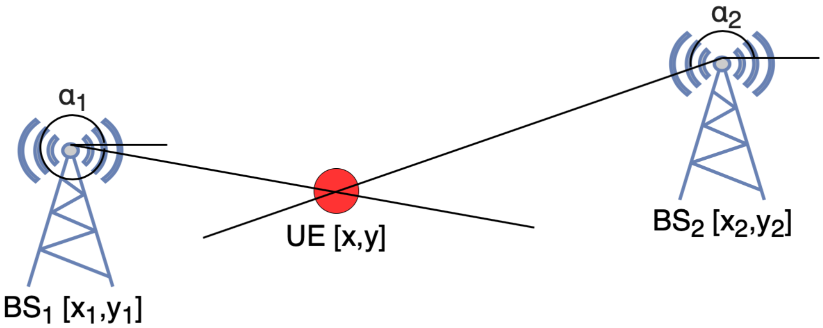


Figure 5.2: The concept of triangulation positioning technique.

The user's position can be calculated with a mathematical formula. This formula is valid when considering 2-D positioning. To obtain the position, this equation has to be solved:

$$y - y_k = \tan \alpha_k (x - x_k), \quad for \quad k = 1, -, k \tag{5.2}$$

we have that x and y references to the user coordinates. In the other hand we have x_k and y_k which are the k-th reference station coordinates, and αk is AoA measured from the user's position to the k-th reference station. Unlike the previous one, this localization technique does not need time synchronization, but it relies on the measurement quality of the AoA. The authors of [WSGD$^+$17] claim that the AoA measurement error is proportional to the inverse number of the antenna array elements. However, this implies that there is a relationship between measurement accuracy and reference station number.

5.3 Proximity

Proximity method is probably one of the easiest methods, as the name indicates it's based on the proximity to a reference station (see figure 5.3). The position is calculated by the closest reference state to the user equipment (UE) [PF18].

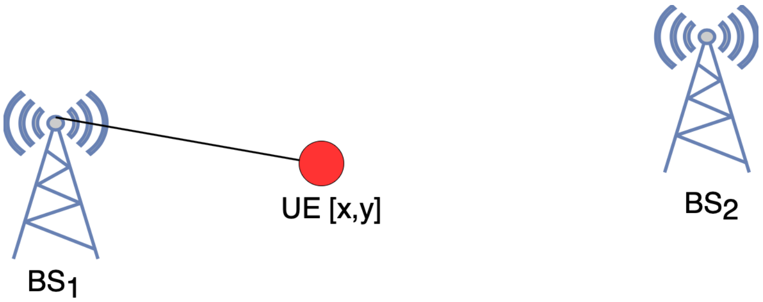


Figure 5.3: The concept of the proximity positioning technique.

5.4 Radio Frequency Pattern Matching

The final localization technique is Radio Frequency Pattern Matching (RFPM), is based on the comparison of the measured value and a saved value in a database. This value can be obtained with the propagation time delay or received signal strength (RSS) [WAR06]. This technique is also called fingerprint. The fingerprint is generally a geographical location related to the signal measurement. The user location is achieved as the best matching fingerprint coordinates in the database [WAR06]. An example is represented in figure 5.4.

In RFPM technique, the estimation of the positioning is performed following three stages [PDDCDS17]:

- Database Building Phase; In this step, the database of fingerprints is created based on performed measurements in the reference points of known locations.
- Acquisition Phase; In this phase, the user device is performing the same measurements as in the previous step, but in an unknown location.
- Matching Phase; In this final step, the values measured by a user are compared (with the use of predefined matching rule) with the ones stored in the database. The user position is claimed to be the position of the best matching fingerprint from the database.

Accuracy of position in fingerprinting is strongly correlated to the quality of the measurements as a function of position, database calibration and grid density [WAR06].

5.5 Cellular Positioning

Position estimation with cellular technology is nothing new. However, this technology has been used where accuracy was not a crucial factor. For instance, if we use 2G, we

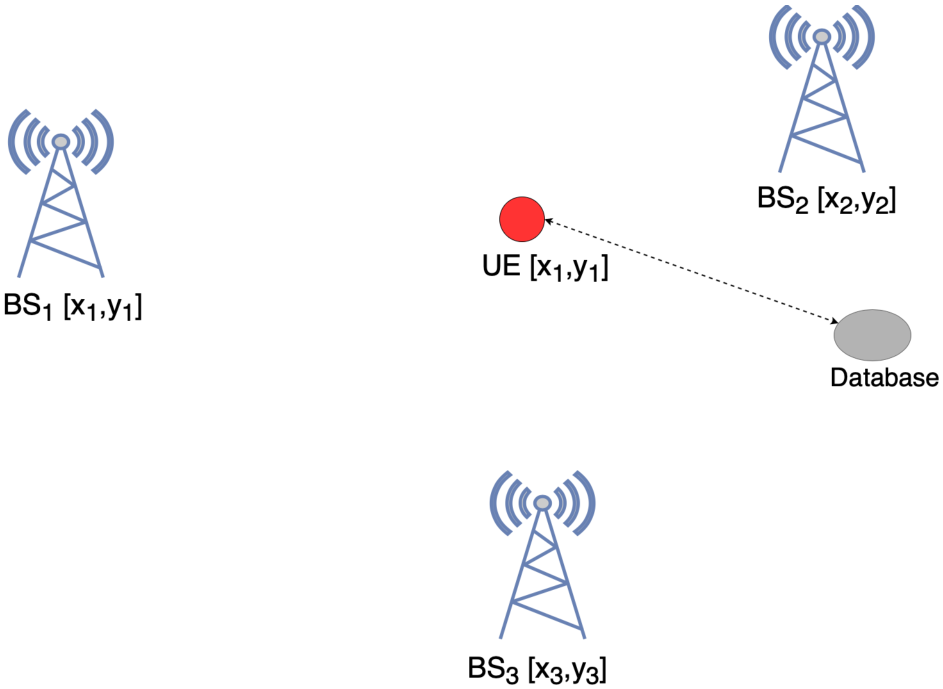


Figure 5.4: The concept of RFPM positioning technique.

have an accuracy in the order of hundreds of meters [dPRRLSSG17]. A new method for measuring distance has been introduced time-difference of arrival (TDOA). This new method improved the performance. We have that in the 3G we have an accuracy of the order of about ten meters [dPRRLSSG17]. In the LTE the accuracy increase even more. Yet none of these cellular generations can presently satisfy the positioning criteria of upcoming vehicular networks. Otherwise, 5G that is a new wireless communications technology will allow meeting these requirements. This implies that 5G can offer positioning services with accuracy beyond Global Navigation Satellite System (GNSS) with low additional costs, using the available infrastructure, and with negligible overheads for communication in terms of time-to-frequency resources. In the following subsections, we present the key elements that enable 5G to achieve high accuracy measurement.

5.5.1 Large number of antennas

When there are a large number of antennas in use, near-pencil signal beams can be used on both transmitter and receiver. Beamforming can enhance the communication quality by enabling higher antenna gain for an increased link budget and by reducing interference with the directive's communications [KA$^+$19]. It also influences the estimated delay accuracy, which depends not only on the signal bandwidth, but also on the SNR, antenna number and amount of multipath interference. Sufficient signal-to-interference-plus-noise ratio (SINR) gain will be achieved through the use of directional or beamforming antennas with relatively large arrays, increasing the impact of bandwidth increase and, ultimately, leading to orders of magnitude improvement in delay estimation accuracy compared to conventional cellular communications [WML$^+$16]. Nevertheless, particularly at high carrier frequencies where the signal is received from fewer directions, a great challenge is the problem of beamforming alignment, without which all the advantages of beamforming will vanish due to the lack of sufficient SNR to establish a link. Therefore, a great deal of research in 5G is dedicated to the design of time-efficient beamforming solutions [HGPR$^+$16]. On the other hand, it is well established that the use of large antenna arrays can significantly enhance the accuracy of the bearing/AoA estimation. The bearing estimation uncertainty is inversely proportional to SNR, antenna number and sample number and thus improve the accuracy of positioning. Initial studies at standard cellular frequencies show that it is possible to reach a positioning accuracy at centimeter level with a bandwidth of 40 MHz when you have several antennas at the base station [ZVK$^+$15].

5.5.2 High Carrier Frequency

As mentioned in the previous chapters, 5G will use mmWave. Therefore, in this technology path loss becomes an important factor. So, we need to use the compensations technique to manatee the quality of the signal. As an example, directive antennas and beamforming are two compensation technique used in the transmitting or/and receiving part. At mmWave, we have that the wavelength is in the order of millimetres. Having a large number of antennas in a small air allows increasing the directional antennas and beamforming. The utilization of mmWave allows having few reflected path. This characteristic of the mmWave permits to have the so-called LOS path. Consequently, the channel becomes:

- Poor in the sense of having only few dominant multipath components.
- Highly dependent on the positions and orientations of transmitter and receiver, as well as the environment [HGPR$^+$16].

Thanks to these channel characteristics it is, therefore, easier to identify and track single specular multipath components that can be used for high accuracy positioning. The scarcity directly translates into an increase in the SINR of the individual components as the disorder of the diffuse part of the channel's impulse response acts as interference. This tight connection between the radio channel and the propagation environment in 5G communication can, therefore, be used for positioning purposes.

This is in sharp contrast to traditional communication below 3 GHz, in which the signal is richer in significant multipath components (concerning interference due to background diffusion noise in the impulse response) and the signals tend to arrive (and start from) several directions, with less dependence on specific elements of the propagation environment. Also, at lower frequencies, it is possible to pack fewer antennas in a given area, limiting the angular resolution. However, signals below 3 GHz are part of 5G and can be a fallback positioning solution with possibly degraded performance. This alternative will be used in the case when mmWave signals are unavailable.

5.5.3 Large Bandwidths

Larger bandwidths can be achieved through the use of higher carrier frequencies in 5G signals. It is expected that 5G will use frequency channels with bandwidths around hundreds of megahertz, far surpassing the 20 MHz channels in LTE and the 100 MHz blocks available in LTE-Advanced (LTE-A) using carrier aggregation [RTG10]. Large bandwidth brings with it two benefits in localization. The first one is the low latency due to the presence of shorter symbol time. This characteristic increase location measurement accuracy due to the lower delay resolution. The capability of 5G to offer time-critical services, providing end-to-end latency below 1 ms, is partly due to the large bandwidths, which enable fast data transmission. In conjunction with advances through the protocol stack and processing near base stations instead of in the cloud, ultra-fast communication and positioning create a mutual synergy. On the other hand, an increase in bandwidth leads also to a corresponding enhancement in the delay estimation, which depends on the so-called effective bandwidth. Time delay estimation is directly translated into distance estimation through measurements of time-off-light (TOF) and TOA, assuming that some form of synchronization between devices is achieved, either through TDOA measurements or through bidirectional packet transmissions. Early research has shown that it is possible to obtain errors in the order of a few centimetres by using a bandwidth of 100 MHz [dPRLSKSG16]. Nevertheless, other technological factors, such as the imperfection of the internal clock-oscillator [dPRLSKSG16], must be taken into account to achieve extreme range accuracy. The accuracy of the delay estimation is not only improved with the bandwidth but also the resolution is improved. This refers to the ability to solve close-range replicas of the signal produced by reflective objects or vehicles in the proximity. However, interference between two replicas does not affect the range estimation when their relative delay is greater than the bandwidth inverse, which, for instance, in the case of 300 MHz bandwidth corresponds to 1 m. As the distance between vehicles and other road elements is normally greater than this value, the delay of the LOS component, as well as the reflectors, can generally be estimated straightforwardly [WML+16].

