

POLITECNICO DI TORINO

Collegio di Ingegneria Elettronica, delle Telecomunicazioni e Fisica

**Corso di Laurea Magistrale
in Computer And Communication Networks Engineering**

Tesi di Laurea Magistrale

Design and performance evaluation of mm-wave vehicular networks



Thesis coordinator:

Carla Fabiana Chiasserini

Thesis advisor:

Francesco Malandrino

Candidate:

Balbo Edoardo

Marzo 2019

Abstract

The rapid increase of mobile data growth and the use of smart phones are creating unprecedented challenges for wireless service providers to overcome a global bandwidth shortage. The millimeter wave (mmWave) frequencies offer the availability of huge bandwidths to provide unprecedented data rates to next-generation cellular mobile terminals. MmWave communications will play a major role in the Fifth generation of mobile networks, they will provide much greater throughput and much lower latency despite the issues related to the higher frequencies; To mitigate these problems the use of directional beams and an high density deployment are needed.

In this thesis we want to test the performance of a mmwave network in a real urban area analyzing the quality perceived by the users (vehicles) considering different scenarios. For this purpose we make use of the ns-3 simulator, using the LTE-5G tight integration architecture provided by the ns-3 mmwave module.

Sommario

La rapida crescita del traffico mobile e dell'uso degli smarthones stanno creando una sfida senza precedenti per gli Wireless internet service provider per superare la carenza di banda globale. Le frequenze millimetriche offrono una enorme disponibilità di banda per fornire dati a velocità senza precedenti ai terminali mobili che sfruttano la prossima generazione cellulare. Le comunicazioni ad onde millimetriche svolgeranno un ruolo chiave nelle reti di quinta generazione. Essi forniranno un throughput molto più elevato e un ritardo molto più basso nonostante i problemi ben noti relative a frequenze più elevate. Per mitigare questi problemi l'uso di fasci direzionali e un spigamento ad alta densità delle celle è necessario. In questa tesi vogliamo verificare le prestazioni di una rete ad onde millimetriche in un'area urbana reale e analizzare la qualità percepita dagli utenti (veicoli) considerando diversi scenari di simulazione. A questo proposito utilizzeremo il simulatore ns3, usando l'architettura di integrazione LTE-5G fornita dal modulo NY mmwave.

Contents

-Abstract	ii
-List Figures	viii
-Acronyms	x
1 Introduction	1
1.1 Internet Traffic Forecast	1
1.2 2G/3G/4G Versus 5G Traffic Forecast	2
1.3 Mmwave Spectrum	3
1.4 Avantages And Disadvantages	4
1.5 Applications	6
2 MmWave Comunications	
2.1 Mmwave Channel	9
2.2 Beamforming	13
2.2.1 Beam Alignment	16
3 5G Mmwave Network	31
3.1 Lte System Architecture	31
3.1.1 User And Control Plain Architecture	34
3.1.2 Layers Functionalities	36
3.2 Toward 5G....	38
3.2.1 Key problems	38
3.2.2 MmWave Module for ns-3	
3.2.2.1 Class Diagram for the end-to-end MmWave module	39
3.2.2.2 Channel Modelling	40
3.3.2.3 Mmwave Layers	44
3.2.2.4 Dual Connectivity Integration	46
3.2.2.5 Initial Dual Connected Access and Handover/Switch procedure	50
3.2.2.6 Bug and problems ns3 in the mmwave NY module	53
4 NS3 Simulator And SUMO	55
4.1 Luxembourg SUMO Traffic (LuST) Scenario	55
5 Simulation Results And Performance Analisis	59
5.1 Simulation	59
5.1.1 Introduction	59
5.1.2 Simulation Parameters	61
5.1.3 Average PDCP throuput and jain's fairness index analysis with RLC UM and ULA 4x4	61
5.1.4 Short term throuput analysis on the map with RLC UM and ULA 4x4	63

5.1.5	Average PDCP throuput and jain's fairness index analysis with RLC UM and ULA 16x16	68
5.1.6	Short term throuput analysis on the map with RLC UM and ULA 16x16	69
5.1.7	Average PDCP throuput and jain's fairness index analysis with RLC UM and ULA 64x64	72
5.1.8	Short term throuput analysis on the map with RLC UM and ULA 64x64	73
5.1.9	Average PDCP throuput and jain's fairness index comparison with RLC UM	76
5.1.10	Average PDCP throuput and jain's fairness index analysis with RLC AM and ULA 64x64	77
5.1.11	Short term throuput analysis on the map with RLC AM and ULA 64x64	79
5.1.12	Average PDCP throuput analysis and jain's fairness index with RLC AM and hibrid scenario	82
5.1.13	Short term throuput analysis on the map with RLC AM and hibrid scenario	83
6 Conclusions And Future Work		87

List of Figures

1.1 Ericsson outlook of Mobile data traffic per active smartphone	1
1.2 Ericsson outlook global mobile traffic 5G and 4G/3G/2G	2
1.3 Mmwave frequency Spectrum	3
1.4 Atmospheric absorption as function of the frequency	5
2.1 Reflection,diffraction,scattering	9
2.2 Fast fading vs slow fading	10
2.3 Delay-angle distribution at low frequencies	11
2.4 Delay-angle distribution at millimeter waves	12
2.5 Analog beamforming architecture	14
2.6 Digital beamforming architecture	15
2.7 Hybrid beamforming architecture	16
2.8 Illustrative example of Agile-Link’s algorithm	20
2.9 Received power at distance of 1 meter vs. misalignment	21
2.10 Beam alignment of the transmitter and receiver with 16 beams.	22
2.11 BeamSpy’s prediction framework	23
2.12 Intuition behind UbiG’s fast beam alignment	25
2.13 UbiG’s AP ranking algorithm	26
2.14 Two fringerprint types	27
2.15 Timing diagram of the proposed inverse fingerprint beam alignment method	28
3.1 Lte system architecture	31
3.2 End-to-End User Plane	35
3.3 End-to-End control Plane	36
3.4 Class diagram for the end-to-end mmWave module	40
3.5 A basic UML diagram that describes the relationships among these classes	42
3.6 Mmwave frame structure	45
3.7 The Dual Connectivity (DC) implementation	48
3.8 Simplified UML of a dual-connected device, an LTE eNB and a MmWave eNB that also support carrier aggregation	50
3.9 (a) Information represented by instances of MmWaveComponentCarrier and extensions	51
(b)UE protocol stack for the integration of multi-RAT dual connectivity and carrier aggregation	51
3.10 Initial dual connected access	52
3.11(a) Switch from LTE RAT to mmwave RAT (b) Switch from mmwave RAT to LTE RAT	53
3.12 Secondary cell Handover procedure(SCH)	54
4.1 Traffic Demand over a day	57

4.2 Lust Scenario Topology	58
4.3 Scheme of the whole the process	59
5.1 Map of the considered area and the relative placement of the antennas	61
5.2 Average throuput per user as a funtion of the numbers of gNB with RLC UM and ULA 4x4	64
5.3 Jain's fairness index as a funtion of the numbers of gNB with RLC UM and ULA 4x4	65
5.4 Colourmap with short-term throuput indication per user using one gNB with RLC UM and ULA 4x4	66
5.5 Colourmap with percentage of short-term throuput LTE and mmwave per user using one gNB with RLC UM and ULA 4x4	66
5.6 Colourmap with short-term throuput indication per user using two gNB with RLC UM and ULA 4x4	67
5.7 Colourmap with percentage of short-term throuput LTE and mmwave per user using two gNB with RLC UM and ULA 4x4	67
5.8 Colourmap with short-term throuput indication per user using three gNB with RLC UM and ULA 4x4	68
5.9 Colourmap with percentage of short-term throuput LTE and mmwave per user using three gNB with RLC UM and ULA 4x4	68
5.10 Colourmap with short-term throuput indication per user using four gNB with RLC UM and ULA 4x4	69
5.11 Colourmap with percentage of short-term throuput LTE and mmwave per user using four gNB with RLC UM and ULA 4x4	69
5.12 Average throuput per user as a funtion of the numbers of gNB with RLC UM and ULA 16x16	70
5.13 Jain's fairness index as a funtion of the numbers of gNB with RLC UM and ULA 16x16	71
5.14 Colourmap with short-term throuput indication per user using one gNB with RLC UM and ULA 16x16	72
5.15 Colourmap with percentage of short-term throuput LTE and mmwave per user using one gNB with RLC UM and ULA 16x16	72
5.16 Colourmap with short-term throuput indication per user using two gNB with RLC UM and ULA 16x16	73
5.17 Colourmap with percentage of short-term throuput LTE and mmwave per user using two gNB with RLC UM and ULA 16x16	73
5.18 Average throuput per user as a funtion of the numbers of gNB with RLC UM and ULA 64x64	74
5.19 Jain's fairness index as a funtion of the numbers of gNB with RLC UM and ULA 64x64	75
5.20 Colourmap with short-term throuput indication per user using one gNB with RLC UM and ULA 64x64	76
5.21 Colourmap with percentage of short-term throuput LTE and mmwave per user using one gNB with RLC UM and ULA 64x64	76
5.22 Colourmap with short-term throuput indication per user using four gNB with RLC UM and ULA 64x64	77
5.23 Colourmap with percentage of short-term throuput LTE and mmwave per user using four gNB with RLC UM and ULA 64x64	77

5.24 Average throuput comparison as a funtion of the numbers of gNB with RLC UM	78
5.25 Jain's fairness index comparison as a funtion of the numbers of gNB with RLC AM and ULA 64x64	79
5.26 Average throuput per user as a funtion of the numbers of gNB with RLC AM and ULA 64x64	80
5.27 Jain's fairness index as a funtion of the numbers of gNB with RLC AM and ULA 64x64	81
5.28 Colourmap with short-term throuput indication per user using one gNB with RLC AM and ULA 64x64	82
5.29 Colourmap with percentage of short-term throuput LTE and mmwave per user using one gNB with RLC AM and ULA 64x64	82
5.30 Colourmap with short-term throuput indication per user using four gNB with RLC AM and ULA 64x64	83
5.31 Colourmap with percentage of short-term throuput LTE and mmwave per user using four gNB with RLC AM and ULA 64x64	83
5.32 Average throuput per user as a funtion of the numbers of gNB with RLC AM and hybrid scenario	84
5.33 Jain's fairness index as a funtion of the numbers of gNB with RLC AM and hybrid scenario	85
5.34 Colourmap with short-term throuput indication per user using one gNB with RLC AM and hybrid scenario	86
5.35 Colourmap with percentage of short-term throuput LTE and mmwave per user using one gNB with RLC AM and hybrid scenario	86
5.36 Colourmap with short-term throuput indication per user using four gNB with RLC AM and hybrid scenario	87
5.37 Colourmap with percentage of short-term throuput LTE and mmwave per user using four gNB with RLC AM and hybrid scenario	87

List Of Tables

3.1 User States Description 36

5.1 Simulation Parameters 63

ACRONYMS

AM	Acknowledged Mode
AQM	Active Queue Management
BS	Base Station
DC	Dual Connectivity
DL	Downlink
EHF	Extremely High Frequency
eNB	evolved Node Base
EPC	Evolved Packet Core
EPS	Evolved Packet System
E-UTRAN	Evolved Universal Terrestrial Radio Access Network
FDMA	Frequency Division Multiple Access
FDD	Frequency Division Duplexing
FS	Fast Switching
FW	Fixed Wireless Access
IR	Infrared waves
GTP	GPRS Tunneling Protocol
HARQ	Hybrid Automatic Repeat reQuest
HH	Hard Handover
IMSI	International Mobile Subscriber Identity
LOS	Line of Sight
LuST	Luxembourg SUMO Traffic
MAC	Medium Access Control
MIMO	Multiple Input Multiple Output
MME	Mobility Management Entity

MTU	Maximum Transfer Unit
NAS	Non access Statum
NLOS	Non Line of Sight
NYU	New York University
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
PDCP	Packet Data Convergence Protocol
PDU	Packet Data Unit
P-GW	Packet Gateway
PHY	Physical
RAN	Radio Access Network
RAT	Radio Access Technology
RLC	Radio Link Control
RRC	Radio Resource Control
SDMA	Space Division Multiple Access
SDU	Service Data Unit
S-GW	Service Gateway
TB	Transport Block
TDD	Time Division Duplexing
TDMA	Time Division Multiple Access
TTI	Trasmission Time Interval
UE	User Equipment
UL	Uplink
ULA	Uniform Linear Array
UM	Unacknowledged Mode

1

Introduction

1.1 Internet Traffic forecast

Traffic is a summary metric that reflects multiple facets of Internet evolution. Internet traffic is the culmination not only of user and device growth, but of growing usage, application trends, and increasing bandwidth availability. There is strong growth coming from mobile (cellular) networks, content is increasingly high-definition, and machines are starting to populate the Internet at a rapid pace.

According to the latest Ericsson Mobility Report, monthly mobile data traffic per smartphone continues to grow in all areas, thanks to improved device capabilities and less costly data plans, as well as an increase in data-intensive content. North America has the highest monthly usage, reaching 8.6 gigabytes (GB) at the end of this year, and is set to reach 50GB by the end of 2024. In North East Asia, traffic per smartphone has grown strongly during 2018 – increasing by around 140 percent year-on-year. The region now has the second-highest monthly usage at 7.3GB and is projected to reach 21GB at the end of the forecast period. Attractive data plans as well as innovative mobile apps and content are driving growth in China.

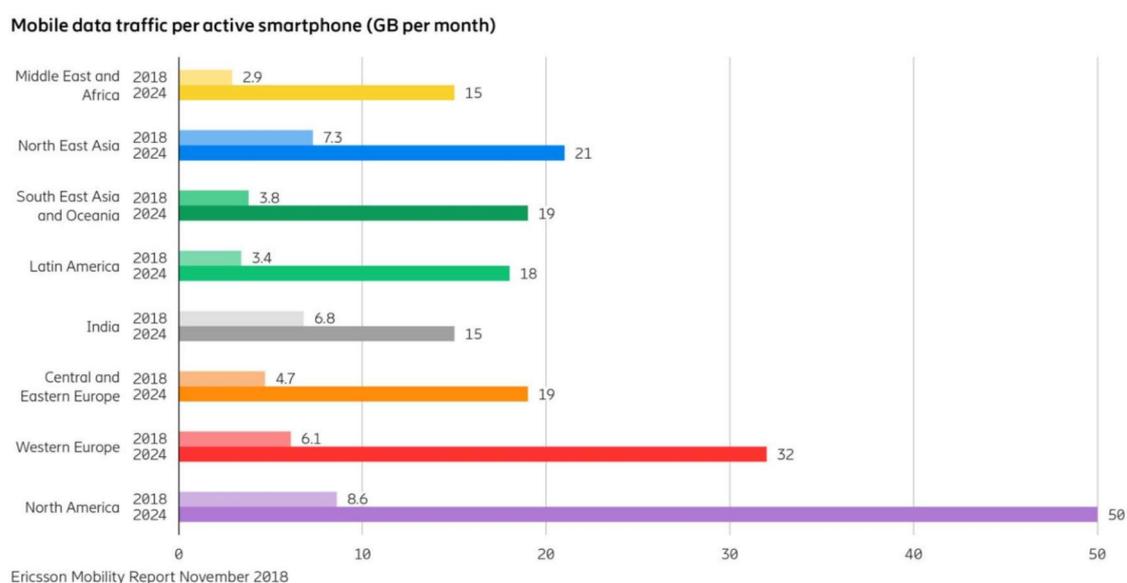


Figure 1.1 Ericsson outlook of Mobile data traffic per active smartphone from [1]

1.2 2G/3G/4G versus 5G traffic forecast

The next generation mobile network (5G) is becoming a reality , driven by an increase in mobile traffic demand and by a variety of use cases that cannot be satisfied by the current LTE networks. According to the latest Ericsson Mobility Report the total mobile data traffic is expected to be five times higher over the next 6years ,it will reach 136EB per month by the end of 2024.

Almost 90 percent of total mobile data traffic is generated by smartphones today and it will reach 95 percent at the end of 2024. As monthly usage per smartphone continues to increase, total mobile data traffic is predicted to rise at a compound annual growth rate (CAGR) of 31 percent over the forecast period, reaching 136 exabytes (EB) per month by the end of 2024. It is expected that 25 percent of mobile data traffic worldwide will be carried by 5G networks at that time. This is 1.3 times more than the total traffic today.

Currently, the 5G traffic forecast does not include traffic generated by fixed wireless access (FWA) services. However, as FWA is one of the early use cases planned for 5G in some regions, it could have a significant impact on the forecast figures, depending on market uptake of the service.

In 2024, 5G networks will carry a quarter of all global mobile data traffic.

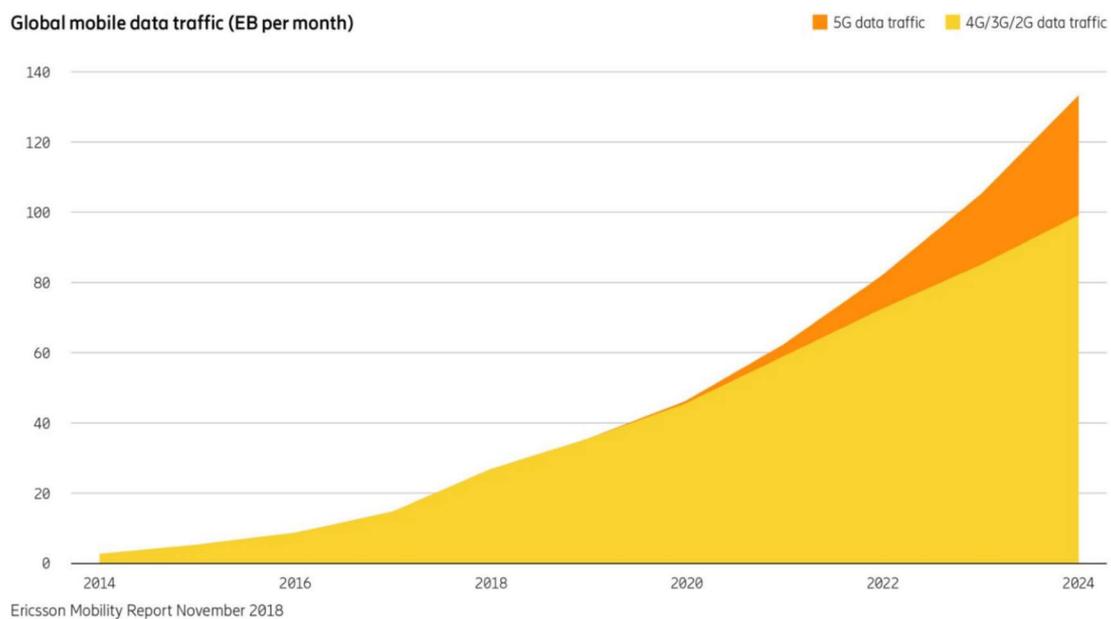


Figure 1.2 Ericsson outlook global mobile traffic 5G and 4G/3G/2G from [1]

1.3 Mmwave spectrum

The adoption of millimeter wave (mmWave) frequencies communications in 5G networks is seen as a way to reach the throughput and capacity increase goals.

Millimeter waves frequencies ranges from 30 GHz to 300 GHz as we can see in the **figure 1.3**. This range of spectrum is located between microwaves (1 GHz to 30 GHz) and infrared (IR) waves, which is normally known as extremely high frequency (EHF). The wavelength interval range from 1-mm to 10-mm.

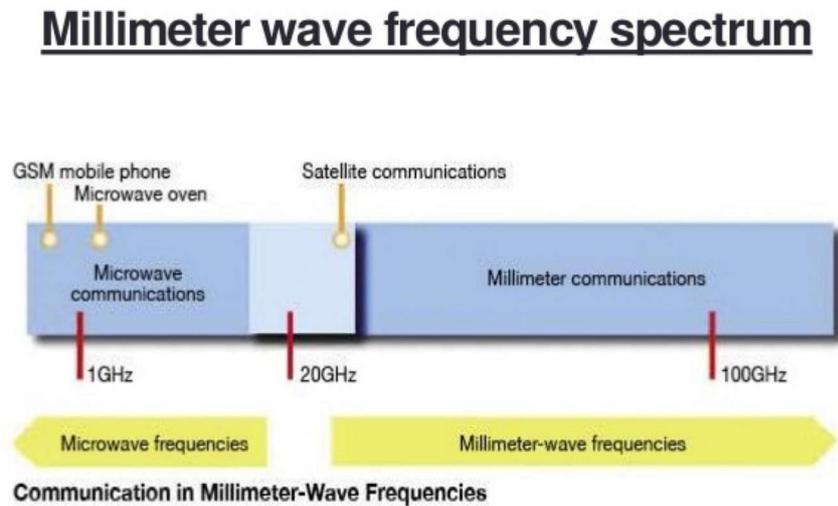


Figure 1.3 Mmwave frequency Spectrum

In the past this interval of the spectrum was basically unused because there was very limited number of electronic devices that could operate in this range of frequencies.

From some years the situation is radically changed. Thanks to research on this range of frequencies we overcome many problems and Millimeter waves are now practical and convenient, and we can find many electronic devices working on this range. Best of all, they take the pressure off the lower frequencies and truly expand wireless communications into the outer limits of radio technology. If we go any higher in frequency, we will be using light.

1.4 Advantages and disadvantages

Millimeter waves offer an huge availability of spectrum. Today, the spectrum from DC through microwave (30 GHz) is really crowded. Government agencies worldwide have allocated all possible spectrum in that range. There are spectrum shortages and conflicts. The expansion of cellular services with 4G technologies like LTE depends on the availability of new spectrum.

For this reason, spectrum it's really expensive. Mmwave frequencies partially cope with this issue because it provides the possibility of expansion.

Millimeter waves also permit very high data throuput. Wireless data throuput in microwave frequencies and below are now limited to about 1 Gbit/s. In the millimeter-wave range, data throuput can reach 10 Gbits/s and more.

The disadvantages is that while this spectrum provide us more spectrum, it isn't useful for all types of wireless applications. It has its limitations. To solve those issue has been the challenge of making this frequency communication practical and convenient. That time has come.

One of the big problem of mmwave frequencies is the limited communication range. We know that the shorter the wavelength this means higher the frequency, the shorter the communication range for a certain level of power.

The free space loss in dB is calculated with:

$$L = 92.4 + 20\log(f) + 20\log(d)$$

d is the line-of-sight (LOS) distance between transmit and receive device

in kilometers, and f is the frequency in gigahertz. For example, the path loss at 10 meters at 60 GHz is:

$$L = 92.4 + 35.6 - 40 = 88 \text{ dB}$$

we can reduce this loss with an improved receiver sensitivity, higher transmit power, and higher antenna gains.

Another contribution to the loss is given by atmosphere, in fact the atmosphere absorbs millimeter waves, reducing more the communication range. Rain, fog in the air provide an important signal attenuation, reducing even more communication distance. Oxygen (O₂) absorption is very high at 60 GHz (**Fig. 1.4**). Water (H₂O) absorption is very relevant in some interval of frequencies providing other peaks of attenuation. We need to choose carefully frequencies within the curve valleys that minimizes the loss. The advantage is that with an high-gain antenna we can push up the effective radiated power (ERP), significantly increasing communication range.

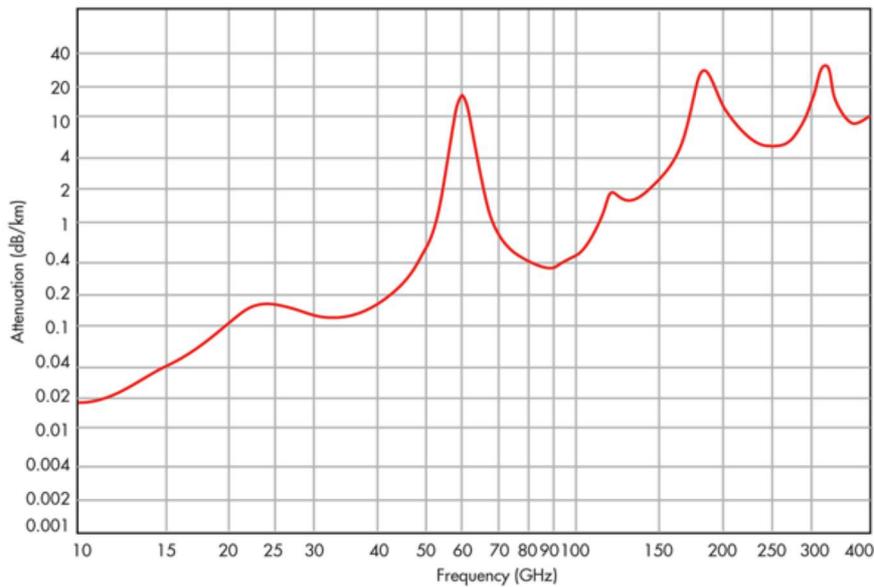


Figure 1.4 Atmospheric absorption as function of the frequency

In the **Figure 1.4** is represented the signal attenuation at sea level and with a temperature of 20°C versus log frequencies, we see that oxygen at 60 GHz and water at the other peaks causes a relevant attenuation of the signal.

In general short range communications can provide some advantage, for example, it minimizes the interferences from other nearby transmitters. The high-gain antennas, this means highly directional, provide another contribution to mitigate interference. In this way using narrow beams we increase power and range as well. From the security point of view it reduces the probability of signal interception.

We know that size of the antenna is related to the working frequency so Small size is another major advantage of millimeter-wave equipment. While ICs keep the circuitry small, the high frequency makes very small antennas necessary and possible. A typical half-wave dipole at a cellular frequency like 900 MHz is six inches long, but at 60 GHz one half-wave is only about 2.5 mm in free space and even less when it's made on a dielectric substrate. This means the entire structure of a radio including the antenna can be very small. It's easy to make multiple-element phased arrays on a substrate chip that can steer and focus the energy for greater gain, power, and range.

Another challenge is making circuitry that works at millimeter-wave frequencies. With semiconductor materials like silicon germanium (SiGe), gallium arsenide (GaAs), indium phosphide (InP), and gallium nitride (GaN) and new processes, though, transistors built at submicron sizes like 40 nm or less that work at these frequencies are possible.

1.5 Applications

The challenge is to cope with the explosion of online video that require the greatest bandwidth and, accordingly, a higher data rate. Speeds of many gigabits per second are needed to transmit 1080p high definition (HD) video, now more and more popular are becoming the Ultra HD video (4K) that it requires even more. The reduction of the data rate can be achieved through video compression techniques before the transmission.

Compression techniques usually reduce the quality to allow available wireless standards like Wi-Fi 802.11n to be used. Standards like 802.11ac that use greater bandwidth in the 5-GHz band are now available to achieve gigabit data rates. Millimeter-wave technologies make gigabit rates commonplace and relatively easy to achieve, making uncompressed video a reality.

Due to its potential for multi-gigabit and low latency wireless links, millimeter wave (mmWave) technology is expected to play a central role in 5th generation (5G) cellular systems.

5G will use small cells. Smaller cells are limited-range basestations that are being deployed to fill in the gaps in macro basestation coverage. With limited range, these small cells will adopt frequency reuse techniques to provide more efficient use of the spectrum available. The small-cell movement, also called heterogeneous networks or HetNets, is becoming the fifth generation (5G) of cellular systems.

Mmwave are perfect candidate for satellite communications because at higher latitude of orbits it operates perfectly with massive data rate and low latency.

Automotive applications is an hot topic; Autonomous driving require detection of passengers or other obstacles; detection and decision are to be taken with very low latency, so mmwave can be used for detection radar in the vehicles.

Another very important application is body scanner, Technology giants like Rohde&Schwarz has recently introduced millimeter wave human body scanner for airport security that it works at frequency range between 70 Ghz and 80 Ghz.

Virtual reality applications are the future of multimedia world. Millimeter waves perfectly fit for virtual reality devices because mmwave can support high bandwidth which is necessary for high definition video and audio transmission.

Common applications include video transmission from a set-top box (STB) to an HDTV set or transmission between a DVD player and the TV set or from a game player to the TV set. Video also can be sent wirelessly from a PC or laptop to a video monitor or docking station.

The thesis is organized in the following way:

- Chapter 2 describe the property of the mmWave channel and the state of the art solutions for beam alignment.
- Chapter 3 introduce briefly the current LTE network architecture and give a detailed description of the mmWave network implemented in the NY mmWave module of ns-3.
- Chapter 4 outlines how to couple ns-3 with the SUMO outputs.
- Chapter 5 describe the simulation scenario and analyze the related results.
- Chapter 6 derive some conclusions and suggests some possible improvements.

Mmwave Comunciations

2.1 Mmwave Channel

In the previous chapter we showed the characteristics of the mmWaves in free space, this means in a infinitely large space that is empty except for the transmitter and the receiver. We saw that another contribution to the attenuation is given by the atmospheric absorption. Fixed point-to-point Microwave are typically designed such that the low order fresnel clearance zone are free of any obstacle.

In this situation LOS behaviour dominates the received signal properties. Moreover the Friis transmission equation don't take into account the potential antenna misalignment (the consequence gain variation) and the polarization mismatch. The Friis transmission equation can be extended to include these problems.

