# Data Traffic Modelling in Mobile Networks for Heterogeneous Types of IoT Services

By

**Mohammed Hassan Dighriri**

A thesis submitted in partial fulfilment of the requirements of Liverpool John Moores University for the degree of Doctor of Philosophy

December 2019

# DECLARATION

I, Mohammed Dighriri, hereby declare that the work in this thesis is original and has been completed by myself, except where reference is made to other works, and has not been submitted for any examination at this university or any other learning institutions.

Mohammed Dighriri
Doctoral Candidate
Department of Computer Science
Liverpool John Moores University

## ABSTRACT

The upcoming 5th Generation (5G) mobile networks will be different from the previous mobile network generations in the fact that it will enable the mobile networks industry, besides offering superior broadband services, to enhance Internet of Things (IoT) industries such as vehicular communication system, factory automation, smart healthcare system and many more. Many of these use cases have challenging and quite often contradicting requirements in terms of data rate, latency, throughput and so on. This suggests that 5G mobile networks need to adopt flexible models that can adapt to different IoT device and traffic requirements. Consequently, a fresh look into how mobile networks are currently designed and deployed is needed. Historically, mobile networks have relied on the axiomatic role of cells as the cornerstone of the Radio Access Networks (RAN). Mobile network systems have witnessed several recent trends such as the increased heterogeneity in heterogeneous types of IoT services infrastructure and spectrum as well as the rise of different traffic types with different Quality of Services (QoS) requirements.

In this direction, this thesis focuses on improving the performance of cell-edge users or IoT devices in 5G mobile networks by initially implementing the network slicing management approach, particularly as, with the fast growth of IoT, billions of devices will join the internet in the next few years. Hence, the latest 5G mobile technologies expected to offer massive connectivity and management ability of high volume of data traffic at the presence of immense interferences from a mobile network of IoT devices. Further, it will face challenges due to congestion and overload of data traffic due to a humongous number of IoT devices. Besides, these devices likely to demand high throughput, low latency and high level of reliability especially for critical real-time smart systems in density and small zone, such as in Vehicular Communication System (VCS), these vehicles mainly rely on connectivity aspects. Furthermore, IoT devices transmit small and large-sized packets with different radio resource requirements. For example, Smart Healthcare System (SHS) devices transmit small-sized of a data with utilizing a small portion of Physical Resource Block (PRB) as the smallest radio resource unit, which is allocated to a single device for data transmission in 5G mobile networks. In the IoT services with transmitting a small-sized data, the capacity of the PRB is not fully utilized, which causes wastage and unfairness of using PRB among these IoT devices or services.

The novelties made in this thesis significantly advance a Slice Allocation Management (SAM) model based on critical services such as (VCS) to satisfy low latency demand. The proposed model aims at providing dedicated slices based on service requirements such as expected low latency for (VCS). To ensure such performance to data traffic of IoT devices in Uplink (UL) of Relay Node (RN) cells in the 5G mobile networks by slicing the RAN, along with assigning the nearest Mobile Edge Computing (MEC) with isolating slices depend on technical and QoS requirements for each IoT nodes. Also, this thesis proposes a Data Traffic Aggregation (DTA) model for efficient utilization of the smallest untie of PRB by aggregating the data traffics of several IoT devices, which can support IoT node throughput such as SHS. Also, this thesis presents a comprehensive comparison of the packet scheduling mechanisms include Priority Queuing (PQ), First-In-First-Out (FIFO) and Weighted Fair Queuing (WFQ) applied based on data traffic slicing model through RN cells.

These thesis models are validated through the OPNET simulator to measure the performance of the SAM and DTA Models along with the assessment of packet scheduling mechanism. The simulation considers IoT devices in various smart systems such as VCS, SHS and smartphones also, different protocols include Simple Mail Transfer Protocol (SMTP), File Transfer Protocol (FTP), and Voice over Internet Protocol (VoIP) and Real-time Transport Protocol (RTP). Simulation results show a significant improvement in IoT nodes packets transmission via RNs and Donor eNodeB (DeNB) cells, in My SAM Model scenario comparing with other scenarios. The model has improved such as End-to-End (E2E) delay in FTP node by reaching 1ms, loading in VoIP node by 80% and throughput of all nodes in the uplink side of networks by 66%. In addition, the results display significant impact of IoT data traffic with different priority, networks E2E performance is improved by aggregating data traffic of several IoT devices with DTA model, which is determined by simulating several scenarios, considerable performance improvement is achieved in terms of averages cell throughput, upload response time, packet E2E delay and radio resource utilization. Finally, the result found PQ packet scheduling mechanism as the appropriate scheduling mechanism in case of supporting several of priorities queuing for data traffic.

# ACKNOWLEDGMENTS

In the name of Allah, the Most Gracious and the most merciful. I am extremely thankful to Almighty ALLAH for his blessings and providing me with the ability to perform this research, without which none of my work would have been done.

My sincerest thanks and profound appreciation go to my supervisor *Dr. Gyu Myoung Lee* for his guidance throughout my PhD study, without which this thesis could not have been produced in the perfect form. His insightful comments, constructive criticism and suggestions have given me invaluable guidance in different stages of my research. He has been very kindly, helpful and supportive for all these years. I greatly appreciate his care and concern throughout my PhD study.

My gratitude and appreciation also go to my second supervisor, *Dr. Thar Baker Shamsa*, and Third supervisor, *Dr. Rubem Pereira*, for their valuable comments, guidance and suggestions.

I take this opportunity to express my gratitude to my father and mother for their support, advice and encouragement during my PhD study.

I will be eternally grateful for the endless love and affection bestowed upon me from my beloved wife for her endless patience, continuous encouragement, and support.

# LIST OF PUBLICATIONS

During the period of my research, the following works have been published:

**Book Chapters:**

1. **Dighriri M,** Lee GM, Baker T. 2018.Big Data Environment for Smart Healthcare Applications over 5G Mobile Network. Applications of Big Data Analytics. (Springer).
2. **Dighriri M,** Lee GM, Baker T. 2017.Measuring and Classification of Smart Systems Data Traffic Over 5G Mobile Networks Technology and Smart Futures. (Springer).
3. **Dighriri M,** Lee GM, Baker T. 2018. Applying Scheduling Mechanisms over 5G Cellular Network Packets Traffic, Third International Congress on Information and Communication Technology in concurrent with ICT Excellence Awards (ICICT 2018). (Springer).

**Conference Proceedings:**

**1.** **Dighriri M**, Lee GM, Baker. 2016. Data Traffic Model in Machine to Machine Communications over 5G Network Slicing. (DeSE-2016).

2. **Dighriri M**, Lee GM, Baker T. 2017. Comparison Data Traffic Scheduling Techniques for Classifying QoS over 5G Mobile Networks the 31st IEEE International Conference on Advanced Information Networking and Applications (AINA-2017)
3. **Dighriri M**, Lee GM, Baker T. 2018.Resource Allocation Scheme in 5G Network Slices. (AINA-2018).
4. Alfoudi A, **Dighriri M**, Lee GM, Baker. 2017. Seamless LTE-WiFi Architecture for Offloading the Overloaded LTE with Efficient UE Authentication (DeSE-2016).
5. Alfoudi A, **Dighriri M**, Lee GM, Baker T. 2017. Traffic Management in LTE-WiFi Slicing Networks (ICIN-2017)
6. **Dighriri M.** 2015.Enhancing 5G Cellular Network Packets Traffic by Scheduling Mechanisms. (IRES-2017)
7. Alfoudi A, **Dighriri M**, Lee GM, Baker T. Mobility Management Architecture in Different RATs Based Network Slicing. (AINA-2018).

8. Omar Aldhaibani1, Mustafa Hamid AL-Jumaili, Faycal Bouhafs1, Michael Makay, Alessandro Raschellà, Ali Alfoudi, **Mohammed Dighriri1**, Ghulam Mohi-Ud-Din1 An SDWN based Architecture for Enormously Dense Wireless Networks to Optimize the Handover Performance in WLAN (ICIDM-2018)

## Poster Presentations

1. **Dighriri M**, Lee GM, Baker T. 2018. Big Data Environment for IoT Devices over 5G Mobile Network. In BAME Network at Liverpool John Moores University sensor city.
2. **Dighriri M**, Lee GM, Baker T. 2017.Data Traffic Models in Machine to Machine Communications over 5G Mobile Network. In Liverpool John Moores University Research Week.

# TABLE OF CONTENTS

# LIST OF FIGURES:

## LIST OF TABLES:

# ABBREVIATION

| | |
|---|---|
| **RLC** | Radio Link Control |
| **RRC** | Radio Resource Control |
| **NAS** | Non-Access Stratum |
| **AS** | Access Stratum |
| **LTE** | Long Term Evolution |
| **LTE-A** | Long Term Evolution-Advanced |
| **IP** | Internet Protocol |
| **VoIP** | Voice over Internet Protocol |
| **UMTS** | Universal Mobile Telecommunications System |
| **PHY** | Physical layer protocols |
| **iDEN** | Integrated Digital Enhanced Network |
| **PRB** | Physical Resource Block |
| **DL** | Downlink |
| **QoS** | Quality of Service |
| **RAN** | Radio Access Network |
| **UL** | Uplink |
| **UE** | User Equipment |
| **Tx** | Transmitter |
| **iDEN** | Integrated Digital Enhanced Network |
| **SINR** | Signal to Interference plus Noise Ratio |
| **SNR** | Signal to Noise Ratio |
| **TCP** | Transmission Control Protocol |
| **ICIC** | Inter-Cell Interference Coordination |
| **TDD** | Time Domain Duplex |
| **CPU** | Central Processing Unit |
| **RN** | Relay Node |
| **SC-FDMA** | Single Carrier Frequency Division Multiple Access |
| **CA** | Carrier Aggregation |
| **PDCP** | Packet Data Convergence Protocol |
| **PPP** | Point-to-Point Protocol |
| **M2M** | Machine-to-Machine |
| **MIMO** | Multiple Input Multiple Output |
| **HSPA** | High Speed Packet Access |
| **HSS** | Home Subscriber Service |
| **HTTP** | Hypertext Transfer Protocol |
| **ICIC** | Inter-Cell Interference Coordination |
| **HetNet** | Heterogeneous Network |

| | |
|---|---|
| **HeNB** | Home Evolved Node-B |
| **HARQ** | Hybrid Automatic Repeat Request |
| **eNodeB** | EPC Evolved Packet Core |
| **EPS** | Evolved Packet System |
| **E-UTRAN** | Evolved Universal Terrestrial Radio Access Network |
| **FTP** | File Transfer Protocol |
| **GBR** | Guaranteed Bit Rate |
| **GPRS** | General Packet Radio Service |
| **GSM** | Global System for Mobile Communications |
| **GPRS** | General Packet Radio Service |
| **ITU** | International Telecommunication Union |
| **MAC** | Medium Access Control |
| **D2D** | Device-to-Device |
| **eICIC** | Enhanced Inter-Cell Interference Coordination |
| **GW** | Gateway |
| **iDEN** | Integrated Digital Enhanced Network |
| **IS** | Interim Standard |
| **5G NR** | 5G New Radio) |
| **1G** | First Generation |
| **2G** | Second Generation |
| **3G** | Third Generation |
| **4G** | Fourth Generation |
| **5G** | Fifth Generation |
| **3GPP** | The 3rd Generation Partnership Project |
| **IoT** | Internet of Things |
| **VCS** | Vehicular Communication System |
| **SHS** | Smart Healthcare System |
| **SMTP** | Simple Mail Transfer Protocol |
| **RTP** | Real-time Transport Protocol |
| **OPNET** | Optimized Network Engineering Tool |
| **MEC** | Mobile Edge Computing |
| **E2E** | End to End |
| **DTA** | Data Traffic Aggregation |
| **SAM** | Slice Allocation Management |
| **mmWave** | millimeter Wave |
| **KPI** | Key Performance Indicator |
| **RATs** | Radio Access Technologies |
| **DeNB** | Donor eNodeB |
| **PDN** | Packet Data Network |

# CHAPTER 1: INTRODUCTION

## 1.1. Background

The current 4th generation of mobile networks, with the demands of the mobile networking ecosystem is changing once again. Mobile network traffic has experienced unprecedented growth in recent years and this trend is expected to continue in the foreseeable future. According to Cisco forecasts that smart device connections are expected to increase from 53% in 2017 to reach 73% in 2022: forecast of the global growth of smart mobile devices and connections per year [1]. This growth in the mobile network traffic is related both to the expected increase in the number of connected smart devices as well as to the type of traffic that these devices are expected to generate. The number of globally connected smart devices is expected to increase from 8 billion in 2017 to 11.6 billion in 2022. At the same time, the share of non-smartphone devices is expected to significantly drop in favour of smartphones, phablets and tablets, while the Internet of Things (IoT) communications are expected to obtain a very large portion of the market share. These new trends in the mobile ecosystem also lead to significant changes regarding the performance requirements that mobile networks need to fulfil [2]. For instance, the increased adoption of smart devices over non-smart devices is followed by a demand for an improvement of the user data rate and the spectral efficiency of the network [1]. This is because the use of the mobile phone, for example is no longer just for voice calls and short messages, but also for social networking, viewing high-quality multimedia content (e.g. HD video streaming and high-resolution images), gaming, etc. Also, the significant rise of IoT communications introduces new requirements that must be met in terms of energy efficiency, latency, throughput etc. Due to the above-mentioned challenges, it has already become apparent that the abilities offered by 4G networks are no longer adequate to cover the newly emerging requirements of the mobile ecosystem. This drives the need for an evolution towards the next generation of 5G mobile networks and has ultimately led to a new wave of research with the main goal of improving network performance. In this context, there have been several indicative new goals to be achieved at an operational level [3], including:

- 1,000 times higher mobile data volume per geographical area
- 10-100 times more connected devices
- 10-100 times higher data rates
- 10 times the lower energy consumption
- End-to-end latency that is below 1ms
- Ubiquitous 5G access

Similarly, to previous generations to move from 4G to 5G and to achieve these Key Performance Indicator (KPI) oriented goals, there is a need to introduce technological innovations in terms of the involved Radio Access Technologies (RATs), network protocols etc. This has given rise to several new hot topics of research in the domain of mobile networks, including network slicing, small cells massive Multiple-Input Multiple-Output (MIMO), Heterogeneous Networks (HetNets) etc. Complementary to this performance-oriented view of 5G networks and following the observations drawn from Table 1.1 there also exists a service-oriented view, based on which, the 5G network is expected to cater to a wide range of services differing in their requirements and types of devices. For example, the International Telecommunication Union (ITU) and 5G- Point-to-Point Protocol (PPP) have identified three broad use case families; enhanced mobile broadband, massive machine-type communications and critical communications. Within those, it is possible to define several specific use cases[4] ranging from general broadband access with global coverage to specialized networks for sensors or extreme mobility. The stark differences between these use cases translate to a set of heterogeneous requirements.