5.5.4 D2D Communication

IEEE 802.11p supports Device-to-Device (D2D) communication for vehicular communication nationally. With release 14 of LTE, Vehicle-to-Everything (V2X) communication will be backed in two additional transmission modes: conventional network-based communication for cloud interaction and direct D2D communication for low-latency, high-speed and high-density communication. In a like manner, 5G D2D will provide ultra-fast, direct, high-speed communication links between vehicles. This brings better coverage, better spectral reuse and high speed, low power connections. As with communication, it is also useful for the dissemination and calculation of location information, using the cooperative positioning paradigm. In collaborative positioning, measurements (relative speeds, distances and angles) are collected by the devices not only concerning reference stations (permanent access points), as well as about other mobile devices. Such measurements can be utilized in cooperative algorithms to increase both positioning coverage (part of devices able to locate themselves) and accuracy. Furthermore, cooperative positioning enables relative positioning, also in the lack of reference stations. This is particularly beneficial for vehicle sensing and planning tasks in vehicles, in addition to sensors on board. Particularly in 5G, D2D communication has the advantage of extremely short latency, which enable tracking of fast-moving devices such as vehicles, thus enhancing the recognition and prediction of dangerous situations.

5.5.5 Network Densification

A definite characteristic of 5G networks is network densification, featuring a hierarchy of base stations combined with cells of different sizes and linked with high-speed backhaul links. Devices in dense networks can connect to multiple access nodes, which provide higher data rates with lower power consumption, as long as interference and mobility problems can be solved. While these challenges can be met and access points can somehow announce their coordinates, ultra-dense networks can provide ultra-accurate positioning [KCW$^+$17]. It is because the accuracy of the positioning is not only dependent on the individual measurement quality (for which high bandwidths and several antennas are advantageous), however, it also depends on the number and diversity of reference stations. Furthermore, in very dense networks, LOS communication is highly probable, and hence a strong signal directly related to the geometry between sender and receiver. Succinctly, the densification of the 5G network can help to support the positioning of vehicles at the accuracy levels [dPRSGKLS19].

5.6 Global Navigation Satellite System

GNSS enables people to estimate their position on the basis of received signals from satellites in Earth orbit. The calculations of the position are derived by the TOA trilateration method [HB19]. Every satellite transmits ephemeris information (the orbital detailed parameters needed to compute its position), almanac information (basic ephemeris for the entire constellation) and clock corrections [HB19]. The signal broadcast by the satellites is modulated in code and can be expressed as [HB19]:

$$s(t) = \sqrt{2P}D(t)c(t)cos(2\pi ft + \phi) \tag{5.3}$$

$\sqrt{2P}$ is the signal amplitude, D(t) is the signal of the navigation data, c(t) is the spread spectrum code, f is the carrier frequency, and ϕ is the phase [HB19]. The internal clocks of the satellites are fully synchronized through the atomic clocks. Based on the ephemeris received and TOA measurements, the receiver evaluates its position. Because of the receiver's clock uncertainty, further measurements are needed to assess the offset of the clock. In order to get the 3D position, a minimum of four satellites must be available, several satellites may improve the accuracy [HB19]. The GNSS currently available are, for example, GLONASS (Russia), BeiDou (China), GPS (USA), Galileo (EU). Nevertheless, the most popular is GPS.

5.6.1 Measurement of Distance

Two types of measurements are used to estimate the distance between receiver and satellite. First approach consists in tracking the pseudo-random code transmitted by the required satellite. Second approach to achieve the distance on the basis of the measurements of the carrier phase.

Code Phase Measurements

The distance between the user and the k-th satellite in this approach is calculated based on the alignment of the replication of the broadcast code with that received by the satellite. The estimation of the distance suffers from some errors, which is why it is called pseudorange. However, the sources of error with the greatest impact on the measurements are: both satellite and receiver clock bias, propagation delays in the ionosphere and troposphere and fading of the multipath [HB19]. Measurements in code phase are common in market product receivers, for examples smart-phones. Accuracy obtained with this method in stand-alone GPS receivers is about 10 m in 95% of cases [HB19].

Carrier Phase Measurements

The pseudorange in this approach is the number of carrier signal wavelengths between the satellite and the receiver [HB19]. The value measured is the phase offset between the carrier replication and that received by the satellite. In the hypothesis of ideally synchronized clocks without other errors, this phase offset is given (in cycles) as [HB19]:

$$\phi(t) = \phi_u(t) - \phi^s(t - \tau) + N \tag{5.4}$$

in which $\phi_u(t)$ is the replica of the carrier phase, $\phi^s(t-\tau)$ is the phase of the signal from the satellite received at time t, τ is the propagation time, and N is called the integer ambiguity (estimation of the resolution of the whole ambiguity, is the most complex part of the measurements of the carrier phase, but when it is done it allows to obtain a precision positioning at cm level [HB19]). Except for the vector of the integer uncertainty, the phase measurements are influenced with the same errors as the phase measurements of the code (multipath fading, clock distortions and propagation delays), but can be eliminated with the double difference measures that are discussed later. Measurement of the carrier phase is much more accurate, but also involves more complex calculations than the code phase estimation method. GNSS positioning accurate to the centimetre with carrier phase measurements is common in surveying, agriculture and geodesy. In the mass market, the major receiver problem is in low quality antennas that cause a suppression of the multipath fading effect not high enough and long time correlations in phase errors, resulting in long initialisation time (tens of seconds) [PJ$^+$15].

5.6.2 Evolution

Since GNSS basic positioning technique for stand-alone receivers is using code phase measurements, we discuss how to improve GNSS accuracy and how to benefit from carrier phase measurements with the Real Time Kinematics (RTK) approach.

The Dual-Frequency GNSS

It is like the upgrade to a stand-alone GPS receiver. The L1 civil signal is transmitted by all GPS satellites at a frequency of 1575.42 MHz with a chip rate of 1.023 Mcps. The latest ones supply the L5 signal [HB19] transmitted at a frequency of 1176.45 MHz and a faster chip rate of 10.23 Mcps [Moo17]. Due to this L5 is less affected by multipath reflections than L1. By putting these two signals together in the receiver, it is possible to remove the ionospheric error [HB19]. An accuracy of about 30 cm [Moo17] is predicted with the dual frequency receiver. In addition, the Broadcom company has launched the BCM47755 chip as the first mass chip using both L1 and L5 signals. Currently, the only phone with a dual-frequency GPS receiver that we have identified is Xiaomi Mi 8, due to the absence of the duty cycle, it provided carrier phase measurements used for relative static single frequency positioning with a twice the distance root mean square (2DRMS) achieved of 1.02 m and 1.95 m in low and high multipath sites, respectively [RBP19]. We can see that the theoretical results don't match the practical ones.

Differential-GNSS

Differential-GNSS (DGNSS) measurements are carried out following the coding approach, however, the concept of error reduction is based on the fact that ionosphere, troposphere, ephemeris and satellite clock errors are correlated in time and space [10]. A reference receiver with known position is used which compares the position derived from satellite signal measurements with the real known position, then calculates pseudorange adjustments which are then sent to a user (called a rover) via the side link. This differential approach gives an accuracy up to 1m [HB19].

Real-Time Kinematics

RTK relies on relative positioning [HB19]. Like DGNSS, note coordinate reference station is required, but it sends (to a rover) its raw measurements rather than calculated corrections. Furthermore, the RTK also takes advantage of the measurements of the carrier phase rather than the coding phase. In order to eliminate clock bias and atmospheric delays, the position estimation (relative to that of the reference station) is carried out on the double difference basis. To achieve a double difference measurement, the carrier phase in the reference station and in the rover must be metered with the same pair of satellites. The measurement result of the rover is deducted from the measurement result of the reference station using the reference station using the same satellite signals in a single difference. This is also done with measurements based on another satellite signal. Lastly, single differences mentioned are subtracted, resulting in a double difference [HB19]. Because the clocks of the satellites and receivers are equal, they cancel each other out in the equation. The same is true for atmospheric delays, if the distance between the rover and the reference station is small enough (below 5 km [PJ+15]). Multipath fading is the main component of the remaining error. In the table 5.1 we summered what we discussed in this section.

Table 5.1: Measurements requirements for GNSS variants. [Adapted from [MMBS+19]]

	Code phase	Carrier phase	Reference station	Dual frequency chip
Stand alone single frequency	✗	✗	✗	✗
Stand alone dual frequency	✓	✗	✗	✓
DGNSS	✓	✗	✓	✗
RTK	✗	✓	✓	✗

5.7 Comparison

Accuracy of measurement is important. This accuracy may vary from application to application. In the vehicle environment it is even more important because an incorrect measurement could cause an accident or unintended braking. In our scenario, we can tolerate up to half a meter of error. This value is based on two factors:

- Driver Presence: The presence of the driver allows braking in case the measurement is too optimistic.
- Speed: Travelling at an average speed of 60 km/h is equivalent to 16 m/s. Keeping the distraction of safety, the half meter is not influential. This is because the human reaction time is one second. Since it will be done automatically in some cases, there is more braking distance, because the automatic rush is faster.