The communication is possible if we have an acceptable SNR, this means that noise and interference generated by the environments and inside the receiver should be considered as well. The beam steering and gain control procedures can compensate within a certain limit the attenuation terms as long as the unwanted signal remain uncorrelated with the transmitted signal.

Unfortunately in practice most of the scenarios include obstacles that are within the fresnel zones, so reflection, diffraction, scattering and penetration must be potentially considered.

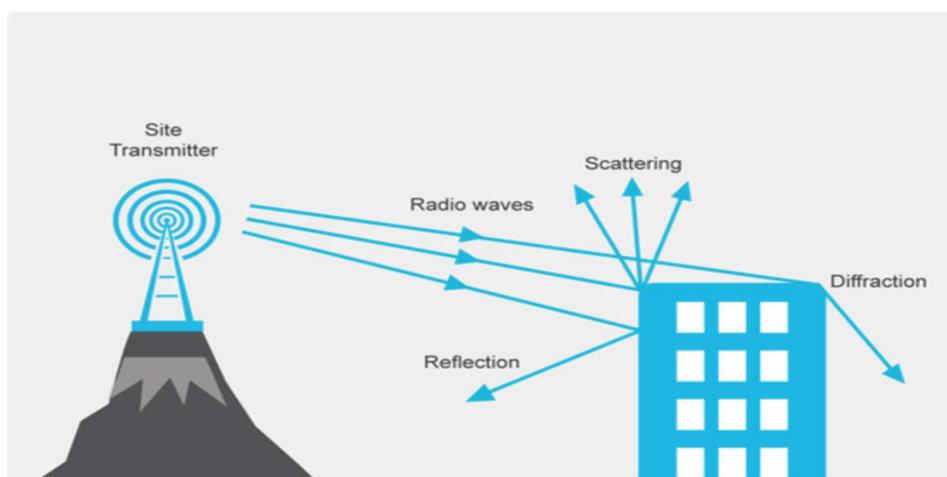


Figure 2.1 Reflection, diffraction, scattering.

First it's possible that the signal reach the receiver end behind the object. Second we can reach the reicever through a multipath propagation. So signal of different paths experience different attenuation, delay, polarization mixing. When different signal from different paths combines together simultaneously, the signal strenght vary based on the receiver location.

The last phenomena is called fading which can be divided in slow or fast fading.

If the receiving end is shadowed by large obstacle (shadowing) we refer as slow fading, so to describe channels in which Channel Coherence Time is greater than the Transmission Symbol Time and the received power distribution is log-normal.

If instead there are many propagation path typically we refer as fast fading, so to describe channels in which Channel Coherence Time is smaller than the Transmission Symbol Time in which the received power present multiple local minima and maxima fluctuating of several decibel over short distances.

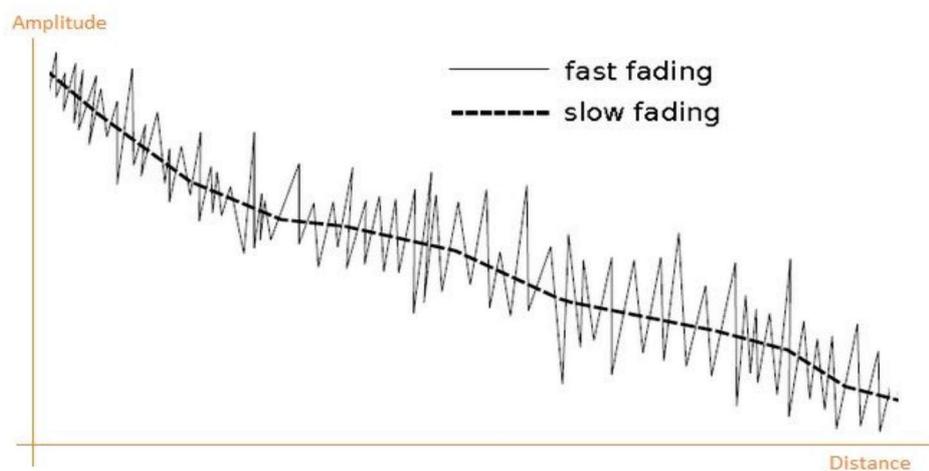


Figure 2.2 fast fading vs slow fading

Furthermore the fading loss may depend on frequency or not, we refer respectively as frequency selective-fading and flat fading. With frequency selective-fading different frequency components of the signal are affected independently, so it shapes the received signal spectrum.

Moreover multi-path propagation causes ISI (intersymbol interference) that is a distortion of the signal in which one symbol interferes with the subsequent symbols. ISI can only be reduced by limiting the bandwidth.

To reduce the impact of this unwanted behavior, wideband systems typically distribute signal to multiple carriers with smaller bandwidth.

By the way the multi-path propagation has not only negative effect. MIMO systems are able to multiply the capacity of a radio link using multiple transmission and receiving antennas exploiting multipath propagation [2].

The radio channel behaves quite differently at millimeter waves compared to conventional frequencies as highlighted in Figures 2.3 and 2.4. The figures illustrate the simplified impulse response of the radio channel in the delay-angle domain, normalized to the LOS component. The figures are drawn for the same exact deployment scenario assuming a LOS

path and a number of non-line-of-sight (NLOS) paths due to objects in the surrounding environment.

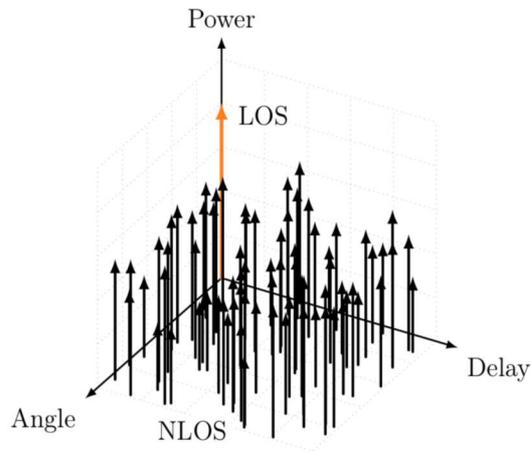


Figure 2.3 Delay-angle distribution at low frequencies.

The Figure 2.3 shows multipath fading at lower frequencies. The situation corresponds to Rice distributed channel, meaning there are a LOS component and a large number of multipath components distributed more or less evenly in the entire delay-angle plane. This is not the case at mmW frequencies, as is evident from Figure 2.4. While there are still both LOS and NLOS components present, the delay-angular spread is completely different. This time the LOS component is clearly dominant, and the multipath components are grouped in both delay and angle domains [3],[4],[5]. Referred to as clusters and marked with shaded circles of different colors, each group consists of a small number of impulses of similar power, delay and angle of arrival. Only azimuth angle is showed in figure 2.3 and 2.4. At lower frequencies we can describe the situation as two dimensional as required gain is obtained with a omnidirectional antennas.

Instead at mmwave frequencies we usually consider a three dimensional space as antenna is directive both in azimuth and elevation planes. The situation in the elevation plane is similar to that in the azimuth plane but the angular spread is smaller and the NLOS components tend to concentrate on elevation angles above the LOS component [6].

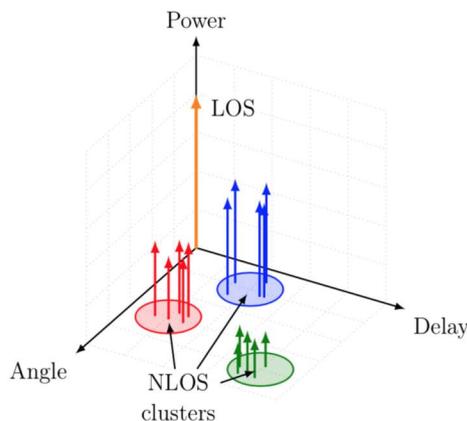


Figure 2.4 Delay-angle distribution at millimeter waves.

Typically the NLOS component is 20dB weaker than the best LOS component,so typically the range of the NLOS components is fairly limited.

The range of the NLOS paths cannot be increased only by power control for different reasons.Mainly the required range for power control is very wide and there are regulations that constraints the maximum trasmitted power.

The LOS component is dominat for different reasons.

First the NLOS components are not direct but are reflected components due to the presence of objects in the sorrounding enviroment.So the trasmit power is not concentrated in the direction of the obstacle since the antennas focus the radiated energy in the main lobe.The power of the indirect components drops very quickly even with small angular offset. The coverage of the indirect components can be improved even with a smaller gain antenna.

Second the higher attenuation of the NLOS components is not only due to the longer distance but also the interaction with the object should be taken into account.In particular the penetration loss is higher due to the higher penetration depth at these frequencies[7][8].

Respect to the lower frequencies at the mmW frequencies the multipath components are not only weaker but also are very reduced in number.Since the components are not approximately uniformly distributed but are clustered the enviroment is considered sparse[9][10][5].

As the distance from the trasmitter and the antenna directivity increase,the less evident is the clustering behaviour(signal less spread over angle-delay domain)[7][6][3].

This clustering behaviour has different consequences.

First the CCI(co-channel interference)is very reduced due to high attenuation,devices operating at the same frequencies not easily interfere each other.Also controlling trasmit power can adress the reduction of the CCI[3][5].

Second since the comunication between tramitter and receiver is possible through a limited number of viable paths,the mmwave channel is denoted as low rank.A limited number of parallel data stream is supported, and low-order spatial multiplexing, i.e., MIMO processing is possible. [9][5]

Third while channel is time-variant, the paths are reciprocal[9].

Moreover, the mmW channel maintains the polarization of the transmitted signal over LOS links. Some polarization mixing is evident in NLOS schemes due to interaction with the environment . For instance, the handedness of circularly polarized signals is changed upon reflection [11]. Furthermore, degradation due to polarization mismatch of up to 20 dB can be expected, suggesting for polarization diversity.

Last but not least the need for beam alignment since there are only a handful of different directions where the signal quality is good enough for a reliable link[12][6].

The high directional antenna and the potential presence of obstacle between trasmitter and receiver pose two problems:deafness and blockage;

Deafness refer the situation in which the main beam of trasmitter and of the receiver are not aligned each other,this prevent the establishment of a high quality mmwave link.

Blockage due to obstacles can cause a severe signal attenuation,even the human body can reduce the signal strength by 20 dB. Thus, an unblocked Line of Sight (LOS) link is highly desirable for mmWave systems.This very high attenuation cannot be compensated by just increasing few dB of trasmission power or using narrower beams.

A mmWave link may have three kinds of blockages, namely, static, dynamic, and self-blockage. Static blockage due to buildings and permanent structures may cause permanent blockage of the LOS link. The dynamic blockage due to mobile humans and vehicles (collectively called mobile blockers) which may cause frequent interruptions to the LOS link. the self-blockage due to a user's body plays a key role in mmWave performance.

Different models has been proposed to study the impact of the three type of blockage[30]. Solving blockage requires a search for alternative directed spatial channels that are not blocked;

2.2 Beamforming

Beamforming is the key technique to compensate the severe channel attenuation and to reduce interference in mmWave networks.

Instead of having one large element, an antenna may also be made more directive using multiple smaller elements placed close to each other.

Feeding these elements simultaneously constitutes an antenna array. Arrays are preferred for their steerability and other benefits, achieved at an acceptable cost of increased feed network complexity. In practice, most arrays are uniform, i.e., comprise similar equispaced elements for easier analysis and synthesis.

Antenna arrays operate based on interference, and steerability is achieved by controlling the directions in which waves interfere constructively and destructively.

A versatile pattern generation is possible if both the phase and amplitude of each individual element are controlled. In addition to being able to control the beamwidth and sidelobes levels, we may use multiple beams and place nulls toward multiple interferers. Still, uniform weighting is favored for its lower complexity and limited power consumption, and such systems are commonly referred to as phased antenna arrays.

Moreover, the phase control is often implemented using a number of predetermined, enumerated values taken from a codebook.

A codebook is table that maps a number of predetermined beams to sets of antenna excitation values. A single set determines the weights for each element corresponding to a certain beam. Such a collection of weights is commonly referred to as a beamforming vector.

Different beamforming approaches may be divided into three categories: analog beamforming, digital beamforming and hybrid beamforming.[13]

Analog beamforming: In these cases, beamforming is performed in radio frequency (RF) through a bank of phase shifters (PSs) – one per antenna element. This architecture reduces the power consumption by using only one pair of analog to digital converters (ADC) and digital to analog converters (DAC) at the Rx and Tx, respectively, per digital stream.

This technique shapes the output beam with only one RF chain using phase shifters, On the positive side, a simple beam-searching procedure can be used here to efficiently find the optimal beams at the transmitter and the receiver, as already established in existing mmWave WPAN and WLAN standards.

With finite size codebooks each covering a certain direction, those standards recommend an exhaustive search over all possible combinations of the transmission and reception directions through a sequence of pilot transmissions. The combination of vectors that maximizes the signal-to-noise ratio is selected for the beamforming.

This procedure alleviates the need for instantaneous CSI, at the expense of a new *alignment-throughput tradeoff*. The tradeoff shows that excessively increasing the codebook size (or equivalently using extremely narrow beams) is not beneficial in general due to the increased alignment overhead, and there is an optimal codebook size (optimal beamwidth) at which the tradeoff is optimized. On the negative side, one RF chain can form only one beam at a time without being able to multiplex within the beam, implying that this architecture provides only directivity gain. For narrow beam operation, pure analog beamforming requires several RF chains to serve UEs that are separated geographically. This diminishes the advantages of this architecture such as low complexity and low power consumption.

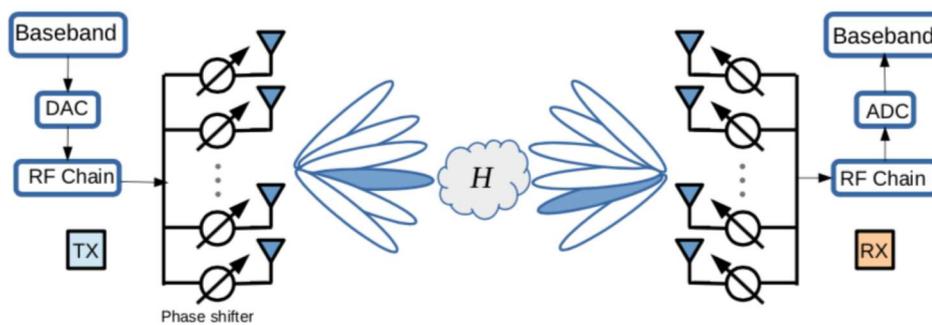


Figure 2.5 Analog beamforming architecture

Digital beamforming: This architecture provides the highest flexibility in shaping the transmitted beam(s), however it requires one baseband-to-RF chain (in short RF chain) per antenna element.

fully digital architectures, shown in Fig. 2.6, beamforming is performed in baseband. Each RF chain has a pair of ADCs at the Rx and DACs at the Tx enabling the transceiver to simultaneously direct beams in theoretically infinite directions at a given time. But, for wide-band systems high precision ADCs and DACs can be very power hungry. To be energy efficient, fully digital beamformers need to use converters with one or few bits of resolution.

This increases the cost and complexity due to the large number of antenna elements operating in very wide bandwidth.

Considering one high resolution analog-digital converter (ADC) per RF chain, digital beamforming also leads to high power consumption both at the BS and at the UEs, which is at odds with the design goals of 5G.

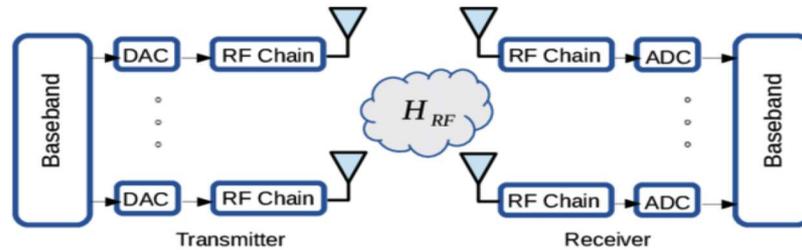


Figure 2.6 Digital beamforming architecture

Moreover, digital beamforming requires estimation of the channel between every pair of antenna elements of the transmitter and the receiver. Apart from a more complicated precoding, the complexity of this estimation scales at least linearly with the number of transmitter antenna elements.

In time division duplexing (TDD) systems, channel state information (CSI) at the transmitter can be obtained using uplink sounding signals. The advantage is that the overhead will be

scaled with the combined number of UEs' antennas that can be much less than the number of BS antennas. However, the limited UE power and the possible lack of beamforming gains for the uplink reference signals may limit the performance of the network. Also, CSI acquisition by uplink reference signals requires the principle of channel reciprocity that holds if the duplexing time is much shorter than the coherence time of the channel. The coherence time in mmWave bands is around an order of magnitude lower than that of microwave bands, as the Doppler shift scales linearly with frequency. Therefore, TDD at mmWave bands needs to be restricted to low-mobility scenarios. In frequency division duplexing (FDD) systems, CSI estimation should be done in both uplink and downlink directions due to the lack of reciprocity. While CSI estimation overhead in the uplink is similar to the TDD case, the overhead in the downlink channel scales with the number of BS antennas, which becomes infeasible as the number of BS antennas grows large. Altogether, for systems operating in very wide spectrum ranges, such as several hundreds of MHz, and employing a large number of antennas, a complete digital beamforming solution using the current requirements (one high resolution ADC per RF chain and channel estimation per antenna element), is hardly feasible and economical. Low-resolution ADCs (ideally with only one bit) and sparse channel estimation are promising solutions for enabling digital beamforming in mmWave systems.

Hybrid beamforming: A promising architecture for mmWave cellular networks is a two-stage hybrid digital-analog beamforming procedure, allowing the use of a very large number of antennas with a limited number of RF chains. With the hybrid solution, digital precoding is applied for the effective channel consisting of the analog beamforming weights and the actual channel matrix.

Analog beamforming provides spatial division and directivity gains, which can be used to compensate the severe channel attenuation, by directing the transmitted signal toward different sectors. Furthermore, digital beamforming may be used to reduce intra-sector

interference and provide multiplexing gain using CSI of an effective channel with much smaller dimension. Exploiting the sparse-scattering nature of mmWave channels, the complexity of hybrid beamforming design can be further reduced

The analysis of shows that, in a single user MIMO system, hybrid beamforming can almost achieve the throughput performance of a fully digital beamforming with 8 to 16 times fewer RF chains, leading to greatly reduced energy consumption and processing overhead with a negligible performance drop. However, analysis and optimization of the the tradeoff between the number of employed RF chains and the achievable network throughput in multiuser MIMO system and in the presence of CSI errors in wideband mmWave systems requires further research.

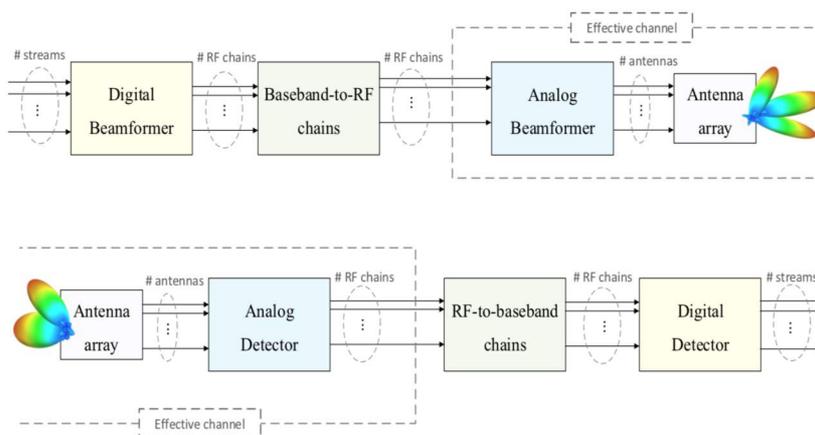


Figure 2.7 Hybrid beamforming architecture

2.2.1 Beam alignment

The finding of optimal beam direction may take long time to complete. As the channel conditions are dynamic, the direction finding may need to be performed constantly before transmissions, which would significantly compromise the transmission capacity due to the high overhead. Therefore, it is in great need to have efficient beam searching schemes or beam alignment methodologies that can be incorporated into the MAC protocol design in mmW networks[13].

Efficient beam alignment is a crucial component in millimeter wave systems.

The design space of beam search proposals in the literature can be divided into three main categories: (1) sequential scanning strategies; (2) adaptive algorithms employing antenna patterns with configurable beamwidth; (3) parallel beam search with simultaneous, multi-direction scanning.

Within the state-of-the-art solutions on this subject, a further subdivision can be made on the basis of the employed mm-wave transceivers. At present analog beamforming (ABF) and hybrid analog-digital beamforming (HBF) represent the more feasible solutions because of

cost and power consumption constraints at mmwave frequencies, even if promising results have been found using digital beamforming (DBF) operating at a low quantization resolution[14].

We present different architectures and protocols within the state-of-the-art solutions that try to find the best beam pair and try to overcome blockage. Different solutions exploit different different strategies with the goal of reducing the beamforming latency.

Exhaustive approach

The naive approach to find the best alignment would have the transmitter and receiver scan the 3D space with their beams to find the direction of maximum power. The receiver has to repeat the scan for each choice of beam direction on the transmitter side. Thus, the complexity of this exhaustive search is $O(N^2)$, where N is the number of possible beam directions.

IEEE 802.11ad

To speed up the search, the 802.11ad standard decouples the steering at the transmitter and receiver[15].

The current scheme for the IEEE 802.11ad requires that the transmitting antenna scan the whole space exhaustively to find the best beam steering direction that will generate the largest receiver SNR through the ACK frame from the receiver.

The IEEE 802.11ad discretizes the search space for beamforming by dividing the antenna azimuth into *virtual sectors*. Beamforming procedure defined by IEEE 802.11ad can take place in two phases. First, coarse-grain antenna sector selection needs to be performed by sector-level sweep (SLS). Each BS transmits probe frames at different sectors with unique identifiers. UE receives these frames from both BSs with a quasi-omni-directional antenna pattern, selecting the BS and its coarse-grained BS transmit sector which has higher SNR, and sending the best transmit sector ID to that BS. The same procedure is inverted to find a list of transmit sectors of UE. This procedure generates high overhead since each probe frame is transmitted from each sector at the lowest PHY rate (MCS 0), which may take up to several seconds (so it can seriously degrade the system performance)[15].

Given an antenna sector, beam refinement phase (BRP) follows. During BRP, the coarse-grained transmit sectors founded in SLS phase is refined. The optimum transmit and receive sectors are determined to optimize the data rate between BS and UE. In fact, BRP is faster than SLS. The frame exchanges and BRP can rely on the directional link built during the SLS phase.

This approach reduces the search complexity to $O(N)$. Still, for a beam of a few degrees, the delay can be hundreds of milliseconds to seconds, which would easily stall realtime applications.

A legacy 802.11ad link transits between 3 states: *Norm*, *Outage*, and *Scan*. An 802.11ad link responds to the blockage event by triggering beam-scanning that searches for the best alternative beam pair. However, whether to use an aggressive or conservative threshold remains an open problem, due to a tradeoff between overhead and responsiveness[15].

For wide adoption, the latency problem in beam alignment at mmWave frequencies must be addressed.

Inferbeam system and protocol

InferBeam protocol can drastically reduce the setup cost for beam alignment for a new environment, and also the latency in acquiring a new beam under intermittent blockage.

InferBeam for inferring millimeter-wave BS and aligning beams is based on conditional random field (CRF), a class of statistical modeling method for structured inference in machine learning. CRF is applied popularly to computer vision tasks such as image segmentation and object recognition. The key idea is to train CRFs to segment a wireless environment modeled as a 3D grid such that via CRF inference every discrete point within the environment is associated with its optimal antenna and sector[16].

InferBeam protocol focuses on minimizing the required samples in inferring best beams.

Example: we assume that the optimal antenna and sector selections are known for the two samples, i.e. UE1 and UE2. We further assume that UE3 is located between UE1 and UE2, the distance between UE1 and UE3 is 0.5m, while the distance between UE2 and UE3 is 0.7m. To infer the antenna and sector selection for UE3, we may take the information from both samples into consideration. Moreover, we need to find the spatial correlation on BS and sector selection between the sample points and the point we want to infer. We use CRF to construct an undirected graphical model that captures a wireless environment. This motivates the InferBeam, which uses CRF to infer optimal BS selection and antenna sector decision, using only small number of samples from the environment's 3D grid. The results indicate that the system can make best beam selection for 98% of locations in test environments, while sampling fewer than 1% of locations[16].

InferBeam has the following advantages:

1. InferBeam enables quick beamforming by skipping the SLS phase.
2. InferBeam enables fast beam adjustment as environment changing (i.e. blockage).

The InferBeam system consists of two cascaded CRFs, CRF1 and CRF2. CRF1 is dedicated for BS selections while CRF2 will be used to select transmit sector.

The algorithm (*Training*) outputs the model parameters for CRF1 and CRF2.

The results show that InferBeam can greatly reduce the latency for beam adjustment.

Agile-link approach

A new protocol that can find the best mmWave beam alignment without scanning the space. Given all possible directions for setting the antenna beam, Agile-Link provably finds the optimal direction in logarithmic number of measurements.

In particular, for highly directional mmWave devices operating under 802.11ad, the delay drops from over a second to 2.5 ms.

it can deliver the best alignment in $O(K \cdot \log N)$ measurements, where K is the number of paths traveled by the signal. Since K is typically 2 or 3 paths, Agile-Link significantly reduces the beam alignment delay[17].

At a high-level, it works as follows: Instead of creating a narrow beam and sampling the power along one spatial direction each time, Agile-Link manipulates the phase shifters to create multi-armed beams, which can sample multiple spatial directions simultaneously (see Fig 2.8(a)). Since a multi-armed beam combines the power along multiple directions, the receiver cannot immediately tell which direction has produced the resulting power. Agile-Link however uses a combination of randomized multi-armed beams, which together provide enough information to identify the signal power along all spatial directions.

Agile-Link works in two stages. First, it randomly hashes the space into bins (using multi-armed beams) such that each bin collects power from a range of directions. Second, it uses a voting mechanism to recover the directions that have the power.

Example: Say for example, that there are 16 possible directions in space, i.e., $N = 16$. Agile-Link can sample all of these directions using 4 multi-armed beams, each covering $N/4 = 4$ directions in space. Fig. 2.8(a) shows four such multi-armed beams, Fig. 2.8(b) shows how together they cover the whole space of directions.

Such set of multi-armed beams operates like a hash function, where $N = 16$ directions are hashed into 4 bins, and each bin covers $N/4 = 4$ distinct directions. The value of the bin represents the combination of the signals that hash into it. For example, if the signal is coming along the 60° direction and 60° hashes to bin number 1, then only bin one will have energy whereas the other bins will have no energy, Thus, one can ignore directions that hash to bins 2,3, and 4, and focus only on directions that hash to bin 1. This significantly reduces the search space to the directions that hash to the first bin.

At this stage, we know that the signal could have come from the directions covered by the first bin i.e., 0° , 60° , 90° and 120° . But we do not know which among them is the correct direction. Thus, we change the hash function and try again. To do so, we use a second set of multi-armed beams which together hash the whole space of directions into a set of bins. The hash however is randomized with respect to the previous hash so that directions that got hashed together are unlikely to hash together again.

The first bin now collects energy along 30° , 80° , 110° and 140° . Since the signal is arriving along 60° , it will be captured by the third bin which is represented in blue in Fig. 2.8(f).

Hence, in this second hashing, only the energy of the third bin will be large. This suggests that the signal arrived along one of the directions that mapped to the third bin which in this case are 40° , 60° , 105° and 150° . Since the 60° direction is the only common candidate from both the first hashing and the second hashing, Agile-Link picks it as the direction of the signal[17].

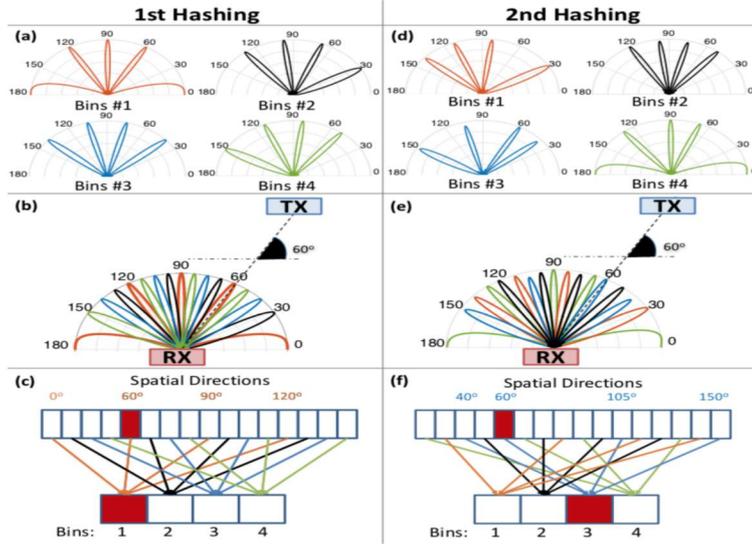


Figure 2.8 Illustrative example of Agile-Link's algorithm

Agile-Link has additional important features. First, Agile-Link is compatible with the 802.11ad protocol, i.e., a Agile-Link device can work with a non-Agile-Link device to find the best alignment while using the 802.11ad protocol.