## 1.2. Research Problem

Service-oriented modelling is a requirement of the 5G mobile network architecture to turn it into a more flexible and programmable fabric that can be used to simultaneously provide a multitude of diverse services over a common generic underlying infrastructure. To achieve this, a concept widely considered as a key feature of the 5G architecture is network slicing, which is the capability to create E2E logical networks spanning both the Radio Access Network (RAN) and the core and tailored for a specific service's needs. In contrast to this, the architecture of previous mobile network generations followed a one-size-fits all approach, with the main goal being the optimization of mobile broadband services. This makes slicing a characteristic unique to the 5G environment and arguably a legacy that 5G networks will leave behind for future generations of mobile network [4][5][6].

Considering the imminent changes expected towards 5G, the RAN part of the mobile network architecture, is expected to be one of the main focuses of attention. This is natural since the RAN is the most complex part of the network infrastructure that significantly differs from conventional IP networks. Among others, it must deal with several very significant issues in terms of coverage, interference and mobility management, the energy efficiency of the connected smart devices etc. To add to this, the RAN has always been one of the major bottlenecks of mobile networks in terms of capacity [7], which constitutes a significant problem when scaling the network to more connected users and devices with increased throughput requirements. Due to the previous reasons, it comes as no surprise that some of the biggest technological innovations when moving from one mobile network generation to the next take place in the RAN, with 5G being no exception to this. In terms of the physical layer and more generally the radio interface, a new RAN called 5G New Radio (5G NR) is being considered to complement the existing radio interfaces [8]. 5G NR is expected to help in reducing the network latency by bringing a new and more flexible radio resource grid. It is also specifically designed to introduce support for technologies such as small cells and massive MIMO, which are expected to greatly boost the network capacity. These new additions will form a part of a wider multi-RAN environment, which is expected to enable ubiquitous access to mobile devices as well as to cater the diverse requirements of the heterogeneous of IoT services that must be supported in the context of 5G.

The 5G RAN is also expected to evolve in terms of its architecture to be able to accommodate the increasing demands that arise from the proliferation of mobile devices. The densification of the RAN using small-cell such as Relay Nodes (RNs) and femtocell deployments for the increase of the network's capacity, as well as the centralization of its processing in a Cloud-RAN (C-RAN) deployment for improved coordination among the cells, are some of the most prominent changes that are expected to take place in 5G. Furthermore, the emerging model of Mobile Edge Computing (MEC) is expected to bring services closer to the edge, allowing them to tap into the data as well as the processing and storage capabilities offered by the RAN [9][10][11]. This will relieve some of the stress posed to the mobile core by serving part of the generated traffic locally in the RAN, while at the same time it can lead to significant latency reductions for critical services. Apart from the performance improvements that the previous changes are expected to bring in various aspects of the RAN (throughput, latency, loading etc), it is equally important for future RANs to provide inherent support for network slicing to accommodate the multi-service capabilities envisioned in the context of 5G. The conventional RAN architecture found in current 4G networks is characterized by a rigid design and cannot efficiently accommodate a diverse set of services with different requirements and characteristics. It is therefore critical to re-shape the RAN into a more flexible and adaptable component of the mobile network architecture[12].

The summary of these challenges includes:

- Support for a large number of IoT devices within a single cell
- Small-sized data transmission after a regular interval of time
- Low latency
- High reliability
- Low power consumption
- Support for the various mobility profiles
- To ensure that the IoT traffic is not affecting the regular HTC traffic since PRBs are an asset and scarcely available.

## 1.3. Research Motivation

To explore the full 4G potential limitations should be resolved by appropriate research. As mentioned in the earlier section one of the critical challenges that the mobile networks industry will face is inter-cell interference due to the expected significant cell densification. Research in this direction can be used to implement and amend new techniques or as a reference point for vendors and service providers to improve and develop their services.

As discussed before, 5G should, at the same time, enable new services, reduce operational costs and improve performance for traditional and smart services. To achieve these strategic goals, there is a need for introducing a set of technical requirements that reflect a multitude of use cases. For example, Smart Healthcare System (SHS) will require providing very low E2E latency with medium to high throughput and reliability requirements, other use cases such as Vehicular

Communication System (VCS) or controlled robots or drones will require low E2E latency, high reliability, and high-device density with moderate throughput requirements. It will be difficult for one 5G network system to be able to satisfy all these requirements at the same time. However, it will be possible for different subsets of requirements to be satisfied with different carriers or sub-systems. For example, reliability demanding applications could be deployed on low-frequency bands whereas data rate demanding applications could be deployed on higher frequency bands where there is more spectrum. This can be extended to the core network where the concept of network slicing could be used where different network slices would be created to satisfy the requirements of different applications or traffic streams.

The existing mobile networks might run out of capacity in future due to significantly increasing IoT traffic, resulting in the performance degradation of regular mobile traffic. IoT devices transmit small and large-sized data with different QoS requirements. For example, SHS devices convey big sized data but are delay-sensitive [13][14][15][16]. The Physical Resource Block (PRB) is the smallest radio resource unit, which is allocated to a single device for data transmission in 5G mobile networks.

Therefore, this thesis is motived by the main research challenges, which are:

- Provide an introduction regarding a novel Data Traffic Aggregation (DTA) model and Slicing Allocation Management (SAM) model, DTA model for 5G radio resources based on efficiently utilizing the smallest untie of PRB by aggregating the data of several IoT devices. Also, a (SAM) model relies on smart systems in a smart city case study, network slices will differentiate smart systems data traffic in term of QoS requirements in each slice for example smartphones, VCS, and SHS.
- Design and develop The Optimized Network Engineering Tool (OPNET) OPNET simulator to assess the performance of the proposed model in case of QoS for each data traffic slice's separation. The simulated 5G mobile network of IoT devices and applications include Simple Mail Transfer Protocol (SMTP), File Transfer Protocol (FTP), and Voice over Internet Protocol (VoIP) and Real-time Transport Protocol (RTP). The results show the impact of the proposed model in terms of assuring different QoS characteristics for the different type of data traffic of 5G networks. The E2E performance of the network is tested by managing data of

different IoT devices in each slice in cases of average cell throughput, SMTP and FTP average upload response time, FTP average packet E2E delay.

## 1.4. Research Aim and tasks

Targeting the challenges stated above, this thesis attempts to cover the following aim:

### 1.4.1. The main aim

The main aim of this thesis is to improve both radio resources capacity and QoS satisfaction of critical mission IoT services in the density connection of uplink 5G mobile networks, which can be achieved by designing and investigating data traffic aggregation and slicing models through dynamic services allocation in RNs cell of MAC and PDCP layers in response to IoT heterogeneous services in the context of reliable throughput and low latency.

### 1.4.2. The main tasks

This can be broken down into the following tasks:

- Develop and design data traffic aggregation model based on IoT devices data packet size that can assist in optimising PRBs allocation.
- Fulfil the slicing allocation management model in a 5G mobile network environment by considering IoT heterogeneous service characteristics and QoS demands such as reliable low latency.
- Conduct investigation and comparison among packet scheduling mechanisms to evaluate and compare their performance in terms of queuing priorities.
- Design OPNET Simulation projects by applying three use cases along with different scenarios and sub-scenarios, which allow assessing the QoS requirements.
- Apply the proposed models on OPNET simulation that considers how it can be used to optimise small cells and MEC.

## 1.5. Research Contributions

The major contributions of this thesis are highlighted as follows:

- A comprehensive state-of-the-art in the three main topics covered, data traffic aggregation, data traffic slicing models and packet scheduling mechanisms in mobile networks in terms of enhancing the performance of networks QoS requirements. Some general discussion introducing literature regarding deployment, backhauling and access methods in small-cell networks. These related works have several advantages such as novel control designs based on SDN and NFV to address key core network issues of data traffic and mobility management and enable mobile networks to scale in the presence of high volumes of data traffic. However, there are major disadvantages when these works are somewhat misleading as it refers only to the UE perceived performance isolation.

- The proposed DTA model is relying on aggregating packet data from several IoT devices at the Packet Data Convergence Protocol (PDCP) layer of the RN and DeNB cells. The PDCP layer performs header compression, retransmission, and delivery of PDCP Session Data Units (SDUs), duplicate detection, etc. In the proposed model, the PDCP layer is used for the aggregation of the IoT devices packet data in the uplink of the 5G RANs network. The main reason for selecting PDCP for aggregation in the uplink is to aggregate data with a minimum number of the additional headers.

- To do investigation and comparison among packet scheduling mechanisms, which is to evaluate and compare the performance of three packet scheduling mechanisms include PQ, FIFO and WFQ designed for 4G and 5G mobile networks in terms of user's suitability to enhance the priorities queuing mechanism, which will reflect on IoT devices and smart systems QoS requirement such as throughput, load, latency and fairness.

- To develop SAM model created on critical services of smart systems such as VCS and SHS to satisfy QoS demands. The proposed model aims at providing dedicated slices based on service requirements such as expected high throughput for SHS and lower latency for VCS. Thereby, to ensure such performance of data traffic of IoT devices in the uplink of RNs cell in the 5G mobile networks by slicing the RAN, along with assigning the nearest Mobile Edge Computing (MEC) with isolating slices depending on technical and QoS criteria of each IoT nodes.

- To assess and investigate the above contribution works by using OPNET Simulation to design and simulate data traffics of IoT heterogeneous services in RNs and DeNB cells in the 5G mobile networks environment in different scenarios.

## 1.6. Thesis Organization

The rest of this thesis is structured as follows:

**Chapter 1** introduces a research problem, motivation, aim and tasks, also the main research contributions.

**Chapter 2** provides a background on the basic concepts behind 5G mobile networks include 5G protocols architecture, 5G enabling technologies, LTE transmission resources, heterogeneous networks and small cell deployments. Also, IoT communication includes IoT network architecture, lastly, network simulation tools comparisons.

**Chapter 3** gives a comprehensive review of the literature related to data traffic aggregation models, data traffic slicing models and packet scheduling mechanisms in DeNB and RNs in term of enhancing the performance of mobile networks QoS requirements. Some general discussion introducing the literature regarding deployment, backhauling and access methods in small-cell networks.

**Chapter 4** presents the Data Traffic Aggregation (DTA) model include related work in terms of data traffics aggregation models and resource allocation schemes, problem statement, also the systems models in both the DTA model and Resource Allocation Scheme (RAS), then I have presented the simulation approach with results.

**Chapter 5** investigation and comparison among packet scheduling mechanisms, includes a problem statement, then comparison scheduling mechanisms in terms of advantages and disadvantages on the packet's transmission and QoS parameters from the users to servers. Lastly, I have presented a simulation approach with results.

**Chapter 6** presents a Slicing Allocation Management (SAM) model along with related work of slicing allocation management models, problem statement, and then the use case vehicular communication system (VCS), also gives service slices Model. In this chapter, I have presented SAM edge cloud by placing the Mobile Edge Computing (MEC) model. Lastly, I have presented a simulation approach with results.

**Chapter 7** concludes this thesis, summarizing the work presented and discussing the limitations of the contributions as well as possible directions for future research.

# CHAPTER 2: BACKGROUND

## 2.1. Introduction

Mobile networks have come a very long way in a very short time, forming a major part of the telecommunications market with reports forecasting a further increase of their significance in the near future [2]. This is due to the constant introduction of new capabilities, which go beyond the basic voice and messaging communication and aim to provide support for many novel use cases in various domains of My lives, including health, entertainment, industrial and home automation, vehicular communication etc. To better understand the direction in which mobile networks are moving, it would be interesting to provide a brief overview of their evolution so far. Mobile networks are distinguished into generations (conventionally denoted by a number preceding a capital G), with each generation characterized by a set of features that either introduces new capabilities to the network or enhances and extends those offered by previous generations. As the baseline, I can consider the first generation of mobile networks or First Generation (1G), which was initially launched in Japan by Nippon Telegraph and Telephone (NTT) in 1979 and only provided voice services based on analogue radio transmission techniques. While 1G was the first true cellular mobile network architecture, it presented major limitations in terms of the number of users that it could support [1][17][18].