In this chapter we have practically described two different methods of localization. The satellite one, that we have seen that represents different versions that guarantee an accuracy in the order of the centimeter. On the other side we presented the new cellular technology that is 5G. On this last one there are still studies for example the standard 3GPP realise 16 that has yet to come out, but there are several papers that discuss this hot topic. Let's take this paper [CGLZ16] as an example, which uses six different mmWave waveforms were considered for timing estimation. In this paper two methods estimations are presented in calculating the position:

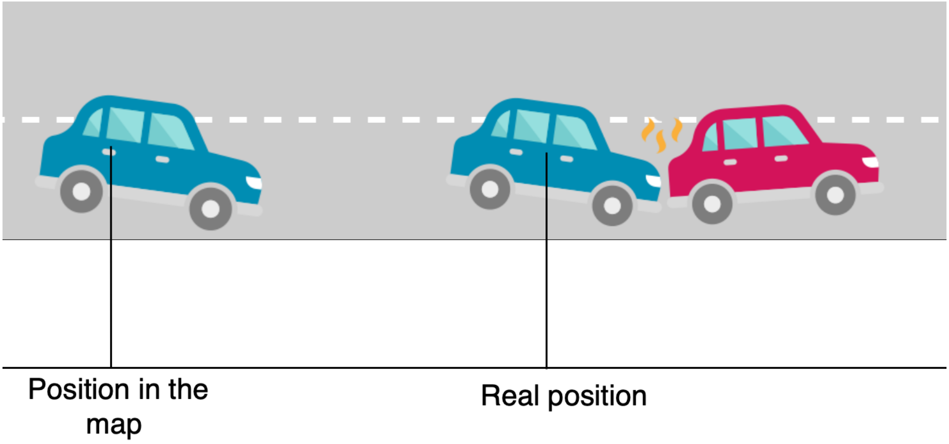


Figure 5.5: In case of erroneous position estimation.

- Correlation Receiver: The goal of range estimation is to obtain an unbiased estimate of the TOA from the received signal.

Table 5.2: Different technology with the correspondent accuracy.

Method	System	Horizontal accuracy	Reference
GNSS	GPS	10 m (95% cases)	[HB19]
DF-GNSS	GPS	$\sim$ 30 cm	[Moo17]
DGNSS	GPS	$\geq$ 1 m (95% cases)	[HB19]
RTK	GPS	$\sim$ 1 cm	[HB19]
TOA	5G	$<$ 10 cm	[WSGD$^+$17]

- Energy Detection: An energy detector consists of a square-law device, an integrator with integration period, and a decision mechanism. The TOA estimate is obtained by comparing the integrator output with a threshold.

In the Energy Detection we have two different kind of threshold. One in fix and the other is dynamic. The dynamic threshold has better accuracy. The correlation Receiver is batter than the Energy Detection, but it require more complexity and it needs the channel estimation.

Looking at the table 5.2 we can see that we have two technologies that guarantee our requirements. One of them is RTK, which has been in use for a long time. The second one is the 5G which has a high accuracy, but it has not yet been tested in real life. Based on this data we will use the RTK system for our system, but in the future it is possible to change the localization technologies since we will use 5G to communicate.

Chapter 6
Mobility

In this chapter, we discuss how the mobility of vehicles is handled. This is a very important issue especially in our case because vehicles are always in movement and also at high speed compared to the pedestrians one.

The chapter is constructed to explain the handover that is the process of migration. In addiction, we will discusses all the different types of handover. It is presented also a comparison between the different handover techniques.

In the ending part, we have a proposal for handover for our application. The choice to use one technology or the other is justified. Also, it is presented how it is possible to reduce the migration time thanks to the architecture of our application.

6.1 Handover

Handover is the process of changing serving base station. The mobile station has an active call of packet data transfer session (in which the radio resource units are assigned). Three factors determine the handover process [EVBH08]:

- Radio link quality: as long as we go further from our base station, the SNR goes down, the measurements reports will sense it.
- Network management: let's assume that in the cell there is a lot of traffic and in the other cells there is no one. To avoid an overloaded cell, it is a good idea to force the mobile stations to perform handover even if the radio link quality is good. This shows clearly that handover has to be performed at the controller side. It shows also the relevance of load control.
- Service options: an operator might deploy his networks with a common set of functionality but then they update the software but sometimes they do not update all the base stations in the same way. So, for a user in a certain cell,

if he requests to set up a new service which is not present in that cell, then a handover will be performed.

The most used are the radio link quality handover. In our case where we have that all small cells provide the same service and are sized to support a large number of users. Therefore, it makes sense to focus on the first point presented above.

Let's consider the situation of two cells and the user is moving over a straight line and the cells overlap each other and there is an area in which we can perform handover figure 6.1. A good question might be "Where is the ideal distance to perform handover?". There is an active session between the mobile station and the serving base station. In the third layer in the control plane, there is an algorithm that works for the decision.

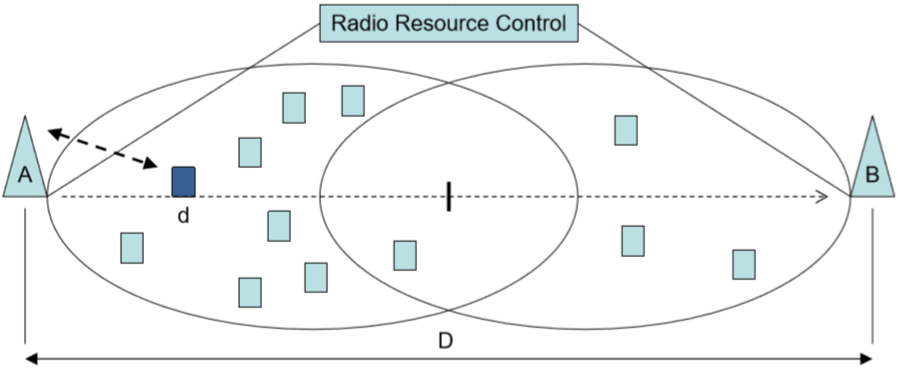


Figure 6.1: Handover scenario.

We have two different handovers the hard handover and the soft/softer handover. These two types will be better explained in the next part. In GSM we have the intra cell (in which we have the change the radio resource inside the same cell) and the inter-cell handover. Therefore we will consider only the inter-cell one. This is because we are more interested in the time needed to change the serving cell.

Handover is a process managed by the network, from the network architecture viewpoint there are three different kinds of handover [EVBH08]:

- Network control handover: at the network side without the assistance of the mobile station.
- Mobile assisted handover: the decision is taken at the network side but based on measurements reports submitted by the mobile station. They are received by the base station including their measurements for the uplink and then

this information is sent to the controller in case of 2G and 3G or Radio Resource Control (RRC) sublayer of layer 3 in 4G. Mobile assisted handover is performed by 2G and 3G. During an active session, the mobile station sends continuously reports of measurements and it is assistance for the base station for the handover.

- Mobile controlled handover. The decision is taken directly by the mobile node and not by the network.

Now let's describe the processes involved before the completion of the handover. The handover process takes place following 4 steps. The following four steps are necessary to make handover [EVBH08]:

- Measurements: since the measurements are made periodically so, the handover process takes place in an active session between the mobile station and a base station.
- Decision: this decision is done based on the measurement reported at the entities that managing handover (RRC sublayer of the control plane) and that decides the handover.
- The radio resource assignment.
- Execution phase: at this stage, we have the execution of the handover process. Meaning it is changing cell.

6.1.1 Hard Handover

Hard handover is a method applied with cellular networks that demand the user's connection to be fully disconnected from an already existing base station before being swapped to another base station figure 6.2. This is why such handovers are also referred to as break-before-make [FA16]. The hard handovers are designed to be instant to minimize disruption of the connection. A difficult handover is experienced by network engineers as something during the exchange of data. This process requires the least amount of processing by the network providing the service. Once the user is located between base stations, then it can switch to any of the base stations, so that the base stations will bounce the connection with the user back and forth. This is called ping pong effect [Tür07]. There are various techniques to reduce this effect. One of them is the introduction of hysteresis margin [Tür07].

6.1.2 Soft Handover

For soft handover, the mobile station is located in the overlapping cell coverage area of two sectors belonging to different base stations. Here the target connection is

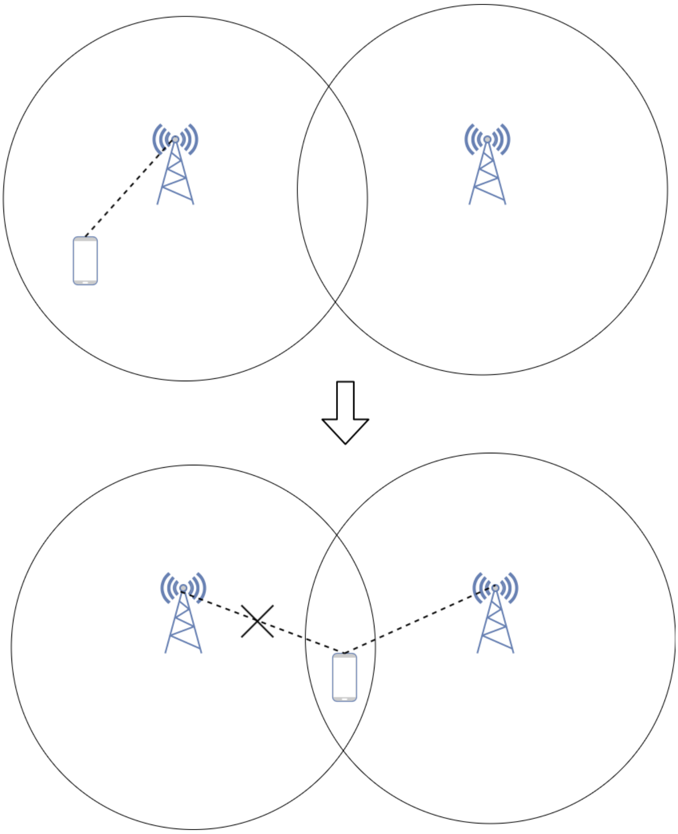


Figure 6.2: Hard handover.

made before the connection to the source is broken figure 6.3, so, this handover is named make-before-break [FA16]. The range, within which the two connections are used in parallel, can be short or long. Therefore, the soft handover is seen by network engineers rather as a state of the exchange data. Soft handover can involve the use of connections with multiple cells. They can be three, four or more cell connections can be held by a user at the same time. Once a user is in a soft handover state, the signal from the best of all channels used can be used for the user at a given time or all signals can be combined to provide a better signal copy. Handover is more advantageous, and when this combination is performed in both downlink and uplink the handover is called softer [Tür07]. A softer handover is achieved when the cells involved in the handover have a single cell site.

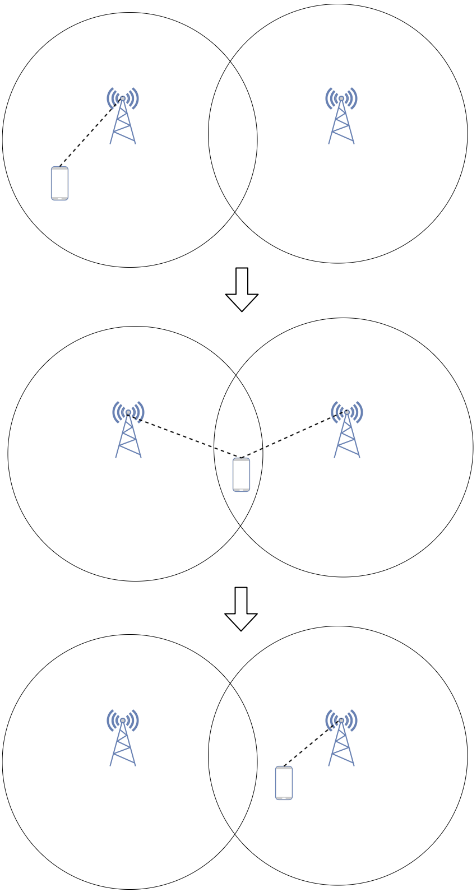


Figure 6.3: Soft handover.