Beam alignment Using Contextual Bandits

Given the outcome of the past beam alignments, it is possible to extract some information and reduce the search space for the subsequent beam alignment procedures.

in particular it is based on the fact that successive beam alignments are stochastically correlated, and thus, outcome of the previous “beam matching” provides *contextual information* for the subsequent matchings, thus eliminating the need to search the entire angular domain. We exploit *correlation* and *unimodality* properties across various beam matching[18].

Specifically, for a given beam matching, we call the difference between the transmitter and receiver direction as *misalignment*. Because of correlation if matching at a larger misalignment is successful (i.e., received energy is above a threshold τ), with a high probability a matching will be successful at a smaller misalignment as well. Furthermore, the directivity gain (or received energy) can be approximated as a unimodal function of the misalignment value.

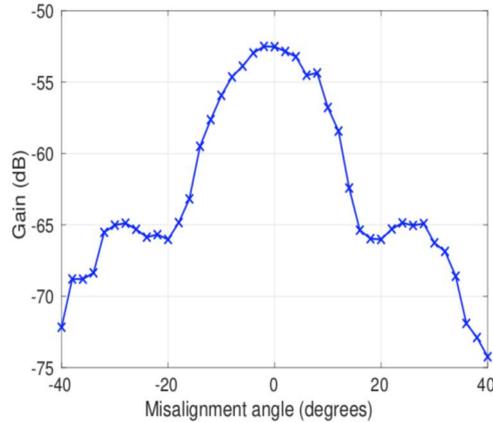


Figure 2.9 Received power at distance of 1 meter vs. misalignment.

We exploit this contextual information in order to obtain a beam search scheme that quickly identifies the best beam direction and maximizes the directivity gain.

We formulate the problem of finding the best beam pair as an online stochastic optimization where the objective is to maximize the expected amount of received energy within a given time period.

It is shown that this problem can be considered as an instance of the Multi Armed Bandit (MAB) model in which each transmit and receive beam pair is considered as a single arm. Thus, the objective is to design a sequential arm selection (or, equivalently, beam alignment) strategy that maximizes the *expected reward* (received energy) over a given time horizon. Performance of MAB models is usually expressed in terms of *regret* that is defined as the total expected reward loss compared with an oracle policy that would know everything.[18]

Example: Let us consider a scenario where the transmitter beam direction is fixed at 75° angle with respect to the receiver. For the sake of exposition, we assume a 2D setting. Assume that there are 16 possible directions at the receiver, as shown in Fig. 2.6. Using the exhaustive beam selection scheme, each of 16 directions will be examined one at a time, and the direction with the largest received energy (from beacon messages) is picked. However, under dynamic conditions (e.g., with mobility), the optimal beam direction can potentially change within a short period of time. In this case, we consider maximizing the received energy within a given period of time. Using our proposed scheme, the receiver assigns an index to each beam direction, and the beam with the highest index will be selected. The important point is that due to the correlation and unimodality properties, the search space will be limited to the neighborhood of the beam with maximum index. As a result, it prevents the need of a uniform exploration over the entire angular domain, thus mitigating the overhead of beam alignment when the number of beam directions becomes large.

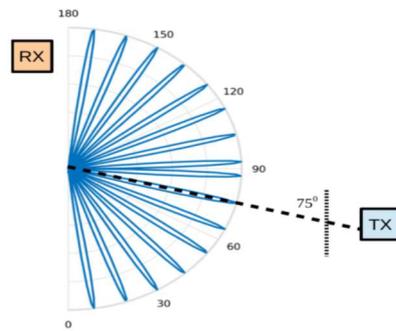


Figure 2.10 Beam alignment of the transmitter and receiver with 16 beams.

They derive a lower bound on the regret of any search-based algorithm, and demonstrate that the regret does not scale with the number of transmission and receive beams. This is a crucial property in MIMO settings in which the number of all combinations of transmit and receive beams grows quickly.

They propose Unimodal Beam Alignment (UBA) algorithm that is shown to be asymptotically optimal and in which the transmitter and receiver are able to refine the search space through successive rounds of beam alignment[18].

Beam spy

We present four measurement observations of 60 GHz channels and phased-array beamforming and these observations constitute the foundation of Beam-Spy's prediction framework. The measurement is conducted using a custom-built 60 GHz software radio platform[19].

Channel sparsity: 60 GHz channels are extremely sparse. The spatial channel response is dominated by a few paths from a few angular directions.

Spatial correlation: Given a 60 GHz phased-array with multiple beam directions, blockage of one beam affects the performance of other beams.

Blockage invariant correlation: The statistical correlation between different beams is invariant to human blockage.

Human blockage does not create additional significant angular clusters.

When a link is deployed, BeamSpy leverages full-beam scanning (such as in 802.11ad) to construct a novel path skeleton model, and extrapolate the blockage-invariant spatial correlation between different beams available on the Tx/Rx's phased-arrays.

Afterwards, whenever beam quality changes due to blockage (indicated by SNR drop), BeamSpy can predict the quality of all other beams by simply measuring the Channel Impulse Response (CIR) of the beam in use.

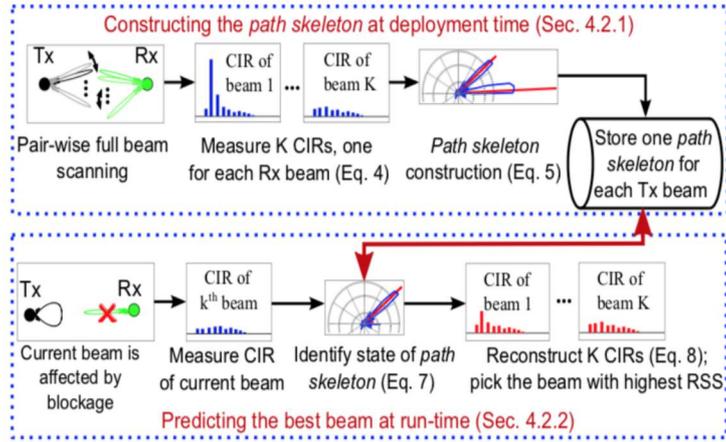


Figure 2.11 BeamSpy’s prediction framework.

For ease of exposition, we focus on a 1-D uniformly spaced antenna array, which has N antenna elements and can generate K beam directions/patterns in total.

$\mathbf{w}(n,k)$ is the beamforming weight applied to the n th element when generating beam k .

Each beam corresponds to one antenna gain pattern, with certain spatial directions amplified whereas others weakened.

For sake of simplicity, suppose the transmitter is omnidirectional and receiver is a directional phased-array. When the receiver steers towards the k th beam, its Channel Impulse Response (CIR), h_k , is a joint effect of the array-factor (gain at spatial direction θ) and the CIR of each antenna element.

Intuitively, the CIR h_k captures the aggregated effect of all paths that arrive at each of the antenna elements, appropriately weighted by $\mathbf{w}(n,k)$ and summed together.

In practice, not only the receiver, but also the transmitter’s phased-array antenna will reshape the channel response, creating directionality effect along different spatial directions.

They formally define a *path skeleton* as the sparse set of dominating paths that can be used to approximate the spatial channel between a 60 GHz transmitter and receiver. Note that, the *path skeleton* only depends on the channel and is independent of the beamforming weights at Tx/Rx[19].

The challenge is that The Tx/Rx can only measure the CIR when using a given beam, and cannot discriminate the channel blockage along each specific path.

But intuitively, since the channel is sparse, the very few number of dominating paths form a *path skeleton* that determines the performance of all receive beams. The core idea of BeamSpy is to “reverse-engineer” the *path skeleton* between the Tx and Rx and, when blockage occurs, estimate the blocked paths within the skeleton, and then predict the CIR of unobserved beams based on their known beamforming weights.

Specifically, given the measured CIR of current beam, BeamSpy estimates which paths within the *path skeleton* are affected by blockage. It then immediately predicts the quality of all alternative beams, based on the *a priori path skeleton* which captures the invariant spatial correlation between beams.

To identify the affected paths, BeamSpy makes the following approximation, inspired by observation human blockage annihilates existing skeleton paths, but does not create new paths.

Given the estimated the blocked/non-blocked states of all the skeleton paths , and the pre-blockage *path skeleton*, BeamSpy can reconstruct the CIR of any unobserved beam. The reconstructed CIR can be straightforwardly converted to link quality metric, like RSS or effective SNR, based on which BeamSpy can identify the best beam.

BeamSpy *does not guarantee a blocked link can be recovered via beam switching* – such feasibility solely depends on Tx/Rx placement and environmental reflectivity. However, BeamSpy can predict how likely a link deployment is to fail completely when blockage occurs.

They call this scheme *risk assessment*. Risk assessment is critical when a 60 GHz link is deployed as a fixture, *e.g.*, from ceiling/wall to a furniture in home.

BeamSpy predicts quality of all beams under each possible blockage pattern, and computes deployment risk κ as fraction of cases where no beam can sustain the minimum bit rate.

BeamSpy acts as a meta-protocol to augment 802.11ad so as to quickly recover from link outage without the high-overhead beam searching.

It achieves comparable throughput performance with an oracle that knows exact beam quality, and outperforms 802.11ad significantly in application tests[19].

UbiG

UbiG implement a mmWave wireless access network that can deliver ubiquitous gigabits per second wireless access consistently to the commercial-off-the-shelf IEEE 802.11ad devices.

It can provide latency guarantees, even using today's cheap COTS hardwares. First, it introduce an efficient beam alignment algorithm that can identify the best beam for a mmWave link with guaranteed latency that is independent of the size of the phase-array antennas. Second, it devise an infrastructure-side predictive AP switching solution to guarantee consistent Gbps connectivity under device mobility and blockage. UbiG needs to implement the algorithms on standard-compliant COTS mmWave hardwares (low hardware cost)[20].

UbiG's fast beam alignment algorithm is built upon a simple observation — while the channel gain and phase change with steering different beam directions, the underlying physical paths along which the signals travel from transmitter to receiver remain the same.

As long as a transmitter can extract the properties (complex gain and directionality) of the paths, it can align the best beam without scanning through the entire space.

UbiG judiciously selects a fixed number of beams to probe and then employs a space-time analysis on the channel measurement of the beams to extract the properties of all dominating paths. The number of probings under this approach is independent of the size of the phased-array antenna and thus the number of beam directions N .

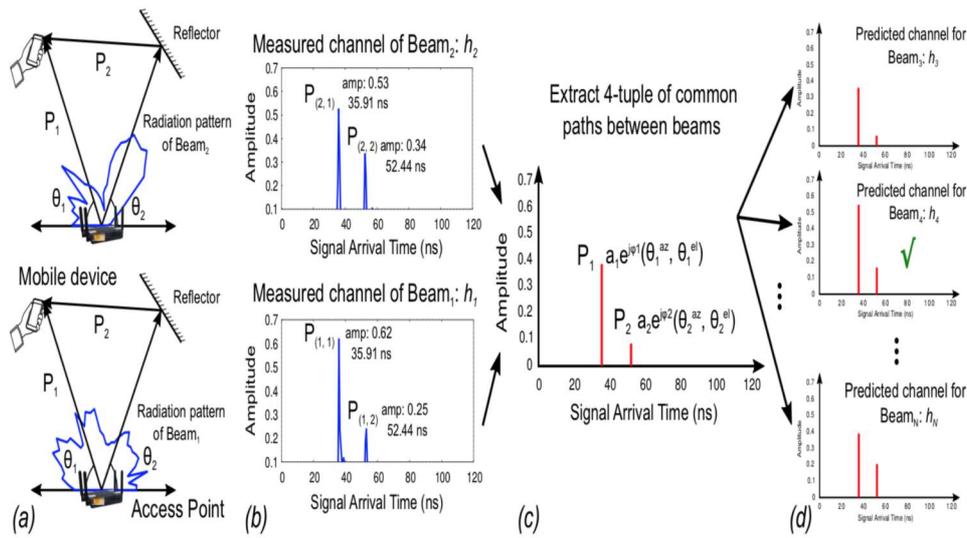


Figure 2.12 Intuition behind UbiG's fast beam alignment

Let's see the Intuition behind UbiG's fast beam alignment:

(a) Two beams with different radiation patterns share the same physical paths to reach the mobile device. (b) Measured channel responses (h_1 and h_2) show the arrival times of the two strong paths under both beams. While the arrival times for both the paths remain exactly same under both beams, the gain of the paths change due to different radiation patterns. (c) Extracting 4-tuple of common paths between the two beams. (d) Mapping the 4-tuple of the paths back to channels of N beams to predict the best (AP can recreate the channels of the rest of the beams since the radiation pattern of each beam is known).

The key goal of UbiG's beam alignment algorithm is to estimate the properties of the K dominant paths using a fixed number of channel measurements.

Efficient beam alignment between a single AP and the user device, however, does not guarantee ubiquitous mmWave connectivity. Rampant blockage from several obstacles, including user's own body in close proximity to the device, may potentially block all the available beams. To enable seamless mmWave connectivity, UbiG leverages cooperation between multiple nearby APs and switch to the unblocked one before the link suffers from a catastrophic outage.

These APs are tightly coordinated through a central controller and can share a high-speed Ethernet or fixed-beam mmWave backhaul [20]. Such multi-AP architecture has been widely deployed in enterprise Wi-Fi networks.

UbiG's AP ranking algorithm allows the controller to keep track of a small set of "good" APs, and quickly switch to the best one before the current AP fails.

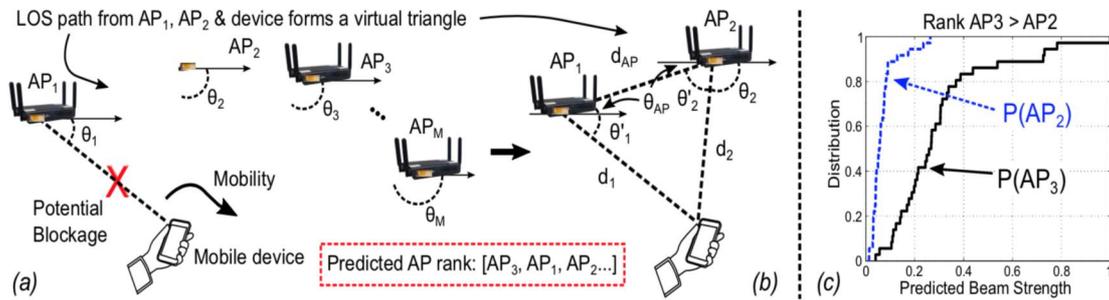


Figure 2.13 (a) UbiG ranks group of potential APs based on the geometrical extrapolation of the dominating LOS path. (b) Extrapolating gain and directionality of LOS path from AP2 using probe packets from only AP1. (c) Distribution of predicted beam strength of AP2 and AP3.

(a) UbiG ranks group of potential APs based on the geometrical extrapolation of the dominating LOS path. (b) Extrapolating gain and directionality of LOS path from AP2 using probe packets from only AP1. (c) Distribution of predicted beam strength of AP2 and AP3.

Suppose that a user device currently uses AP1 for data transmission, but has $M - 1$ alternative APs in its vicinity. The UbiG controller can rank the $M - 1$ APs without additional probing based on two key observations:

- (1) When the LOS between the AP and user device is open, all the beams' performance mostly depends on it. Thus, by simply estimating the gain and direction of the dominant path, UbiG can predict the potential link performance of an AP.
- (2) The LOS dominating paths between two APs and the device form a virtual triangle in space (Figure 2.13(b)). This allows the controller to geometrically estimate the gain and direction of the dominating LOS path from the second AP without probing it.

First, it leverages the channel measurements from the fast beam alignment to estimate the gain g_1 and direction θ_1 of the LOS path from the current AP1 to the user device. Then, g_1 and θ_1 information is further sent to the controller that uses it to extrapolate the LOS path from the other $M - 1$ APs to the user device without probing[20].

The experimental results show that UbiG performs close to an "Oracle" solution which instantaneously knows the best AP-user pair and the strongest communication beam. The algorithms in UbiG can be the key building blocks for the next-generation 5G networks consisting of dense mmWave deployments.

UbiG's AP ranking and fast switching, unfortunately, may not guarantee robust connectivity under certain cases where all APs in a device's vicinity are blocked [20].

INVERSE FINGERPRINT BEAM ALIGNMENT for a vehicle-to-infrastructure (V2I) setting.

The main idea of the proposed approach is to leverage prior knowledge to identify promising beam directions and only train those directions. The prior knowledge is obtained from past observations in the database, and some of these paths may not exist in the current channel due to blockage by a truck for example. Therefore, beam training among the beam directions identified from the database is still required.

In general, a fingerprint refers to some characteristics of the channel at a given location.

In this work, a fingerprint refers to a set of received powers of different pairs of transmit and receive beams at a given location.

In particular two type of fingerprints are proposed, which differ by how measurement data are collected and stored.

The first type requires that the contributing vehicle perform a full exhaustive beam measurements over all beam pairs. This ensures that the measurements over the different pairs happen within a channel coherence time so that the spatial channel does not change. This way, the fingerprint captures the correlation between the different beam pairs, i.e., whether they tend to have similar received power or not[21].

TABLE I

AN EXAMPLE OF TYPE A FINGERPRINTS. FOR EACH CONTRIBUTING VEHICLES, THE MEASUREMENTS OF THE TOP- M BEAM PAIRS ARE STORED. IN EACH CELL, THE INTEGER AT THE TOP IS THE BEAM PAIR INDEX AND THE NUMBER AT BOTTOM IS THE RECEIVED POWER.

Observation No.	Best	2nd best	...	M -th best
1	5 -64.5 dBm	159 -69.2 dBm	...	346 -95.8 dBm
2	159 -70.4 dBm	263 -72.6 dBm	...	354 -97.1 dBm
...
N	5 -66.4 dBm	258 -68.1 dBm	...	2 -82.6 dBm

TABLE II

AN EXAMPLE OF TYPE B FINGERPRINTS AT A LOCATION BIN. THE AVERAGE RECEIVED POWER FOR EACH BEAM PAIR IS RECORDED.

Beam pair index	1	2	...
Average received power	-92.3 dBm	-73.5 dBm	...

Figure 2.14 Two fingerprint types

If memory storage is of concern, only the measurements of the top- M beam pairs (ranked by the received power) can be stored. This makes sense because most of the beam pairs have negligible received power (since their main beam directions do not point along any propagation paths).

The second type of fingerprint do not require that the measurement of all the beam pair combinations be completed within a channel coherence time. This less restrictive data collection reduces the burden on individual vehicles contributing to building the database.

A simple method would be to do a round-robin over the beam pair combinations. For example, assume that each vehicle can do only 1/4 of the full exhaustive search. The roadside unit (RSU) then divides the set of all beam pairs into four disjoint sets and assign these sets sequentially to subsequent contributing vehicles to collect the measurement data. The disadvantage is that now the correlation between the beam pairs cannot be easily captured. Only the average received powers (the average is computed in linear scale) of the beam pairs are stored[21].

The procedure is initiated by a communicating vehicle (CV) by transmitting a beam training request using a low-frequency system such as DSRC or mmWave communication with a large spreading factor. Within this request packet, the CV also includes its current position, which is available from some localization sensors equipped on the vehicle (e.g., GPS or other more advanced positioning method based on LIDAR and 3D map). Upon receiving the request, the RSU will use the position information to query its database to obtain the fingerprint associated with that location. Using the fingerprint, the RSU determines the candidate beam pairs that are likely to provide a satisfactory link connection. The RSU then responds back with an acknowledgment to allow the beam training and also provides a list of candidate beam pairs.

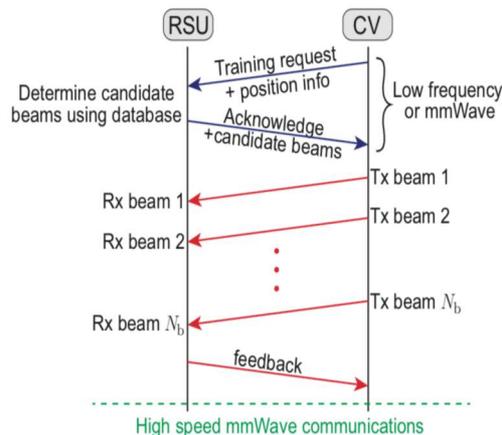


Figure 2.15 Timing diagram of the proposed inverse fingerprint beam alignment method.

The CV can now proceed to perform the beam training following the list provided by the RSU. Note that the list consists of beam pairs, and there is no need to do exhaustive measurements over all combinations of transmit and receive beams in the list. Upon completing the list, the RSU provides a feedback indicating the best beam index to the CV. This feedback ends the beam alignment and high data rate mmWave communication can start.

The beam training here has much lower overhead than conventional methods because a large number of unlikely directions have already been eliminated using the database.

Regarding the overhead of the beam training, the approach requires training less than 30 beam pairs when 16×16 arrays are used and the overhead increases roughly linearly with the number of antenna elements.

The overhead calculation under the mobility context using the beam coherence time shows that while the proposed approach consumes less than a few percent of the beam coherence

time for training, the IEEE 802.11ad beam training duration can exceed the beam coherence time for large arrays such as 32×32 [21].

3

5G mmwave Networks

The 5th generation of the mobile network will bring the revision of the current legacy LTE networks with the new mmWave technologies. The integration of the two networks helps in the robustness in terms of link outage and service availability. In this chapter first we introduce briefly the current LTE network architecture and then we see a possible choice of integration that is the one implemented in the NY mmwave module of ns-3;

3.1 Lte System architecture

The LTE standard provides specifics on the Evolved Universal Terrestrial Radio Access Network (E-UTRAN), which is the radio access part and is used in conjunction with the Evolved Packet Core (EPC) network. Together they form the EPS (Fig. 2.1).

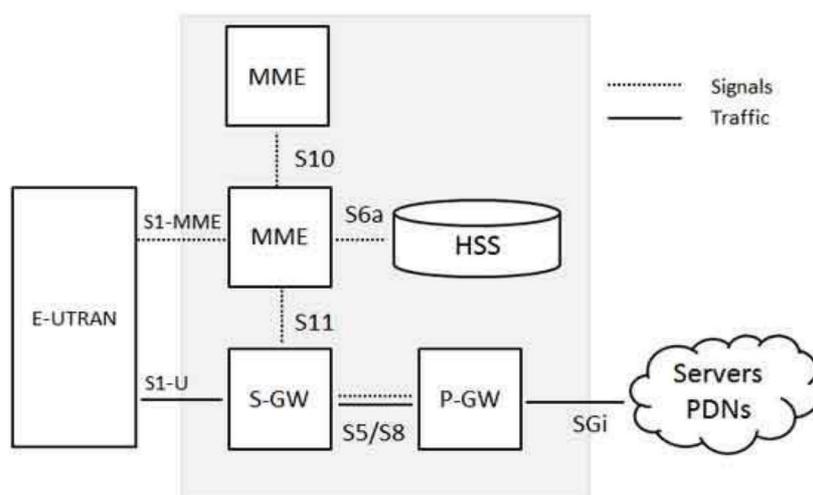


Figure 3.1 LTE system architecture

While the CN(EPC) consists of many logical nodes, the access network(E-UTRAN) is made up of essentially just one node, the evolved NodeB (eNodeB), which connects to the UEs. Each of these network elements is interconnected by means of interfaces that are standardized in order to allow multi-vendor interoperability.

The core network

The core network (called EPC in SAE) is responsible for the overall control of the UE and establishment of the bearers. The main logical nodes of the EPC are:

- PDN Gateway (P-GW)
- Serving Gateway (S-GW)
- Mobility Management Entity (MME)

In addition to these nodes, EPC also includes other logical nodes and functions such as the Home Subscriber Server (HSS) and the Policy Control and Charging Rules Function (PCRF).

The logical CN nodes are shown in Figure 1 and discussed in more detail below:

P-GW: The PDN Gateway is responsible for IP address allocation for the UE, as well as QoS enforcement and flow-based charging according to rules from the PCRF. It is responsible for the filtering of downlink user IP packets into the different QoS-based bearers. This is performed based on Traffic Flow Templates (TFTs). The P-GW performs QoS enforcement for guaranteed bit rate (GBR) bearers. It also serves as the mobility anchor for interworking with non-3GPP technologies such as CDMA2000 and WiMAX® networks.

S-GW: All user IP packets are transferred through the Serving Gateway, which serves as the local mobility anchor for the data bearers when the UE moves between eNodeBs. It also retains the information about the bearers when the UE is in the idle state (known as “EPS Connection Management — IDLE” [ECM-IDLE]) and temporarily buffers downlink data while the MME initiates paging of the UE to reestablish the bearers. In addition, the S-GW performs some administrative functions in the visited network such as collecting information for charging (for example, the volume of data sent to or received from the user) and lawful interception. It also serves as the mobility anchor for interworking with other 3GPP technologies such as general packet radio service (GPRS) and UMTS.

MME: The Mobility Management Entity (MME) is the control node that processes the signaling between the UE and the CN. The protocols running between the UE and the CN are known as the Non Access Stratum (NAS) protocols.

The main functions supported by the MME can be classified as:

- *Functions related to bearer management*—This includes the establishment, maintenance and release of the bearers and is handled by the session management layer in the NAS protocol.

- *Functions related to connection management*—This includes the establishment of the connection and security between the network and UE and is handled by the connection or mobility management layer in the NAS protocol layer.

HSS: The Home Subscriber Server contains users' SAE subscription data such as the EPS-subscribed QoS profile and any access restrictions for roaming. It also holds information about the PDNs to which the user can connect. This could be in the form of an access point name (APN) (which is a label according to DNS naming conventions describing the access point to the PDN) or a PDN address (indicating subscribed IP address(es)). In addition the HSS holds dynamic information such as the identity of the MME to which the user is currently attached or registered. The HSS may also integrate the authentication center (AUC), which generates the vectors for authentication and security keys.

PCRF: The Policy Control and Charging Rules Function is responsible for policy control decision-making, as well as for controlling the flow-based charging functionalities in the Policy Control Enforcement Function (PCEF), which resides in the P-GW. The PCRF provides the QoS authorization (QoS class identifier [QCI] and bit rates) that decides how a certain data flow will be treated in the PCEF and ensures that this is in accordance with the user's subscription profile.

The access network

The access network of LTE, E-UTRAN, simply consists of a network of eNodeBs, as illustrated in Figure 3. For normal user traffic (as opposed to broadcast), there is no centralized controller in E-UTRAN; hence the E-UTRAN architecture is said to be flat.

The eNodeBs are normally interconnected with each other by means of an interface known as "X2" and to the EPC by means of the S1 interface — more specifically, to the MME by means of the S1-MME interface and to the S-GW by means of the S1-U interface.

The protocols that run between the eNodeBs and the UE are known as the "AS protocols."

The E-UTRAN is responsible for all radio-related functions, which can be summarized briefly as:

- *Radio resource management (RRM)* – This covers all functions related to the radio bearers, such as radio bearer control, radio admission control, radio mobility control, scheduling and dynamic allocation of resources to UEs in both uplink and downlink.
- *Header Compression* – This helps to ensure efficient use of the radio interface by compressing the IP packet headers that could otherwise represent a significant overhead, especially for small packets such as VoIP.
- *Security* – All data sent over the radio interface is encrypted.
- *Connectivity to the EPC* – This consists of the signaling toward MME and the bearer path toward the S-GW.

An important feature of the S1 interface linking the access network to the CN is known as “S1-flex.” This is a concept whereby multiple CN nodes (MME/S-GWs) can serve a common geographical area, being connected by a mesh network to the set of eNodeBs in that area. An eNodeB may thus be served by multiple MME/S-GWs, The set of MME/S-GW nodes that serves a common area is called an MME/S-GW pool, and the area covered by such a pool of MME/S-GWs is called a pool area. This concept allows UEs in the cell or cells controlled by one eNodeB to be shared between multiple CN nodes, thereby providing a possibility for load sharing and also eliminating single points of failure for the CN nodes. The UE context normally remains with the same MME as long as the UE is located within the pool area.

In ns3-simulator the P-Gw and S-Gw are located in the same node.

3.1.1 User and control plain architecture

The radio protocol architecture of LTE is subdivided in a control plane and a user plane architecture.

In the user plain architecture, the application creates data packets that are processed by protocols such as TCP, UDP and IP, while in the control plane, the radio resource control (RRC) protocol writes the signalling messages that are exchanged between the base station and the mobile. In both cases, the information is processed by the packet data convergence protocol (PDCP), the radio link control (RLC) protocol and the medium access control (MAC) protocol, before being passed to the physical layer for transmission.

User plane

The E-UTRAN user plane protocol stack between the e-Node B and UE consists of the following sub-layers:

- PDCP (Packet Data Convergence Protocol)
- RLC (radio Link Control)
- Medium Access Control (MAC)

On the user plane, IP packets for a UE in the core network (EPC) are encapsulated in a specific EPC protocol and tunneled between the P-GW and the eNodeB. Different tunneling protocols are used depending on the interface. GPRS Tunneling Protocol (GTP) is used on the S1 interface between the eNodeB and S-GW and on the S5/S8 interface between the S-GW and P-GW.