As a result, the second generation of cellular technologies (or 2G) was released at the beginning of the 90s. This generation was characterized by the digitization and compression of speech, supporting a larger number of mobile users connected to the network. Moreover, 2G networks introduced for the first-time data services for mobiles, initially with the hugely popular feature of SMS text messages and later with the implementation of General Packet Radio Service (GPRS) (the so-called 2.5G), which introduced a new packet-switched domain. The major rise of personal computers and the Internet during the 90s created a need for the further evolution of mobile networks to support high-speed data transfers and inter-communication of mobile devices with the Internet. As a result, the third generation of mobile networks (3G) appeared in the early 2000s, providing higher data rates that could reach up to 21.6Mbps. This evolution enabled the appearance of several applications over 3G networks, including mobile Internet access, video calls, mobile TV, GPS etc. The applications enabled by 3G and the emergence of smartphones in the 2000s had a major impact on the telecommunications market, leading to a very high adoption rate of mobile devices by users. This increase in the scale of mobile networks, along with the need for significant improvements in the Quality of Service (QoS) of users led to Long-Term Evolution (LTE), which formed the fourth generation (or 4G) of mobile networks, which was released commercially just before 2010 and is still the most widespread mobile network architecture. Among others, 4G was the first generation to introduce all-IP packet-switched networks that supported peak Downlink (DL) data rates of more than 100Mbps in mobility conditions and greatly improved the spectral efficiency of the radio interface to support more simultaneous users per cell [2][1][19].

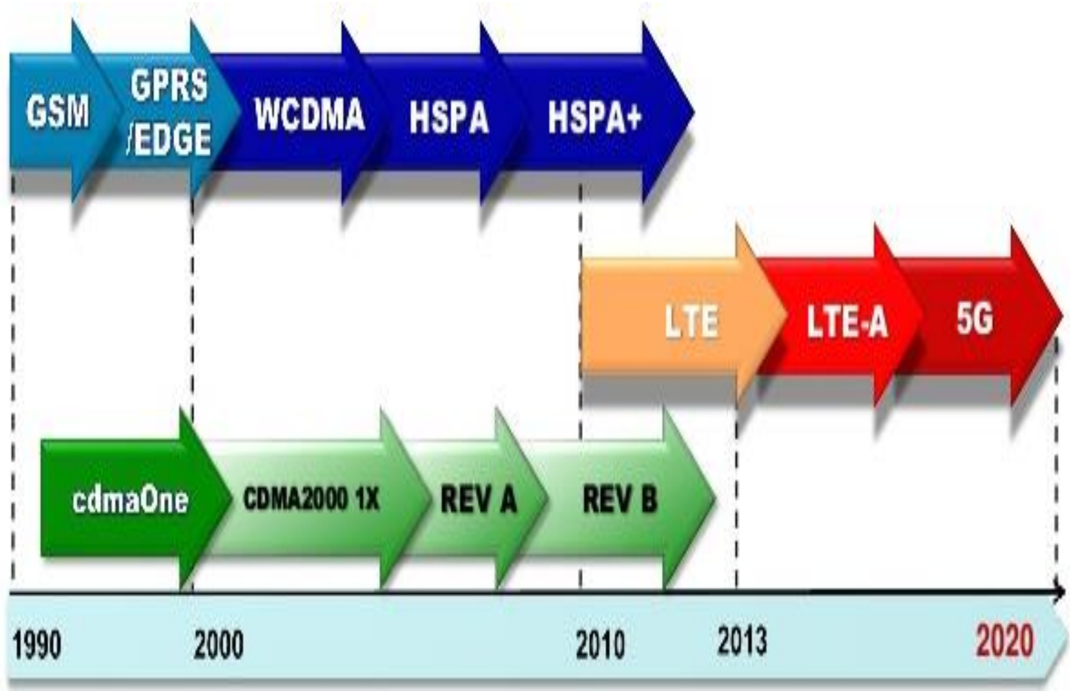


Figure 2. 1**: Mobile Networks Evolution**

In the world of mobile networks, IoT communication is a term that reflects the automatic operation of smart devices, connected to each other, with or without human intervention [20]. Besides the tremendous increase in the mobile traffic, it is also expected that the IoT traffic will rise quickly due to growing use of the IoT devices (e.g. smart meter reader, traffic control and blood pressure sensor) in numerous applications. The applications areas of IoT contain for example smart office, smart traffic monitoring, smart alerting system, smart healthcare system, and logistics system [21][22]. Furthermore, IoT communication offers ubiquitous connectivity between

IoT devices that allows the interconnection of devices for instance laptops, smart sensors, computers, etc. to perform several automatic operations in various IoT applications. Unlike traditional Human-to-Human (H2H) communication, for example, voice call, video chat, and email, etc. IoT data traffic has distinctive features such as a large number of IoT devices, group-based communication between devices, dynamic mobility scenarios, small-sized data transmission, and extra-low power consumption[23].

## 2.2. 5G Protocols Architecture

5G mobile network interfaces between User Equipment (UE), Evolved Universal Terrestrial Radio Access Network (E-UTRAN), EPC and SGWs are linked with protocol stacks that are used by the network component to conversation signalling messages and data [24][25][26]. Consequently, the 5G-protocol stack could be separated among the following groups, which include user plane protocols and control plane protocols that are defined as follows:

- **Signalling protocols**

  Signalling protocols are utilized for exchanges signalling data between numerous devices within the network.

- **User plane protocols**

  These protocols enhance routing of user's messages between UEs and S-GWs.

- **Transport protocols**

Transport protocols are accountable for the conveying of data and signalling messages between several networked devices on the air (Uu) interface, the UE high-level functionalities are controlled by the MME. Nevertheless, there is no straight connection route between the UE and the MME. The connection path between UE and the MME is established using E-UTRAN eNB that somehow supports the level of hardware complexity in the network. To decrease this complexity, the Uu interface is additionally separated into two levels of protocols. One is named Access Stratum (AS) while the second is the Non-Access Stratum (NAS) level protocols. The MME high-level signalling lies at the NAS level, however, is transported within the network using AS protocols. The control and user plane termination protocols are maintained by the eNB. A high-level outline of the protocol stack is portrayed in Figure 2.2 The user plane protocols contain the Packet Data Convergence Protocol (PDCP), the Medium Access Control (MAC) and the Physical (PHY) layer protocols. Besides, the protocols of the control plane include the Radio Resource Control (RRC) protocols. According to [27], the main functionalities of these protocols are as shown below:

- **Access Stratum (AS)**

The AS protocols create UE able to access the abilities and services of the communication networks. There are numerous functionalities of the radio AS protocols like dynamic allocation of radio resources, controlling bearers, radio admission control, traffic management, scheduling techniques, throughput, and bit error rate.

- **Non-Access Stratum (NAS)**

The NAS functionalities consist of the establishing of radio links between the network and the UE. Furthermore, registration, authentication, local registration management are also involved in the functionalities of the NAS protocols.

- **Radio Resource Control (RRC)**

In RRC the key functionalities of RRC contain distribution system information, an establishment of RRC connections, mobility functions, security management, QoS management and direct connection between UE and NAS.

- **Packet Data Convergence Protocol (PDCP):**

PDCP layer re-transmit the Session Data Units (SDUs). Moreover, the PDCP layer also performs header compression, ciphering and detection of the duplicate data.

- **Radio Link Control (RLC)**

In RLC there are several functionalities of the RLC, which include the segmentation of data packets according to the available size of the transport block. Besides, adjustment of errors through Automatic Repeat reQuest (ARQ) and re-segmenting for the retransmission functionalities of the RLC. Other functions contain a concatenation of same bearer SDUs, packet transfer and error discovery in the protocols.

- **Medium Access Control (MAC)**

The main functionalities of the MAC protocol are RLC SDUs multiplexing/de-multiplexing, error corrections using Hybrid Automatic Repeat Request (HARQ), scheduling, prioritization of local channel and padding.

- **Physical (PHY) Layer**

In the PHY layer, Transmission Time Interval (TTI) of a 1ms period is used by the PHY layer for sharing of a channel to the upper layer of the protocols stack. The frequency and time variation in cellular are exploited over OFDM and Single Carrier Frequency Division Multiple Access (SC-FDMA) techniques. For example, the physical layer sub-carrier spacing in LTE is 15 kHz [24][27].

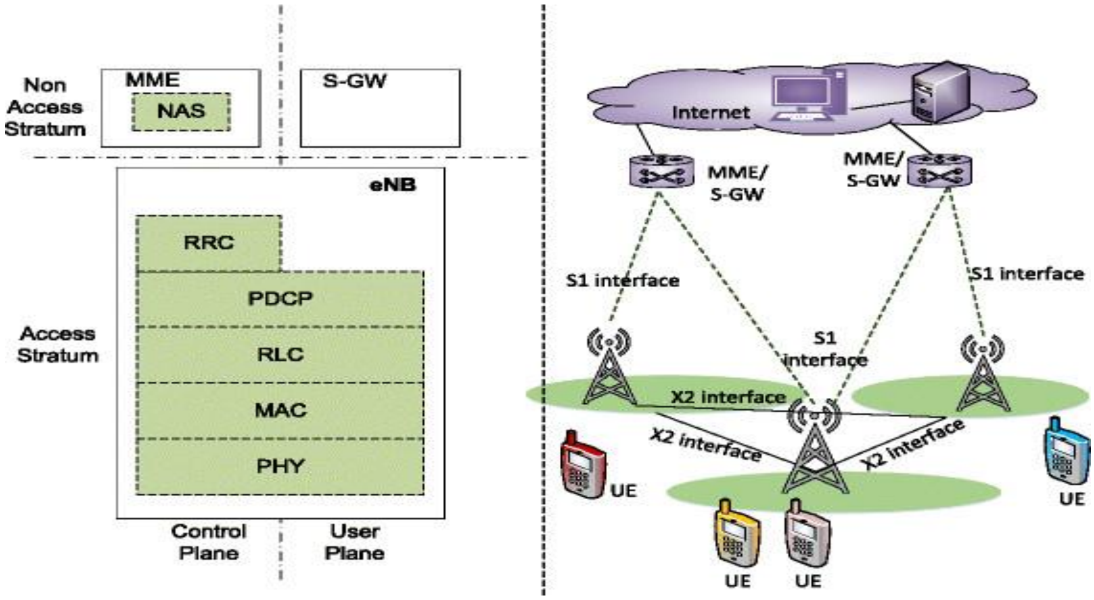


Figure 2. 2**: 5G protocols stake**

### 2.2.1. Evolved Packet Core (EPC)

The high-level architecture of EPC is described in Figure 2.2 EPC generally contains Packet Data Network (PDN) Gateways (P-GWs), Home Subscriber Server (HSS), Serving Gateways (S-GWs), Mobility Management Entity (MME), Policy Control etc.[28]. The functions of EPC include the overall User Equipment (UE) control in addition to the establishment of signalling. A simple explanation of various EPC components is shown as following:

- The function of HSS is to keep all the data of the UEs irrespective of their operators. It is transported forward from the Global System for Mobile Communications (GSM) and Universal Mobile Telecommunications System (UMTS) system architecture.
- P-GW creates a connection with the PDN over an SGi interface. Every PDN is known by Access Point Name (APN). The role of P-GW is similar to GPRS

Support Node and Serving GPRS Support Node (SGSN) in present cellular standards such as GSM and UMTS.

- The S-GW acts as a data router between evolved Node B (eNB) and the PDN.
- MME is capable of controlling high-level operations of the UEs and regularly uses the Non-Access Stratum (NAS) protocol. It processes signalling between the UEs and EPC.
- The functionalities of Policy and Charging Rule Function (PCRF) consist of policy control decision and flow-dependent charging, which lies in Policy Control Enforcement Functions (PCEF) which are present in P-GWs.
- S-GWs connect with the P-GWs among S5 and S8 interfaces. The slight difference between the two interfaces is that the two devices communicate using the S5 interface if both the UEs are using the same network operator services. If the devices belong to two different networks, then the interface between S-GW and P-GW is S8 [29].

### 2.2.2. **Evolved UMTS Terrestrial Radio Network (E-UTRAN)**

According to [30] [28], the E-UTRAN contains eNBs with the same role as that of Base Station (BS) in GSM, however with further functionalities. The E-UTRAN deals with communication between the UE and the EPC, while the eNB handles UE in one or more cells. The Serving eNB (SeNB) is the base station that presently offers services to UEs. In the downlink, eNB transmits data to all the corresponding UEs, whereas in uplink eNB receives transmissions from all the UEs to launch communication between the UE and the EPC. Additionally, the eNB controls low-level processes for instance handover functions. The eNB connects with the core network over the S1 interface. The numerous eNBs are also interconnected with each other through the X2 interface, which is primarily used for switching data such as packet forwarding in the time of handover. The S1 interface between the eNB and EPC could also make handover functions the cause of more delay. S1 and X2 interface functions use an IP-based transport network. E-UTRAN also delivers Home eNB (HeNB) services. HeNB is a base station that is precisely optimized for a user within a home or street for RN coverage. It could be linked directly either to the EPC similar to the way a normal eNB is connected or using some middle terminals such as gateways that forward data from several eNBs [31][32][33].

### 2.2.3. User Equipment

The inner structure of a UE, which could be given from [34][35]. The internal UE structure resembles the structure of UE in the 3G and 4G mobile networks. In UE, the Mobile Equipment (ME) is essentially responsible for communication purposes. The Mobile Termination (MT) within UE can control all the tasks, which are linked to create communication. Whereas the Terminal Equipment (TE) accomplishes the established communication link. The Universal Integrated Circuit Card (UICC) is an intelligent module that offers valuable data associated with a UE, cellular standard identifications and security connected information with the assistance of an application namely Universal Subscriber Identity Module (USIM) [24]. 5G delivers services to all types of IP connectivity such as IP version 4 (IPv4), IP version 6 (IPv6) and dual IPv4/IPv6. In order to connect with PDNs, every UE uses an Internet Protocol (IP). IP can be IPv4 or IPv6. According to[36], UEs have an extensive range of abilities, which include handling of maximum data rates, multiple radio access technologies, carrier frequencies for transmission and reception etc. which are carried out using signalling messages controlled by the eNBs [29].