6.1.3 Comparison

One advantage of hard handover is that a call uses only one channel at any given time [Tür07]. The hard handover event is very short and normally not perceived by the user. In old analogue systems, it could be audible as a very quick click or beep. With digital systems it is imperceptible. A further benefit of hard handover is that the hardware of the phone must not be able to receive two or more channels in parallel, which makes it cheaper and easier. A drawback is that if a handover fails the call can be temporarily interrupted or even terminated unexpectedly. Hard handover technologies normally have procedures that can restore the connection to the source cell if the connection to the target cell cannot be made. Nevertheless, re-establishing this connection may not happen at all times (in that case the call will be disconnected) and even when it does, the procedure may cause a momentary interruption of the call.

A benefit of soft handover is that the source cell connection is broken only when a reliable connection to the target cell has been made, and therefore the probability of the call ending unexpectedly due to a failed handover is reduced. However, a significantly higher advantage results from the simple fact that channels are kept in multiple cells at the same time and that the call may only terminate if all channels are disturbed or disappear at the same moment. Fading and interference in the various channels are not related and hence the chance of them happening at the same moment in all channels is quite low. Therefore the connection reliability is higher when the call is in a soft handover. Since in a cellular network most handovers occur in places with poor coverage, where calls would often become unreliable when their channel is interfered with or fading, soft handovers bring a major enhancement to call reliability in these places by making interference or fading in a single channel non-critical [FA16]. This benefit is achieved at the cost of more complex hardware at the user part, which must be able to process several channels in parallel. A further price to pay for soft handover hardware is the use of several channels on the network to support a unique call. However, this decreases the number of remaining free network channels and thus decreases the capacity of the network. By regulating soft handover duration and the size of the areas where they happen, network engineers can offset the benefit of extra call reliability with the price of reduced capacity.

As it is just been described, however, the two technologies have the upside and the downside. It is not that one is better than the other, but what makes one better than the other is what it will be used for. Some applications required less restriction in term of reliability than others application.

6.2 5G inter gNB handover

In this section is presented the handover between two different cells. The cells in 5G are called gNodeB (gNB). In LTE they were called eNodeB (eNB). The handover in the same cell is also presented, this is because we will use directional antennas to transmit. Therefore we will switch from one bream to another. This type of handover will not be treated, only the one between different gNBs will be treated. This is because we are interested in how we can reduce this migration time.

The figure 6.4 depicts 5G inter gNB handover type. Now we are going to present the procedure described by the standard [3GP19a]:

- The source gNB initiates handover and issues a Handover Request over the Xn interface.

- The target gNB performs admission control and provides the RRC configuration as part of the Handover Acknowledgement.
- The source gNB provides the RRC configuration to the UE in the Handover Command. The Handover Command message includes at least cell ID and all information required to access the target cell so that the UE can access the target cell without reading system information. For some cases, the information required for contention based and contention free random access can be included in the Handover Command message. The access information to the target cell may include beam specific information.
- The UE moves the RRC connection to the target gNB and replies the Handover Complete.

6.3 Propose

The forthcoming steps for 3GPP, releases 16 and 17, will bring further functionality to improve even further support for new use cases related to smart generation, connected vehicles, power distribution and more, as drones that are monitored by the network.

Such potentially crucial use cases require URLLC, implying high reliability and availability, as well as very low end-to-end latency within milliseconds. The 5G system was designed with this goal in mind and the continued evolution of 5G will continuously improve mobility performance. A key part of this is the decrease in downtime between cells in the 5G network.

We have seen how the handover process is carried out both for the previous technologies and for the current technologies. But our goal is to optimize this migration by proposing solutions that are also facilitated by our architecture.

starting to analyze which handover technique is suitable for our system, we have mainly two types of handover. In the previous sections we have shown the difference between the two types of handover. For our system we believe that it is suitable to use soft handover. This is for the following

- Reliability: since our application has to be feasible it is better not to interrupt the connection before creating a new connection.
- Packet loss: a continuous connection reduces the number of lost packets and therefore reduces latency.

One of the weaknesses of soft handover is that it involves more cells and therefore uses more resources at the same time. This makes it impossible for other users to

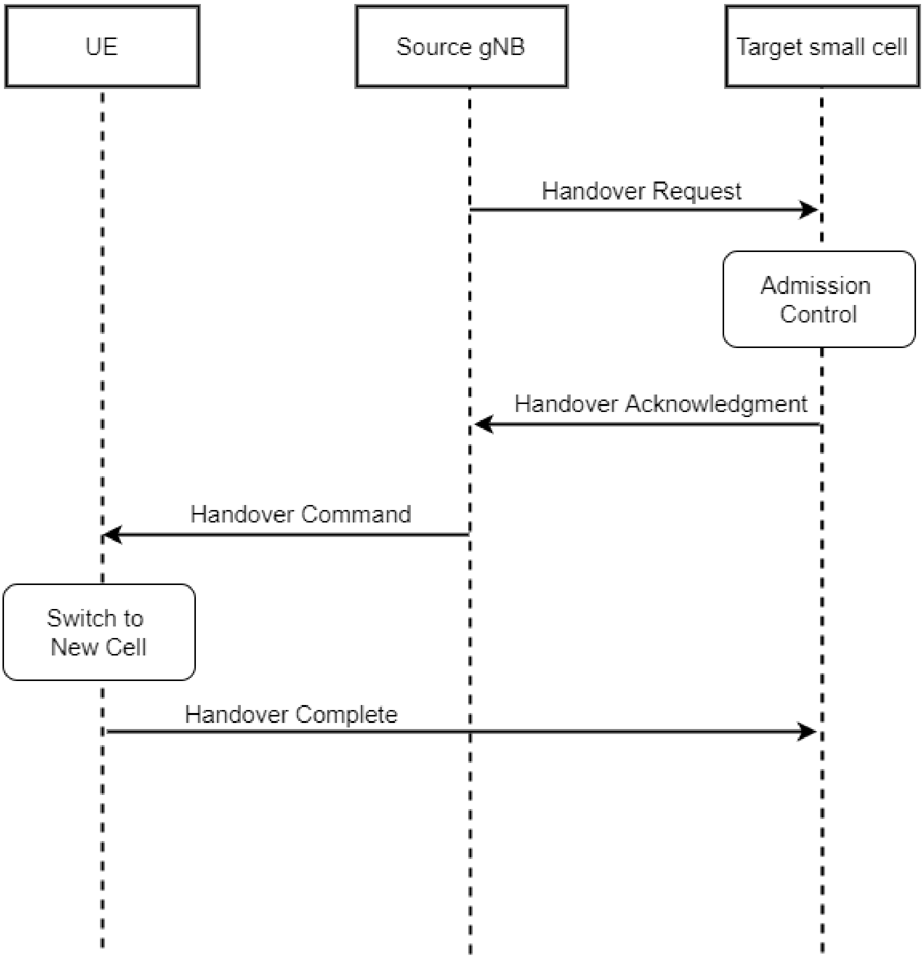


Figure 6.4: Inter gNB handover [3GP19a].

use that resource. But with the reduction of migration time, the time to use more than one resource will be short.

In our application it is expected that we know the destination of the vehicles. This makes that every Fog node knows which small cell the user has to connect to. This information will significantly reduce the handover time. This is because the two message exchanges (figure 6.4) can be done even before the user needs to change small cells. This makes the migration time start when the small cell asks the vehicle to change small cell.

So, we can conclude by saying that thanks to the links between the nodes and the information of the vehicles it is possible to reduce the migration time. This will certainly help to reduce disconnection problems and increase the performance of the application.

Chapter 7

Dependability and Performance

In this chapter we present two fundamental concepts that are dependability and performance. These are two important concepts for all types of systems. It is necessary to make sure that both are present in order to allow a good functioning of the system. This is not easy to achieve. In many cases, obtaining systems that are both reliable and with excellent performance requires high costs.

The chapter is divided in two sections where in each case it deals with one of the two topics. We start with performance and end with dependability, trying to touch all the factors involved.

7.1 Latency

A delay is synonymous with latency. In telecommunications, low latency is associated with a good QoE, whereas high latency is linked to poor QoE. Network latency can be measured by determining the outward and return time of a data packet to travel to a destination and back. This measure is also called round-trip time (RTT).

One of the main reasons for low latency is geography. Highly distributed Internet Protocol (IP) networks travel great distances, adding transmission time that can cause a delay in an application. In any situation where the latency between detection and response must be extremely low. For example, some actions in autonomous driving. So, it makes sense to put the processing computer as close as possible to the data source.

In figure 7.1 we have a representation of the latency that we want to estimate. As can be seen, the communication can start from the vehicle or the Fog node. But either way, it has to go through the small cell. This can be through the propagation medium that is air when starting from the vehicle. In the other case, it goes through the fiber optic. The latency is calculated as the time it takes for the signal to go back to where it started.

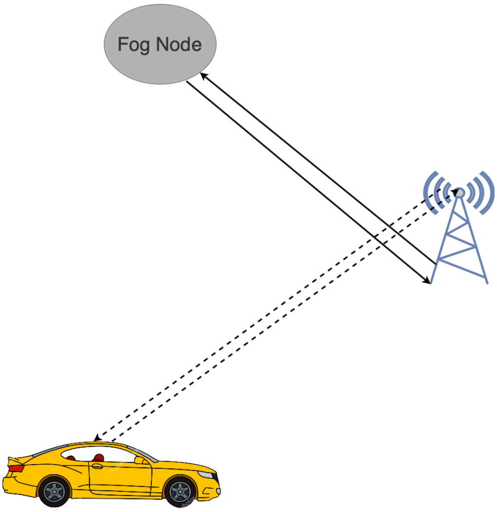


Figure 7.1: Route on which latency is calculated

7.1.1 Causes of Latency

Latency grows according to several facts, including the following:

- Transmission media: latency may be influenced by the type of media used to transmit data. For instance, packets travelling on air can be expected to have lower latency than packets travelling on a copper cable.
- Packet size: the size of a packet affects the time needed to transport it. A large packet will take longer to send round trip compared with a small one.
- Packet loss: latency can also be introduced by a high number of packages failing to reach their destination or by too much time variation of some packages to travel from one system to another.
- Signal strength: in case the signal is low and needs to be boosted by a repeater, this operation can introduce latency.
- Processing and storage delays: some of the packets requires information from memory or processing of the received data before re-transmitting. This operation increases latency.