E-UTRAN and EPC Protocol Stacks

End-to-End User Plane

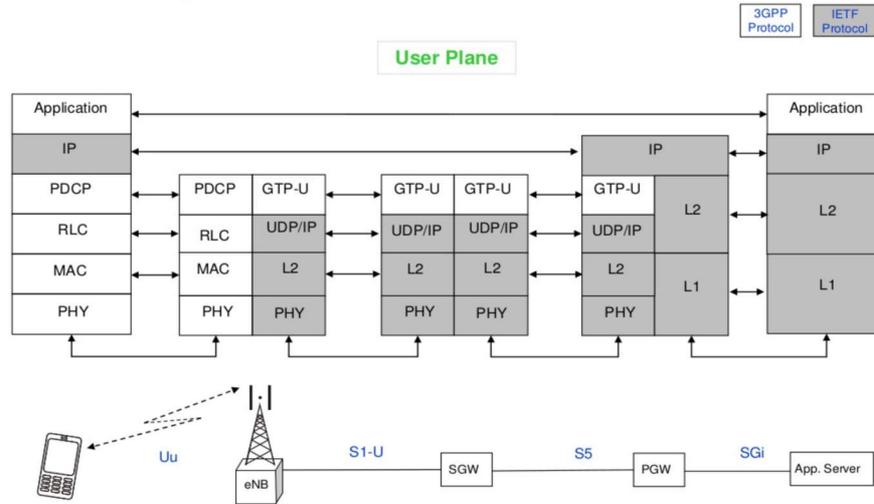


Figure 3.2 End-to-End User Plane

Control plane

The protocol stack for the control plane between the UE and MME is shown in Figure 7. The blue region of the stack indicates the AS protocols. The lower layers perform the same functions as for the user plane with the exception that there is no header compression function for the control plane.

The Radio Resource Control (RRC) protocol is known as “layer 3” in the AS protocol stack. It is the main controlling function in the AS, being responsible for establishing the radio bearers and configuring all the lower layers using RRC signaling between the eNodeB and the UE.

The Control Plane handles radio-specific functionality which depends on the state of the user equipment which includes two states: idle or connected(**Table 2.1**).

Idle

The user equipment camps on a cell after a cell selection or reselection process where factors like radio link quality, cell status and radio access technology are considered. The UE also monitors a paging channel to detect incoming calls and acquire system information. In this mode, control plane protocols include cell selection and reselection procedures.

Connected

The UE supplies the E-UTRAN with downlink channel quality and neighbour cell information to enable the E-UTRAN to select the most suitable cell for the UE. In this case, control plane protocol includes the Radio Link Control (RRC) protocol.

Table 3.1 User states description

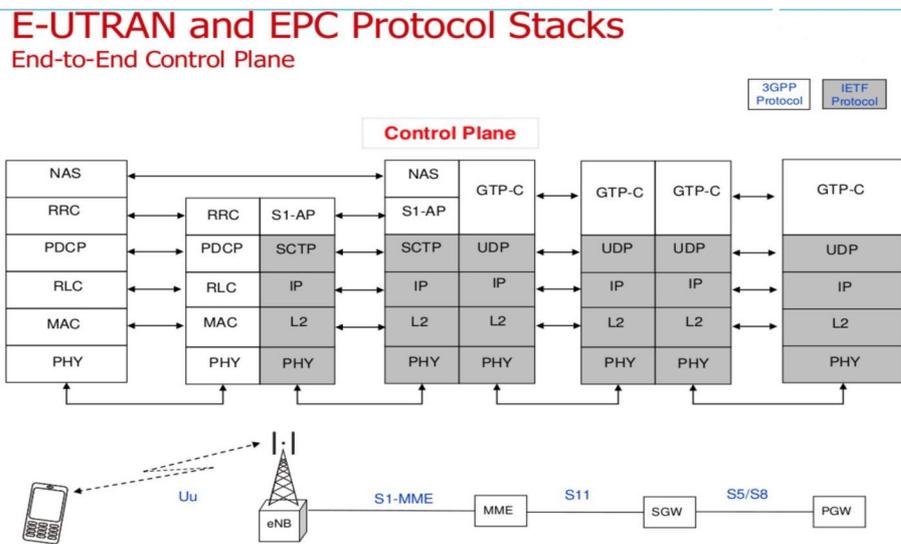


Figure 3.3 End-to-End Control Plane

3.1.2 Layers Functionalities

Physical Layer (Layer 1)

Physical Layer carries all information from the MAC transport channels over the air interface. Takes care of the link adaptation (AMC), power control, cell search (for initial synchronization and handover purposes) and other measurements (inside the LTE system and between systems) for the RRC layer.

Medium Access Layer (MAC)

MAC layer is responsible for Mapping between logical channels and transport channels, Multiplexing of MAC SDUs from one or different logical channels onto transport blocks (TB) to be delivered to the physical layer on transport channels, de multiplexing of MAC SDUs from one or different logical channels from transport blocks (TB) delivered from the physical layer on transport channels, Scheduling information reporting, Error correction through HARQ, Priority handling between UEs by means of dynamic scheduling, Priority handling between logical channels of one UE, Logical Channel prioritization.

Radio Link Control (RLC)

RLC operates in 3 modes of operation: Transparent Mode (TM), Unacknowledged Mode (UM), and Acknowledged Mode (AM).

RLC Layer is responsible for transfer of upper layer PDUs, error correction through ARQ (Only for AM data transfer), Concatenation, segmentation and reassembly of RLC SDUs (Only for UM and AM data transfer).

RLC is also responsible for re-segmentation of RLC data PDUs (Only for AM data transfer), reordering of RLC data PDUs (Only for UM and AM data transfer), duplicate detection (Only for UM and AM data transfer), RLC SDU discard (Only for UM and AM data transfer), RLC re-establishment, and protocol error detection (Only for AM data transfer).

Radio Resource Control (RRC)

The main services and functions of the RRC sublayer include broadcast of System Information related to the non-access stratum (NAS), broadcast of System Information related to the access stratum (AS), Paging, establishment, maintenance and release of an RRC connection between the UE and E-UTRAN, Security functions including key management, establishment, configuration, maintenance and release of point to point Radio Bearers.

Packet Data Convergence Control (PDCP)

PDCP Layer is responsible for Header compression and decompression of IP data, Transfer of data (user plane or control plane), Maintenance of PDCP Sequence Numbers (SNs), In-sequence delivery of upper layer PDUs at re-establishment of lower layers, Duplicate elimination of lower layer SDUs at re-establishment of lower layers for radio bearers mapped on RLC AM, Ciphering and deciphering of user plane data and control plane data, Integrity protection and integrity verification of control plane data, Timer based discard, duplicate discarding, PDCP is used for SRBs and DRBs mapped on DCCH and DTCH type of logical channels.

Non Access Stratum (NAS) Protocols

The non-access stratum (NAS) protocols form the highest stratum of the control plane between the user equipment (UE) and MME.

NAS protocols support the mobility of the UE and the session management procedures to establish and maintain IP connectivity between the UE and a PDN GW.

3.2 Toward 5G....

5th generation mobile networks or 5th generation wireless systems, abbreviated 5G, are the proposed next telecommunications standards beyond the current 4G/LTE.

Actually researchers and engineers from entire world are working on 5G, while 3GPP (Third Generation Partnership Project) is trying to synthesize the 5G standard, foreseeing which 5G will be on air around 2020.

The design of an end-to-end cellular system to achieve very high throughput and ultra-low latency requires innovations across all the layers of the communication protocol stack.

Research has been done on channel modelling, beamforming and other physical layer procedures; The use of high directional beams increases the complexity of a number of MAC layer procedures such as synchronization, control signalling, cell search and initial access that translates in delay and robustness.

3.2.1 Key problems:

Adaptive beamforming and beam tracking:

RX and TX must continuously track the channel in order to align their antennas to achieve the maximum directional gain. At the mmwave frequency the shadowing becomes critical, so the signal can be blocked also by small obstacles. Fortunately recent research demonstrates that the communication is possible through reflected power that could be sufficient in NLOS case, so by steering the beam we can recover through an alternate NLOS path.

The User Equipment (UE) and base station must then jointly initiate a procedure to search for and select another path to reestablish the link.

Directional synchronization and broadcast channel:

The cell discovery and the initial access procedures require an innovative approach. Traditional cells periodically broadcast synchronization signals (e.g. PSS in LTE) omnidirectionally which are received by all devices under the coverage range and used to initially connect to the cell. Here we cannot adopt this approach because the signal would not benefit from directional gain and might not have adequate range to be detected by many UEs. The gNBs and the UE must perform an angular search so that users can detect the PSS and adjust on the optimal TX/RX beamforming angles.

Issues in MAC, Network and transport layer:

The high variability of the mmwave channel and the high susceptibility to shadowing will require very frequent handover between 4G or mmwave RATs.

In case we make use of TCP protocol, the high variability in the mmwave channel capacity requires some innovations in the TCP congestion control to better exploit mmwave channel capacity without create too much congestions in the network.

An End-to-End network simulator for mmWave cellular networks is an invaluable tool that can help address these challenges.

In particular the solutions proposed by the **ns3-mmWave module** implements custom PHY and MAC layers.

In particular we can analyze briefly the innovations to cop with the previuos issues.

3.2.2 MmWave module for ns-3

The ns-3 discrete-event network simulator is a very powerful tool available to communication and networking researchers for developing new protocols and analyzing complex systems.

The ns-3 simulator is organized into multiple folders. The src folder provides a collection of C++ classes, which implement a wide range of modular simulation models and network protocols. The different modules can be aggregated and instantiated to build diverse simulated network scenarios.

In the following sections, we will describe in detail the mmWave module for ns-3.

The ns-3 mmWave module is designed to perform end-to-end simulations of 3GPP-style cellular networks. It leverages the detailed implementation of LTE/EPC protocols.

The structure, high-level functions and naming scheme of each class closely follow the LTE LENA module.

3.2.2.1 Class diagram for the end-to-end mmWave module

The *MmWaveEnbNetDevice* and *MmWaveUeNetDevice* classes, which represent the mmWave eNB and UE radio stacks, respectively, along with a perspective on the end-to-end structure of the simulator.

The ns-3 mmWave module also includes a *McUeNetDevice*, which is a NetDevice with a dual stack (LTE and mmWave), i.e., a device capable of connecting to both technologies.

The *MmWaveEnbMac* and *MmWaveUeMac* MAC layer classes implement the LTE module Service Access Point (SAP) *provider* and *user* interfaces, which enable the interoperation with the LTE RLC layer.

Support for RLC Transparent Mode (TM), Saturation Mode (SM), Unacknowledged Mode (UM), Acknowledged Mode (AM) is built into the MAC and scheduler classes (i.e., *MmWaveMacScheduler* and derived classes).

The MAC scheduler also implements a SAP for configuration at the LTE Radio Resource Control (RRC) layer (*LteEnbRrc*).

Hence, every component required to establish Evolved Packet Core (EPC) connectivity is available.

The *MmWavePhy* classes handle directional transmission and reception of the DL and UL data and control channels based on control messages from the MAC layer.

Similar to the LTE module, each PHY instance communicates over the channel (i.e., *SpectrumChannel*) via an instance of the *MmWaveSpectrumPhy* class, which is shared for both the DL and the UL (since the design of the mmWave PHY layer is based on Time Division Duplexing (TDD)).

Instances of *MmWaveSpectrumPhy* encapsulate all PHY-layer models: interference calculation (*MmWaveInterference*), Signal to Interference plus Noise Ratio (SINR) calculation (*MmWaveSinrChunkProcessor*), the Mutual Information (MI)-based error model (*MmWaveMiErrorModel*), which computes the packet error probability, as well as the Hybrid Automatic Repeat reQuest (HARQ) PHY-layer entity (*MmWaveHarqPhy*) to perform soft combining.

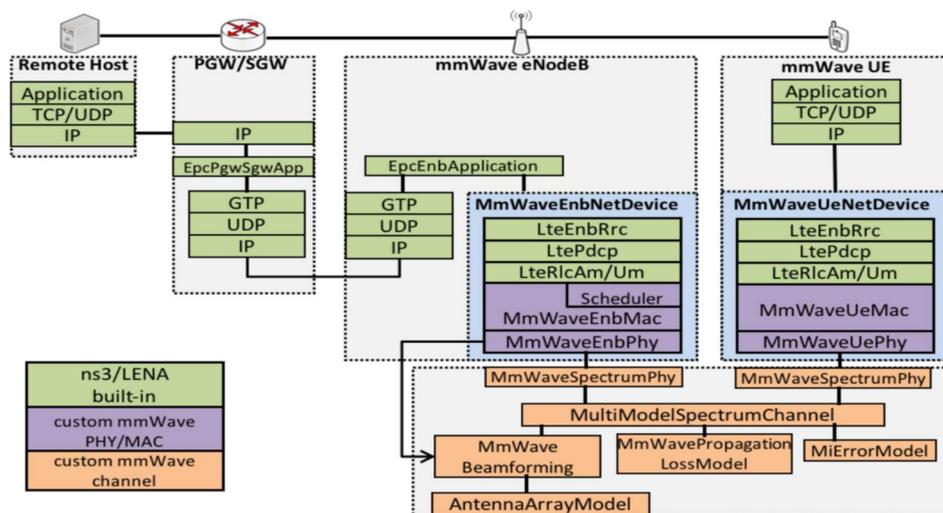


Figure 3.4 Class diagram for the end-to-end mmWave module

3.2.2.2 Channel modelling

The ns-3 mmWave module allows the user to choose among different channel models, which provide a trade-off between computational complexity, flexibility and accuracy of the results.

The most flexible and detailed channel model is based on the official 3GPP channel model for the 6-100 GHz frequency band.

The second model is based on traces from measurements or third-party ray-tracing software. This makes the channel model detailed and realistic, but constrains the simulation to limited measurements/ray-tracing routes. The third is the statistical channel model introduced in [26] and based on MATLAB traces, which makes the computation less demanding, but is available only for the 28 and 73 GHz frequencies.

The 3GPP model is a *Statistical Channel Model* for the 6-100 GHz band, It provides several optional features that can be plugged into the basic model, in order to simulate, for example, spatial consistency (i.e., the radio environment conditions of close-by users are correlated), oxygen absorption, and random blockage. It is applicable for bandwidth up to 10% of the carrier frequency f_0 , and accounts for the mobility of one of the two terminals (usually in a cellular network the BS is fixed and the UE moves)

The model defines different scenarios, which describe different possible cellular network deployments: urban (with macrocells and microcells), rural and indoor. In particular we have UrbanMicrocell (UMi), UrbanMacrocell (UMa), Rural Macrocell (RMa), Indoor Office (In). Different scenarios present different characteristics in term of Inter-Site Distance (ISD) and height of the BSs respect to the rooftop of the buildings.

In particular in the simulated scenario we propose in **chapter 6** we use Umi in which the Bss are below the rooftop of the buildings surrounding the streets and with small ISD.

The ns-3 modules that handle the wireless channel are the Propagation module, which defines the PropagationLossModel interface, and the Spectrum module with the Spectrum-PropagationLossModel interface.

By extending the first, it is possible to implement different propagation models, while the second is the basis for modeling small scale fading in terms of a Power Spectral Density(PSD).

In the mmWave module, the physical layer at the BS or the UT side uses *MmWaveSpectrumPhy* to simulate actual transmissions with noise and interference. It is this class which is in charge of computing the SINR of the transmissions, given the propagation conditions specified in the mmWave module extensions of *PropagationLossModel* and *SpectrumPropagationLossModel*.

A ray-tracing model implemented in the class MmWaveRayTracing and the NYU statistical model implemented in MmWavePropagationLossModel and MmWaveBeamforming.

Here we introduce the implementation for ns-3 of the 3GPP channel model.

The propagation and shadowing are computed in the MmWave3gppPropagationLossModel, or, if the ns-3 buildings module is used, in the MmWave3gppBuilding-PropagationLossModel. The channel matrix $H(t,\tau)$ is computed in the class MmWave3gppChannel, which computes also the beamforming vector and interacts with the AntennaArrayModel class. A basic UML diagram that describes the relationships among these classes is showed in **Figure 3.5**.

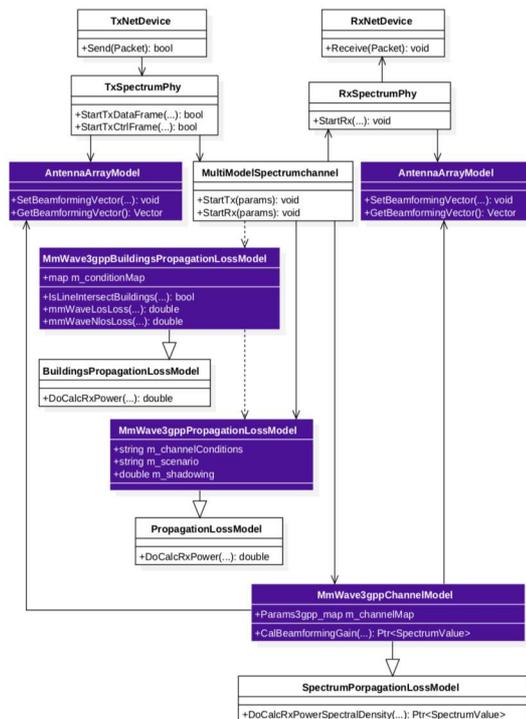


Figure 3.5 A basic UML diagram that describes the relationships among these classes

The antenna is modeled in the class *AntennaArrayModel* as a Uniform Planar Array (UPA), In the current implementation we support a single rectangular panel per BS/UT, with C columns and R rows.

The number of antenna elements for each device can be set using the *AntennaNum* attribute of the *MmWaveEnbNetDevice* and *MmWaveUeNetDevice* classes. This allows to specify a different number of antennas for the BSs and the UTs, as it would happen in a realistic deployment. The default values are 64 antennas for the BS, arranged in a square with 8×8 elements, and 16 for the UT, in a square with 4×4 elements. The horizontal and vertical spacing dH and dV are stored in two private variables of the *AntennaArrayModel* class, and expressed as multiples of the wavelength $\lambda = c/fc$. The default value is 0.5, but it is possible to update it using the *AntennaHorizontalSpacing* and *AntennaVerticalSpacing* attributes.

The method *GetAntennaLocation* which computes the coordinates of an antenna element, the method *GetRadiationPattern* returns the antenna radiation pattern, given the angle of departure or arrival.

In the implementation, once one of the scenarios is set, it is possible to select the LOS condition and the indoor/outdoor state statistically, geometrically or deterministically. The pathloss then depends on the UT status, the distance on the horizontal plane (2D distance) and the 3D distance from the UT to the BS.

The class *MmWave3gppPropagationLossModel*, which extends the *PropagationLossModel* interface, handles O2O or indoor to indoor transmissions and performs the pathloss computation in the *GetLoss* method. In this class it is possible to deterministically set the LOS condition, using the *ChannelCondition* attribute, so that the UT is always in a LOS or NLOS condition, or opt for the statistical approach proposed by the 3GPP model. With the latter, given the selected scenario, the 2D distance and the height of the BS and UT, it is possible to compute a LOS probability PLOS. Another optional component is the shadowing, which is enabled by the *Shadowing* attribute. For a moving UT, the shadowing is correlated in space. The pathloss and the shadowing (if enabled) are updated at every transmission, while the LOS condition is fixed for the whole duration of the simulation.

The class *MmWave3gppBuildingPropagationLossModel* provides an alternative model for pathloss computation. This class takes advantage of the ns-3 Buildings module in order to geometrically model the LOS/NLOS and the indoor/outdoor conditions. In particular, it is possible to place buildings in the simulation scenario, randomly or deterministically, and, once the UTs and BSs positions are specified, a 3D ray tracing algorithm determines whether there is LOS or not. Once the LOS/NLOS condition is set and the possible penetration loss is computed, this class relies on the shadowing and pathloss computation performed in the *MmWave3gppPropagationLossModel* class.

The fast fading modeling is implemented in the class *MmWave3gpp-Channel*.

The fast fading modeling represents the bottleneck of system level simulations with wireless channels, so assumptions and simplifications are introduced.

The main method of this class is *DoCalcRxPowerSpectralDensity*, which implements the interface specified in *SpectrumPropagation-LossModel* and returns a *SpectrumValue* object, i.e., a vector with the values of the PSD for each subcarrier of the OFDM symbol transmitted. In order to get this PSD, two steps are required: the first is the actual computation of the

channel matrix $H(t, \tau)$ with the large scale and small scale fading effects, and the second is the generation of the beamforming vectors and of the beamforming gain.

The channel matrix is generated for all the possible BS-UT links, but the beamforming vectors are computed only for the links involved in a communication, i.e., those for which a UT is attached to a BS.

The analog MIMO beamforming scheme implementation can be used together with the channel model. Once pathloss, shadowing and channel coefficient matrix are generated, the beamforming vector at transmitter and receiver and the gain are computed and the latter is applied to the PSD object handled by the `DoCalcRxPowerSpectralDensity` method of `MmWave3gppChannel`.

The framework assumes analog beamforming and a TDD physical layer frame. Since with analog beamforming a single stream is transmitted, each antenna element only applies a phase shift to the transmit signal, so that it is concentrated in one direction.

The beamforming gain is obtained by multiplying the transmitter and receiver beamforming vectors with the channel coefficients matrix $H(t, \tau)$.

We have two ways to compute the beamforming vector, The two default methods are: power and cell scan method, it is possible to select one of the two using the `CellScan` attribute.

The power method computes the optimal beamforming vector for the transmitter and the receiver, assuming a perfect knowledge of the channel matrix, and it is implemented in the `PowerMethodBeamforming` method.

The cell scanning method does not assume that the transmitter and receiver know the channel information. Basically, the cell scanning method divides the coverage area of the cell into several sectors and steers the beam to one of the directions by selecting the beamforming vector from a predefined codebook. This means that, if the transmitter wants to transmit in sector M , it selects from a list (i.e., the codebook) the beamforming vector corresponding to that sector and the signal will be directed in that sector. The method `CellScan` of the class `MmWave3gppChannel` currently implements a simplified cell scanning approach, based on a brute force approach, i.e., it tries all the possible BS-UT sector combinations and selects the one that returns the maximum received power.

For 5G requirements and applications, conventional drop-based modeling is inadequate.

Closely spaced channel snapshots should have similar delays, AOA and AOD, scattering environment etc. should be similar, spatially consistent models are required to capture smooth transitions of channel characteristics over relatively closely spaced locations.

For simulations in which the mobility is an important factor, the spatial consistency of the channel throughout the path on which the UT moves can be simulated by enabling this option in the `MmWave3gppChannel` class.

In the current implementation, it is supported spatial consistency for both LOS and NLOS communications.

3.2.2.3 Mmwave Layers

PHY layer

It's implemented a TDD frame structure and subframe structure similar to the TDD-LTE but that allow for more flexible allocation and placement of the control and data channels within the subframe and is suitable for the variable TTI MAC scheme we explain later.

5G target TDD because it offers improved utilization of wider bandwidths and the opportunity to take advantage of channel reciprocity for channel estimation.

It is introduced a shorter symbol periods and/or slots length so that we reduce radio link latency.

The ns-3 mmWave module therefore implements a TDD frame structure which is designed to be configurable for evaluating different potential designs and numerologies. These parameters, are accessible through the attributes of the common *MmwavePhyMacCommon* class, which stores all user-defined configuration parameters used by the PHY and MAC classes.

As LTE each frame is divided in a number of subframes of fixed length(length as multiple of OFDM symbols). Within each subframe, a variable number of symbols can be assigned by the MAC scheduler and designed for either control or data channel transmission.

Each variable-length time-domain data slot can be assigned by the scheduler to different users for either the uplink or the downlink.

In the following image we see the a possible frame structure that is the one implemented in the mmwave module(**Fig. 3.6**)

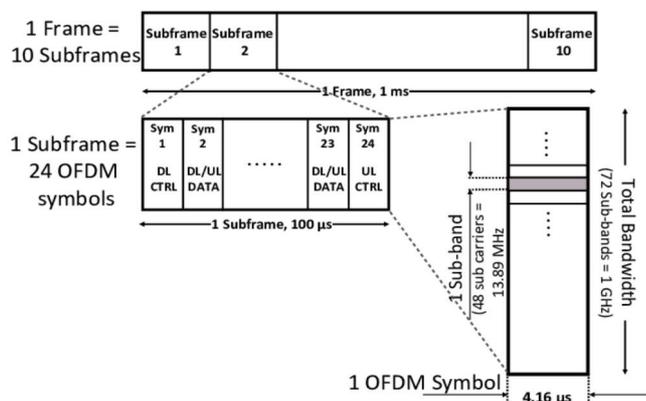


Figure 3.6 Mmwave frame structure

Each frame of length 1ms is split in time into 10 subframes of duration 100 μs each, representing 24 symbols of 4.16 μs. In the proposed structure the downlink and uplink control channel are always fixed in the first and last symbol of the subframe.

A guard period of one symbol period is introduced every time we change direction from UL to DL. In the frequency domain the total bandwidth of 1GHz is split in 72 subbands of width 13.89 Mhz(each composed of 48 subcarriers).

The *MmWaveEnbPhy* and *MmWaveUePhy* classes model the physical layer for the mmWave eNB and the UE. These objects (i) handle the transmission and reception of physical control and data channels(ii) simulate the start and the end of frames, subframes and slots, and (iii) deliver received and successfully decoded data and control packets to the MAC layer.

MAC layers

The proposed scheme make use of analog beamforming (where the transmitter and receiver align their antenna arrays to maximize the gain) so we are constrained to the TDMA scheme; With digital beamforming(base station transmit and receive in multiple direction at the same time) would be possible the use of SDMA or FDMA. The key performance indicator of the 5G MAC layer is the latency, the over-the-air latency have been proposed by bodies as International Telecommunication Union is 1ms.

The well-known drawback of the TDMA is that TTIs or fixed slots lead to a non optimal resource utilization and latency ,very dramatic when we have many intermittent,small packets must be transmitted or received.

Based on these considerations, variable TTI-based TDMA frame structures and MAC schemes have been proposed.

So this approach permit to adapt slot sizes according to the length of the packet or TB to be transmitted,so that different kind of traffic are targeted(from bursty to file transfer).

The MAC layer implementation can be found in the *MmWaveEnbMac* and *MmWaveUeMac* classes, whose main role is the coordination of procedures such as scheduling and retransmission. Moreover, they interact with the RLC layer to receive periodic reports on the buffer occupancy, i.e., the Buffer Status Reports (BSRs), and with the physical layer classes for the transmission and reception of packets.

We have an Adaptive Modulation and Coding(AMC) technique to adapt the modulation and coding scheme to the channel quality. The AMC is implemented in the *MmWaveAmc* class. In the simulator, this translates into (i) mapping the CQI(Channel Quality Information report) into the Modulation and Coding Scheme (MCS), using the error model implemented in the *MmWaveMiErrorModel* and described in Section V-D, and (ii) computing the available TB size for a subframe given the MCS. This information is then used by the scheduler to perform radio resource management.

As in the LTE network we have support for the HARQ which enables fast retransmissions with incremental redundancy.

the HARQ mechanism is based on multiple stop and wait retransmission processes, and a maximum of 8 simultaneous HARQ processes can be active at any given time . The HARQ retransmissions have priority with respect to new transmissions, thus the available resources are given first to HARQ processes and then to the data queued in the RLC buffers.

Despite being fundamental in protecting from the losses of packets due to rapid variations in the channel quality, the HARQ mechanism introduces additional latency [19], [21], therefore the optimization of its performance is necessary to enable the target of sub-1-ms latency for ultra-low-latency communications.

The *MmWaveHarqPhy* class along with the functionalities within the different scheduler classes are based heavily on the LENA module code.

The scheduler at the eNB uses the information provided by HARQ feedback messages to assign new resources to the HARQ processes that require retransmissions. Each transport block is granted a maximum number of transmission attempts, which is set to 3.

The multiple HARQ processes per user can be created not only for the downlink but also for the uplink, and the maximum number of processes per user is not fixed (it can be configured through the NumHarqProcesses attribute in MmWavePhyMacCommon), this permit to control the simultaneous stop and wait retransmissions processes to optimize the bandwidth utilization.

Finally using a flexible TTI physical layer allows reduction in the latency of the retransmissions.

RLC layer

The RLC layer is basically very similar to the one implemented in the LTE module of ns3, except for the RLC AM retransmission entity that is modified to be suitable for the mmwave PHY and MAC layers; An Active queue management (AQM) for the RLC buffer is added. In RLC AM we have retransmission and reordering. Since in the mmwave module we have a shorter frame structure we have to reduce the timers of the RLC entity. In the RLC AM proposed setup is implemented the segmentation also for the retransmissions, its important especially with a high variable mmwave channel where the channel capacity change suddenly, otherwise in order to transmit the segment we need to wait that the transmission opportunity is big enough so we can halt the retransmission.

Active Queue Management techniques are used in the buffers of routers, middleboxes and base stations in order to improve the performance of TCP and avoid the manual tuning of the buffer size. AQM strategies allow the network to avoid congestion at the buffers, because they react early to the increase in the buffer occupancy by dropping some packets before the buffer is full.

In the mmWave module, the RLC layer can use either the default Drop-tail approach or more sophisticated AQM techniques, that can be enabled by setting the *EnableAQM* attribute to true. The default AQM is the CoDel (Controlled Delay) scheme. CoDel works by monitoring the minimum queue delay in every 100 ms interval, and only drops packets when the minimum queue delay is more than 5 ms.