## 2.3. 5G Enabling Technologies

5G specified the next-generation network requirements and components in its Release 8. Those main objectives include LTE and SAE for the specification of EPC, E-UTRAN, and E-UTRA. The communication between UE and E-UTRAN is accomplished using IP, which is delivered by the Evolved Packet System (EPS). In 5G, air interface and radio access networks are modified while the architecture of EPC is kept almost the same. The EPS is the basis for Long Term Evolution (LTE), Long Term Evolution-Advanced (LTE-A), and 5G networks. The main 5G features include Carrier Aggregation (CA), Enhanced MIMO technology, Coordinated Multi-Point (CoMP), and small cells such as RN cell. The next sections give more details about each technology such as CA, MIMO techniques, CoMP and small cell RN [24].

## 2.4. LTE Transmission Resources

The radio frame structure in LTE uplink and downlink is similar. According to The 3rd Generation Partnership Project  (3GPP) LTE, the system bandwidth is divided into six equal portions which range from 1 to 20 MHz and are allocated for signals transmission in uplink and downlink [29]. In the downlink, each LTE frame is further split into 10 subframes. Every single subframe is comprised of two equal time slots of 0.5 ms. each slot is further divided into 6 extended Cyclic Prefix (CP) or 7 (normal CP) OFDM symbols. For uplink and downlink communications in LTE, the smallest radio resource which is allocated to each UE is named as PRB. A single PRB is further composed of 12 consecutive sub-carriers each 0.5 ms. In normal CP, there are 84 resource elements whereas extended CP is composed of 72 resource elements because the number of individual sub-carriers in extended CP is 6. One PRB has a maximum bandwidth of 180 kHz. A physical channel possibly may comprise of multiple contiguous PRBs each of 180 kHz.

### 2.4.1. **Carrier Aggregation (CA)**

CA is among the most significant attributes of LTE-A introduced by 3GPP in its Release 10 [31]. In CA, multiple uplinks or downlinks LTE carriers are bundled either in a contiguous or non-contiguous way. The component carriers themselves are backwards compatible in order to support Release 8 and Release 9 UEs. These carriers provide signals for synchronization purposes as well as for transmitting information. The goal of CA is to increase data rates for which wider bandwidth is required for data transmission. According to [29], up to five carrier components can be aggregated in the LTE-A network.

### 2.4.2. **Component Carrier (CC) Aggregation**

The major difference in LTE and LTE-A networks is that LTE terminals operate on single carriers i.e. transmission in uplink and downlink will be a single carrier. In LTE-A, uplink and downlink transmission can take place using multiple carriers due to CA. In the frequency domain, the component carriers further consist of 112 PRBs which are backwards compatible [29]. The channel bandwidth for carriers is 1.4 MHz, 3 MHz, 5 MHz, 10 MHz, 15 MHz and 20 MHz. The number of different PRBs against the above-mentioned channel bandwidths is listed in Table 2.1.

Table 2. 1: **Number of PRBs in E-UTRA channel bandwidth** [29]

| **Spectrum in MHz** | 1.4 | 3 | 5 | 10 | 15 | 20 |
|---|---|---|---|---|---|---|
| **Number of physical resource blocks** | 6 | 15 | 25 | 50 | 75 | 100 |
| **Number of occupied carries in PRB** | 72 | 180 | 300 | 600 | 900 | 1200 |

This discussion concludes that a single UE is capable of aggregating multiple CCs in uplink or downlink direction. Figure 2.4 shows the uplink and downlink carrier aggregation. Release 8 and 9 UEs communicate in uplink and downlink using a single component carrier, while Release 10 UEs can communicate in both directions using one or more component carriers. Both contiguous and non-contiguous CA schemes are used in uplink and downlink communication. According to [29], 3GPP also proposes to use non-contiguous CA in which the aggregated components carriers (CCs) are non-contiguous in similar or distinct bands. It is since the availability of a complete spectrum of 100 MHz for some operators is sometimes not possible. For this reason, the CA in LTE-A is further categorized into the following three types:

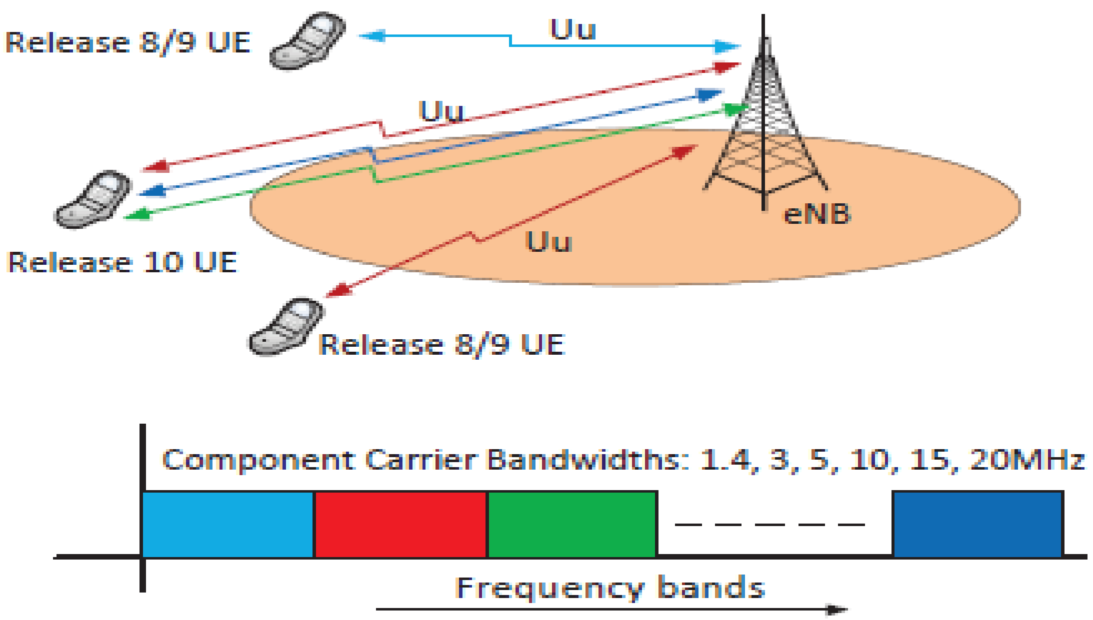


Figure 2. 3: **Component Carrier aggregation in LTA-A**

### 2.4.3. Multi-Users MIMO Scheme

The MIMO technique is based on transmission and reception of radio signals using multiple antennas. If there are uncorrelated communication channels, then multiple data transmissions can share the same frequency resources. If the multiple transmissions/receptions take place for a single UE, then it is termed as Single User MIMO (SU-MIMO). On the other hand, if the transmissions and receptions are carried out for multiple users, then it is called Multiple Users MIMO (MU-MIMO) systems. The capacity gain of the communication systems depends upon how efficiently the communication channels are used for multiple transmission and receptions. The transmission data rates depend upon several factors which include the number of transmitter and receiver antennas, the available bandwidth and how efficiently radio parameters like allocation of control channel resources are configured. MU-MIMO systems provide spatial multiplexing that is used to improve throughput, cell coverage and increases the reliability of communication systems. Among several features introduced in LTE-A, the MU-MIMO scheme is among the most essential attributes for achieving higher data rates up to 3 Gbps in downlink [37]. 3GPP LTE standard supports MU-MIMO technology with a maximum of 4 antennas at the transmitters and receivers in the downlink, while it supports one antenna in the uplink. 3GPP Release 10 extents the MU-MIMO technology. In LTE-A, the number of transmitter and receiver antennas in the downlink is 8 each, whereas in uplink it is 4 transmit and 8 receiver antennas. The reference signal design according to user-specific symbols also has been enhanced in Release 10. The enhanced design of the reference signals works well in multiple antenna transmissions receptions environment.

In the uplink direction, the data rates and spectral efficiency is significantly improved using MU-MIMO. SU-MIMO improves spectral efficiency mostly in lightly loaded networks while MU-MIMO promises for significantly improved spectral efficiency even in case of single antenna UE and improves the capacity of the network at a very low cost. LTE-A systems are capable of operating simultaneously in both modes using dynamic user-specific MIMO transmission configuration. In downlink directions, MU-MIMO has already been introduced in 3GPP Release 8. The reference signal design and codebook introduced by 3GPP in Release 8 are considered optimum for 2 and 4 transmit antennas. But the Channel State Information (CSI) could have been more accurate from UE to DeNB. The higher the number of antennas, the greater is the gain provided by Release 10 in the downlink. Besides these, the compatibility with Release 8 and Release 9 is also important. The capacity gain introduced in Release 10 also generates overhead for UEs by introducing new reference signals. Additionally, the performance of the common control channel is also degraded. Consequently, the introduction of new features and system configurations is also necessary for the optimization of the LTE network performance[29][38].

## 2.5. Heterogeneous Networks and Small Cell Deployments

With the term HetNet, I refer to RANs that present heterogeneity in the size and the placement of the cells within the same radio access technology. Generally, I can distinguish the cells in macro and small cells. A macrocell is a high-powered cell that has traditionally been used to cover large areas with a range of up to 20km. Small cells are low-powered radio access nodes operating in licensed and/or unlicensed spectrum and have a much smaller range of up to 2km. In the context of LTE, these are usually deployed in an unplanned manner in densely populated areas to increase the spectral capacity by offloading traffic or by patching areas with bad signal quality. Small cells can be further distinguished in RN (2.5km), micro (2km range), Pico (200m range) and Femto (10m range) cells. The latter are usually found indoors in homes and offices and are closed-type cells allowing access only to a predefined group of subscribers. The co-existence of various small cells within the area of a macro cell is one of the most common causes of interference in HetNets. This leads to the need for interference management among the cells, through various coordinated techniques commonly known as Enhanced Inter-Cell Interference Coordination (eICIC). In the context of this thesis and for the remaining chapters, the term small cell will be used to refer to RN, Pico and Femtocells with very small ranges[39].

### 2.5.1. Infrastructure based RNs cell

The RNs cell is categorised into fixed and mobile cells depending upon the infrastructure. RNs cell is used in distinct scenarios to improve data rates, coverage and to facilitate UEs indoor and outdoor movements. The RNs can provision UEs movements from indoor to outdoor. Also, UEs experience satisfactory coverage through mounted RNs such as at the top of a bus or train. The further classifications of the infrastructure based RNs are given below [40][41]:

Fixed RNs are mainly used to advance the coverage for those UEs, which are not close to the regular Donor eNB (DeNB) or base station which usually exits at the corner of the cells. Furthermore, the coverage holes due to shadowing are also improved. Fixed RNs can extend the cell coverage for the users outside the coverage of the regular Base stations. In Figure 2.4 the functionalities of fixed RNs are shown. The fixed RNs contain comparatively small antennas as compared to the antennas at the DeNB. The RNs antennas are normally positioned at the top of a building, tower, poles, etc.

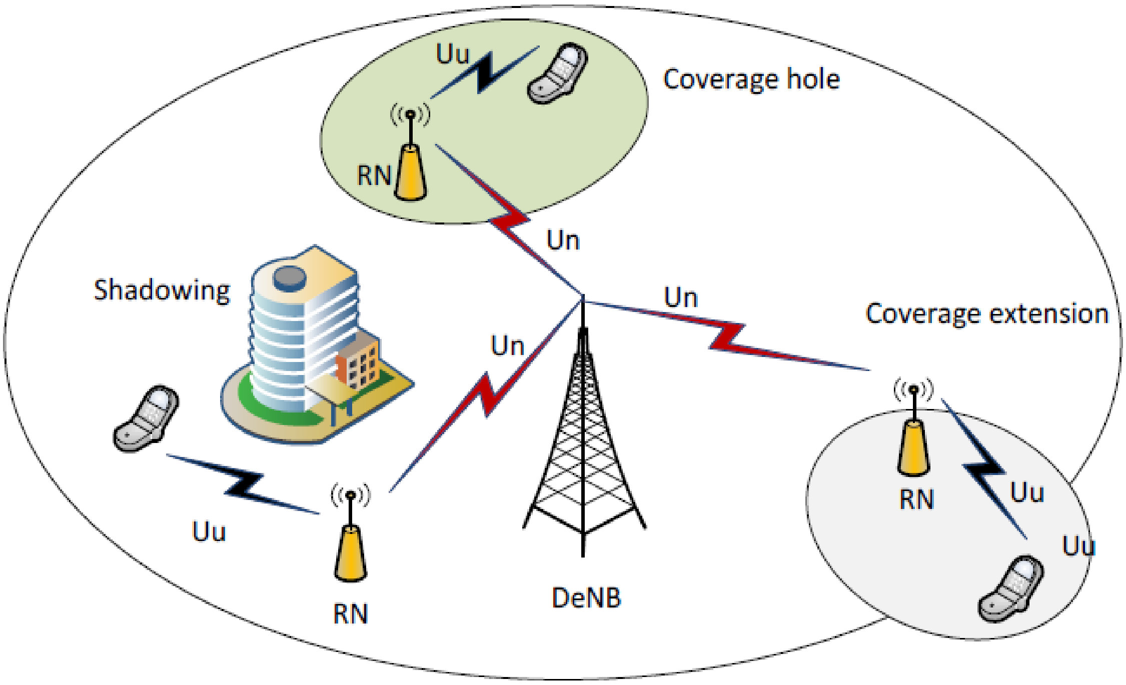


Figure 2. 4: **Fixed RN**

According to [42], 3GPP has considered mobile RNs to provide satisfactory services to the users in fast-moving trains. But, in the recent literature, it has been shown that mobile RNs can also professionally improve the services in public vehicles for instance buses and trams. The purpose of mobile RNs is to offer coverage within a moving environment. The mobile RNs are positioned on the vehicle, train, etc. and create a communication path between the mobile UEs and the DeNB. The RNs communicate with the cell through the mobile relay link (backhaul), whereas using an access link with the mobile UEs. Due to the vehicle restrictions and other safety measures, the antenna size of the mobile RNs is kept small; the functionalities of mobile RNs are shown in Figure 2.5.