Transmission Media

Based on theoretical calculations, Maxwell concluded that the propagation speed of electromagnetic waves in a vacuum is about 300,000 km/s, equal to the speed of light in vacuum. This result led him to think that light was a particular type of electromagnetic wave. Maxwell's intuition was exact regarding the electromagnetic waves, also light, include a wide range of types of waves that differ among them for the wavelength λ or, what is the same, for the frequency f, where λ and f are linked by the formula:

$$\lambda = \frac{v}{f} \tag{7.1}$$

where v is the speed of the wave. This can also be calculated using the following formula:

$$v = \frac{1}{\sqrt{\mu\epsilon}} = \frac{1}{\sqrt{\mu_r \mu_o \epsilon_r \epsilon_o}} \tag{7.2}$$

where we have that μ is the permeability that is equal to $\mu_r \times \mu_o$. Where μ_r is the relative magnetic permeability of the material and μ_o is the vacuum permeability. Also, we have ϵ that is permittivity that is equal to $\epsilon_r \times \epsilon_o$. Where ϵ_r is the relative permittivity of the material in a good conductor is one. We have that ϵ_o is commonly called the vacuum permittivity. However, we have that the speed of light is calculated in the following way:

$$c = \frac{1}{\sqrt{\mu_o \epsilon_o}} \tag{7.3}$$

because of this result we can see that in formula 7.2 we can break it in two terms as follows:

$$v = \frac{1}{\sqrt{\mu_r \epsilon_r}} \frac{1}{\sqrt{\mu_o \epsilon_o}} = \frac{c}{\sqrt{\mu_r \epsilon_r}} \tag{7.4}$$

the denominator is also common to refractive index n. In our case, we are interested in the refractive index of the air, which is equal to 1,00029. placing this value in the equation above, we obtain the propagation speed of the electromagnetic waves in the air. So we have that:

$$v = \frac{c}{n} = \frac{299,792,458\, m/s}{1.00029} = 2.997 \times 10^8\, m/s \tag{7.5}$$

note that we do not have a great variation compared to the speed of an electromagnetic wave in a vacuum. Now let's consider the worst-case scenario. The one that occurs when the vehicle is at the furthest distance from the small cell. So, when it is 100 meters away. Now, thanks to the value found above, it is possible to calculate how long it will take the signal to reach the small cell, by using the following formula:

$$t_t = \frac{v}{d} = \frac{2.997 \times 10^8\ m/s}{10^2\ m} = 2.997 \times 10^6\ s^{-1} \sim 3\ \mu s \tag{7.6}$$

we can see that the transmission time is very low, but for the sake of accuracy, we will take that into account.

In our architecture, we have two propagation media. One is air and the other is optical fiber. we know that the optic fibber has a refractive index more than air.

This obviously leads to delay.. Considering a distance of 100 meters from the small cell we can use the same formula 7.6 with an index equal to 1.46. We will obtain approximately around 50 microsecond one way.

Thanks to these two very important results for the latency calculation, we can estimate the maximum time taken by the signal to reach the fog node. Summing the two results, we obtain that it takes 53 microseconds. However, this time must be multiplied by two because the latency will lead to a round trip.

Packet size

As we have seen, the length of the packet affects the transmission delay. This is because the more data we have, the longer it will take to transport it. For example, if we consider shopping at the supermarket, the amount of things bought affects the time needed to load everything into the car. In fact, in the field of communication, we have a bit rate that represents the number of bits that can be transported in a second. The higher the bit rate the better is, because it will allow us to transport the same amount of information at the same time.

In our case, we have that all vehicles will transmit the same amount of information. This is because they all have to transmit the same information (speed, direction, position). This is an advantage because the data transmitted will be very small. This brings us to analyze some of the frames structured in 5G. Henceforth we will refer to the standard 3GPP release 15 [3GP19b].

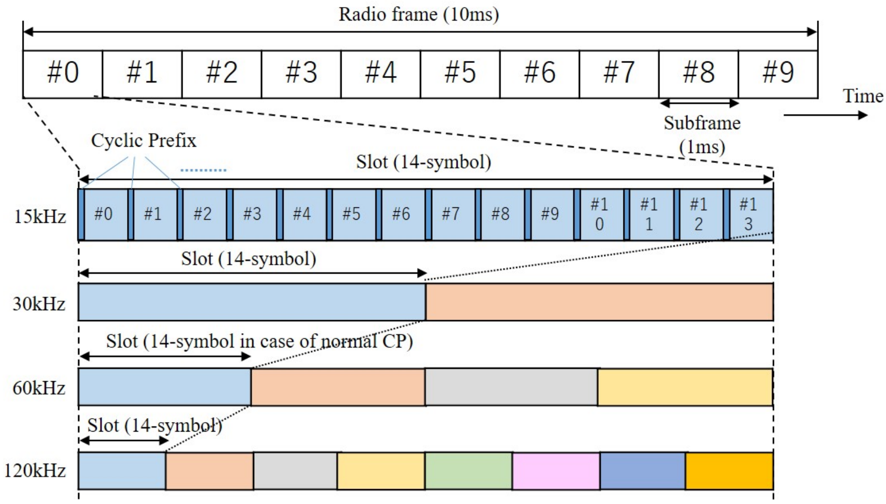
In the table 7.1 we have represented the two frequencies that are going to be used in 5G. We can notice an important thing that will help us to reduce the latency which is the bandwidth. In our application, we will use the high frequencies and so we will have a wide bandwidth. We will have an increase of the bit rate because in the formula to calculate the bit rate we have that the bandwidth is at the numerator. Given the order of magnitude of the bandwidth, it will affect the result of the bit rate.

Table 7.1: 5G frequency [3GP19b].

Frequency range	Frequency range [MHz]	Supported channel bandwidth [MHz]
FR1	410 - 7125	5, 10, 15, 20, 25, 30, 40, 50, 60, 80, 90, 100
FR2	24250 - 52600	50, 100, 200, 400

Here we will present the frame structure since we can not broadcast as we want, but we have to respect standards.

Similar to LTE, orthogonal frequency-division multiplexing (OFDM) with cyclic prefix (CP) is utilized as a downlink waveform for 5G. Contrary to LTE, OFDM can additionally be utilized in the uplink direction NR. As a complementary waveform with a lower peak-to-average power ratio (PAPR) to improve uplink coverage, OFDM with discrete Fourier Transform precoding can be used in the uplink, however, limited to single-layer transmission only. To allow this flexibility, 5G uses a flexible frame structure with several subcarrier spacings (SCS). An SCS is the spacing between the centres of any two consecutive sub-carriers, and the allowable values for the SCS are (in kHz): 15; 30; 60; 120 and 240. A series of 12 consecutive sub-carriers forms a single resources block (RB). An NR channel bandwidth is made up of several RBs. A resource element is referred to as a unit of a sub-carrier (frequency domain) and an OFDM (time-domain) symbol. As far as the time domain is concerned, it is divided into 10 ms radio frames, each of which consists of 10 subframes of 1 ms each, as shown in figure 7.2. In turn, each subframe is composed of 1, 2, 4, 8 or 16 slots (in figure 7.2 they are showed with different colours) according to the SCS selected. The figure 7.2 shows cases of SCS values of 15, 30, 60 and 120 kHz. According to the figure for the 15 kHz case, each slot is made up of 14 OFDM symbols preceded by a CP.



,

Figure 7.2: Frame structure in 5G [3GP19b].

The duration of the OFDM symbol and the length of the CP are inversely proportional to the SCS. The table shows the number of values which are important for the latency estimation.

Table 7.2: Scalable OFDM numerology in 5G [3GP19b].

Subcarrier Spacing Δf [kHz]	15	30	60	$15{\times}2^n$
OFDM Symbol suration [μs]	66.67	33.33	16.67	$66.67/2^n$
Cyclic prefix duration [μs]	4.69	2.34	1.17	$4.69/2^n$
Slot duration (14 symbols) [μs]	1000	500	17.84	$1000/2^n$
Slot in subframe	1	2	4	2^n

However, the transmissions are normally carried out on a single slot. In certain cases, transmissions can be made on only a portion of a slot, with a minimum of two symbols. Such very short transmissions are primarily intended for use cases requiring low latency, such as some URLLC services.

Now let's calculate how much data we can transmit in two symbols. This information will be essential because through it we will know if it is enough to transmit our data and also we will be able to calculate the time taken. To do this calculation we need to make assumptions. These are the assumptions:

- SCS = 120 kHz. We have chosen this value because we will use the high frequency (FR2) and therefore we will have a wide bandwidth, so, it is not a problem to have a large SCS.
- Bit rate = 100 Mbit/s corresponding to the user experienced data rate. This value is proposed by the International Mobile Telecommunications 2020 (IMT-2020).

Using a SCS we have that the slot duration is equal to 62.5 μs. this value is taken from the table 7.2. Let's calculate the number of bytes that we can carry in two symbols:

$$C_d = \frac{Slot\ duration}{7} Bit\ rate = \frac{62.5 \times 10^{-6}\,s}{7} 10 \times 10^{7}\,bit/s \sim 900\,bit \sim 112\,B \quad (7.7)$$

where C_d is the carried bytes in due symbols of a slot. We can see that 112 B are enough for our application both for the uplink and for the downlink.

We can conclude that the time taken will correspond to 2 slot symbols plus twice the CP. The result obtained will then be multiplied by two, since we are calculating the latency. We can neglect the CP since it is a very low value and will not affect the final calculation. numerically speaking we have a value corresponding to approximately 18 μs.

Loss packets

The loss of packets happens when one or more data packets travelling through a network and they fail to reach their destination. Packet loss is due to errors in data transmission, typically over wireless networks, or network congestion. Packet loss is typically determined as a percentage of lost packets compared to sent packets. The Transmission Control Protocol (TCP) senses packet loss and performs retransmissions to provide reliable messaging. Packet loss in a TCP connection is used to avoid congestion and therefore produces intentionally reduced throughput for the connection.

The method used by the TCP communication protocol to identify a packet loss is to use acknowledgement. This means that once the package is received, it is generated. This acknowledgement lets the transmitter know that the packet has been received correctly. In our application, it is important to know if we receive and transmit correctly. This is to prevent our vehicle from disappearing from the map. This could cause big problems for example accidents. In our application, we do not even tolerate the disappearance of the vehicle for a short period. Because this would result in providing incorrect information (sending to the other vehicles the last position received and also, we estimate how much the vehicle has advanced since we had the speed of the vehicle before it disconnected. This estimate starts after 10 milliseconds).