RRC layer

the RRC layer was expanded in order to add the dual connection control procedure. Specifically the uplink measurements framework, the gNBs receive the uplink reference signal, they compute the SNR for each UE and send the infos to the LTE eNB through the X2 link. Thanks to this Framework the LTE act as coordinator for the nearby gNB and find the best association in term of SNR between UEs and gNBs.

3.2.2.4 Dual Connectivity integration

The layer at which the 5G mmWave stack and the legacy 4G stack converge is called integration layer. There are several proposal with advantages and disadvantages; The integration layer chosen in the simulator is the PDCP layer.

This layer transmit packets to different RATs and receive traffic from the different lower layers.

This choice offers different advantages, permit to have a custom PHY, MAC and RLC layers so that we can satisfy the requirements of the mmWave 5G system. Since PDCP doesn't require strict synchronization so it permits to have non co-located placement. The choice of using PHY and MAC as integration layer is almost impracticable because at this layer is required good synchronization and this is not possible with non co-located placement, and because to address the 5G requirements needs a new custom MAC layer.

RLC as integration layer presents some limitations that would prevent a non co-located deployment. Indeed, the RLC layer receives from the MAC layer scheduler indications on the transmission opportunities, i.e., how many bytes are available for transmission during the next slot. This communication cannot be subjected to the additional latencies of a MAC- RLC communication between remote locations. Moreover, segmentation and reassembly would work only in the presence of a common scheduler.

The ns-3 mmWave module is also capable of performing simulations with dual-stack UEs connected both to an LTE eNB and to a mmWave eNB.

The Dual Connectivity (DC) implementation of this simulation module assumes that the core networks of LTE and of mmWave will be integrated. Therefore the LTE and the mmWave eNBs share the same backhaul network, i.e., they are connected to each other with X2 links and to the Mobility Management Entity (MME)/Packet Gateway (PGW) nodes with the S1 interface.

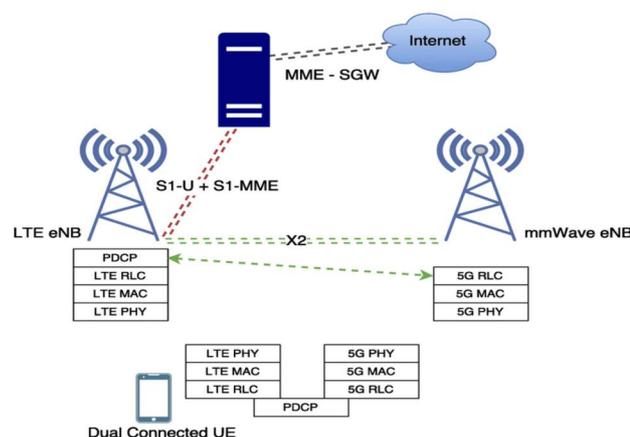


Figure 3.7 The Dual Connectivity (DC) implementation

In particular, a single bearer per DC flow is established, with a connection from the core network to the LTE eNB, where the flow is split and forwarded either to the local stack or to the remote mmWave stack.

A basic diagram for a DC UE device, an LTE eNB and a mmWave eNB is shown in **Figure 3.7**. The core of the DC implementation is the McUeNetDevice class, which is a subclass of

the ns-3 NetDevice and provides an interface between the ns-3 TCP/IP stack and the custom lower layers.

The McUeNetDevice holds pointers to the custom lower layer stack classes, and has a Send method that forwards packets to the TCP/IP stack.

This method is linked to a callback on the DoRecvData of the EpcUeNas class, which as specified by the 3GPP standard acts as a connection between the LTE-like protocol stack and the TCP/IP stack.

The McUeNetDevice describes a dual connected UE with a single EpcUeNas, but with a dual stack for the lower layers, i.e., there are separate LTE and mmWave PHY and MAC layers. Moreover, there is an instance of the RRC layer for both links.

Besides, the LTE RRC manages both the LTE connection and the control plane features related to DC, while the mmWave RRC handles only the mmWave link.

The usage of a secondary RRC, dedicated to the mmWave link, avoids latency in control commands (i.e., the mmWave eNB does not have to encode and transmit the control Packet Data Unit (PDU)s to the master LTE eNB). The EpcUeNas layer has an interface to both RRC entities to exchange information between them.

The LTE RRC manages also the data plane for the DC devices. In particular, for each bearer, a dual connected PDCP layer is initialized and stored in the LTE RRC.

The classes describing the DC PDCP layer are McEnbPdcP and McUePdcP, respectively at the eNB side and at the UE side.

However, while McUePdcP simply has to communicate with a local RLC in the UE, the implementation of McEnbPdcP requires new interfaces to the class describing the X2 links between eNBs (i.e., EpcX2). In particular, in downlink the eNB PDCP has to send packets to the X2 link and the mmWave RLC layer has to receive them, and vice versa in uplink.

With FS, the UE is in the RRC CONNECTED state with respect to both eNBs, but only transmits data to one of the two.

As to the physical layer, the two stacks rely on the mmWave and LTE channel models.

The RRC layer implementation of the original LTE ns-3 module was extended in order to account for DC-related control procedures. In particular, the multi-connectivity uplink-based measurement framework. The MmWaveEnbPhy instance simulates the reception of uplink reference signals, computes the SINR for each UE in the scenario, and sends this information to the LTE eNB on the X2 link.

the control packets with the SINR values must be transmitted on an ns-3 PointToPointLink, which adds a certain latency and has a certain bitrate.

Thanks to this framework, the LTE eNB is able to act as a coordinator for the surrounding mmWave eNBs, and learns which is the best association (in terms of SINR) between UEs and mmWave secondary eNBs. This enables automatic cell selection for mmWave eNBs at the beginning of a simulation, and the control of mobility-related operations. The DC module is indeed capable of simulating FS procedures between mmWave and LTE links and Secondary Cell Handover (SCH) (i.e., handovers between mmWave eNBs that do not involve the MME in the core network) initiated by the central controller in the LTE eNB.

Different handover (either inter-RAT or SCH) algorithms can be tested, by implementing them in the *LteEnbRrc* class.

In order to model the additional latency given by the interaction with the MME for inter-RAT handovers for standalone UEs, the link between the eNBs and the MME is modeled in this module as a PointToPointLink.

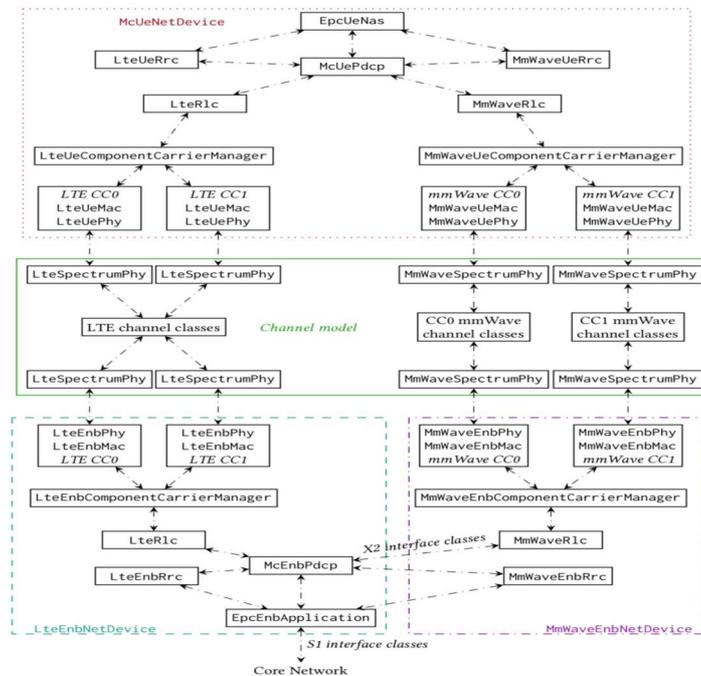


Figure 3.8 Simplified UML of a dual-connected device, an LTE eNB and a MmWave eNB that also support carrier aggregation.

In the current implementation the **McUeNetDevice** class is extended in order to support carrier aggregation.

At the MAC layer, in fact multi connectivity is usually achieved with CA, which is already widely used in LTE-Advanced networks.

Different CCs can use different frequencies, and can be adapted to the channel independently (i.e., use different Modulation and Coding Schemes (MCSs), and/or retransmission processes), CA increases the available datarate for the user, since it aggregates the spectrum across multiple bands.

In the implementation CA has a single RLC for each bearer (independently of the number of CCs) and can support joint scheduling at the MAC layer across the different carriers, thus, until the bearer data is actually scheduled on one of the available carriers, in principle it could be transmitted on any of them.

Fig. 3.8 shows a simplified UML diagram of the integration of the DC and CA implementation for an **McUeNetDevice** and an LTE and a mmWave base stations.

It can be seen that the mmWave and the LTE CA implementations are used respectively in the **MmWaveEnbNetDevice** and **LteEnbNetDevice** classes, while they coexist in the **McUeNetDevice**.

The example in Fig. 3.9 shows two CCs per RAT, but it is actually possible to configure independently the number of CCs in the LTE and mmWave RATs. Then, given that different RRCs are in control of the LTE and mmWave links, it is possible to set the carriers after the UE has attached to either of the two RATs, using RRC connection reconfiguration messages.

The basic class of the CA implementation is the `MmWaveComponentCarrier` class and its `MmWaveEnbComponentCarrier` and `MmWaveUeComponentCarrier` extensions. An instance of this class represents a single carrier, and contains pointers to the associated protocol stack layers and relevant configurations. In particular, in the implementation, a `MmWaveComponentCarrier` object contains a reference to a `MmWavePhyMacCommon` object, which is used to specify the numerology, frequency and bandwidth information for the carrier.



Figure 3.9 (a) Information represented by instances of `MmWaveComponentCarrier` and extensions
(b) UE protocol stack for the integration of multi-RAT dual connectivity and carrier aggregation.

The different `MmWaveComponentCarrier` objects in the UEs and base stations are managed by a single CC manager, i.e., an object that implements respectively the `LteUeComponentCarrierManager` or the `LteEnbComponentCarrierManager` interfaces.

A typical use case for CA in the mmWave band would be the aggregation of a CC at relatively low carrier frequency, with a smaller bandwidth, but with better propagation properties (i.e., lower pathloss), and other CCs at much higher frequencies with larger bandwidths.

Another important aspect is the usage of different channel model objects for the different carriers.

Therefore, we use different `MmWave3gppChannel` objects for each carrier, and use the `MmWavePhyMacCommon` of the carrier to set up the necessary parameters.

3.2.2.5 Initial dual connected access and handover/switch procedure

Initial access is the procedure by which a mobile UE establishes an initial physical link connection with a cell, a necessary step to access the network.

In the ns-3 module the initial dual access procedure complete the setup of the remote RLC layer in the mmWave eNB, and of the associated RLC in the UE, and to the end the Initial Access with a switch to the mmWave RAT. In the Fig. 3.10 is represented all the procedures.

A UE first searches for synchronization signals from conventional 4G cells. This detection is fast since it can be performed omni-directionally and there is no directional scanning.

Under the assumption that the 5G mmWave cells are roughly time synchronized to the 4G cell, and the round trip propagation times are not large, an uplink transmission from the UE will be roughly time aligned at any closeby mmWave cell.

A UE desiring initial access thus broadcasts a random access preamble (RAP) scanning different angular directions.

Each of these RAPs will arrive roughly time-aligned in the random access slots of all potential neighboring mmWave cells. The mmWave cells will scan for the presence of RA preambles, and when the RA preambles are detected, the LTE Cell performs the best attachment decision, based on the received RTs, feeding back the choice to the UE through the 4G-LTE link.

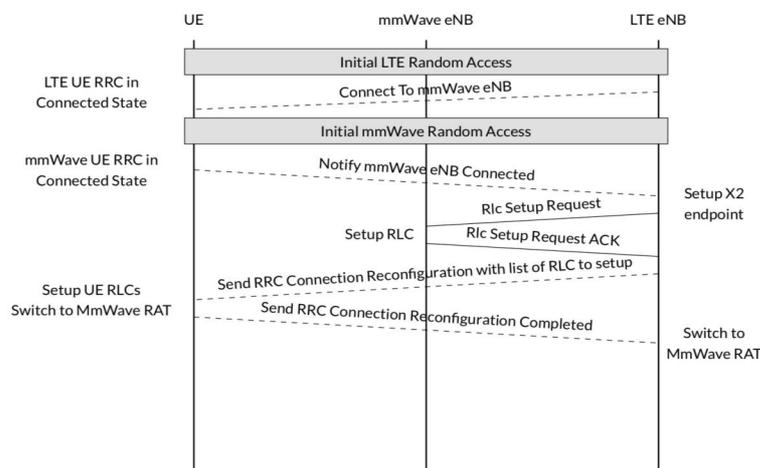


Figure 3.10 Initial dual connected access

In our mmwave network we use dual connectivity with Fast Switching mechanism, in which both the LTE and the 5G radios are in connected mode, but only one of the two at a time is actually used for transmission. If the quality of the signal on the link that the UE is currently using degrades too much, we have a switching procedure.

The fast switching procedure is used when all the mmWave eNBs for a certain UE are in outage. Since the handling of the state of the user plane for both the mmWave and the LTE RATs is carried out by the LTE RRC, it is possible to correctly modify the state of the PDCP layer and perform a switch from the mmWave to the LTE RAT. The proposed switch procedure, simply requires an RRC message (RRC Connection Switch command) to the UE, sent on the LTE link, and a notification to the mmWave eNB via X2 if the switch is from mmWave to LTE, in order to forward the content of the RLC buffers to the LTE eNB.

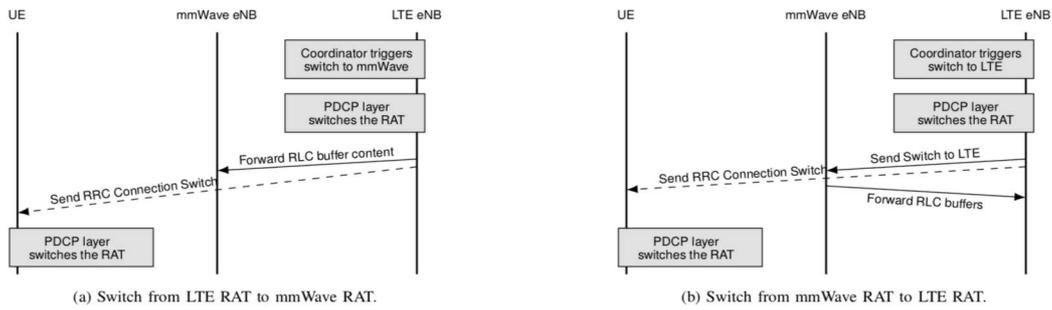


Figure 3.11 (a) Switch from LTE RAT to mmwave RAT (b) Switch from mmwave RAT to LTE RAT

The DC solution therefore allows to have an uninterrupted connection to the LTE anchor point. However, it is possible to switch from a secondary mmWave eNB to a different mmWave eNB with a procedure which is faster than a standard intra RAT handover, since it does not involve the interaction with the core network. i.e., the path switch message that a classic handover procedure has to send to the MME is not needed, since the bearer has a single endpoint in the LTE eNB.

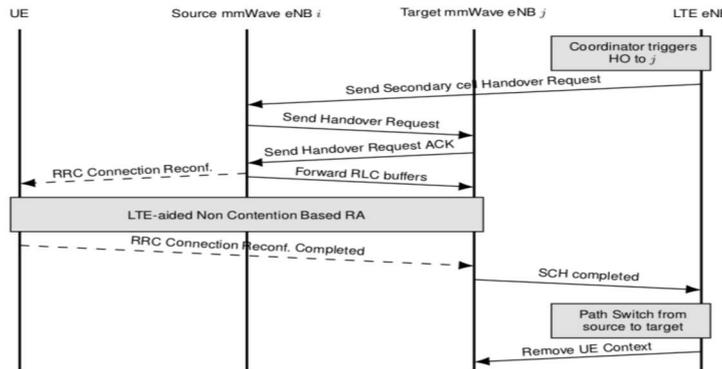


Fig. 3: Secondary cell Handover procedure (SCH).

Figure 3.12 Secondary cell Handover procedure(SCH)

In particular the switch/handover mechanism is very simple;The coordinator is placed in the LTE eNB.

In particular the LTE eNB receive a report table from a given gNB where is stored the SNR's measured between the gNB and all the UE's(considering the best direction).The LTE coordinator based on that received report update a complete report table(CRT) where for each pair UE-gNB is associated the related SNR.When all the entries are updated the LTE coordinator select the best gNB for each UE. Now we can describe the swith/handover algorithm used in this thesis;the parameters are the SNR threshold Δ_{LTE} and the histeresis Δ_{his} .

Consider a U_e connected to a given LTE gNB,and currently attached to a given gNB with a current SNR(Γ_{curr});

In the updated CRT I search for the best gNB(Γ_{opt}) for the UE in term of SNR,Let's call $\Delta = \Gamma_{opt} - \Gamma_{curr}$ the gain in SNR to change gNB.

if $\Gamma_{opt} < \Delta_{LTE}$, we switch to LTE stack or continue to use it.

Otherwise if the $\Gamma_{opt} > \Delta_{LTE}$ we can have three cases:

-if the current attached gNB is not already the optimal one and the $\Delta > \Delta_{his}$ and we are using the mmWave stack, we switch initially to LTE one (to guarantee service continuity), then we trigger the switch to the optimal mmWave.

-Otherwise if the current attached gNB is not already the optimal one and the $\Delta > \Delta_{his}$ but we are using the LTE stack we trigger directly the switch to the optimal mmWave.

-In the case in which the current attached mmWave is already the optimal and we are using the LTE stack, we trigger a switch to the mmWave gNB.

3.2.2.6 Bug and problems in the ns-3 mmWave NY module

-the handover between LTE Cells is not implemented.

-A bad channel condition during a handover/switch procedure is not handled so it causes a crash in the simulation.

- scalability issue with respect to the size of the simulated scenario (number of nodes, buildings, application rate,....).

- the Ns3 simulator does not contain libraries to handle polygons as obstacles (problems solved by introducing a new module in ns3 package created by the MIT university).

4

Ns3 Simulator and SUMO

In this thesis we make use of ns-3 simulator (in particular the NY mmwave module of Ns3) as previously described and SUMO; SUMO is an open source, highly portable, microscopic and continuous road traffic simulation package designed to handle large road networks.

4.1 Luxembourg SUMO Traffic (LuST) Scenario

The first step is to use a vehicular traffic simulator with an appropriate scenario in order to reproduce realistic mobility patterns.

In the interest of building a realistic scenario, we decided to start from a real city with a standard topology common in mid-size European cities, and real information concerning traffic demands and mobility patterns. We use the Luxembourg SUMO Traffic (LuST) Scenario (Fig. 4.2). The traffic demand over a day of the running (R) and waiting (W) vehicles at each given time is represented in Figure 4.1. In this scenario for our purpose we extract information about roads (of any kind), traffic lights, locations of bus stops; another retrieved information was the buildings that are represented as polygons.

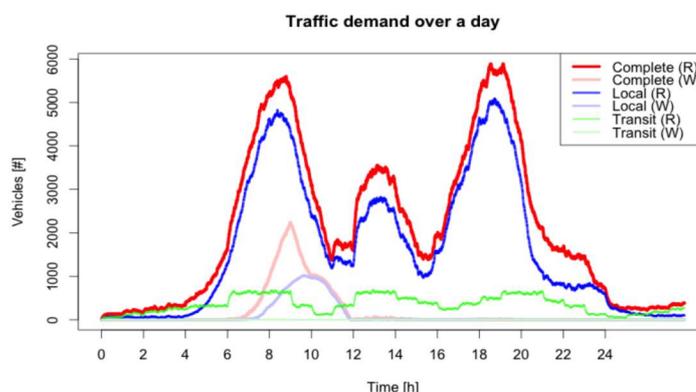


Figure 4.1 Traffic Demand over a day

In particular we are interested in the position and behaviour of the single vehicle and how this impacts on the communications and the related applications and on the position of the buildings that acting as obstacle change the communication channel.

For this purpose the vehicular mobility has to be coupled with a network simulator, therefore a micromobility simulator such as SUMO is appropriate for our needs.

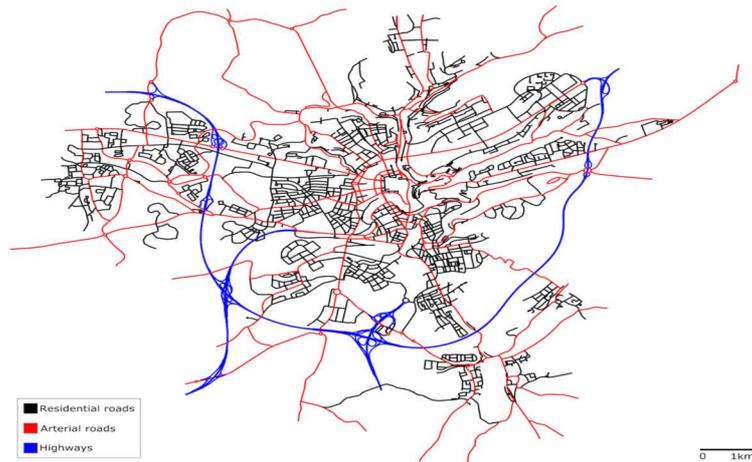


Fig. 1. LuST Scenario Topology.

Figure 4.2 Lust Scenario Topology

Generating mobility traces

SUMO trace files can be generated with the help of a .cfg file (configuration file), as follows:

```
$sumo -c first.sumo.cfg -fcd-output trace.xml
```

Traces of the vehicles' position are dumped into the XML file. These SUMO traces can be used to provide mobility to nodes in the ns-3 simulator. The ns-3 simulator provides the ns-2 mobility helper class, which can be used to read the movements of nodes from trace files. To make use of the SUMO traces generated from the SUMO simulator, the XML file needs to be converted to a TCL file, which then is provided to the ns-2 mobility helper. SUMO provides tools like traceExporter.py and traceExporter.jar to convert SUMO trace to different file formats.

```
$python $SUMO_HOME/tools/traceExporter.py -fcd-input  
trace.xml -ns2mobility-output mobility.tcl
```

mobility.tcl generated above contains the position and the speed of the nodes (vehicles).

With SUMO-GUI which is basically an extension of SUMO we can simulate a scenario by a graphical user interface.

In the following picture is represented the whole process(Fig. 3.2)

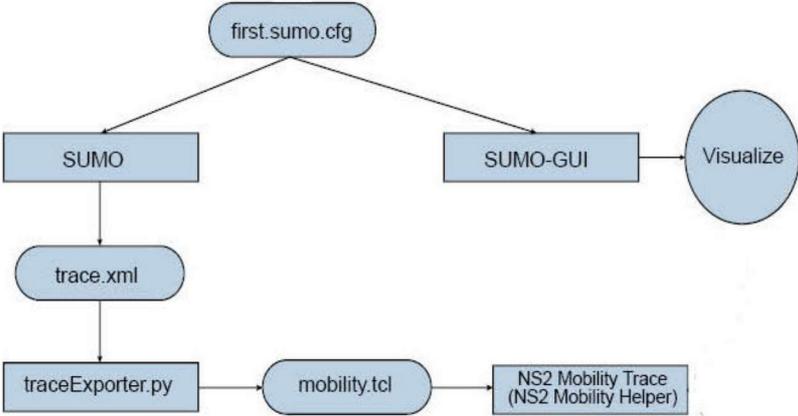


Figure 4.3 Scheme of the whole the process

For what concerns instead the buildings,we said they are rappresented as polygons in the SUMO scenario,we extracted the coordinates of this polygons and provides to the ns3-simulator as obstacles to test the LOS/NLOS condition to be used in the building propagation model(3GPP). Another important information we extracted are the coordinates of the semaphores and of the busstop to be used as potential locations of the antennas.

Simulation Results And Performance Analysis

5.1 Simulation

5.1.1 Introduction

The simulations that will be obtained in this chapter use the ns-3 simulator with the NY mmWave module described in Chap. 3 and Chap. 4.

In particular we analyze the behaviour of a mmWave vehicular network;

We consider a portion of Luxembourg city and we locate a different number of gNB's. For this purpose, we exploit some predefined locations that are semaphores and bus stops; We use an intuitive positioning of the gNB's, starting to consider first all the semaphores (crossings) that are traversed by the largest amount of vehicles and then, in decreasing order the other semaphores, last we consider the bus stops taking into account always the traffic.

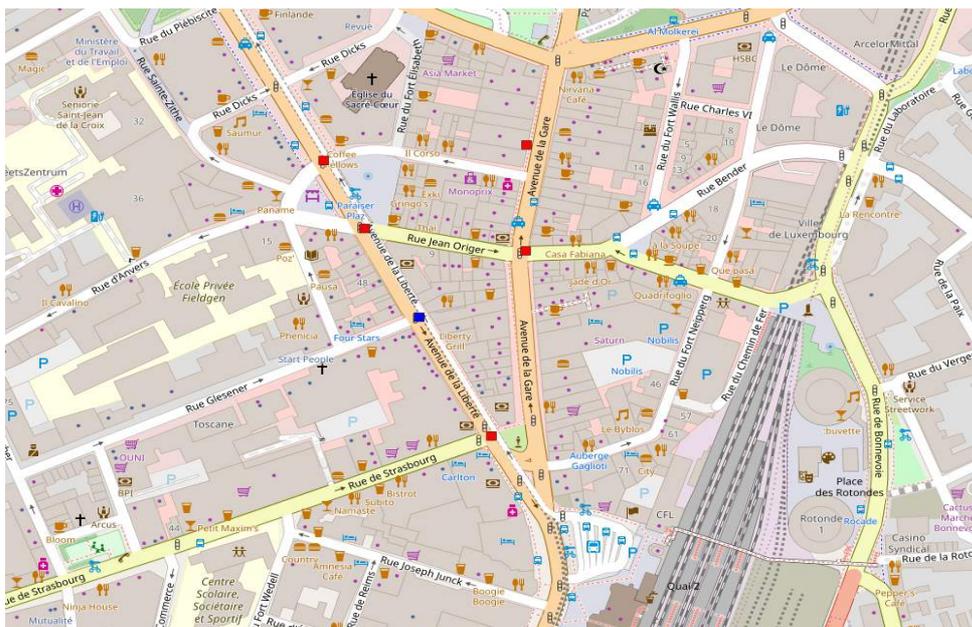


Figure 5.1 Map of the considered area and the relative placement of the antennas

We see in the **Figure 5.1** the placement of the antennas, the red squares represent the gNB's and the blue one represent the LTE eNB.

Since mmWave cells are expected to have a coverage radius of at most 200 m, they will be deployed with a density higher than that of LTE cells (which are already installed) .

This simulation we test the dual connectivity mmwave with fast switching technology(FS);

The simulations are performed with two different RLCs, AM and UM;

Then, in order to get the mean throughput PDCP per user , we account the received packets per user averaged over the total simulation time, then we averaged the user's throuputs.

The PDCP throughput is mainly a measure of the rate that the radio network can offer, given a certain application rate. In this particular scenario after some tests we saw that in ns-3 we cannot increase too much the application rate because of the scalability issue;

Since we a UDP packet every one ms ,the UDP application rate at which every vehicle should receive a new packets at PDCP is $S_{in} = 8.432$ Mbit/s. Using RLC AM If the PDCP throughput S_{PDCP} is higher than S_{in} , it means that the contribution given by unneeded retransmissions (that increase the throughput) is more significant than that of packet losses (that decrease the throughput), and vice versa if $S_{PDCP} < S_{in}$. The number of unnecessary retransmissions decreases as the number of switches decreases(we explain later).

The NYU mmWave module provides a antenna model for a Uniform Linear Array (ULA) in the *AntennaArrayModel* class, that supports analog beam forming.In the NS-3 framework, analog BF is implemented through Uniform Linear Array (ULA) .

In ULA, a set of two dimensional antenna arrays is used at both BS and UE. The array can be comprised of 8×8 , 4×4 or 2×2 elements. The spacing of the elements is set at $\lambda/2$, where λ is the wavelength. (for instance, at 28 GHz, a 4×4 array will have a size of roughly $1.5 \text{ cm} \times 1.5 \text{ cm}$). It must be noticed that there exists a strong correlation between beamwidth, number of antenna elements and BF gain. we can see the variation of beamwidths and gains by varying the number of elements in the ULA;The more antenna elements in the system, the narrower the beams, the higher the BF gain, the more precise and directional the transmission.For an 8×8 array beams can be at most 22.5° wide;for 4×4 system,they are 45° wide;for a 2×2 configuration,beams are 90° wide.

We concentrate our analysis on the average throughput perceived by the vehicles for a different number of mmwave gNB and changing their beamwidth;Let's see first the simulation parameter that we adopt to get the results.