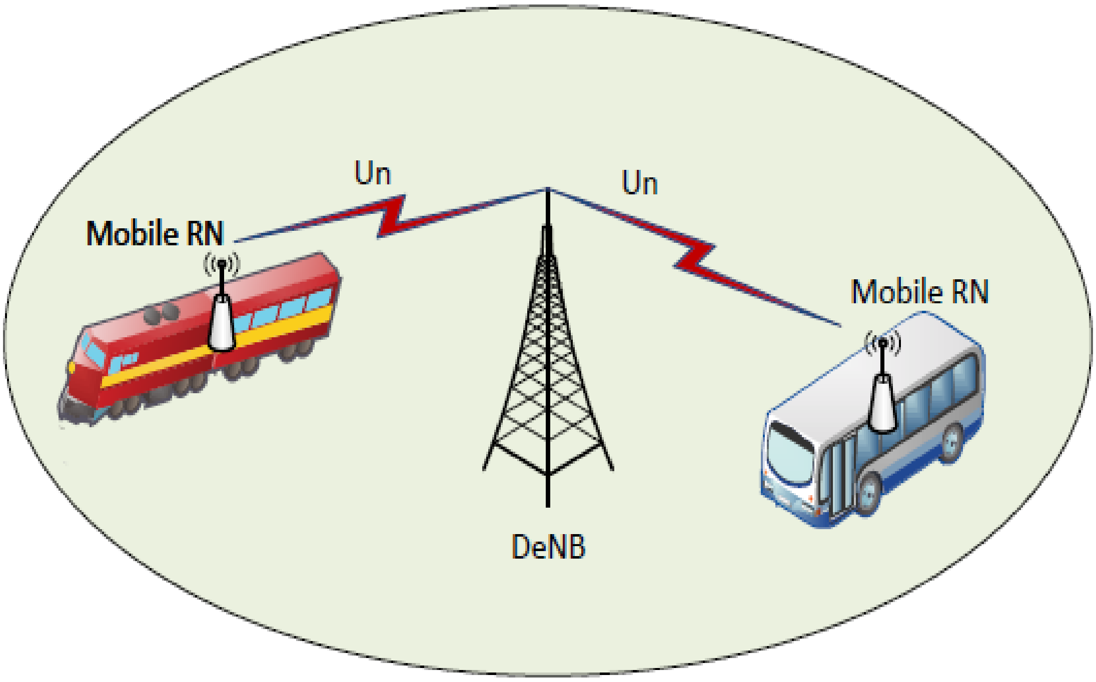


Figure 2. 5: **Mobile RN**

## 2.6. 5G Network Slicing

5G as a new generation of the mobile network is being actively discussed in the world of technology; network slicing surely is one of the most deliberated technologies nowadays. Mobile network operators such as China Mobile, and SK Telecom, and merchants such as Nokia, and Ericsson all know it as a model network architecture for the coming 5G period [43]. This novel technology allows operators to slice one physical network among numerous, virtual, E2E networks, each rationally isolated counting device, access, transport and core networks such as separating a HDD into C and D drives and devoted for diverse kinds of services with different features and QoS requirements. Every network slice, committed resources for example resources within Network Functions virtualization (NFV), Software Defined Networking (SDN), cloud computing, network bandwidth, QoS, and so on are certain as seen in Figure 2.6[44][45][46].

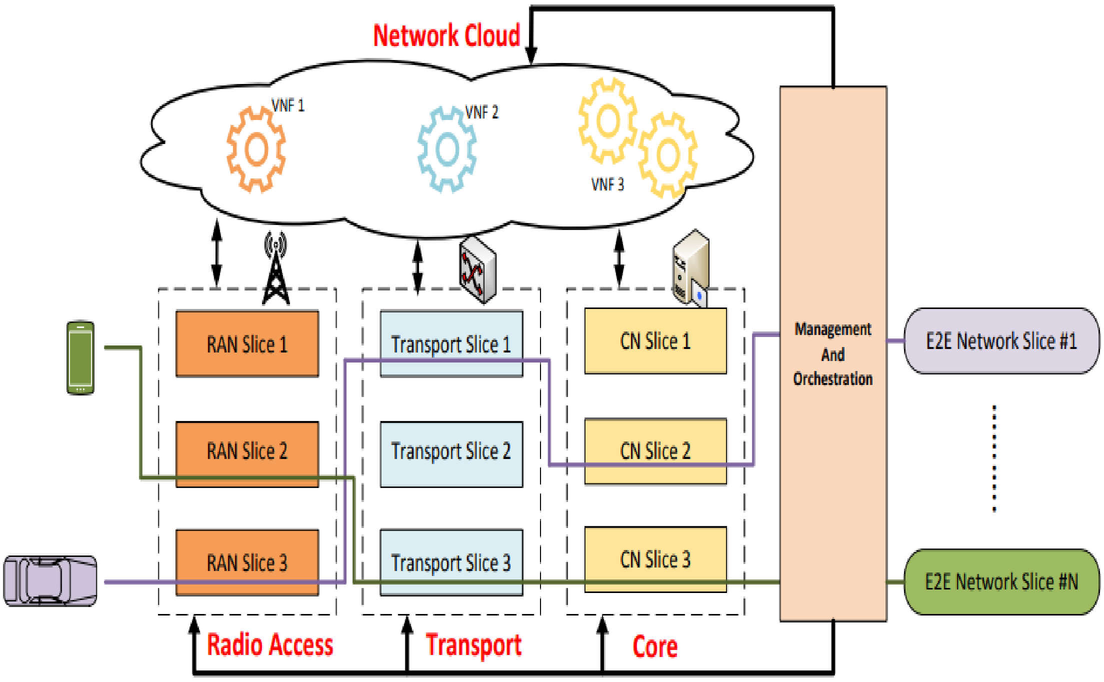


Figure 2. 6**: 5G Network Slicing**

## 2.7. IoT Technologies communication:

In recent years, research has achieved significant importance for collaborative communication between numerous intelligent systems through either wireless or wired networks. One of the resultant emerging domains in collaborative communication is the M2M communication also called Machine-Type-Communication (MTC) and is a pattern which identifies the evolving paradigm of inter-connected intelligent IoT devices communicating with each other with or without limited human interaction [47]. IoT devices' communication has achieved substantial endorsement since intelligent devices are considered as a crucial part of the communication systems having abundant application domains such as smart healthcare system, vehicular communication system, transportation, logistics system, and smart metering. There are several types of IoT devices which are already available on the market. The most important of these smart devices include sensors, actuators, embedded Digital Signal Processors (DSPs), Field-Programmable Gate Arrays (FPGAs), etc. The disparate application domains employing these smart devices include intelligent transport systems, smart metering and monitoring, home automation, e-healthcare, safety and emergency, logistic processes and many more [48][49]. The logic behind the emerging scope of M2M communication is based upon the following two points [50][51]:

- System performance is improved using networked devices as compared to isolated, stand-alone devices.
- More autonomous operations are possible in using networked devices with the advancement in the Internet and wireless communications, the IoT plays a prominent role in facilitating human life standards through providing various services.

## 2.8. IoT Network Architecture

The high-level IoT architecture According to the European Telecommunications Standards Institute (ETSI), the IoT architecture consists of the following key elements [52]:

The IoT devices are responsible for the collection and autonomous transmission of a sensor's data such as pulse rate of a patient, temperature of a room, humidity level, etc. The IoT devices are usually connected to the small local networks termed as subnets for the transmission or reception of data to or from the IoT application domains (backend servers) [52][53].

### 2.8.1. IoT Area Networks:

The IoT area networks are responsible for establishing the communication path between the IoT devices and the IoT gateways. These networks are usually called subnets which collect and route information from the IoT devices to IoT gateways. There are several subnets which are used for generating the communication link between the IoT devices and IoT gateways. Some of the major subnets are presented in Table 2.2. Generally, the use of subnets is dependent on network technology. In *fully distributed networks,* all the IoT devices are connected as peers to the network. One of the nodes which are connected to the network through the Internet acts as a router. *In client-server networks,* all nodes or devices communicate directly with the servers. Whereas, in the *cooperative networks*, all the nodes communicate with each other using some intermediate gateways[52] [53].

• **IoT Gateways:**

The intelligent sensors which are deployed to collect information communicate with the communication network with the help of IoT gateways. Hence, IoT gateways act as a bridge between sensors and communication network.

• **Communication Networks:**

The main role of the communication network is to create a communication path between the IoT devices and application servers either through wired or wireless communication networks.

• **IoT Application Domain:**

The IoT application domain consists of a middleware layer where the collected packets pass through several application services and they are used by the related agencies.

Table 2. 2 **: List of the major IoT area Networks**

| **IoT Area Networks** | |
|---|---|
| 1. | Power Line Communication (PLC) |
| 2. | Short Range Device (SRD) |
| 3. | Ultra-Wide Band (UWB) |
| 4. | ZigBee |
| 5. | Meter bus (M-BUS) |
| 6. | Wireless meter bus (M-BUS) |
| 7. | Bluetooth |
| 8. | WiMAX |
| 9. | Satellite |
| 10. | Hybrid Fiber Coax |

## 2.9. Network Simulation Tools

### 2.9.1. OMNeT++ Simulation Environment

Objective Modular Network Testbed in C++ (OMNeT++) is considered one of the open-source component-based separate event network simulators [54]. The simulator is used to support standard wired and wireless IP communication networks, however, some extensions for Wireless Sensor Networks (WSN) exist. OMNeT++ is a popular simulator, extensible and actively maintained by its user community. OMNeT++ uses C++ language for simulation models. One of the significant advantages in OMNET++ is the support of graphical tools for simulation building [54][55][56]. The disadvantage of using this simulator is a limitation of the available number of protocols.

### 2.9.2. OPNET Modeler Simulation Environment

Optimized Network Engineering Tool (OPNET) is a powerful simulation network that allows the simulation of a large-scale network. In addition, the OPNET modeller is an emerging environment suitable for the design, simulation and analysis of networks. Due to that OPNET comprises a large number of Ethernet, Asynchronous Transfer Mode (ATM), Internet Protocol (IP), and Transmission Control Protocol (TCP), as well as the easy way of simulating and flexible analysis of all parts of the data network. I have used OPNET Modeler to assess My models for several reasons. The first reason is the simulation works with Finite State Machines (FSM). The second reason it offers a graphical editor interface that is used to build models for diverse network entities from the physical layer to application processes. There are three graphical editors used in OPNET Modeler, the network editors, node editor, and process editor as shown in Figure 2.7. [57] [55][56].

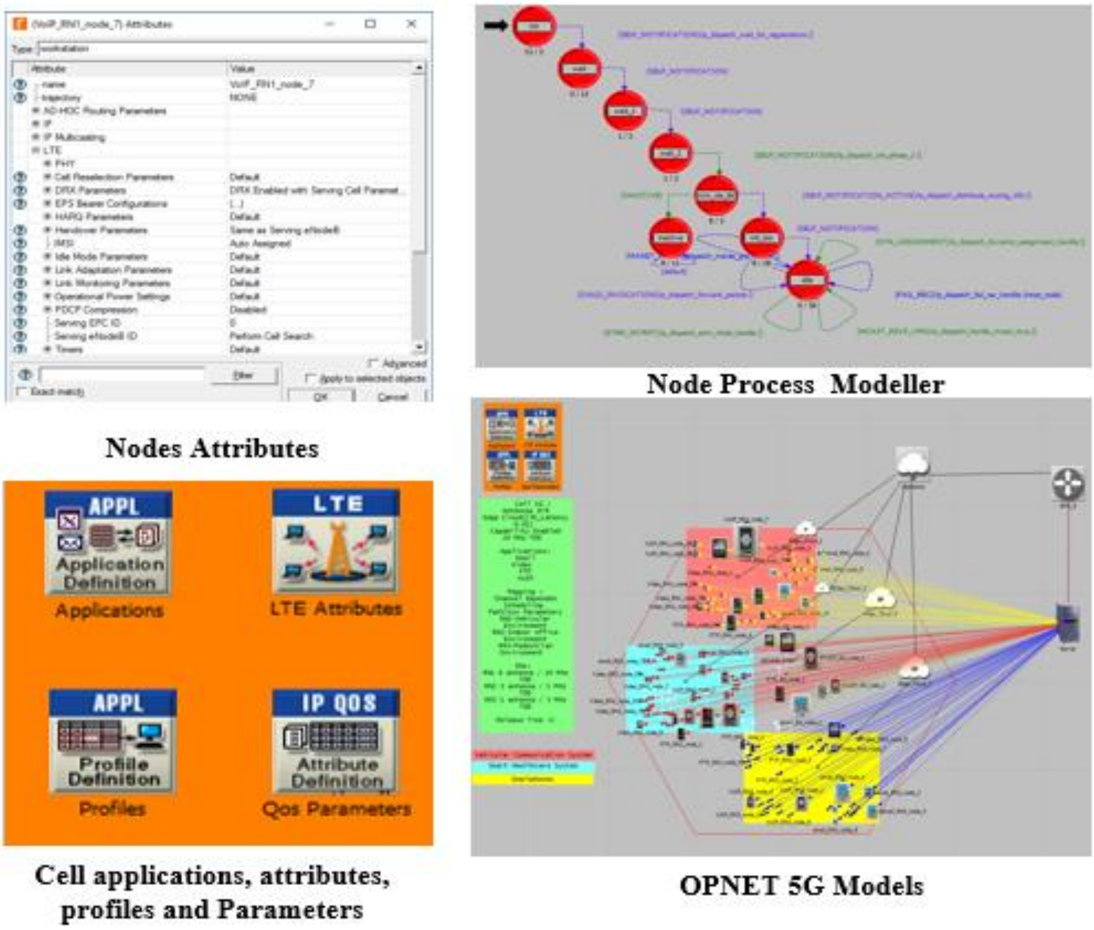


Figure 2.7**: Graphical editor used in OPNET**

The third reason is the option for connecting between OPNET Modeler and the external system. The connection is called co-simulation, using External System Interface (ESYS) to connect between the OPNET Modeler and the external system. This interface is used to ensure the control and data exchange between each system. The fourth reason is the development environment for the simulation and analysis of networks from several fields and many levels of complexity. OPNET is used to model devices, protocols, and behaviours with nearly 400 modelling functions. OPNET has a special feature to let the user adjust the existing system and produce new ones and the flexibility of generating several scenarios in the same project to create a good design for the network, using Proto-C language to program the OPNET model [58].