An import value to be established is how often we have to send our information to the Fog node. We decided to send the information every 5 milliseconds (Δt). This choice is dictated by mathematical calculations. Let's assume that the average speed (Δv) of the vehicles is 60 km/h which is about 17 m/s. Assuming all goes well, we have that between transmissions, the vehicle has advanced 8.5 centimetres. This has been calculated using the following formula:

$$d = \Delta v \times \Delta t = 0.017 \times 10^3\, m/s \times 5 \times 10^{-3}\, s = 0.085\, m = 8.5\, cm \tag{7.8}$$

Where d is the distance travelled in 5 milliseconds at a speed of 17 m/s. We can see that this distance meets the safety requirements. That is because an inaccuracy like that can not cause an accident.

Now let's see in case you do not get an acknowledgement. This choice will also be justified mathematically speaking. After not receiving two acknowledgement. The third transmission will transmit the data through both of the two frequencies used. This means our vehicle has not updated its correct location for 15 milliseconds. Using formula 7.8 we can obtain that the vehicle can be advanced by about 25 centimetres. This value can still be tolerated.

So far we have considered the transmission from the vehicle to Fog node. Now let's treat the one from the Fog node to the vehicle. These two links are disjointed, but they share the same numerology. It only changes that if we do not receive two acknowledgements from one vehicle then the next transmission will be transmitted by both small cell and macrocell.

Signal strength

In our case, we will not need repeaters. So we will not be delayed. This is because the 5G uses small cells that provide excellent coverage and in our application, we can benefit also from the macro cell coverage.

Processing and storage delays

Processing data and storing in memory certainly takes time. In the next section, we can see how we can make it possible that this time is not included in the latency calculation.

7.1.2 Uncorrelation

Our idea is based on the possibility of uncorrelated the two messages. This means that the messages sent from the fog node to the vehicles and the message sent from the vehicle to the fog node are not correlated. This is possible because we (users) will transmit data related to their position, orientation and velocity and they will receive data relative to the other vehicle's position. We can see that neither message needs the other to be transmitted.

This idea will allow us to greatly reduce latency. Because the data processing is not involved in the latency calculation. This is because the vehicle that sends its data does not expect the processing of the transmitted data. The latency will be calculated in this way:

- Time needed to reach the Fog node.
- The time needed to create the acknowledgement.
- Time needed to reach the user.

This is the latency when we consider when transmitting data to the fog node. We can see that even when we consider the Fog node sending data to the other vehicles, we have the same process only with destination and source reversed.

The possibility of sending messages that did not depend on those received was not a trivial thing to find. Thinking deep down we found that we could send the position, direction and speed and at the same time receive the position of the other vehicles. In the next section, we will see that this technique has helped us to reduce the latency considerably.

7.1.3 Latency calculation

The latency is not only the sum of the values found so far. Therefore, if this were possible, the latency would have been considered negligible due to the very low value. Now let's consider the latency of the user plan. This corresponds to the radio network contribution to the time from the point when the source sends a packet to when the target receives it. This is the return time required to successfully deliver a packet of the application layer from the 2/3 service data unit (SDU) layer radio protocol layer input point to the 2/3 SDU radio protocol layer output point of the radio interface at either uplink or downlink in the network for a certain service under unloading conditions, supposing that the small cell is in the status active.

Alignment delay is the time needed after we are ready to send until the transmission can start. Primarily, this time involves waiting for the start of slots/sub-slots/mini-slots. The worst possible latency is assumed, which means that the alignment delay is assumed as long as possible. For NR, it is taken that the EU processing delay is 3 symbols OFDM for 15 and 30 kHz SCS and 9 symbols for 120kHz SCS [SKA+18]. The shortest response time to a small cell (between hybrid automatic repeat request (HARQ) acknowledgement and re-transmission) is assumed to be 1 transmission time interval (TTI). For a higher SCS and a mini-slot with fewer symbols, the TTI is shorter, and more TTIs are considered in processing.

In the section above (packets loss) we have identified that among the causes of delay is the loss of packages. We now see that thanks to certain techniques it is possible to retransmit in a very short time. For example, we have that two HARQ-based transmissions are possible within the latency bound of 1 ms for the 120 kHz SCS with 14 symbols [SKA+18]. Considering another re-transmission technique we have 5 re-transmissions in the same configuration [SKA+18]. This technique is called automatic repetitions.

In this section of the results, we looked at user latency. That consider one-way latency. Thanks to the idea of uncorrelation presented in the section above, this one-way latency corresponds to round trip latency. This is because the data received must not be processed. This allows us to have low latency even in case of packet loss. We know that the longer times spend is while processing the data. That is true because in our case we do not have control plane latency due to the presence of fog nodes.

For the sake of completeness, we also need to talk about data processing both by the vehicle and by the fog node. For the information received to be used, it must be processed. Now let's consider the two types of processing:

- Vehicle: The data received from the vehicle is processed locally. These data are the positions of the other vehicles that must be mapped on the screen. In addition, in case of emergency, it sends warnings or intervenes by braking or steering.
- Fog node: All the data received from the vehicles are processed and then select the positions to be sent to each vehicle.

These two operations do not require a large computational power. The fact is that you will receive new data every 5 milliseconds. So they have the necessary time to meet the latency requirements of under 5 milliseconds for the vehicle environment.

The concept of uncorrelation is still valid because when we transmit a data we don't need to wait for a response. This also applies when we receive data from the network. So during the period of processing at the locate (vehicle) level it is possible to send data to the network. As far as the fog node processing is concerned, it is possible to receive data from the vehicles during the processing. This is because the two actions are uncorrelated.

7.2 Reliability

A key design requirement of the new generation 5G wireless network infrastructure is the guarantee of close to 100% uptime and ultra-consistent network services. Most of the services that rely on 5G include autonomous driving vehicles, health services, traffic orchestration, power network management, and other services directly responsible for public safety.

Reliability should not be considered only at the link level. Network links, devices such as switches, routers, firewalls, application delivery controllers, servers, storage systems and others should be reliable; also component of these devices needs to be reliable.

In the next two subsections we'll simulate the failure of one layer by layer at a time. We will try to find a solution that allows us to claim that our system is reliable.

7.2.1 Fog node layer

I think its a good idea to have the figure 7.3 of the architecture. That is to make the explanation easier to follow. We know the components can fail. This should obviously not block our whole application. In this sub-section we will simulate malfunction/non-operation situations of some components in the fog node layer. In this layer we have two main components, which are Fog node and small cell. Now let's consider the failure of one at a time:

- Fog node fail: This means that the transmitted signal cannot reach the fog node. however, there are two solutions to this problem. The first solution is to connect to the next small cell. But it is to be expected that this small cell is not connected to the same fog node. As we discussed earlier, we will have three or more small cells connected to the same fog node. In case this next small cell is connected to a different fog node it is possible to start transmitting and receiving information since the fog node is connected to the BBU which has a big view.

  The second option is dictated by the presence of two types of connection. So we could use the secondary connection. That means we connect to the macro cell. Then we will have direct access to the data in the BBU. So it will be like connecting to the small cell only with a higher latency.

- Small cell: That means we cannot connect to the small cell and consequently not even to the Fog node. We still have two solutions. The first solution is to connect to the next small cell. This unlike before is not important if we are connected to the same Fog node or a new one, so it's easier to manage. This is because I would be handled in the same way.

  In the second case, we consider that this is not possible or too late. The second solution is the same as the second solution already presented. Which is to connect to the macro cell.

7.2.2 Baseband layer

In this sub-section we will consider the BBU layer. One of the layers of great support in case of problems with the fog node layer. So the failure of this layer involves serious consequences if we consider a heterogeneous scenario. Instead, if we consider our case there are no changes if the fog node layer works nominally. The only problem is that we won't get any information about what the next fog node is, but this can be managed by performing calculations inside the fog node.

In the BBU layer we have two entities can fail, which are macro cell and BBU. Now let's see what happens if one of the two components fails:

- BBU: This component is the brain of our network. So in case of failure, we have several problems. One of the biggest ones is that you won't get a big

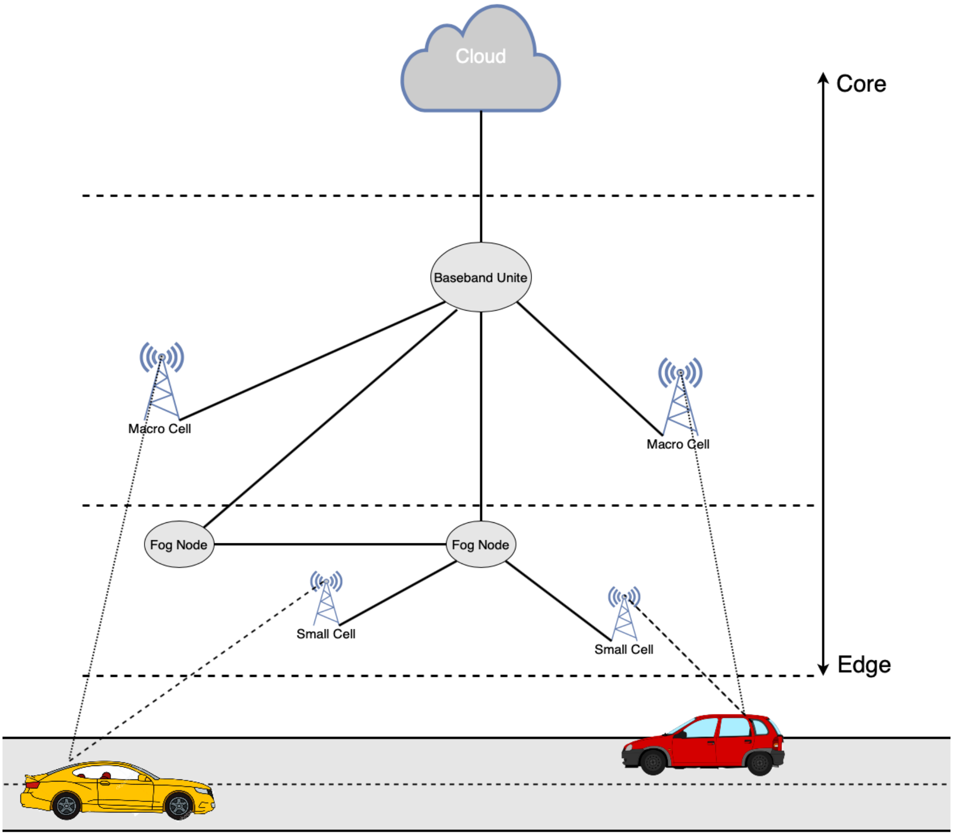


Figure 7.3: Architecture.

picture of the whole town. This information is useful in an emergency. But as we said before, this is not going to make a big change to our application. In case we fail to pass the information to the BBU, we will send the information to the nearest fog to a new BBU. This will allow us to solve the problem since we will have a bigger picture of our scenario in close to the one that failed.