5.1.2 Simulation Parameter

Simulation time	50 seconds
DL-UL Ptx	30dbm
UDP application data rate	8.432 Mbit/s
Frequency Band LTE	20MHz
Frequency Band mmwave	1GHz
LTE carrier Frequency	2.1GHz
Mmwave carrier Frequency	28GHz
Area Radius	200m
Mobility trace size	34 vehicles
Beam type	analogic
Δ_{LTE}	-5dB
Δ_{hys}	3dB
B_{RLC}	20MB
D_{X2}	0.5ms
D_{MME}	10ms

Table 5.1 Simulation Parameters

$\hat{\phi}$

The B_{RLC} represent the buffer size of the RLC entity, the D_{MME} is the delay to the MME node and model both the propagation delay to the MME node and processing delay of the MME node. The D_{X2} represent the one-way delay of the X2 link; We've already discussed about the other parameters.

5.1.3 Average PDCP throughput and jain's fairness index analysis with RLC UM and ULA 4x4

The first simulations are performed with RLC UM (Unacknowledged mode in RLC is meant for delay sensitive applications such as VOIP);

At RLC layer no retransmission is performed, and packets are simply declared lost (even if a single segment of the entire packet is missing). Anyway the MAC layer offers a retransmission mechanism (HARQ), but is limited by a maximum number of retransmissions (in our case 3).

During a switch/handover the packets in the RLC buffer are not forwarded to the target eNB, and thus are lost.

The simulation outcomes is shown in the **figure 5.2**, we used a beamwidth of 90° both at the gNB's and at the UE's;

We have the bigger step passing from one to two mmwave gNB's until we reach the maximum at five gNB's; We see that using five gNB's we reach a throughput very close to the UDP application rate, so this could be an optimal design choice in this specific scenario.

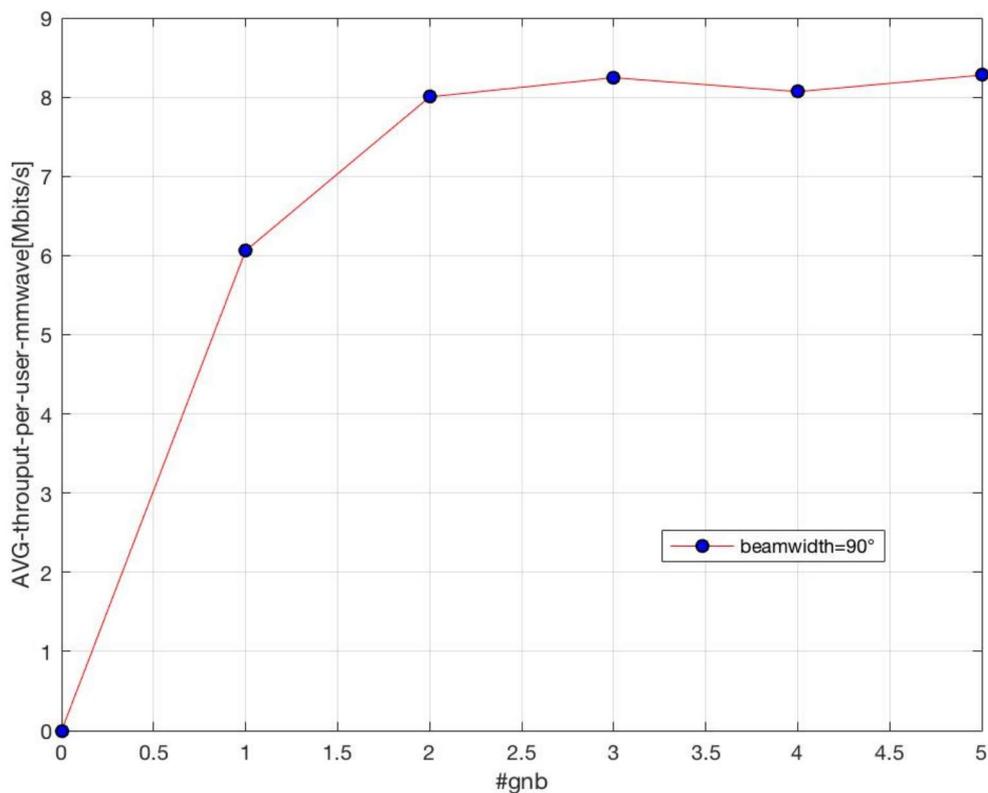


Figure 5.2 Average throughput per user as a function of the numbers of gNB's with RLC UM and ULA 4x4

As we see from the plot, it can happen that adding one more gNB we perceive a smaller throughput, this can be due to the fact that we have an increase of interference that translates in a lower SNR, so we receive more packets with error to be discarded and with high probability we increase the number of handover and/or switch, this could decrease the mean throughput.

For the same simulated scenario we can analyse the Jain fairness index plotted (**Fig. 5.3**), we see that with a number of gNB's greater or equal than two we get a value very close to one (best case);

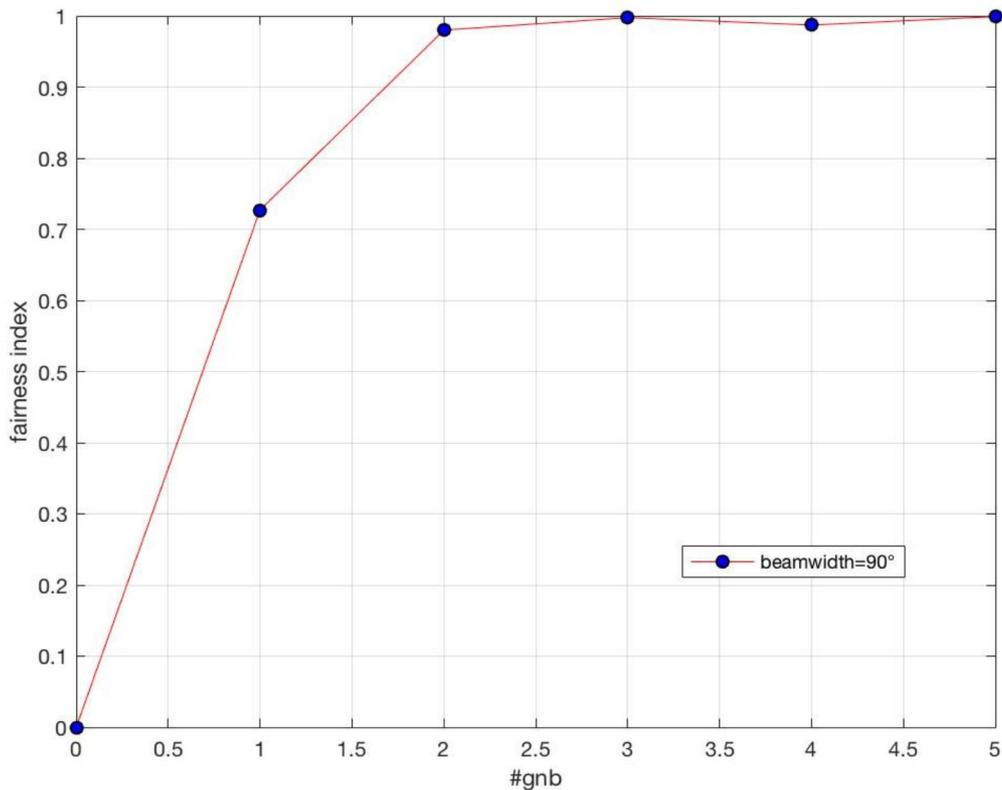


Figure 5.3 Jain's fairness index as a function of the numbers of gNB with RLC UM and ULA 4x4

5.1.4 Short term throughput analysis on the map with RLC UM and ULA 4x4

Let's see some maps where we show the nearly instantaneous mmWave throughput (over a interval of one second) perceived by the vehicles correlated to their positions that is represented the **Figure 5.4** with a colourbar from 0 to Tmax(application rate previously called Sin) and the percentage of mmWave traffic with respect to the LTE one (**Fig. 5.5**). The green polygons in the map represent the mmWave cell towers, the pink one represent the LTE cell tower.

The above description is valid for all the images in this chapter regarding the short term throughput analysis with UM setup.

So we see the evolution of the maps throughput increasing the number of gNB's.

Let's see the map with one gNB:

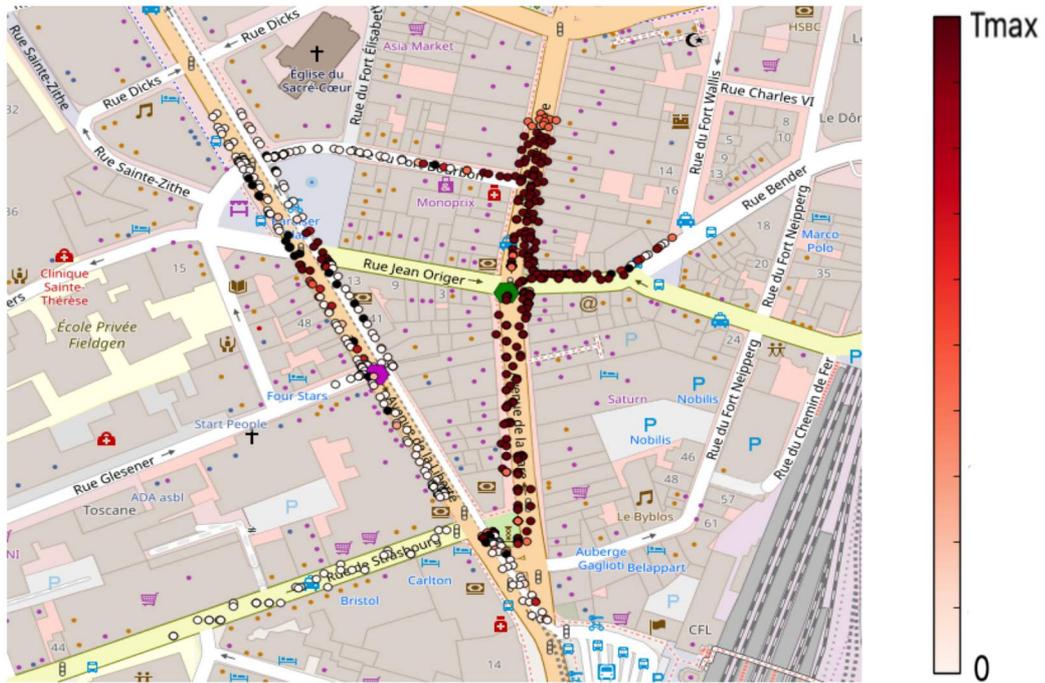


Figure 5.4 Colourmap with short-term throughput indication per user using one gNB with RLC UM and ULA 4x4

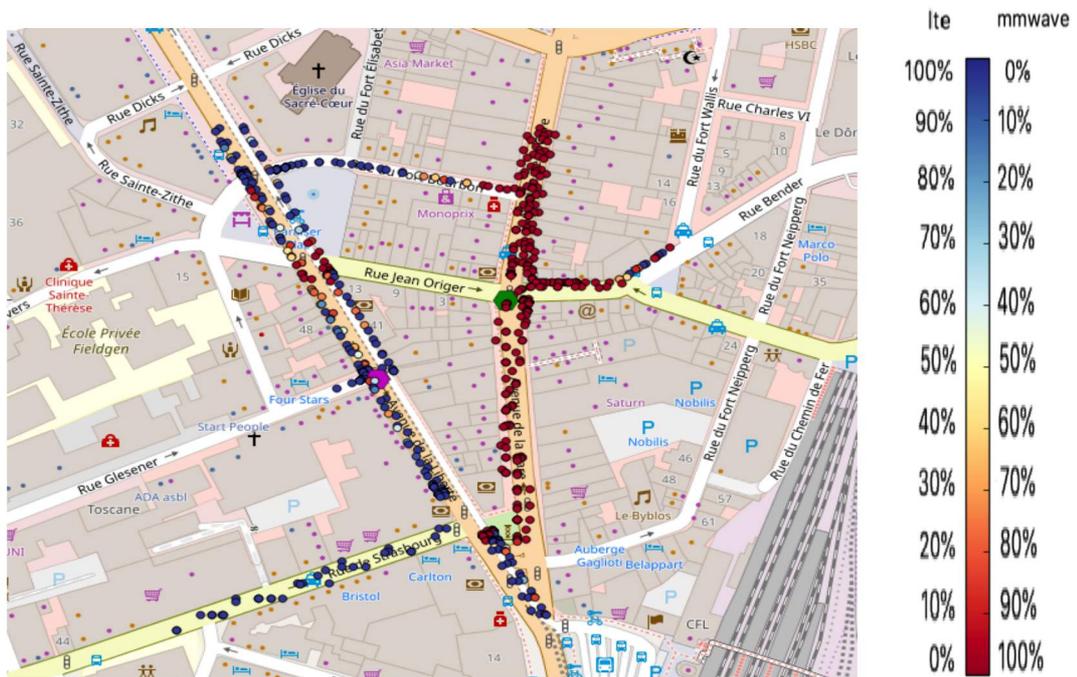


Figure 5.5 Colourmap with percentage of short-term throughput LTE and mmwave per user using one gNB with RLC UM and ULA 4x4

We see that the mmwave cell coverage can be very relevant (over 200m in LOS condition) and another important observation is that the contribution to the mmwave throughput given by the NLOS communications (through reflections) is not negligible.(Fig 5.4). Let's see what happens with two gNB:



Figure 5.6 Colourmap with short-term throughput indication per user using two gNB with RLC UM and ULA 4x4



Figure 5.7 Colourmap with percentage of short-term throughput LTE and mmwave per user using two gNB with RLC UM and ULA 4x4

We see that adding one more gNB the areas where we make use of the LTE RAT is very limited in some NLOS areas from the mmwave perspective(Fig 5.7).

Adding one more gNB:



Figure 5.8 Colourmap with short-term throughput indication per user using three gNB with RLC UM and ULA 4x4



Figure 5.9 Colourmap with percentage of short-term throughput LTE and mmwave per user using three gNB with RLC UM and ULA 4x4

with four gNB:

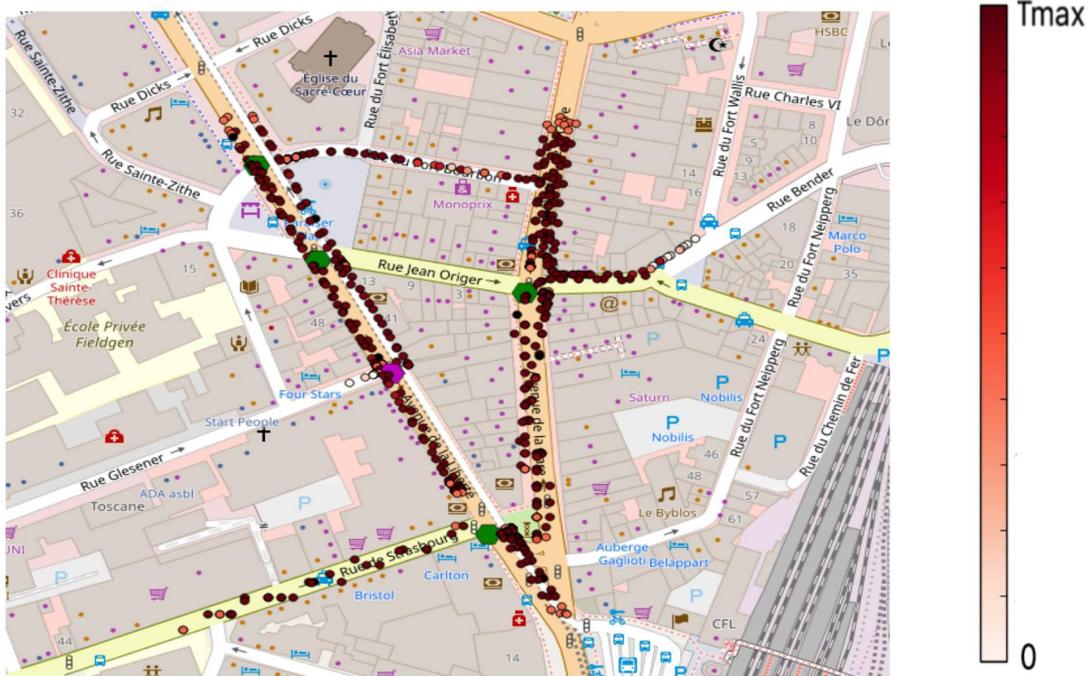


Figure 5.10 Colourmap with short-term throughput indication per user using four gNB's with RLC UM and ULA 4x4

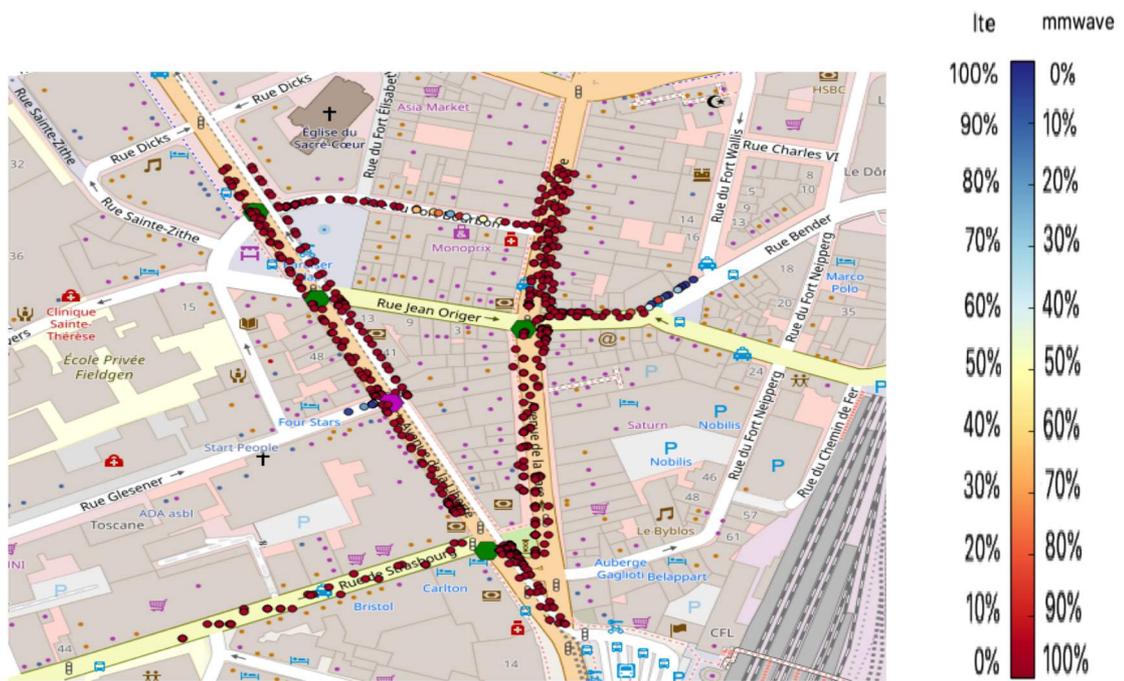


Figure 5.11 Colourmap with percentage of short-term throughput LTE and mmwave per user using four gNB's with RLC UM and ULA 4x4

Now we see that there's small difference respect to the previous case, basically almost all the traffic is mmWave , the use of LTE RAT is negligible and we see from the **Figure 5.10** through the colorbar that the perceived throughput is very high in almost all the area.

5.1.5 Average PDCP throughput analysis and jain's fairness index with RLC UM and ULA 16x16

The second simulated scenario is shown in **Figure 5.12**, we used a beamwidth of 45° both at the gNB's and at the UE's; Again we have the bigger step passing from one to two gNB's, the maximum is around two, where the perceived throughput is close to the application rate and a similar performance around four gNB's; In this case using two gNB's we reach a throughput performance very similar to the case with ULA 4x4 and five gNB's , for deployment cost it's preferable ULA 16x16.

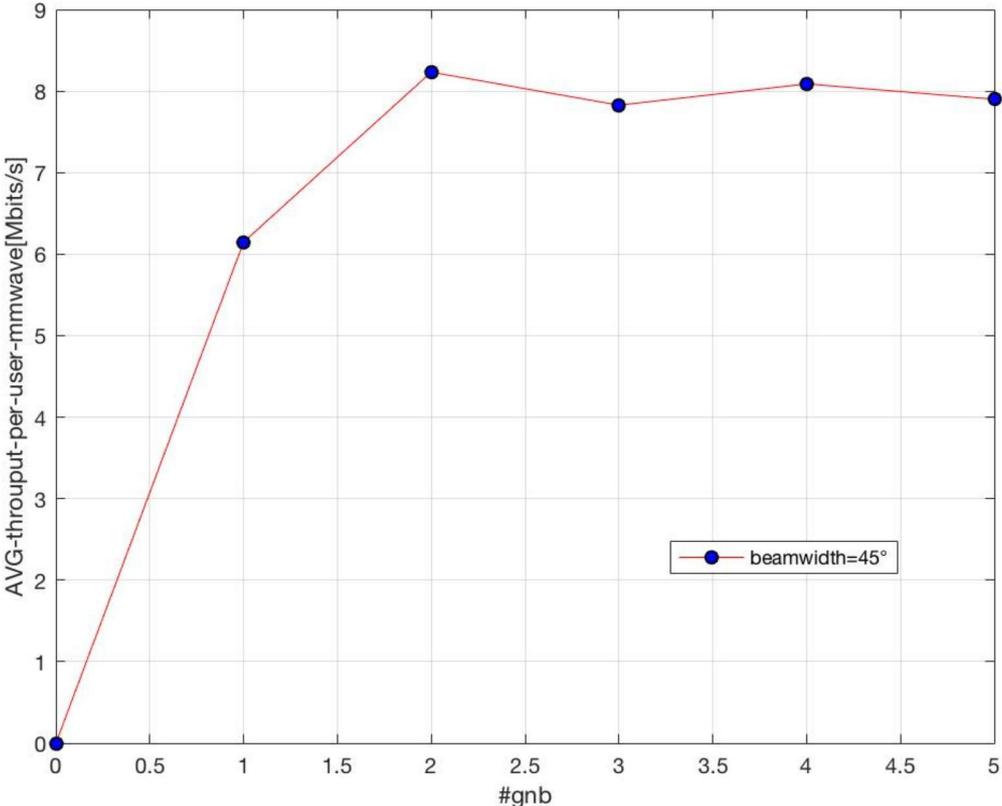


Figure 5.12 Average throughput per user as a funtion of the numbers of gNB's with RLC UM and ULA 16x16

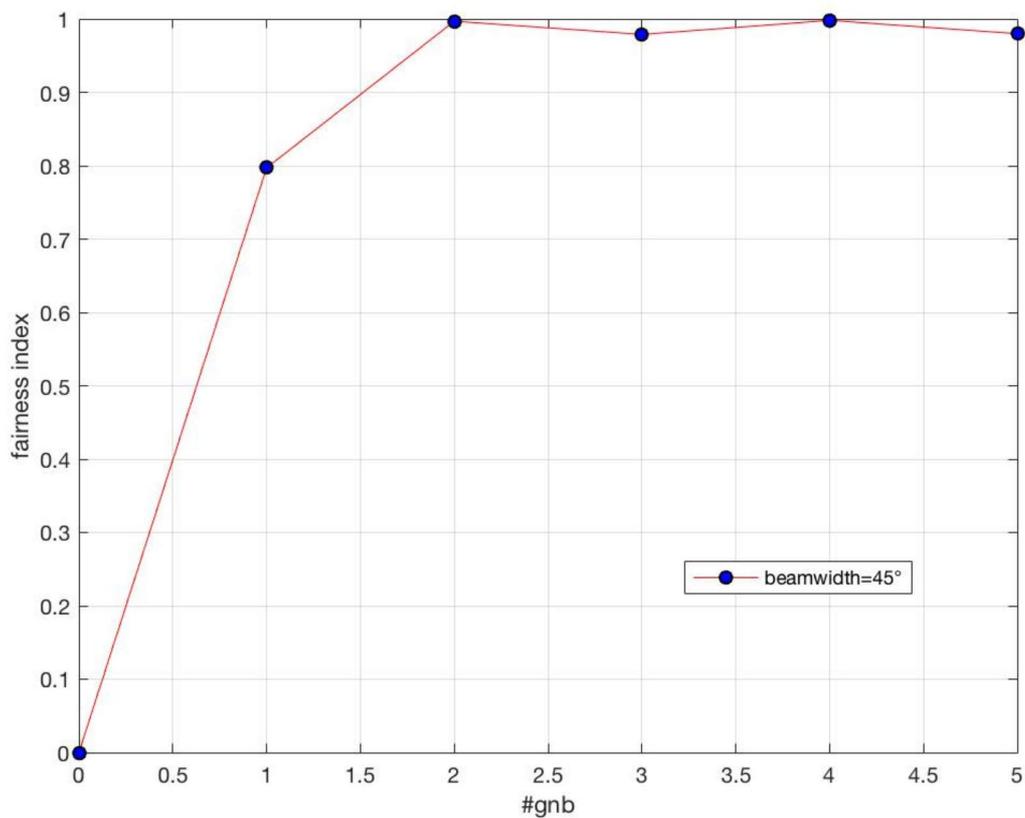


Figure 5.13 Jain's fairness index as a function of the numbers of gNB with RLC UM and ULA 16x16

The Jain's fairness index is very similar to the previous simulated scenario (**Fig. 5.13**).

5.1.6 Short term throughput analysis on the map with RLC UM and ULA 16x16

We see again the short term throughput that is perceived by vehicles correlated to their positions on the map, now shrinking the beamwidth respect to the previous case.

The maps with one gNB are depicted (**Fig. 5.14** and **Fig. 5.15**)



Figure 5.14 Colourmap with short-term throughput indication per user using one gNB with RLC UM and ULA 16x16

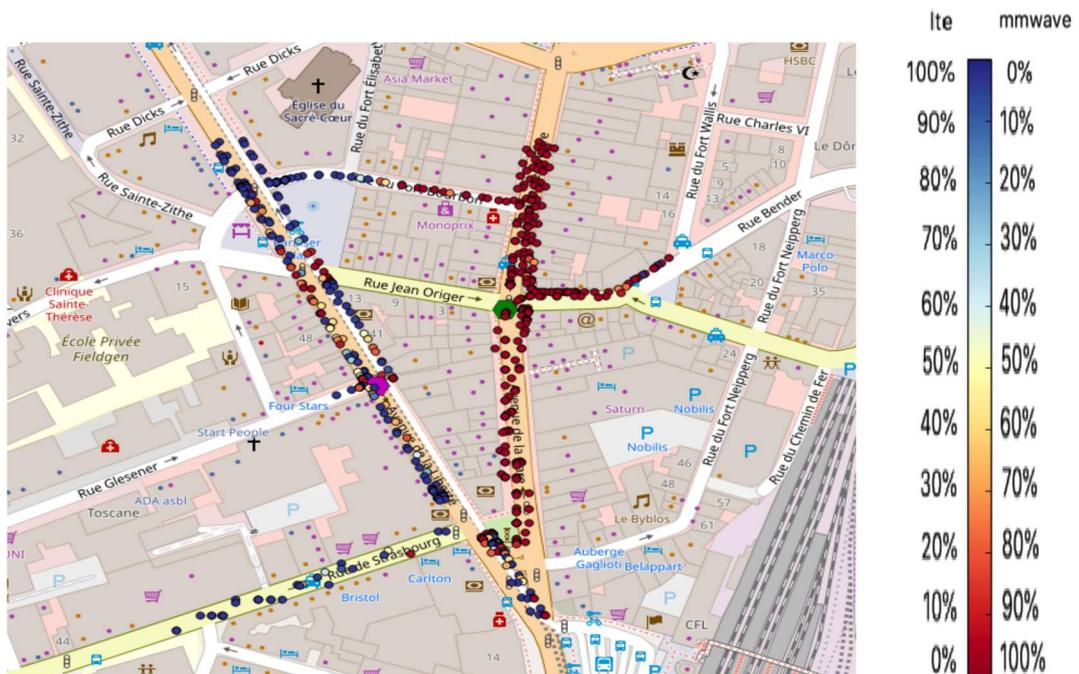


Figure 5.15 Colourmap with percentage of short-term throughput LTE and mmwave per user using one gNB with RLC UM and ULA 16x16

Again with one gNB we have similar radius coverage and similar amount of NLOS communications coming from the mmwave RAT with respect to case with ULA 4x4 and one gNB.