OPNET Modeler can use topologies created manually as well as those imported or selected from the pool of predefined ones. There is a vast number of protocol models available in the program suite. For a wireless network's simulation, OMNET uses an extension called OPNET Modeler Wireless Suite [46]. Modeler has an advanced GUI interface used for creating models, simulation execution and data analysis as shown in Table 2.3. OPNET Modeler provides a good manual, there is also dedicated technical support for commercial use of the simulator. There are also specialized training sessions provided the manufacturer to help you use the software. The system can be run under Windows or Linux [55][56][46].

Table 2 .3: **Comparison of network simulation tools**

| | **NS3** | **OPNET** | **OMNET++** | **QualNet** | **GloMoSim** |
|---|---|---|---|---|---|
| **Interface** | C++/Python | C/C++ | C++ | Parsec | Parsec (C) |
| **Graphical Support** | Limited | Yes | Yes | Yes | Limited |
| **Scalability** | Large | Yes | Yes | Yes | Yes |
| **Parallelism** | Yes | Yes | Yes | Yes | Yes |
| **Documentation and user support** | Excellent | Excellent | Good | Good | Poor |
| **Extendibility** | Excellent | Excellent | Excellent | Excellent | Excellent |
| **Emulation** | Yes | Not direct | Limited | Yes | Not direct |

## 2.10. Chapter Summary

In this chapter, I introduced the features and concepts of technologies around 5G mobile networks and IoT services related to the research area. I have presented the 5G emerging technologies and the simple features of radio resources. In addition, I have presented the heterogonous networks of small cells in 5G mobile networks such as femtocell and RN cells and how these cells can support the shapes of the 5G mobile networks and explained the network slicing concept in 5G mobile networks, how it benefits, as well as reviewed the IoT services and devices architecture and the area of the IoT networks. Besides, I have discussed diverse types of simulation tools and environments and the reasons for choosing the OPNET Modeller 18.5 to build and assess My proposed models.

# CHAPTER 3: LITERATURE REVIEW

## 3.1. Introduction

This chapter provides a comprehensive overview of the existing literature review related to HetNet consisting of macro-cells and small-cells operation in mobile networks. Three main topics are covered in this chapter, data traffic aggregation, slices resource management and packet scheduling mechanisms in DeNB and RNs in terms of enhancing the performance of networks' QoS requirements. There is some general discussion introducing literature regarding deployment, backhauling and access modes in small-cell networks. These related works have several advantages such as novel control designs based on SDN and NFV to address key core network issues of traffic and mobility management and enable mobile networks to scale in the presence of high volumes of data traffic. However, there are major disadvantages when these works are somewhat misleading as they refer only to the UE perceived performance isolation (i.e. throughput) among operators sharing the RAN radio resources and not on the functional isolation and corresponding performance isolation of the slices' virtual network functions in terms of the required computing resources processing, memory, networking, I will explain these research studies in the next sections.

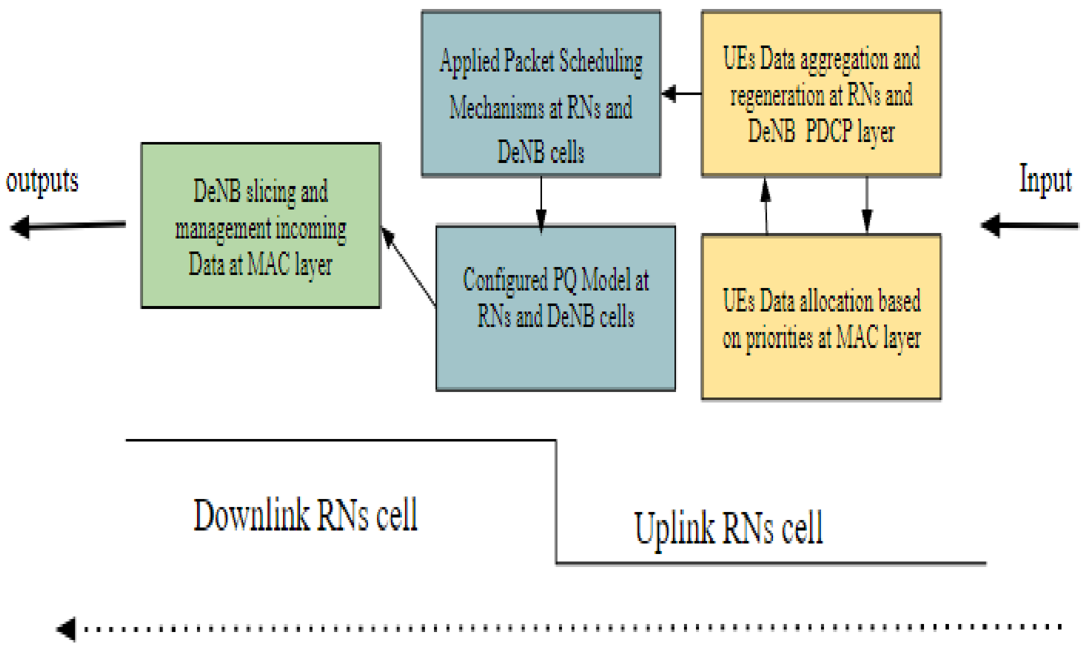


Figure 3. 1 : **Overview of System Models**

### 3.1.1. **Data Traffic Aggregation models in Mobile Networks**

The authors in [59][60] have presented unique data aggregation algorithms in mobile networks. The data aggregation method increases the lifetime of the IoT network by reducing the number of packets to be sent to sink or DeNB cell. This study shows the importance of exploring the data aggregation algorithms on the base of network topology, then they have explored several trade-offs in data aggregation algorithms and finally they have highlighted security issues in data aggregation [59][60]. This research has stated that the purpose of the data aggregation is as a good method to save the valuable energy of IoTs devices, when regularly in mobile networks thousands of IoTs are organised for area monitoring most of the intelligence and the situation and convey the data traffic to the DeNB cell and they have to gather all the data for the desired output. They have presented the aggregated data before reaching the DeNB, and this can reduce the number of packets in the network so they will have to transfer fewer packets to the DeNB cell and that can decrease the energy use of the IoT [59] [60]. These kinds of data aggregation are called In-Network data aggregation where packets are combined before reaching the DeNB. They proposed the data aggregation method as follows, data aggregation techniques discover how the data is to be sent in the network, in addition to the processing method that is used on the packets received by a Node. They have an excessive impact on the energy consumption of nodes and thus on network productivity by decreasing the length of packet or number of transmission [59][60]. This proposal describes the network aggregation method as follows: *"In-network aggregation is the global process of gathering and sending data traffic over a multi-hop network, handling data at intermediate nodes to decrease radio resource consumption (in specific energy), thus increasing network lifetime."* [59][60]. This research has limitations, there has not been work on exploring the influence of heterogeneity and mode of communication single-hop versus multi-hop on the performance of the data traffic aggregation protocols. However, there is no consideration of combining aspects such as data latency, and a network lifetime in the context of data aggregation is worth exploring [59][60].

The authors [61][62] have discussed several issues of the data aggregation framework on WSNs, they have offered a study on many energy-efficient algorithms for data aggregation. The framework works as a pathway for aggregating data measured by the number of nodes within a network**.** This study has given the data communication in sensor networks such as data aggregation, understood how communication in sensor networks is diverse from other wireless networks, and how they are energy-controlled in the network, when most of the energy is spent on conveying and receiving data, the procedure of data aggregation becomes a significant issue and a customized solution is desired. Well-organized data aggregations not only offer energy saving, however, they also eliminate redundancy data and hence deliver valuable data only[61][62].

Once the data from the source node is conveyed to DeNB cells over nearest nodes in a multichip fashion by decreasing conveying and reception power, the energy consumption is low compared to that of sending data straight to the sink of the DeNB; that is: Aggregation uses less energy than conveying data without aggregation'. Also, they have proposed energy efficiency techniques for data aggregation in wireless sensor networks; their scheme integrates energy-efficient and data storage mechanisms. This research has shown that these techniques not only decrease power consumption but also extend the lifetime of a network. However, the limitation of this research is in the case of facing challenges with the issue of topology structure, data routing, loss tolerance by containing several optimization techniques that further decrease message costs and improve tolerance to failure and loss[61][62].

The authors [63][64] have developed a data aggregation analytical model which is an effective method to reduce the energy consumption of the IoT network. They have developed an analytical model to compute the transmission delay and energy efficiency in data aggregation. Then they upgraded an extensive simulation to validate their proposed analytical model. Statistical results showed that addition of buffering time results in a major reduction in energy consumption[63][64]. The study offers guidelines to set up the buffering time for data aggregation in IoT networks. Smart devices communication technologies offer abilities for devices to communicate with each other through wired and wireless systems. Major IoT applications including alerting and habitat monitoring are considered as the scale of IoT networks become extremely large, a significant amount of transmission overhead could be observed in IoT devices such as the smart meter sensor. Since an IoT device can retrieve sensing data from the same kind of devices, reducing transmission overhead by data aggregation is an efficient way to prolong the lifetime of the IoT network [63][64].

They have utilized data aggregation, IoT devices need to exercise message buffering, and a time-based mechanism utilizes a buffering timer. Upon expiration of a buffering timer, an IoT device retrieves data collected in the buffering period and aggregates them into a new message. Once a device extends the buffering time, both the aggregated volume in a message and energy efficiency improves; however, the transmission delay from the source to the application server (or IoT gateway) rises [63] [64]. To address this issue, they have proposed an analytical model to investigate the configuration on buffering time to balance the trade-off. Nevertheless, there are disadvantages in this work, as they have investigated the effect of buffering time in the aggregation mechanism through an analytical model and simulation experiments. The disagreements between their proposed analytical model and the simulation results are within 1%. Simulation results show that an increase in buffering time (e.g., 1/50λR) results in a significant reduction in energy consumption on a message (e.g., 88%), also the research did not consider the overall delay to IoT networks as a result of the buffering time[63][64].

In this research [65][66] they evaluated an LTE-A mobile network performance in the RNs cell-based solution to integrate IoT devices into mobile networks. They developed an IoT scenario consisting of both e-healthcare and logistic devices along with normal LTE-A users in their simulation model using OPNET modeler. The simulation results indicate the better performance of the LTE-A network if the RN cell is used for the multiplexing and aggregation of IoT data traffic. LTE-A is the most current standard of wireless communication proposed by the 3GPP to serve its cellular mobile users by providing high data bandwidth. 3GPP has specified several objectives including throughput enhancement of the LTE-A cell-edge users. Many solutions are proposed by 3GPP to solve this issue including low power heterogeneous nodes and low power RN. RNs are a low power device, mostly used for coverage extension of cells [65][66]. RN connects wirelessly over the air interface to its DeNB, the communication between the DeNB and the users in the edge takes place through the RN, the link between DeNB and RN is through a downlink (Un Interface) while the link between RN and UE is through an uplink (Uu Interface).

In addition, this research has applied a layer 3 in the RN cells which, as PDCP layer, is the focus of this research because of its better throughput and Signal-to-Interference Noise Ratio (SINR). Based on the resource usage and support of two types of RNs i.e.

inbound and outbound relaying, the inbound relaying uses the same carrier frequency on both backhaul link and access link while outbound relaying uses different carrier frequencies on backhaul and access links[65][66]. However, the main limitations of this research were that it did not provide a high level of scalability compared to the huge number of IoT devices in density connection and it did not assess how the performance was affected when a huge number of IoT devices are deployed along with normal LTE-A users[65][66].

The authors [67][68] have presented a data aggregation tree model based on a reliability model, schedule data transmissions for the links on the tree and assign transmitting power to each link accordingly. Since the reliability of a link is highly related to its SINR, the SINR of all the currently used links on the data aggregation tree should be greater than a threshold to guarantee high reliability. They have formulated the joint problem of tree construction, link scheduling and power assignment for data gathering into an optimization problem, intending to minimize data gathering latency[68].