- Macro cell: In case you cannot connect to the macro cell. It is always possible to try to connect to the next one. Since in normal situations this link will only be used for entertainment, it will not be a problem. Since we can tolerate being logged out for entertainment.

Chapter 8

Conclusions and Future Research Directions

In this thesis, we have presented a vehicular system that allows reducing the number of accidents and besides, facilitates traffic coordination. The objective of the thesis is to guarantee good performance in three aspects:

- Latency: here we exploit the uncollelation of output data with input data concept. Therefore, thanks to that we obtained a low latency even in case of errors during the transmission as explained in the above chapter.
- Mobility management: to ensure a fast and effective handover we rely on the cooperation between the various small cells through fog nodes. The exchange of information helped also by the knowledge of the vehicle's destination, accelerates the handover process.
- Reliability: to ensure high system reliability, we used two ways to get it. The first thanks to the use of two different frequencies. It is possible to rely on the other connection in case the first one fails. The second is thanks to the connection present between the various nodes and between the fog nodes the BBU.

We can observe that thanks to our idea we have considerable improvements. These improvements can be big or small depending on the components used. For example, in case both the processor in the fog node and the one in the vehicle had a high computing power then we will have the best case with the highest performance.

8.1 Future Research Directions

The thesis deals with a proposal that can be used from now since all the features to make it work are present. Therefore, this idea can be geared to the future with the use of new technologies. For example, the possibility of introducing V2V to allow direct communication between vehicles. So, it could help in the best way the accuracy in calculating the distance between one vehicle and another. This technology is possible to integrate thanks to the fact that it is present in the 5G standard.

The introduction of sensors/cameras to detect the presence of pedestrians or objects. This becomes even more relevant when talking about autonomous vehicles where the processor inside the vehicle is the decision-maker. So, knowing what is present around the vehicle is important to avoid accidents. Another key point is to reduce the time it takes to access the data, which is even more critical when travelling at high speeds.

A future research is to receive suggestions from fog computing on the shortest and freer route to get to your destination. To redistribute the traffic to have about the same number of vehicles on all roads. This idea can be extended and therefore used for rescue vehicles. Then it will be possible to warn the vehicles of the presence of a rescue vehicle and the rescue vehicles will take the shortest route.

A interesting topic might be to remove all the traffic lights in the city. This will be possible thanks to the fact that everything will be connected and therefore there will no longer be any need for traffic lights. This, of course, requires a lot of research to create such a situation.

References

[3GP17] 3GPP. LTE. Technical Specification (TS) 36.141, 3rd Generation Partnership Project (3GPP), 04 2017. Version 13.7.0.

[3GP19a] 3GPP. NR and NG-RAN Overall Description. Technical Specification (TS) 21.915, 3rd Generation Partnership Project (3GPP), 12 2019. Version 16.0.0.

[3GP19b] 3GPP. Release 15 Description. Technical Report (TR) 21.915, 3rd Generation Partnership Project (3GPP), 09 2019. Version 15.0.0.

[Ant14] S Antonio. Cisco delivers vision of fog computing to accelerate value from billions of connected devices. *Cisco*, 2014.

[ARS16] Mamta Agiwal, Abhishek Roy, and Navrati Saxena. Next generation 5g wireless networks: A comprehensive survey. *IEEE Communications Surveys & Tutorials*, 18(3):1617–1655, 2016.

[BMLS17] Damir Bećirbašić, Marijana Molnar, Željko Lukač, and Dragan Samardžija. Video-processing platform for semi-autonomous driving over 5g networks. In *2017 IEEE 7th International Conference on Consumer Electronics-Berlin (ICCE-Berlin)*, pages 42–46. IEEE, 2017.

[BS15] Tarek Bouali and Sidi-Mohammed Senouci. A fuzzy logic-based communication medium selection for qos preservation in vehicular networks. In *Proceedings of the 5th ACM Symposium on Development and Analysis of Intelligent Vehicular Networks and Applications*, pages 101–108, 2015.

[CGLZ16] Xuerong Cui, Thomas Aaron Gulliver, Juan Li, and Hao Zhang. Vehicle positioning using 5g millimeter-wave systems. *IEEE Access*, 4:6964–6973, 2016.

[dPRLSKSG16] José A del Peral-Rosado, José A López-Salcedo, Sunwoo Kim, and Gonzalo Seco-Granados. Feasibility study of 5g-based localization for assisted driving. In *2016 International Conference on Localization and GNSS (ICL-GNSS)*, pages 1–6. IEEE, 2016.

[dPRRLSSG17] José A del Peral-Rosado, Ronald Raulefs, José A López-Salcedo, and Gonzalo Seco-Granados. Survey of cellular mobile radio localization methods: From 1g to 5g. *IEEE Communications Surveys & Tutorials*, 20(2):1124–1148, 2017.

[dPRSGKLS19] José A del Peral-Rosado, Gonzalo Seco-Granados, Sunwoo Kim, and José A López-Salcedo. Network design for accurate vehicle localization. *IEEE Transactions on Vehicular Technology*, 68(5):4316–4327, 2019.

[DR$^+$58] Wilbur B Davenport, William L Root, et al. *An introduction to the theory of random signals and noise*, volume 159. McGraw-Hill New York, 1958.

[DRC$^+$16] Kakan Chandra Dey, Anjan Rayamajhi, Mashrur Chowdhury, Parth Bhavsar, and James Martin. Vehicle-to-vehicle (v2v) and vehicle-to-infrastructure (v2i) communication in a heterogeneous wireless network–performance evaluation. *Transportation Research Part C: Emerging Technologies*, 68:168–184, 2016.

[Eri13] David Erickson. The beacon openflow controller. In *Proceedings of the second ACM SIGCOMM workshop on Hot topics in software defined networking*, pages 13–18, 2013.

[ETS] ETSI. Multi-access edge computing (mec).

[ETS13] GSNFV ETSI. Network functions virtualisation (nfv): Architectural framework. *ETsI Gs NFV*, 2(2):V1, 2013.

[EVBH08] Jörg Eberspächer, Hans-Jörg Vögel, Christian Bettstetter, and Christian Hartmann. *GSM-architecture, protocols and services.* John Wiley & Sons, 2008.

[FA16] Borko Furht and Syed A Ahson. *Long Term Evolution: 3GPP LTE radio and cellular technology.* Crc Press, 2016.

[Fri46] Harald T Friis. A note on a simple transmission formula. *Proceedings of the IRE*, 34(5):254–256, 1946.

[GB13] Y Jay Guo and Stephen K Barton. *Fresnel zone antennas.* Springer Science & Business Media, 2013.

[GKJ$^+$13] Aaron Gember, Anand Krishnamurthy, Saul St John, Robert Grandl, Xiaoyang Gao, Ashok Anand, Theophilus Benson, Vyas Sekar, and Aditya Akella. Stratos: A network-aware orchestration layer for virtual middleboxes in clouds. *arXiv preprint arXiv:1305.0209*, 2013.

[HB19] Marcin Hoffmann and Hanna Bogucka. Localization techniques for 5g radio environment maps. In *International Conference on Cognitive Radio Oriented Wireless Networks*, pages 232–246. Springer, 2019.

[HBPT15] Oliver Hahm, Emmanuel Baccelli, Hauke Petersen, and Nicolas Tsiftes. Operating systems for low-end devices in the internet of things: a survey. *IEEE Internet of Things Journal*, 3(5):720–734, 2015.

[HGJL15] Bo Han, Vijay Gopalakrishnan, Lusheng Ji, and Seungjoon Lee. Network function virtualization: Challenges and opportunities for innovations. *IEEE Communications Magazine*, 53(2):90–97, 2015.

[HGPR⁺16] Robert W Heath, Nuria Gonzalez-Prelcic, Sundeep Rangan, Wonil Roh, and Akbar M Sayeed. An overview of signal processing techniques for millimeter wave mimo systems. *IEEE journal of selected topics in signal processing*, 10(3):436–453, 2016.

[HLC⁺16] Xueshi Hou, Yong Li, Min Chen, Di Wu, Depeng Jin, and Sheng Chen. Vehicular fog computing: A viewpoint of vehicles as the infrastructures. *IEEE Transactions on Vehicular Technology*, 65(6):3860–3873, 2016.

[HSE⁺12] Rasheed Hussain, Junggab Son, Hasoo Eun, Sangjin Kim, and Heekuck Oh. Rethinking vehicular communications: Merging vanet with cloud computing. In *4th IEEE International Conference on Cloud Computing Technology and Science Proceedings*, pages 606–609. IEEE, 2012.

[HZXH15] Debiao He, Sherali Zeadally, Baowen Xu, and Xinyi Huang. An efficient identity-based conditional privacy-preserving authentication scheme for vehicular ad hoc networks. *IEEE Transactions on Information Forensics and Security*, 10(12):2681–2691, 2015.

[Int14] GSMA Intelligence. Understanding 5g: Perspectives on future technological advancements in mobile. *White paper*, pages 1–26, 2014.

[JAQ90] J-J JAQUIER. Etsi-european telecommunications standards institute-institut européen de normalisation dans le domaine des télécommunications. i: Historique, mission, structure et méthodes de travail. *Technische Mitteilungen PTT*, 68(7):276–284, 1990.

[JMJZ19] Mohammad Reza Jabbarpour, Alireza Marefat, Ali Jalooli, and Houman Zarrabi. Could-based vehicular networks: a taxonomy, survey, and conceptual hybrid architecture. *Wireless Networks*, 25(1):335–354, 2019.

[KA⁺19] Abla Kammoun, Mohamed-Slim Alouini, et al. Elevation beamforming with full dimension mimo architectures in 5g systems: A tutorial. *IEEE Communications Surveys & Tutorials*, 21(4):3238–3273, 2019.

[KAH⁺19] Wazir Zada Khan, Ejaz Ahmed, Saqib Hakak, Ibrar Yaqoob, and Arif Ahmed. Edge computing: A survey. *Future Generation Computer Systems*, 97:219–235, 2019.

[KCW⁺17] Mike Koivisto, Mário Costa, Janis Werner, Kari Heiska, Jukka Talvitie, Kari Leppänen, Visa Koivunen, and Mikko Valkama. Joint device positioning and clock synchronization in 5g ultra-dense networks. *IEEE Transactions on Wireless Communications*, 16(5):2866–2881, 2017.

[KLL⁺17] Yu-Jen Ku, Dian-Yu Lin, Chia-Fu Lee, Ping-Jung Hsieh, Hung-Yu Wei, Chun-Ting Chou, and Ai-Chun Pang. 5g radio access network design with the fog paradigm: Confluence of communications and computing. *IEEE Communications Magazine*, 55(4):46–52, 2017.

[LWJ$^+$17] Jianqi Liu, Jiafu Wan, Dongyao Jia, Bi Zeng, Di Li, Ching-Hsien Hsu, and Haibo Chen. High-efficiency urban traffic management in context-aware computing and 5g communication. *IEEE Communications Magazine*, 55(1):34–40, 2017.