With two gNB's (Fig. 5.16 and Fig. 5.17):



Figure 5.16 Colourmap with short-term throughput indication per user using two gNB's with RLC UM and ULA 16x16

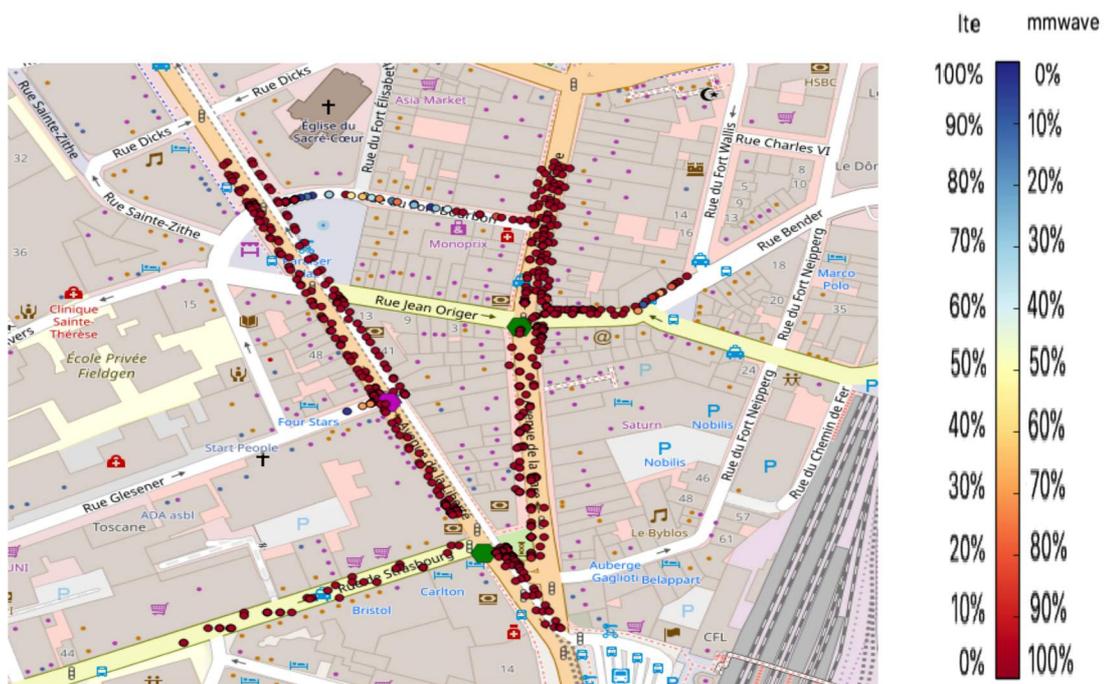


Figure 5.17 Colourmap with percentage of short-term throughput LTE and mmwave per user using two gNB's with RLC UM and ULA 16x16

The map with two gNB's is very similar to the one with four gNB's(almost same mmwave throughput) that is the one with the best overall performance.

5.1.7 Average PDCP throughput and jain's fairness index analysis with RLC UM and ULA 64x64

The third simulated scenario is shown in the **Figure 5.18**, we used a beamwidth of 22.5° both at the gNB's and at the UE's; Again we have the bigger step passing from one to two gNB, the maximum is at four gNB's where the perceived throughput is closed to the application rate, this could be the optimal choice;

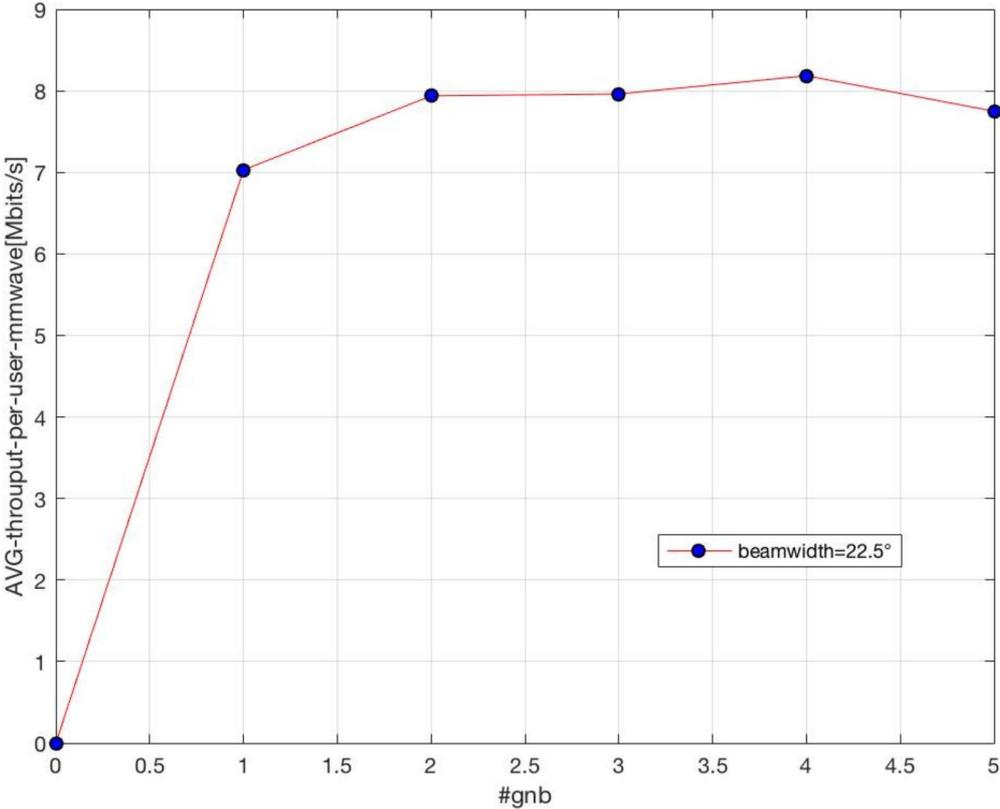


Figure 5.18 Average throughput per user as a function of the numbers of gNB's with RLC UM and ULA 64x64

We plot the jain's fairness index(**Fig. 5.19**):

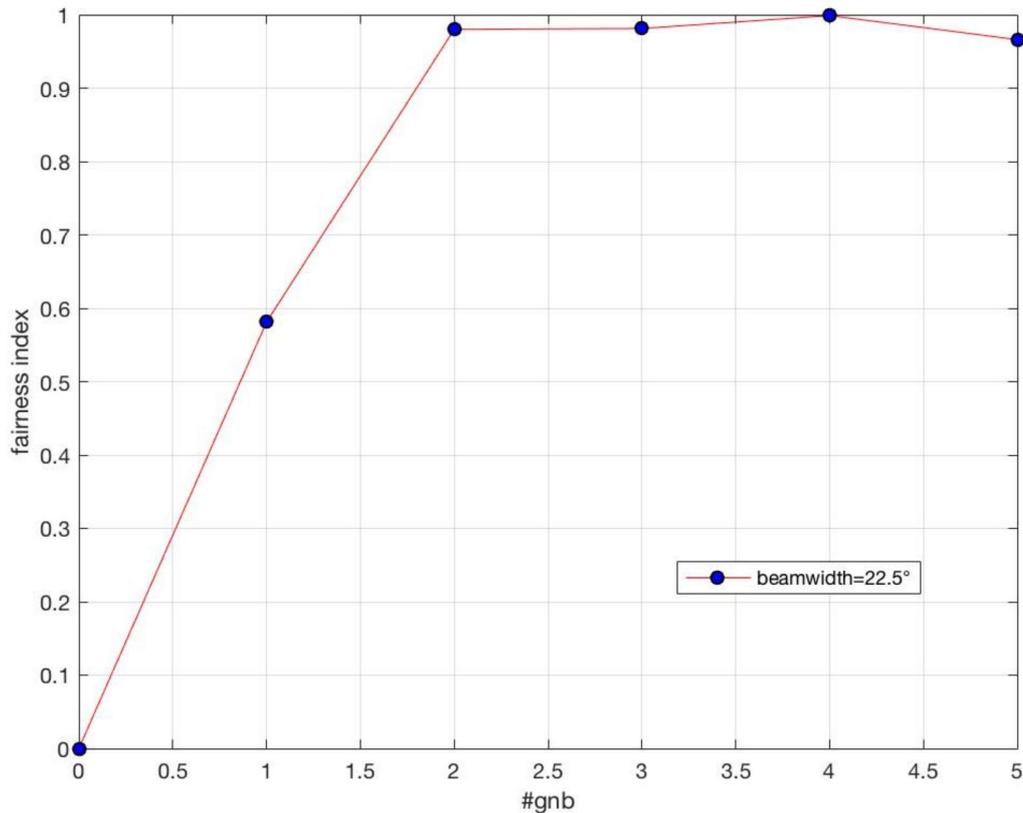


Figure 5.19 Jain's fairness index as a function of the numbers of gNB's with RLC UM and ULA 64x64

Again as before, the case with one gNB is closed to the worst case (very different throughput allocation), with more than two gNB's approximately all the users receive the same allocation.

5.1.8 Short term throughput analysis on the map with RLC UM and ULA 64x64

We see again the short term throughput that is perceived by vehicles correlated to their positions on the map, now shrinking the beamwidth to 22.5°.

The maps with one gNB are depicted (Fig. 5.20 and Fig. 5.21):

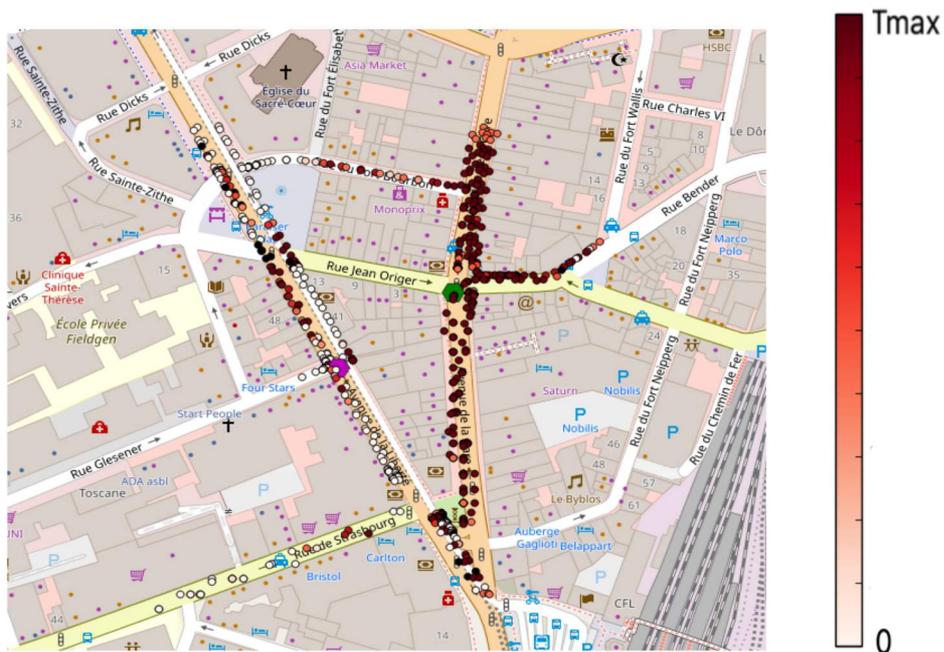


Figure 5.20 Colourmap with short-term throughput indication per user using one gNB with RLC UM and ULA 64x64



Figure 5.21 Colourmap with percentage of short-term throughput LTE and mmwave per user using one gNB with RLC UM and ULA 64x64

Also in this case the map present almost the same characteristics of the previous cases with one gNB in term of areas where we are served by mmwave RAT respect to the LTE one (in this limited scenario), but as expected we have an higher throughput at higher distance

coming from the mmwave RAT respect to the previous case and also the NLOS communications account an higher contribution.

The best perceived throughput is obtained with four gNB's ,so let's see the maps in this case:

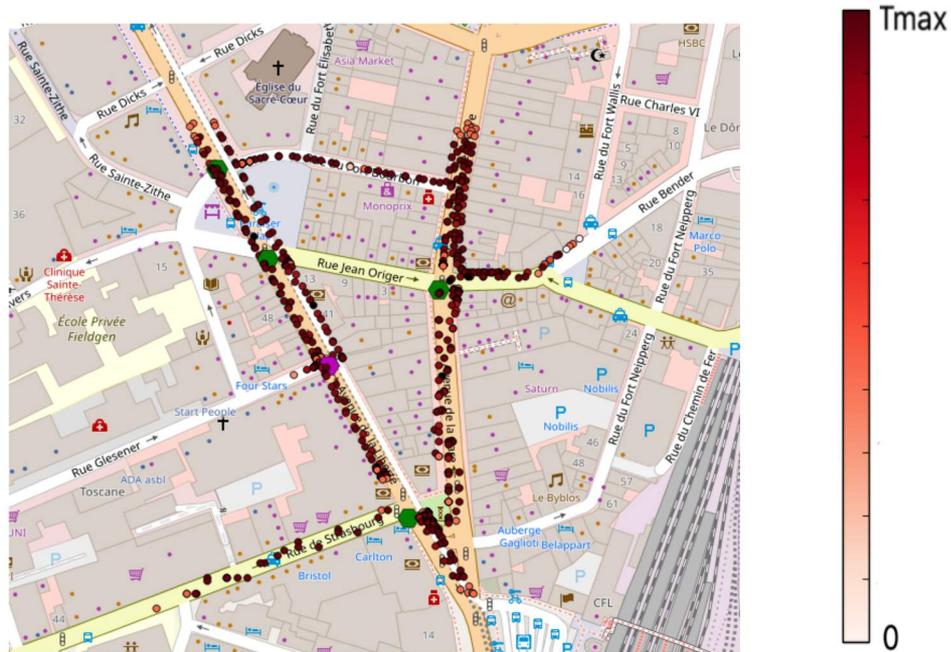


Figure 5.22 Colourmap with short-term throughput indication per user using four gNB with RLC UM and ULA 64x64

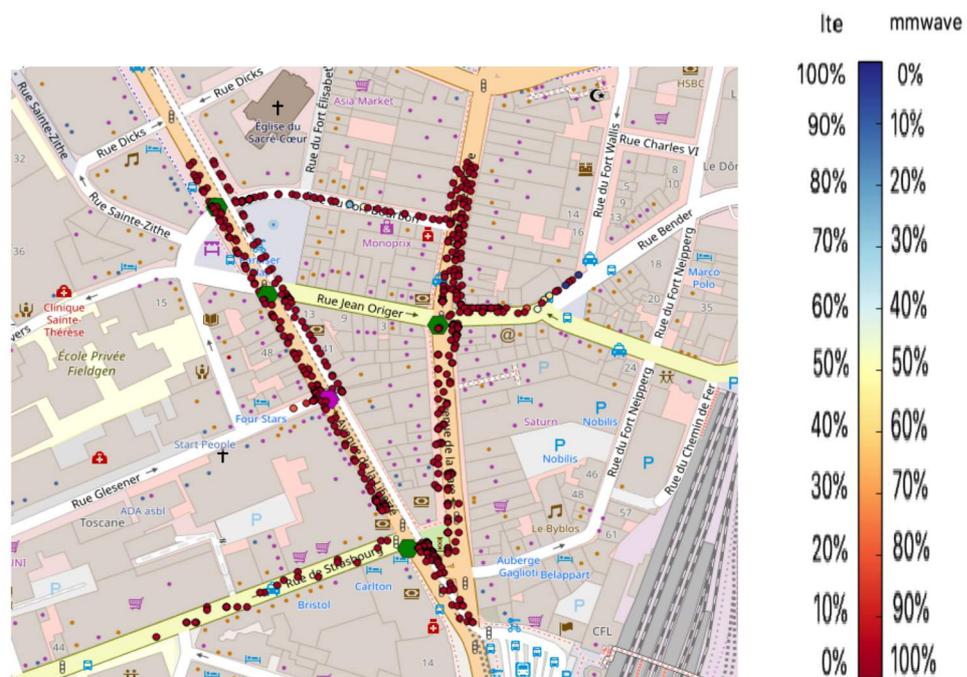


Figure 5.23 Colourmap with percentage of short-term throughput LTE and mmwave per user using two gNB with RLC UM and ULA 64x64

We see from the second image that all the traffic is coming from the mmwave RAT(Fig. 5.23), the first depicted image show an high throughput uniformly all over the map from the mmWave RAT's(Fig. 5.22).

5.1.9 Average PDCP throughput comparison with RLC UM

If we compare the three curves(Fig. 5.24) we see that there's no a clear winner depending on the number of antennas and the beamwidth and the scenario we are studying we experience a different throughputs and therefore a difference design choice; With one gNB we already discussed that with beamwidth of 90° and 45° we get almost the same performance,instead its preferable a beamwidth of 22.5°.

In this specific scenario an optimal choice could be with two antennas and a beamwidth of 45° or 5 antennas with 90° or 4 antennas with 22.5°,all the these three cases experiences a throughput close to the application rate,For cost deployment the second could be the best one.

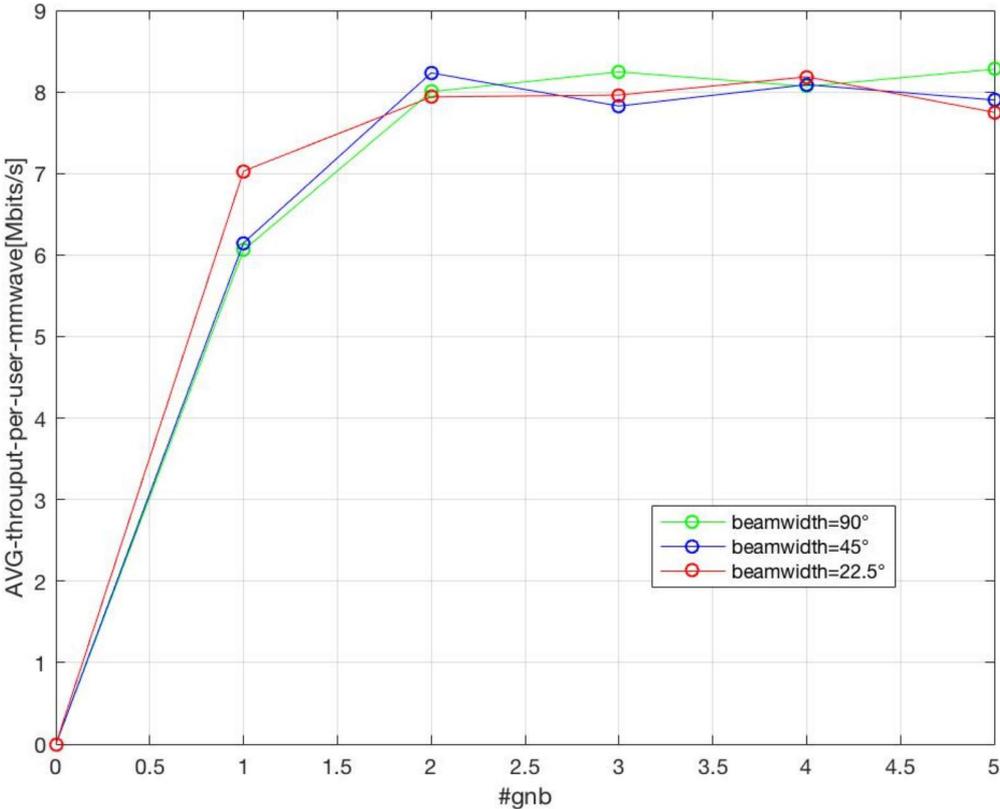


Figure 5.24 Average throughput comparison as a funtion of the numbers of gNB's with RLC UM

In the **Figure 5.25** we compare the jain fairness indexes of the three curves ,we see that with a number of gNB's greater or equal than 2 we get a value very close to one so we have a fair allocation in term of throughput (best case);

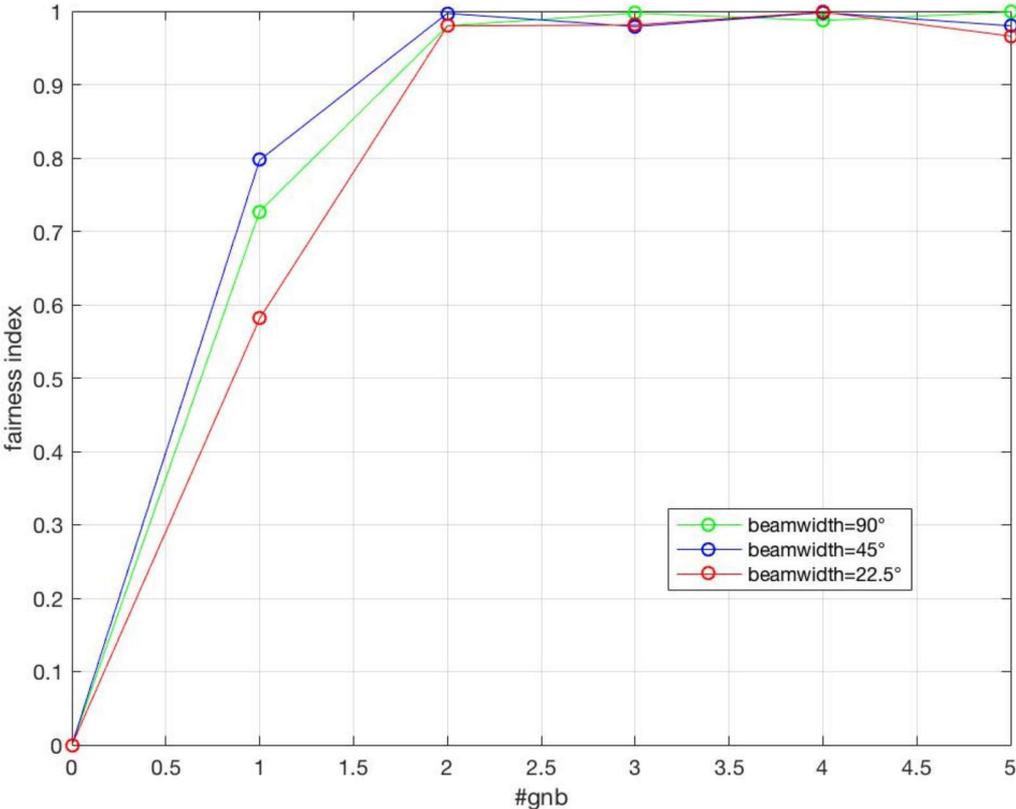


Figure 5.25 Jain's fairness index comparison as a funtion of the numbers of gNB's with RLC UM

5.1.10 Average PDCP throughput analysis and jain's fairness index with RLC AM and ULA 64x64

Now we can see some example with RLC AM setup, Notice that normally the PDCP throughput is mainly made up of the transmission of new incoming packets, but it may also account for retransmission of already transmitted packets. Indeed, in the RLC AM setup, if a packet was transmitted, but not already ACKed, it is stored in the RLC AM retransmission buffer. Then, when a handover or switch happens, the retransmission buffer is forwarded to

the target eNB (RAT) and transmitted again. Therefore, if at the first time it was received successfully, it is wastefully retransmitted.

In the **Figure 5.26** we used a beamwidth of 22.5° both at the gNB and at the Ue; With the NY mmWave module, the UE experiences a lot of switches, because of the simple switch/handover algorithm and of the high variability of the mmWave channel; this phenomena is more remarkable using just one gNB so we experience a lot of switches, this translates in a giant amount of retransmissions this explains the high value of experienced throughput.

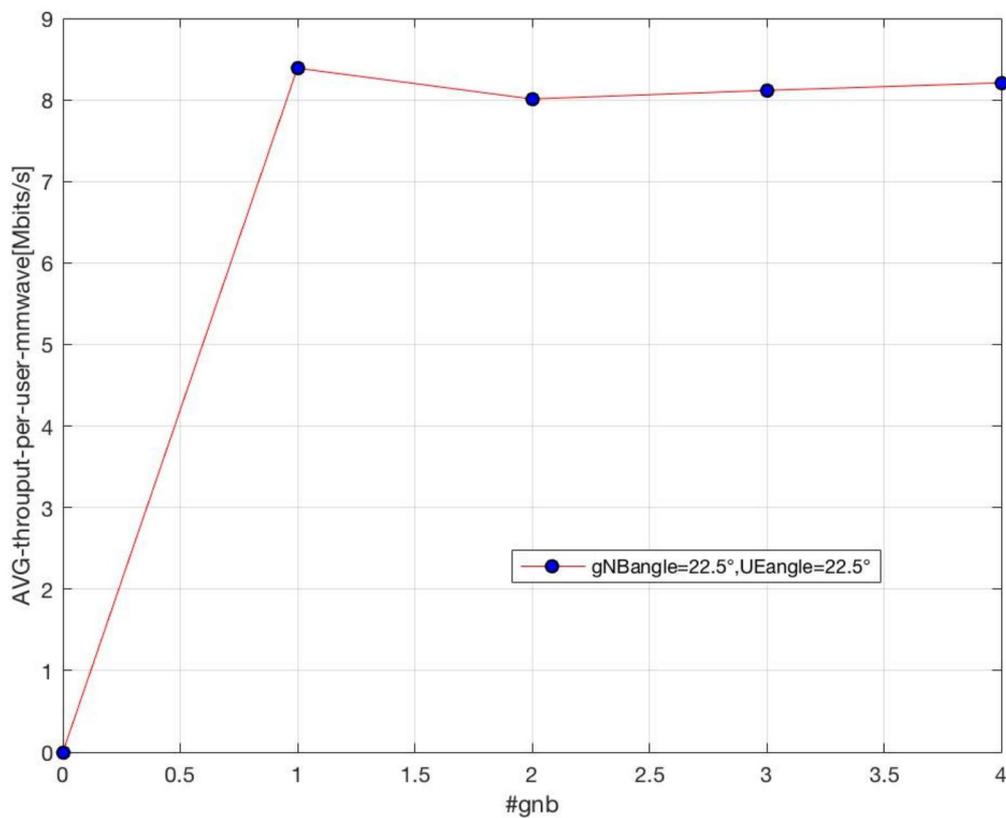


Figure 5.26 Average throughput per user as a function of the numbers of gNB's with RLC AM and ULA 64x64

In the **Figure 5.27** represent the jain fairness index:

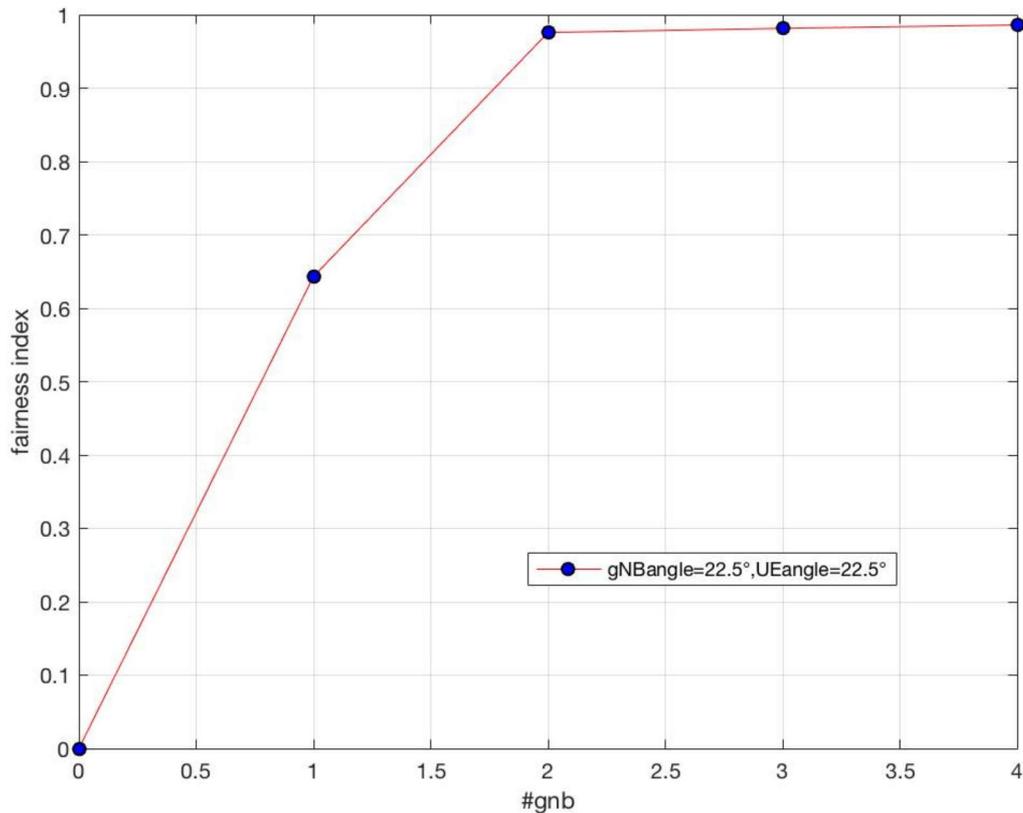


Figure 5.27 Jain's fairness index as a function of the numbers of gNB with RLC AM and ULA 64x64

5.1.11 Short term throughput analysis on the map with RLC AM and ULA 64x64

As we said before with one only gNB we experience a lot of Switches, in fact we see from the map in the black points the throughput overcome the application rate (in this case of a very relevant value), this is due by giant amount of retransmissions, it happens where the UE's are not in LOS with the mmwave tower or where the distance between the two is relevant so in general the SNR is lower. In the **Figure 5.28** we depicted the case with one gNB; we don't draw some points where the throughput that is perceived by the vehicles is 0 from both the LTE and mmwave RAT's.

In the areas where we have a lot of missing points, we expect that the vehicles are served by the LTE RAT, but this is not true; During the switch the target RLC AM transmission buffer receives both the packets sent by the UDP source, and the packets that were in the source RLC buffer, if the source RLC buffer is full, then the target buffer may overflow and discard packets. The LTE RLC buffer is heavily stressed by the giant amount of retransmissions needed in this scenario. Moreover, the high number of Switches lead to large amount of control signalling related to user plane on the LTE connection that increase consistently the traffic. Since the coordinator is co-located that can become critical in the design of the network.