This research has studied tree-based data aggregation in WSNs. The main objective is to gather data from all sensors with low latency and high reliability, by carefully constructing a data gathering tree, scheduling links on the tree, and assigning transmitting power levels to active links in each time slot, also they proposed a joint link scheduling and transmitting a power assignment algorithm which gives high priority to links that have heavy relaying traffic load or experience severe interference. They have conducted extensive simulations to evaluate the proposed algorithms[67][68]. According to this research, results demonstrate that the proposed algorithms can significantly reduce the data gathering latency under various node densities, SINR thresholds, and traffic demands while guaranteeing that the SINR of each active link is above the edge. In addition, algorithms can distribute the relaying traffic well onto the data aggregation tree and have low requirement on the buffer size on sensor nodes. Therefore, this research has a limitation, the authors do not provide any mechanism to deal with low and middle priorities queueing data traffic, which can lead to the neglect of normal and sensitive data traffic in case of low latency and high reliability without sacrificing the energy efficiency and lifetime of the network compared to other algorithms [67][68]

In this research [69] they proposed that a network-integrated D2D communication model under channel uncertainties is investigated where D2D data traffic is carried via RNs cell. The authors have discussed the multi-user and multi-relay network and proposed a model in which they created a robust distributed solution for resource allocation to maximize the network sum-rate when the user faces the issue of interference from other RNs cell and the link gains are ambiguous. They were customizing issue is framed for allocating radio resources at the RNs cell to assess E2E rate. Also, they have considered enhancing the performance QoS requirements for LTE-A mobile network cells and D2D user equipment under total power limit, each of the undefined parameters is modelled by a restricted distance between its estimated and bounded values [69]. Furthermore, in this research they have shown due to the time-varying and random nature of wireless channels, they have formulated a robust resource allocation problem with an objective to maximize the E2E rate (i.e., minimum achievable rate over two hops) for the UEs while maintaining the QoS (i.e., rate) requirements for cells and D2D UEs under total power constraint at the RN. However, there are some limitations in this research, the authors did not consider simulation to show the results of data traffic delay through RNs cell in case of with or without as the main parameter of QoS requirements.[69].

This research [70] has presented an investigation of the issue of building a shortest path data traffic aggregation tree with maximum lifetime for WSNs. They found one shortest path data traffic aggregation tree for each sensor network that would maximize network lifetime. They have explored that this issue could be reduced to a general version of *semi-matching* issues and show that it can be solved in multinomial time. Their significant contributions are as follows:

- They have studied the issue of maximizing the lifetime for the shortest path aggregation tree
- They have restricted to shortest-path trees comes from the requirement of delay
- They have presented a centralized algorithm that runs in $O(|E|\sqrt{N} \log N)$ time, this algorithm to the best of their knowledge is the fastest in the related work, in addition, this algorithm is designed according to their results that the issue could be separated into numerous sub-issues, each of which is a general form of a semi-matching problem[70].

According to this research result to WSNs, they have offered a distributed protocol. Since it is hard to transform the fastest centralized algorithm to a distributed one, they have used another algorithm distinct from the first one centralized algorithm for the semi-matching issue, which has slightly high time complexity but is more appropriate for distributed implementation. These research results showed that their approaches significantly recover the lifetime of the network. On the other hand, in this study there is a disadvantage, the main issue of this research is the previous results showed that aggregation ability is critical in generating a multinomial time algorithm in their study, without this ability, it is challenging to discover the shortest path tree with a maximum lifetime [70].

In this research [71] the authors have developed a data aggregation framework on WSNs, the framework works as a middleware for aggregating data measured by several nodes within a network. The research proposed to compare the performance in terms of energy efficiency in comparison with and without data aggregation in WSNs and to assess the suitability of the protocol in an environment where resources are limited, and sensor networks are a collection of sensor nodes which supportively send sensed data to DeNB cell. As sensor nodes are battery-driven an efficient utilization of power is essential to use networks for a long duration, hence it is needed to reduce data traffic inside sensor networks, reduce the amount of data that needs to be sent to the DeNB cell [71]. The main goal of data aggregation algorithms is to aggregate data in an energy-efficient manner so that network lifetime is enhanced. However, this research has some limitations, for example, the framework algorithm did focus on routing the data from the source to the sink, which can show the difficulties of topology construction, data routing, loss tolerance by including several optimization techniques that further decrease message costs and improve tolerance to failure and loss [71].

### 3.1.2. **Packet Scheduling Mechanisms Comparison in Mobile Networks**

In this part of related work, I have done investigation and comparison among packet scheduling mechanisms researches, this part evaluates and compares the performance of three scheduling mechanisms, PQ, FIFO and WFQ applied based on cells of 4G and 5G mobile networks in terms of user's suitability enhance the priorities order, which will reflect on IoT devices and smart systems QoS requirements such as throughput, load, latency and fairness. The results from My performance evaluation was provided to draw conclusions on the performance of the three packet scheduler mechanisms and point out the advantages and disadvantages that are common to schedulers under study. This would help design the scheme of the schedule at RNs and DeNB cells appropriately.

In this research [72] has shown different transport network packet scheduling mechanisms for resource allocation and their influence on real-time traffic in LTE mobile networks. The research has useful information in terms of considerate basics of LTE mobile networks and packet scheduling mechanisms for additional deep studies. The authors have presented how packet scheduling mechanisms can manage resources in the transport network, showing a key role in guaranteeing to enhance E2E performance for both VoIP and FTP applications. They have introduced a packet-switched system becoming more popular for carrying data, the Internet Protocol Multimedia Subsystem (IMS) was introduced to ensure QoS for several multimedia services including VoIP. This improves the issue with VoIP calls and guarantees that a voice call quality might be as good as the circuit-switched based voice call [72]. This study has presented different packet scheduling mechanisms having various impact on different services under different traffic scenarios, therefore, they came with the choice of individual scheduling schemes based on the parameters that are used to define the QoS of the several services in the LTE mobile network. The authors [72] have indicated other reviewed research, "It has been observed that FIFO and PQ are not appropriate for high-speed networks due to their tendency to drop a large number of packets and poor reception of VoIP data, the bursty nature of WFQ does not make it receive any voice traffic", they have concluded with the purpose of providing satisfactory user performance for services such as VoIP, though at the same time offering high user capacity, the allocation of resources to users must be carefully managed, however, the packet scheduling mechanisms that allocate time slots to users must be carefully designed. However, in this research there are some limitations when the authors did not consider simulation to show the results of the density of data traffic at uplink direction of the access point via small or large cells with high capacity, which can allow them to assess to packet scheduling mechanisms and at the same time guaranteeing VoIP QoS requirements.[72].

The authors [73] have discussed and demonstrated the performance of different packet scheduling mechanisms, they have considered a hypothetical network topology in the OPNET simulations. They have designed with two routers connected by a DS1 link and all other links through cells, they have developed three different data traffic scenarios, one for each packet scheduling mechanism, it includes three types of protocols : FTP, VoIP and RTP also, a separate server for each data traffic type, they have evaluated and compared the performance of different packet scheduling mechanisms, they determined the results to collect the average queuing delay, the average E2E delay, the average packet drop-rate and the average delay jitter for each scheduling mechanism scenarios. In addition, they have focused on the impact of using Random-Early Drop (RED) scheme and tested its dominance over drop-tail policy, they included for the node with protocol FTP data traffic, they used exponential distribution for packet arrival, constant packet size and best-effort type of service, for nodes with protocol RTP data traffic, they used low-resolution video starting at 10 Frames Per Sec (FPS) arrival rate and 128x120 pixels and keep increasing this rate and size as load increases, for nodes with protocol VoIP data traffic, the voice encoder scheme is G.711, the silence and talk spurt lengths are exponentially distributed. All these settings were made using OPNET simulation attributes profile, these sets include the three packet scheduling mechanisms specification, and they set the maximum queue size to be 500 packets. They have applied minimum and maximum threshold as 100 and 200 respectively while keeping the mark probability denominator (the fraction of packets dropped when the average queue size is at maximum threshold) as 10. On the other hand, this study has a disadvantage in terms of the OPNET simulation designed when the authors do not examine the data traffic at the edges of the LTE mobile network, where the massive density can affect the packet delay and packets dropping. [73].

The authors [74] have presented comparisons of packet scheduling mechanisms performance, they have considered the OPNET simulations to assess the responses obtained from the Fibre to the home standards. The simulation base model was created with different protocols include FTP, SMTP, Database, VoIP, RTP servers and user work stations connected to the network across the passive optical splitter in the core networks. The authors have designed the OPNET simulation based on three scenarios of packet scheduling mechanisms consisting of FIFO, WFQ and Deficit Weighted

Round Robin (DWRR), they have developed all these scenarios to go across the links in the core network [74]. The results have been evaluated and showed for the Ethernet delay, VoIP packet E2E delay, FTP traffic received, and SMTP server traffic received. The total simulation was run for 180 seconds and it has been detected that the packet scheduling mechanisms have different performances for dissimilar simulation times under the specified configuration. Nevertheless, this research has a limitation when the authors do not consider an OPNET designed simulation and applied the PQ packet scheduling mechanism as one of the important packet scheduling mechanisms.[74]

Table 3. 1 : **A comparison of Aggregation Models and Packet Scheduling Mechanisms**

| **Data Traffic Aggregation models** | | | | |
|---|---|---|---|---|
| **Research** | **Contributions** | **RAT Features** | **Methodo logy** | **Limitations** |
| Pandey, [38] | Data aggregation method growths the lifetime of IoT network by reducing the number of packets to be sent to sink or DeNB cell | LTE-A, IoT DeNB cells | OPNET and MATLA B | There has not been worked on exploring the influence of heterogeneity and mode of communication single hop versus multi-hop on the performance of the data traffic aggregation protocols. |
| (Massad et al.,[39] | Data aggregation framework on WSNs, to measure by the number of nodes within a network**.** | LTE-A, DeNB cells near to nodes | NS2. | Facing challenges with the issue of topology structure, data routing, loss tolerance by containing several optimization techniques that further decrease message costs and improve tolerance to failure and loss |
| Tsai, [40] | Data aggregation analytical model is an effective method to reduce the energy consumption of the IoT network. | LTE-A, IoT gateway, DeNB cells | OPNET | The research did not consider the overall delay IoT networks as a result of the buffering time |
| Ahmad et al.[41] | They evaluated an LTE-A mobile network performance in the RNs cell-based solution to integrate IoT devices into mobile networks | LTE-A, RNs cell IoT | OPNET | The main limitations of this research did not provide a high level of scalability compared to the huge number of IoT devices in density connection |
| (Gong and Yang, [42] | Data aggregation tree model based on a reliability model, schedule data transmissions for the links on the tree and assign transmitting power to each link accordingly | LTE-A, WSNs | OPNET | It does not provide any mechanism to deal with low and middle priorities queueing data traffics, which can be led to neglect of normal and sensitise data traffics in case of low latency |

| | | | | |
|---|---|---|---|---|
| Hasan, Hossain and Kim,[43] | Network integrated D2D communication model under channel uncertainties is investigated where D2D data traffic is carried via RNs cell. | LTE-A, RNs cell IoT | OPNET | It does not consider simulation to show the results of data traffics delay through RNs cell in case of with or without as the main parameter of QoS requirements |
| (Shan et al.[44] | Research of the issue of building a shortest path data traffic aggregation tree with maximum lifetime for WSNs | LTE-A, RNs cell IoT | OPNET | The main issue of this research is the previous results showed that aggregation ability is critical in generating multinomial time Model in their study. |
| Dagar and Mahajan, [45] | Data aggregation framework on WSNs, the framework works as a middleware for aggregating data measured by several nodes within a network | LTE-A, DeNB cell, WSNs | NS3 | The framework algorithm did focus on the routing the data from the source to the sink, which can show the difficulties of topology construction |
| **Packet Scheduling Mechanisms Comparison** | | | | |
| **Research** | **Contributions** | **RAT Features** | **Methodo logy** | **Limitations** |
| Sroya,[46] | Comparisons packet scheduling mechanisms for resource allocation and their influence on real-time traffic in LTE mobile networks. | LTE-A | OPNET | The authors did not consider simulation to show the results of the density of data traffics at uplink direction of the access point via small or large cells with high capacity. |
| Velmurugan, Chandra and Balaji,[47] | Comparisons packet scheduling mechanisms, they have considered a hypothetical network topology | LTE-A | OPNET | The authors do not examine the data traffics at the edges of the LTE mobile network, where the massive density can effect on the packet delay and packets dropping |
| (Pradhan and Gregory,[48] | Comparisons of packet scheduling mechanisms performance, they have considered the OPNET simulations to assess the responses obtained from the Fibre to the home standards | Fibre networks | OPNET | When the authors do not consider in OPNET simulation designed and applied the PQ packet scheduling mechanism as one of important packet scheduling mechanisms. |

### 3.1.3. Data Traffic Slicing Models in Mobile Networks

In this part of related work, the 5G mobile network slicing requirement assessment is an important topic, which was discussed in the previous research to meet the various requirements from numerous types of use cases. Simultaneously, some of the 5G business services originate usually hard, and expensive to meet, especially QoS demands in relations of latency and throughput. It is a natural understanding that one system cannot suit all and there is a requirement to make a special network path to cover requirements of slicing services. Network slicing is presented to divide a common physical network to diverse slices to be configured for offering dissimilar QoS as demanded by the slice's operator and required by the slice's use cases or tenants. Meanwhile, these slices will be resolved by the businesses, e.g. verticals, allocating physical radio resources to the network slices; it is no longer an individual issue of performance, however, also an issue of revenue and uses case models.

In this research [75][76] the authors have developed a novel auction model; the major objective of this research was to tackle a combined resource and revenue optimization. They have concluded a general simulation study; they have demonstrated the proposed auction model can allocate network resources to network slices for offering:

- Higher satisfaction of requirements for signal network slice
- Improved network revenue.

This study was to clarify several aspects of focusing on a network slicing approach relying on a novel auction model in order to maximize the network revenue, it worked very well compared with the other potential model. The authors [75][76] said the network chunks and slices can work together in this proposed auction model and be observed as an actual method for analysis of cooperating decision making. This model was applied to the issue of price decision in their proposed system model. In addition, their proposed auction model can be performed as a significant model for resource allocating to network slices to take advantage of the network revenue. This research of the auction model studies the volume of network resources the network slices are demanding from the network. As a result, their proposed auction Model can be applied to improve the satisfaction of demands of network slices and to maximize the network revenue. In summary, the main contributions of this model are briefed as follows:

- They have formulated a novel business network model for offering 5G network slices computational and storage resources to satisfy their resource demands optimally.