[LWZ$^+$17] Jianqi Liu, Jiafu Wan, Bi Zeng, Qinruo Wang, Houbing Song, and Meikang Qiu. A scalable and quick-response software defined vehicular network assisted by mobile edge computing. *IEEE Communications Magazine*, 55(7):94–100, 2017.

[Mat17] Sven Mattisson. Overview of 5g requirements and future wireless networks. In *ESSCIRC 2017-43rd IEEE European Solid State Circuits Conference*, pages 1–6. IEEE, 2017.

[MCK19] Leo Mendiboure, Mohamed-Aymen Chalouf, and Francine Krief. Edge computing based applications in vehicular environments: Comparative study and main issues. *Journal of Computer Science and Technology*, 34(4):869–886, 2019.

[ME06] Pratap Misra and Per Enge. Global positioning system: signals, measurements and performance second edition. *Global Positioning System: Signals, Measurements And Performance Second Editions*, 206, 2006.

[MMBS$^+$19] Ingrid Moerman, Johann Marquez-Barja, Adnan Shahid, Wei Liu, Spilios Giannoulis, and Xianjun Jiao. *Cognitive Radio Oriented Wireless Networks: 13th EAI International Conference, CROWNCOM 2018, Ghent, Belgium, September 18–20, 2018, Proceedings*, volume 261. Springer, 2019.

[Moo17] Samuel K Moore. Super-accurate gps coming to smartphones in 2018 [news]. *IEEE Spectrum*, 54(11):10–11, 2017.

[MR17] George R MacCartney and Theodore S Rappaport. Rural macrocell path loss models for millimeter wave wireless communications. *IEEE Journal on selected areas in communications*, 35(7):1663–1677, 2017.

[MRSD15] George R Maccartney, Theodore S Rappaport, Shu Sun, and Sijia Deng. Indoor office wideband millimeter-wave propagation measurements and channel models at 28 and 73 ghz for ultra-dense 5g wireless networks. *IEEE access*, 3:2388–2424, 2015.

[MSG$^+$15] Rashid Mijumbi, Joan Serrat, Juan-Luis Gorricho, Niels Bouten, Filip De Turck, and Raouf Boutaba. Network function virtualization: State-of-the-art and research challenges. *IEEE Communications surveys & tutorials*, 18(1):236–262, 2015.

[NHW19a] Zhaolong Ning, Jun Huang, and Xiaojie Wang. Vehicular fog computing: Enabling real-time traffic management for smart cities. *IEEE Wireless Communications*, 26(1):87–93, 2019.

[NHW$^+$19b] Zhaolong Ning, Jun Huang, Xiaojie Wang, Joel JPC Rodrigues, and Lei Guo. Mobile edge computing-enabled internet of vehicles: Toward energy-efficient scheduling. *IEEE Network*, 33(5):198–205, 2019.

[NKW+12] Bram Naudts, Mario Kind, Fritz-Joachim Westphal, Sofie Verbrugge, Didier Colle, and Mario Pickavet. Techno-economic analysis of software defined networking as architecture for the virtualization of a mobile network. In *2012 European workshop on software defined networking*, pages 67–72. IEEE, 2012.

[NWH18] Zhaolong Ning, Xiaojie Wang, and Jun Huang. Mobile edge computing-enabled 5g vehicular networks: Toward the integration of communication and computing. *ieee vehicular technology magazine*, 14(1):54–61, 2018.

[OSP+15] A Osseiran, J Sachs, M Puleri, et al. Manufacturing reengineered: robots, 5g and the industrial iot. *Ericsson Business Review*, (4), 2015.

[PDDCDS17] Giovanni Pecoraro, Simone Di Domenico, Ernestina Cianca, and Mauro De Sanctis. Lte signal fingerprinting localization based on csi. In *2017 IEEE 13th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pages 1–8. IEEE, 2017.

[PDFF17] Stefan Parkvall, Erik Dahlman, Anders Furuskar, and Mattias Frenne. Nr: The new 5g radio access technology. *IEEE Communications Standards Magazine*, 1(4):24–30, 2017.

[PF18] Troels Pedersen and Bernard Henri Fleury. Whitepaper on new localization methods for 5g wireless systems and the internet-of-things. 2018.

[PFD11] Stefan Parkvall, Anders Furuskar, and Erik Dahlman. Evolution of lte toward imt-advanced. *IEEE Communications Magazine*, 49(2):84–91, 2011.

[PJ+15] Kenneth Mark Pesyna Jr et al. *Advanced techniques for centimeter-accurate GNSS positioning on low-cost mobile platforms.* PhD thesis, 2015.

[Rap18] Theodore S Rappaport. 5g millimeter wave wireless: Trials, testimonies, and target rollouts. In *IEEE Infocom*, 2018.

[RBP19] Umberto Robustelli, Valerio Baiocchi, and Giovanni Pugliano. Assessment of dual frequency gnss observations from a xiaomi mi 8 android smartphone and positioning performance analysis. *Electronics*, 8(1):91, 2019.

[RMSS15] Theodore S Rappaport, George R MacCartney, Mathew K Samimi, and Shu Sun. Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design. *IEEE Transactions on Communications*, 63(9):3029–3056, 2015.

[RMZ+16] Rapeepat Ratasuk, Nitin Mangalvedhe, Yanji Zhang, Michel Robert, and Jussi-Pekka Koskinen. Overview of narrowband iot in lte rel-13. In *2016 IEEE conference on standards for communications and networking (CSCN)*, pages 1–7. IEEE, 2016.

[RTG10] Rapeepat Ratasuk, Dominic Tolli, and Amitava Ghosh. Carrier aggregation in lte-advanced. In *2010 IEEE 71st Vehicular Technology Conference*, pages 1–5. IEEE, 2010.

[RXM⁺17] Theodore S Rappaport, Yunchou Xing, George R MacCartney, Andreas F Molisch, Evangelos Mellios, and Jianhua Zhang. Overview of millimeter wave communications for fifth-generation (5g) wireless networks—with a focus on propagation models. *IEEE Transactions on Antennas and Propagation*, 65(12):6213–6230, 2017.

[Sat17] Mahadev Satyanarayanan. The emergence of edge computing. *Computer*, 50(1):30–39, 2017.

[SCZ$^+$16] Weisong Shi, Jie Cao, Quan Zhang, Youhuizi Li, and Lanyu Xu. Edge computing: Vision and challenges. *IEEE internet of things journal*, 3(5):637–646, 2016.

[Sha48] Claude E Shannon. A mathematical theory of communication. *Bell system technical journal*, 27(3):379–423, 1948.

[SKA$^+$18] Alexey Shapin, Kittipong Kittichokechar, Niklas Andgart, Marten Sundberg, and Gustav Wikström. Physical layer performance for low latency and high reliability in 5g. In *2018 15th International Symposium on Wireless Communication Systems (ISWCS)*, pages 1–6. IEEE, 2018.

[SKB$^+$13] Yuya Saito, Yoshihisa Kishiyama, Anass Benjebbour, Takehiro Nakamura, Anxin Li, and Kenichi Higuchi. Non-orthogonal multiple access (noma) for cellular future radio access. In *2013 IEEE 77th vehicular technology conference (VTC Spring)*, pages 1–5. IEEE, 2013.

[SPO] EA SPORTS. Fafa20.

[SRT$^+$16] Shu Sun, Theodore S Rappaport, Timothy A Thomas, Amitava Ghosh, Huan C Nguyen, István Z Kovács, Ignacio Rodriguez, Ozge Koymen, and Andrzej Partyka. Investigation of prediction accuracy, sensitivity, and parameter stability of large-scale propagation path loss models for 5g wireless communications. *IEEE Transactions on Vehicular Technology*, 65(5):2843–2860, 2016.

[Sta15] William Stallings. *Foundations of modern networking: SDN, NFV, QoE, IoT, and Cloud.* Addison-Wesley Professional, 2015.

[Tür07] Ulrich Türke. *Efficient methods for WCDMA radio network planning and optimization.* Springer, 2007.

[TWS$^+$17] Oumer Teyeb, Gustav Wikström, Magnus Stattin, Thomas Cheng, Sebastian Faxér, and Hieu Do. Evolving lte to fit the 5g future. *Ericsson technology review*, 1:1–16, 2017.

[VDP19] Mojtaba Vaezi, Zhiguo Ding, and H Vincent Poor. *Multiple access techniques for 5G wireless networks and beyond.* Springer, 2019.

[VLG$^+$17] Ricard Vilalta, Victor López, Alessio Giorgetti, Shuping Peng, Vittorio Orsini, Luis Velasco, Rene Serral-Gracia, Donal Morris, Silvia De Fina, Filippo Cugini, et al. Telcofog: A unified flexible fog and cloud computing architecture for 5g networks. *IEEE Communications Magazine*, 55(8):36–43, 2017.

[WAR06] Yong Wang, Garhan Attebury, and Byrav Ramamurthy. A survey of security issues in wireless sensor networks. 2006.

[WML+16] Klaus Witrisal, Paul Meissner, Erik Leitinger, Yuan Shen, Carl Gustafson, Fredrik Tufvesson, Katsuyuki Haneda, Davide Dardari, Andreas F Molisch, Andrea Conti, et al. High-accuracy localization for assisted living: 5g systems will turn multipath channels from foe to friend. *IEEE Signal Processing Magazine*, 33(2):59–70, 2016.

[WSGD+17] Henk Wymeersch, Gonzalo Seco-Granados, Giuseppe Destino, Davide Dardari, and Fredrik Tufvesson. 5g mmwave positioning for vehicular networks. *IEEE Wireless Communications*, 24(6):80–86, 2017.

[XZ17] Yu Xiao and Chao Zhu. Vehicular fog computing: Vision and challenges. In *2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pages 6–9. IEEE, 2017.

[ZML+17] Ke Zhang, Yuming Mao, Supeng Leng, Yejun He, and Yan Zhang. Mobile-edge computing for vehicular networks: A promising network paradigm with predictive off-loading. *IEEE Vehicular Technology Magazine*, 12(2):36–44, 2017.

[ZVK+15] Meifang Zhu, Joao Vieira, Yubin Kuang, Kalle Åström, Andreas F Molisch, and Fredrik Tufvesson. Tracking and positioning using phase information from estimated multi-path components. In *2015 IEEE International Conference on Communication Workshop (ICCW)*, pages 712–717. IEEE, 2015.

[ZYD+17] Ming Zeng, Animesh Yadav, Octavia A Dobre, Georgios I Tsiropoulos, and H Vincent Poor. Capacity comparison between mimo-noma and mimo-oma with multiple users in a cluster. *IEEE Journal on Selected Areas in Communications*, 35(10):2413–2424, 2017.