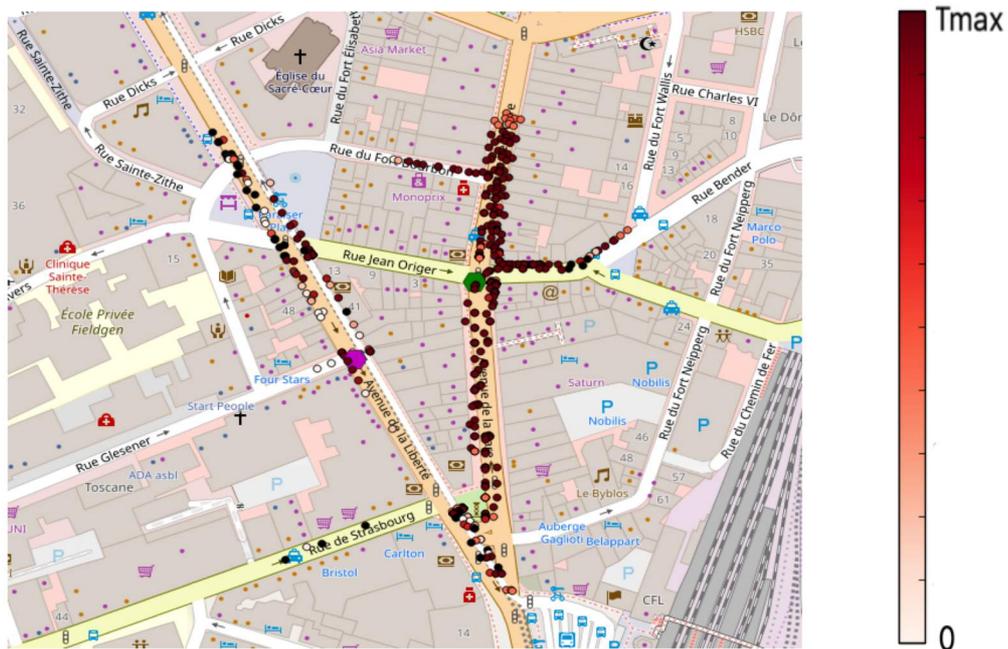


Figure 5.28 Colourmap with short-term throughput indication per user using one gNB with RLC AM and ULA 64x64



Figure 5.29 Colourmap with percentage of short-term throughput LTE and mmwave per user using one gNB with RLC AM and ULA 64x64

The case with four gNB's could be the best design choice in term of perceived throughput, the retransmissions are present as in the case with one gNB but less frequently and now we serve much more vehicles with the mmwave RAT's; In the second Figure we can see that all the vehicles are served by the mmWave antennas, so we have uniformly covered the considered area. (Fig. 5.31)

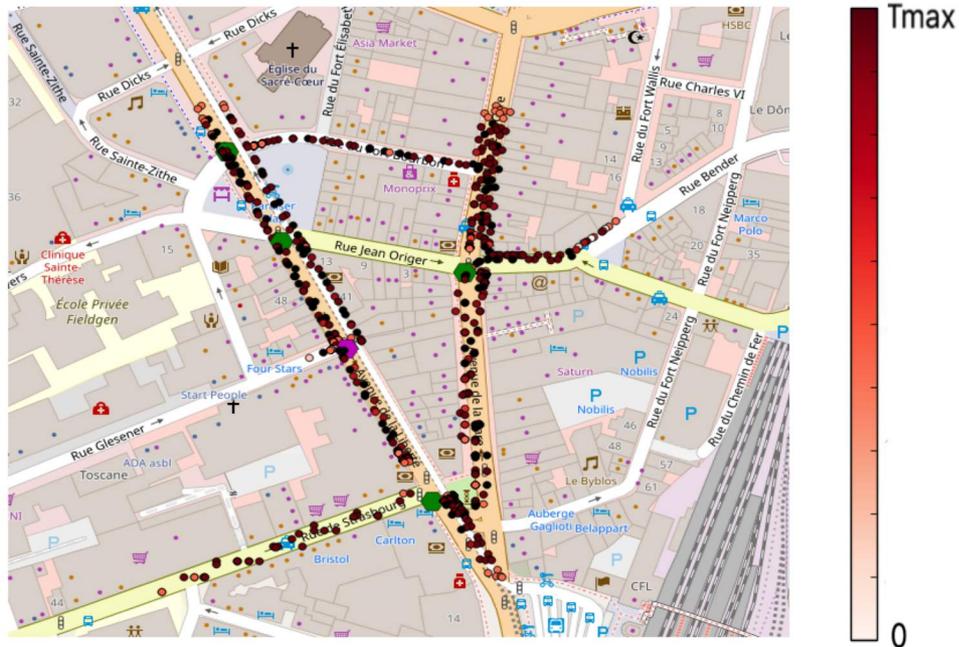


Figure 5.30 Colourmap with short-term throughput indication per user using four gNB's with RLC AM and ULA 64x64

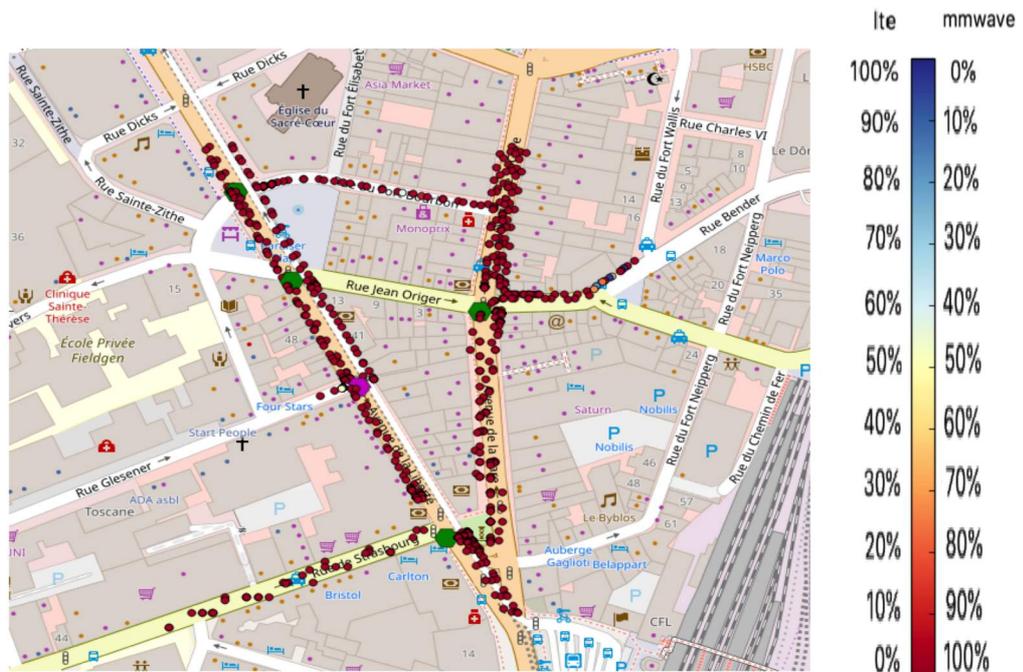


Figure 5.31 Colourmap with percentage of short-term throughput LTE and mmwave per user using four gNB's with RLC AM and ULA 64x64

5.1.12 Average PDCP throughput analysis and jain's fairness index with RLC AM and hibrid scenario

Another simulated scenario is shown in the **Figure 5.32**, we used a beamwidth of 22.5° at the gNB's and 90° at the UE's so we have a sort of hybrid scenario; We see an increasing throughput until four gNB's;

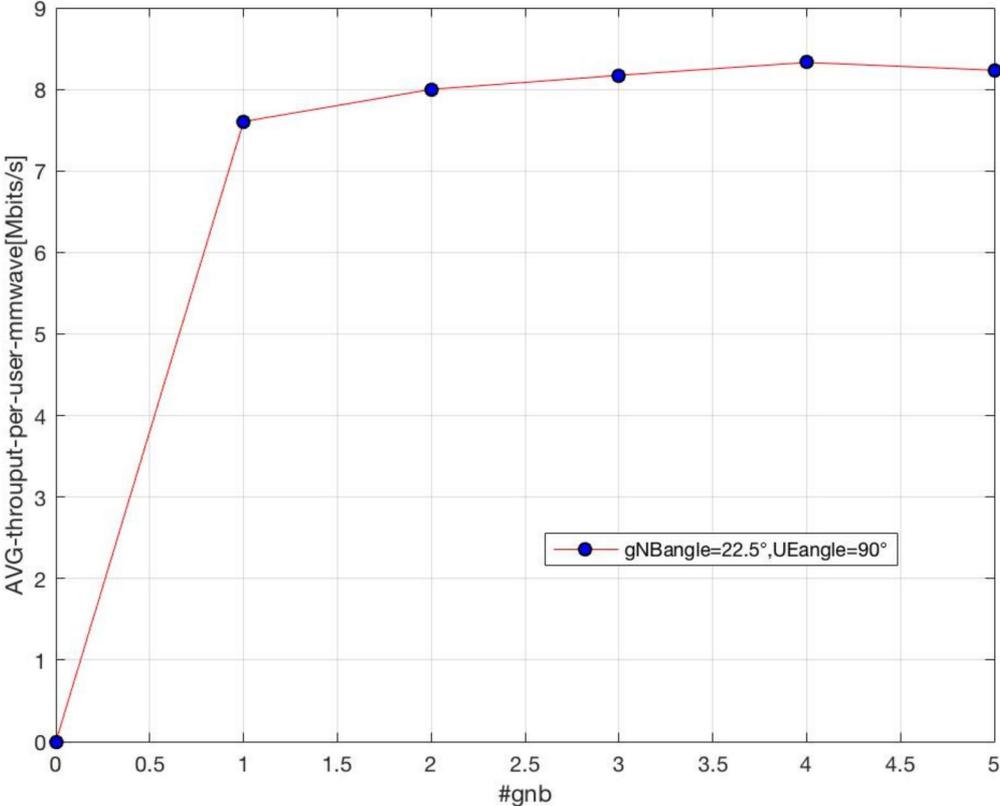


Figure 5.32 Average throughput per user as a funtion of the numbers of gNB with RLC AM and hybrid scenario

Again the Jain's fairness index plotted in the **Figure 5.33**:

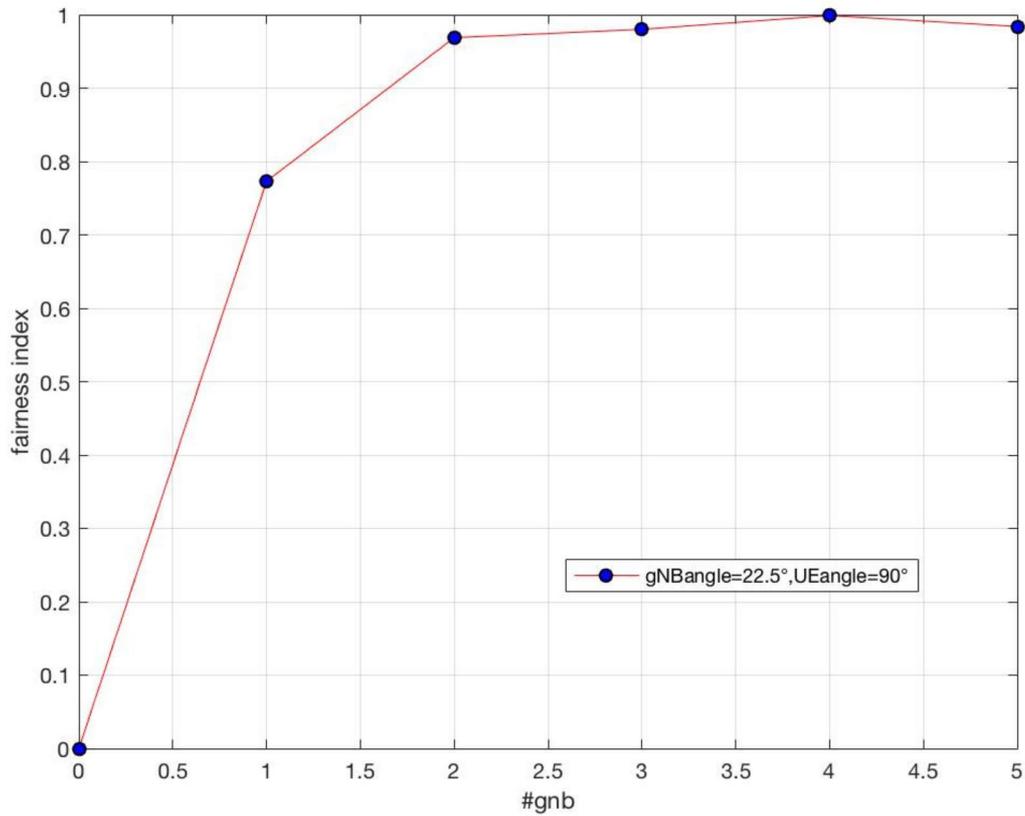


Figure 5.33 Jain's fairness index as a function of the numbers of gNB with RLC AM and hybrid scenario

5.1.13 Short term throughput analysis on the map with RLC AM and hybrid scenario

In the case with one gNB the amount of retrasmissons is a little reduced respect to the previous case(ULA 64x64) but are still relevant so are valid the same considerations made for ULA 64x64(**Fig 5 34** and **Fig 5 35**)

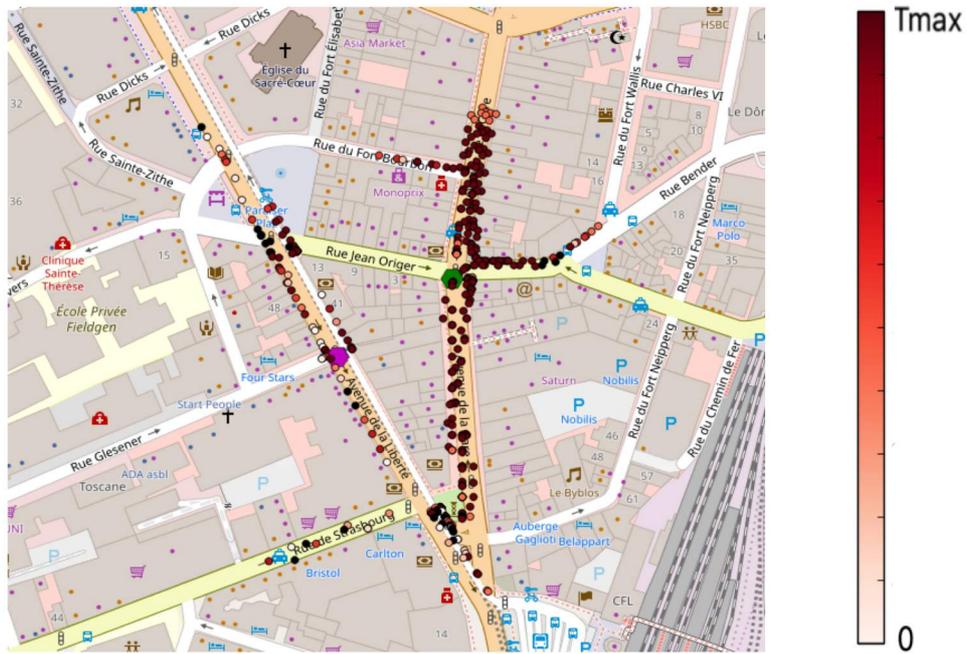


Figure 5.34 Colourmap with short-term throughput indication per user using one gNB with RLC AM and hybrid scenario

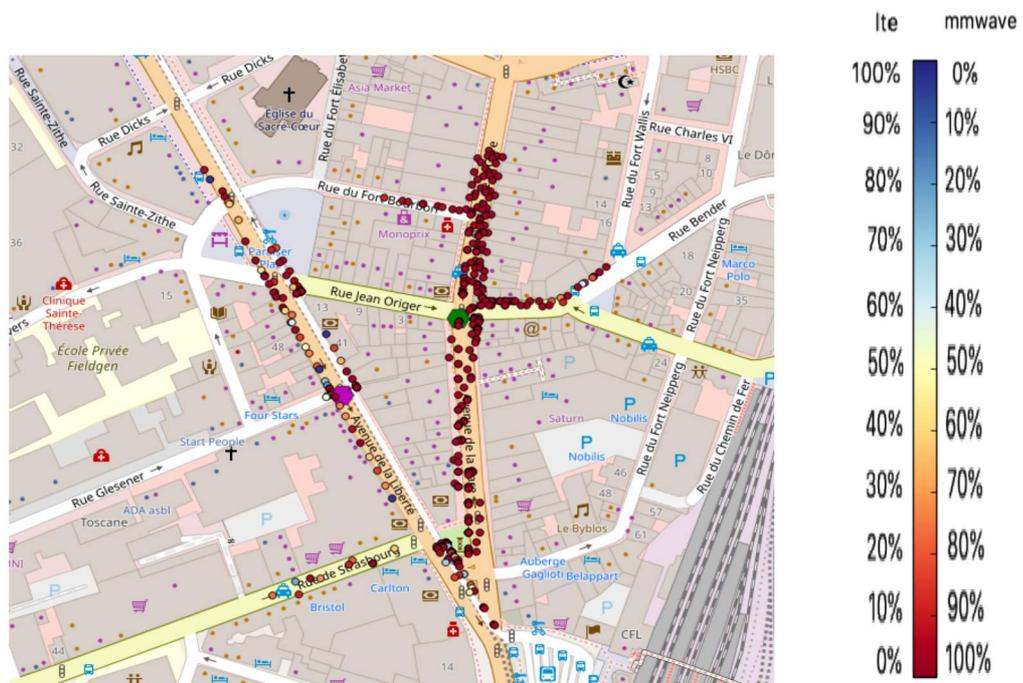


Figure 5.35 Colourmap with percentage of short-term throughput LTE and mmwave per user using four gNB's with RLC AM and hybrid scenario

Again as before we have the best performance with four gNB's so it's similar to the case with ULA 64x64, again we see that in some points (black) we have retransmissions but all the vehicles are served by the mmWave cell antennas (Fig. 5.36 and Fig. 5.37)

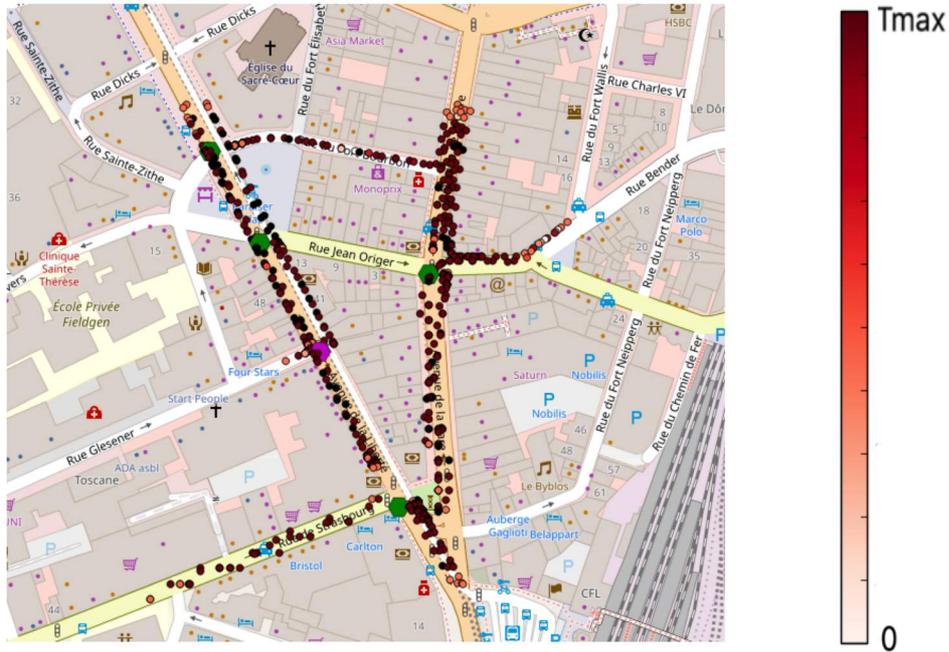


Figure 5.36 Colourmap with short-term throughput indication per user using four gNB's with RLC AM and and hybrid scenario



Figure 5.37 Colourmap with percentage of short-term throughput LTE and mmwave per user using four gNB's with RLC AM and hybrid scenario

6

Conclusions And Future Work

In this thesis, first we extensively discussed about the property of the mmWave channel and all the derived consequences. The need of beamforming and in particular we introduce some of the state-of-the-art solutions for beam alignment.

After a brief review of the current LTE network architecture we describe the proposed mmwave architecture developed by NYU and implemented in ns-3 simulator.

This architecture present a LTE-5G tight integration at the PDCP layer.

The dual connection with Fast switching setup permit to have the best performance in term of throughput and latency.

In particular we tested a mmWave vehicular network using the ns-3 simulator and we concentrated our analysis on the performance perceived by a relevant number of vehicles in a real urban area.

The simulation results underline different aspects of the mmwave communications; In particular the coverage range can reach relevant distances especially in LOS scenario and the communications in an NLOS condition are not negligible.

Different scenarios for delay-sensitive applications and with delay-insensitive Application were tested.

Moreover we modify the beamwidth of the UE's and the gNB's to test the performance indexes and evaluate the possible design choices like an optimization problem (i.e. Minimize the number of antenna to be deployed at similar level of performance). We showed some maps to underline the areas where the vehicles are served by the mmwave technology or by the LTE one.

The simulation with the NY mmwave module present some weaknesses, the hysteresis is not sufficient to cope with the high number of switches/handovers, so it will be necessary to introduce a more sophisticated algorithm.

In our simulation we make use of UDP protocol, in case we need to work with the TCP protocol we need an algorithm to follow in a proper way the high fluctuations in the channel capacity.

In the current implementation of the module is not implemented the handover between LTE cells. The scalability of the current implementation of ns-3 is not able to simulate very large scenario, we used a size scenario to get a reasonable simulation time.

The mmwave module still need an improvement across all the layers, as an example at the PHY layer is used OFDM;

The choice of OFDM can be reasonable for different reasons: it's computationally simple, high compatibility with MIMO, simple modulation and demodulation but on the other hand presents disadvantages like the high OOB emissions, requires cyclic prefix and perfect synchronization and does not support different numerologies. One proposal is the filtered OFDM that overcomes most of the OFDM issues by turning down OOB emissions and to have coexistence of different numerologies, and relaxed synchronization requirements.

References

- [1] “Mobile data traffic growth outlook”,
Available: <https://www.ericsson.com/en/mobility-report/reports/june-2018/mobile-data-traffic-growth-outlook>

- [2] Lara Codeca, Raphael Frank, Thomas Engel, “Luxembourg SUMO Traffic (LuST) Scenario :24 Hours of Mobility for Vehicular Networking Research”
Available: <https://ieeexplore.ieee.org/document/7385539>

- [2] Larsson, Erik G.; Edfors, Ove; Tufvesson, Fredrik; Marzetta, Thomas L.
“Massive MIMO for next generation wireless systems”
Available <https://ieeexplore.ieee.org/document/6736761>

- [3] T. S. Rappaport, F. Gutierrez, E. Ben-Dor, J. N. Murdock, Y. Qiao, and J. I. Tamir,
“Broadband millimeter-wave propagation measurements and models using adaptive-beam antennas for outdoor urban cellular communications”
Available: <https://ieeexplore.ieee.org/document/6387266>

- [4] T. S. Rappaport, E. Ben-Dor, J. N. Murdock, and Y. Qiao, “38 GHz and 60 GHz angle-dependent propagation for cellular & peer-to-peer wireless communications”
Available: <https://ieeexplore.ieee.org/document/6363891>

- [5] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, “Millimeter wave channel modeling and cellular capacity evaluation”
Available: <https://ieeexplore.ieee.org/document/6834753?arnumber=6834753>

- [6] T. A. Thomas, H. C. Nguyen, G. R. MacCartney, and T. S. Rappaport, “3D mmWave channel model proposal”
Available: <https://ieeexplore.ieee.org/document/6965800>

- [7] M. Kyrö, “Radio wave propagation and antennas for millimeter-wave communications”
Available: <https://aaltodoc.aalto.fi/handle/123456789/7580>
- [8] S. Wyne, K. Haneda, S. Ranvier, F. Tufvesson, and A. F. Molisch, “Beamforming effects on measured mm-wave channel characteristics”
Available: <https://ieeexplore.ieee.org/document/6015598>
- [9] S. Sun, T. S. Rappaport, R. W. Heath, A. Nix, and S. Rangan, “MIMO for millimeter-wave wireless communications: Beamforming, spatial multiplexing, or both?” Available: <https://ieeexplore.ieee.org/document/6979962>
- [10] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, “Channel estimation and hybrid precoding for millimeter wave cellular systems,”
Available: <https://ieeexplore.ieee.org/document/6847111>
- [11] A. Maltsev, E. Perahia, R. Maslennikov, A. Sevastyanov, A. Lomayev, and A. Khoryaev, “Impact of polarization characteristics on 60-GHz indoor radio communication systems” Available: <https://ieeexplore.ieee.org/document/5451154>
- [12] W. Roh, J. Y. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, “Millimeter-wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results,”
Available: <https://ieeexplore.ieee.org/document/6736750>
- [13] Hossein Shokri-Ghadikolaie, Carlo Fischione, Gabor Fodor, Petar Popovski, Michele Zorzi, “Millimeter Wave Cellular Networks: A MAC Layer Perspective”
Available: <https://ieeexplore.ieee.org/document/7156092>
- [14] Joan Palacios, Danilo De Donno, Joerg Widmer,
“Tracking mm-Wave Channel Dynamics: Fast Beam Training Strategies under Mobility” Available: <https://ieeexplore.ieee.org/document/8056991>

- [15] Thomas Nitsche, Carlos Cordeiro, Adriana Flores, Edward W. Knightly, Eldad Perahia and Joerg C. Widmer “IEEE 802.11ad: Directional 60 GHz Communication for Multi-Gbps Wi-Fi”
Available:<https://ieeexplore.ieee.org/document/6979964>
- [16] Sai Qian Zhang, H.T. Kung and Youngjune Gwon,
“*InferBeam*: A Fast Beam Alignment Protocol for Millimeter-wave Networking”
Available:<https://arxiv.org/abs/1802.03373>
- [17] Haitham Hassanieh, Omid Abari, Michael Rodriguez, Mohammed Abdelghany, Dina Katabi “Fast Millimeter Wave Beam Alignment”
Available:<https://synrg.csl.illinois.edu/papers/agile-sigcomm18.pdf>
- [18] Morteza Hashemi, Ashu Sabharwal, C. Emre Koksall, and Ness B. Shroff
“Efficient Beam Alignment in Millimeter Wave Systems Using Contextual Bandits” Available:<https://arxiv.org/abs/1712.00702>
- [19] Sanjib Sur, Xinyu Zhang, Parmesh Ramanathan, Ranveer Chandra
“BeamSpy: Enabling Robust 60 GHz Links Under Blockage”
Available:<https://dl.acm.org/citation.cfm?id=2930625>
- [20] Sanjib Sur, Ioannis Pefkianakis, Xinyu Zhang, Kyu-Han Kim
“Towards Scalable and Ubiquitous Millimeter-Wave Wireless Networks”
Available:<https://dl.acm.org/citation.cfm?id=3241579>
- [21] Vutha Va, Junil Choi, Takayuki Shimizu, Gaurav Bansal, and Robert W. Heath, Jr.
“Inverse Multipath Fingerprinting for Millimeter Wave V2I Beam Alignment”
Available:<https://ieeexplore.ieee.org/document/8240727/versions>
- [22] Marco Mezzavilla, *Member, IEEE*, Menglei Zhang, Michele Polese, *Student Member, IEEE*, “End-to-End Simulation of 5G mmWave Networks”
Available:<https://ieeexplore.ieee.org/document/8344116>

- [23] *P. Pablo Garrido Abenza, Manuel P. Malumbres and Pablo Piñol Peral* “*GatcomSUMO: A Graphical Tool for VANET Simulations Using SUMO and OMNeT++*”
- [24] *Menglei Zhang, Michele Polese, Marco Mezzavilla, Sundee Rangan, Michele Zorzi* “*Ns-3 Implementation of the 3GPP Frequency MIMO Channel Model Spectrum above 6 GHz for frequency spectrum above 6GHz*”
Available: <https://arxiv.org/abs/1702.04822>
- [25] Università di Padova “*Performance Study of LTE and mmWave in Vehicle-to-Network Communications*”
Available: <https://arxiv.org/abs/1805.04271>
- [26] Roberto Garello, Di Stasio Francesco “*Radio access techniques on 5G*”, Politecnico Torino
- [27] North Carolina State University “*An Obstacle Model Implementation for Evaluating Radio Shadowing with ns-3*”
Available: <https://dl.acm.org/citation.cfm?id=2756512>
- [28] Tommaso Zugno, Michele Polese, Michele Zorzi
“*Integration of Carrier Aggregation and Dual Connectivity for the ns-3 mmWave Module*” Available: <https://arxiv.org/abs/1802.06706>
- [29] Michele Polese, Marco Mezzavilla, Michele Zorzi
“*Performance Comparison of Dual Connectivity and Hard Handover for LTE-5G Tight Integration*” Available: <https://arxiv.org/pdf/1607.05425.pdf>
- [30] Ish Kumar Jain, Rajeev Kumar, and Shivendra Panwar
“*Limited by Capacity or Blockage? A Millimeter Wave Blockage Analysis*”
Available: <https://arxiv.org/abs/1808.01228>