- The model network slice manager has been considered to calculate diverse prices of network chunks and offer a crucial view of the network information when required.

Limitations, in this research the authors do not consider the priority of slices in the case of the model being able to minimize the satisfaction of slices with high priority while meeting the requirements of slices with lower priority; as a result of Priced-Based Network Slicing (PB-NS) the model has lower satisfactory levels as it does not consider the priority. Thus, PB-NS presents higher dissatisfaction for slices with higher priority, and this consequently decreases the overall satisfaction. [75][76].

The authors [77][78] have presented an architecture for network-slicing-based 5G mobile network systems, and introducing an architecture for handling mobility between diverse access networks, along with shared power and a subchannel allocation scheme in spectrum-allocation two-tier systems based on network slicing, where both the co-tier interference and cross-tier interference are considered. The authors aimed in architecture that 5G mobile networks would meet several uses case QoS demands in diverse application scenarios such as the case of data delivered rate and latency. Moreover, in this study authors have developed an understanding of scenarios where seamless wide area network coverage is required, 5G mobile network systems should deliver users with seamless high-data-rate services anywhere and anytime, even at cell edges or with high-speed up to 500 km/h mobility, in urban areas where the density and volume of wireless data traffic requirement are both very high, 5G networks should offer dense hotspot coverage with high radio resource capacity [77][78]. In this research scenario they consider the reliable links of a large number of extensive low-power nodes such as wireless sensors are needed, 5G mobile networks should be capable of connecting millions of IoT devices under the restraints of low power consumption and cost per device, particularly low latency and high reliability of 5G mobile networks are essential to meet the performance requirements of reliable, real-time and secure communications in some vertical use cases such as connected vehicles and manufacturing production control[77][78]. However, there is a drawback, the research has enhanced the slices use cases requirements in term of latency and data rate in density data traffic of IoT devices, unfortunately, the authors did not consider solving the issue of data traffic overload in density environment connection when the number of IoT devices is rapidly increased in the mobility connection [77] [78]

This study [79][80] has developed a logical architecture for network-slicing-based 5G systems, as well as an overview of the important concepts of network slicing. According to the proposed network architecture, the main purpose of this study is to develop an understanding of mobility management and virtualized radio resource allocation technologies in network slicing based on 5G mobile network systems. Because of the diversity and difficulty of 5G mobile network scenarios, these studies essentially provide suitable mobility management for diverse mobility scenarios. Therefore, the author proposed a handover management scheme for handovers between different access networks relying on SDN and NFV technologies. The authors focused on virtualized resource management as being responsible for inter-slice and intra-slice allocation of network resources actively and effectively. Also, the authors have proposed a novel combined power and subchannel allocation scheme for network-slicing-based spectrum-sharing two-tier networks, where both the co-tier interference and cross-tier interference are taken into account [79][80]. In a study simulation, the results display that the proposed resource allocation scheme can flexibly allocate network resources between diverse slices, thus understanding the effective allocation of network resources in 5G mobile networks. Moreover, the authors showed the future challenges and open issues on network slicing in 5G mobile networks such as the security issue. Nevertheless, this study has the main limitation that the authors do not provide customized slicing demands in both the RAN and core E2E network slicing. [79][80].

The authors [81][82] have presented a mathematical formulation of an optimization problem for an E2E network slices deployment for diverse 5G-based use-cases. Each use case, for example, video streaming, intelligent transport, e-Health and public safety, has its availability, reliability and delay tolerance demand. For the mobile network slicing, issues of the admission control and resources allocation rise. Their research is related to this context. They are attracted in the deployment of E2E network slices for 5G use cases. Therefore, they have introduced a 5G use cases classification relying on demands in terms of reliability, availability and latency [81][82]. Along with that, they have developed a heuristic algorithm which aimed to customize the slice on the most suitable resources to the use case needs, the targeted resources must be selected according to the demands of different use cases such as

reliability, availability and latency demands. In addition, in order to accomplish the highest performance and flexibility for the mobile network, 5G aimed to break the "one size fits all" approach to move to the one size per service approach [81][82].
This research addresses the optimization of slices deployment within the 5G context while taking into consideration the most important restraints for any use case. According to the literature they have provided, the main use case requirements are reliability, availability and latency. In the following, those parameters are defined:

- Reliability: The reliability is the capability to provide a service correctly according to its demand without interruption. The reliability rate required by each 5G use case based on its classification. Some mathematical standards of reliability rates for different 5G use cases are existing in the Ericsson white paper. For example, the 5G network should deliver a reliability rate that is equal to or higher than 99.999% for the ultra-reliable communication use case. For the use cases which are non-reliability critical such as pervasive video, the reliability rate may be less than 99%.
- Availability: As the 5G will be applied for the public safety and the communication of emergencies, the network must cover a required level of availability. In fact, availability is the ability to deliver a service when requested in order to achieve the required functions.
- Latency: The latency parameter based on E2E delay and data plane latency. 5G networks must assure in general 10 ms E2E latency for the non-latency critical use cases and at most 1 ms for the ultra-low latency use cases.

However, this research has disadvantages, the authors have not considered the management of the deployed slices in order to maintain the required QoS. Also, they have not tried to develop the investigation of the network performances by handling the interference between the several deployed slices in the uplink of the network [81].

The authors [83][84] have presented a designed Flexible Radio Access Network (FlexRAN) approach by efficient radio resource scheduling algorithms while considering certain guarantees for operators. The fact that radio resource sharing is relevant for efficient RAN slicing is reflected in the more recent algorithmic work in this thread [83][84]. This work is complementary to My focus on systems support for RAN slicing. They have employed the NVS scheduling algorithm in their prototype to highlight the efficient radio resource use feature of orientation. From a systems perspective, RAN sharing oriented slicing with no functional isolation among slices has been explored by the FlexRAN platform; network slicing is enabled with FlexRAN by programmatically defining how the radio resources need to be allocated among the connected UEs based on the requirements of the slice they belong to. A unified control plane, which is controlled by a single entity usually the infrastructure provider is responsible for performing the corresponding control operations. However, this approach has a limitation, since the capabilities of the slices are fully dependent on the types of control functions that are bundled in the control plane of the FlexRAN approach [83].In addition, a common and key limitation of the stated FlexRAN approach is that there is no consideration to implement associated challenges such as real-time control in their designs and they do not tackle the issue of separating the control and data planes in the RAN practically and concretely . Moreover, this work does not offer Models to take control in the RAN adaptive and flexible by allowing a dynamic functional split such as centralized to distributed scheduling and vice versa depending on the deployment scenario and the constraints posed by the underlying network conditions at any given point in time. [83][84].

The research [85][86] has presented the RAN approach of allowing independent and fully customizable control planes for each slice so that slice owners can flexibly introduce their functionality in the RAN and tailor their slice as per the needs of their service. The approach has accommodated this need for slice customizability, the RAN slicing approach seeks full isolation by running the virtual cell instance of a slice by assuming dedicated radio hardware and spectrum per slice, bearing some similarity to the RAN form of RAN sharing in that resource sharing among slices is limited at best to computing, memory and storage resources [85][86]. In general, this research is into wireless virtualization overview and position functional isolation for slices, which supports isolated and customizable control plane for each slice. In terms of radio resource virtualization to the domain of RAN slicing, the research has encouraged the need for abstractions that decouple the control plane decisions from the physical radio resource grid [85][86]. However, this approach has limitations when the authors do not consider the restraints that can be compulsory by the physical layer such as frequency-dependencies in scheduling, also, this has the disadvantage of inefficient use of radio resources and foregoing potential statistical multiplexing gains although the focus here is on the idea of a network store for VNFs to aid in dynamic network slicing.[85][86].

Table 3. 2 : **A comparison of Slicing Models**

| **Data Traffic Slicing Models** | | | | |
|---|---|---|---|---|
| **Research** | **Contributions** | **RAT Features** | **Methodology** | **Limitations** |
| Jiang, Condoluci and Mahmoodi, [49] | Novel auction model the major objective of this research to tackle a combined resource and revenue optimization. | 5G, edge network | MATLAB | This research the authors do not consider the priority of slices in case if the model is able to minimize the satisfaction of slices with high priority while meeting the requirements of slices with lower priority |
| Zhang et al.,[50] | Architecture for network-slicing-based 5G mobile network systems and introducing an architecture for handling mobility between diverse access networks. | 5G, IoT | OPNET | The authors did not consider the solve the issue of data traffic overload in density environment connection when the number of IoT devices are rapidly increased in the mobility connection |
| (Afolabi et al.,[51] | Logical architecture for network-slicing to develop an understanding of mobility management and virtualized radio resource allocation technologies in network slicing | 5G, core network SDN and NFV | OPNET | This study has the main limitation when the authors do not provide customized slicing demands in both the RAN and core E2E network slicing |
| (Kammoun et al.,[52] | Mathematical formulation of an optimization problem for an E2E network slices deployment for diverse 5G-based use-cases. | 5G | OPNET | The authors have not considered the management of the deployed slices in order to maintain the required QoS |
| (Kamel, Le and Girard,[53] | Flexible Radio Access Network (FlexRAN) approach by efficient radio resource scheduling algorithms while considering certain guarantees for operators | 5G, NVS | MATLAB | This work does not offer Models to make control in the RAN adaptive and flexible by allowing a dynamic functional split such as centralized to distributed scheduling and vice versa depending on the deployment scenario |
| Nikaein et al.,[54] | RAN approach of allowing independent and fully customizable control planes for each slice so that slice owners can flexibly introduce their own functionality in the RAN and tailor their slice as per the needs | 5G | OPNET | The authors do not consider the restraints that can be compulsory by the physical layer such as frequency-dependencies in scheduling. |

## 3.2. Chapter Summary

This chapter discussed the related work about the models that are covered in this thesis research area, including data traffic aggregation models, slices resource management and packets scheduling mechanisms in DeNB and RNs cells of mobile networks in term of supporting the IoT nodes and the cells' overall performance. Some of the related work discussed and introduced deployment to reduce the nodes' energy consumption, other models to work as brokerage to offer pre-slicing prices for the tenants, also the other models focus on the downlink or the core of mobile networks and other RAN or access modes in a small-cell zone. These related works have several advantages such as novel control designs based on SDN and NFV to solve key core network issues of traffic and mobility management and enable mobile networks to scale in the presence of high volumes of data traffic. However, there are many disadvantages when these works are somewhat misleading as they refer only to the UE supposed performance separation such as throughput or latency among operators sharing the RAN radio resources and not on the functional isolation and corresponding performance isolation of the slices' virtual network functions in terms of the required computing resources include processing, memory, and networking.

# CHAPTER 4: DATA TRAFFIC AGGREGATION MODEL OVER 5G MOBILE NETWORKS

## 4.1. Introduction

In this chapter, I presented the 5G mobile networks as representing a promising mobile network to efficiently support emerging smart systems with a large volume of data traffic. Over the last few years, IoT devices have seen exponential growth over mobile networks, which resulted in the need to increase the capacity due to generating higher data rates. These mobile networks are expected to face challenges, such as the support of significant data explosion and various QoS requirements. The chapter started with a theoretical and conceptual background on the data traffic models over different mobile network environments with the overall implications of the data size. It also discusses selected related work on data traffic aggregation models. This chapter demonstrates a novel Data Traffic Aggregation (DTA) Model along with a Resource Allocation Scheme (RAS) based on slicing technique. Simulation results are demonstrated to evaluate the performance of the proposed Model and Scheme.

Due to the fast growth of mobile networks such as 5G mobile network and IoT devices' increasing demand for services with high QoS requests [87], the managing of network resources becomes a permanently more challenging step that requires being correctly designed in order to advance network performance. The applications' areas of IoT devices contain for example smart office, VCS, smart alerting system, SHS, and logistics system [21][22][64][66]. Therefore, the 5G mobile network is getting importance as an effective generation of mobile networks to introduce flexibility in the radio resources management, as 5G slice is a gathering of the selected network resource to satisfy the service(s) QoS demands.[88][89][90].

An assisting aspect of network slicing is the virtualization of network resources, which allows network operators to share the common physical resources in a flexible, dynamic manner in order to utilize the existing resources in a more effective approach [91]. In My proposal, 5G radio resources are efficiently utilized as the smallest unit of a Physical Resource Blocks (PRBs) in a RNs or DeNB cells by allocating the data traffic of several devices as separate slices based on QoS requirements for each smart system [91][92][93]. Due to the various QoS demands and the limitation of network resources, competently allocating network resources between service slices and users or IoT devices is a major challenge [94][95].

### 4.1.1. IoT Device Challenges

In 4G and 5G mobile networks, there is an expectation of massive access such as smart devices, and IoT devices can lead to serious wireless challenges in term of RAN overload and congestion. Since radio resources are an essential component and hardly exist, therefore, the efficient utilization of these radio resources is required. The mobile networks, such as 4G and 5G; make use of multiple carrier schemes to offer better data rates and to ensure high QoS. The smallest resource unit allocable in the 5G system to a single device is the PRB as illustrated in Figure 4.1. Under favourable channel conditions, PRB is capable of transmitting numerous kilobytes of data. These multiple carriers' schemes are capable of transmitting a large amount of data. However, in the case of smart or IoT devices communication, both narrowband and broadband applications have to be considered to enhance QoS requirements. Especially, these applications have different sizes of data traffic, which need QoS specifications such as real-time, accuracy and priority. If one PRB is allocated to a single smart device for data transmission of just a few bytes, then it might cause severe wastage of radio resources, also, the different types of data traffic should be considered in the 5G slices approach. Therefore, the full radio resources utilization and data traffic classification should be a brilliant solution to the data traffic explosion and the fairness of services in the near future[96][60].

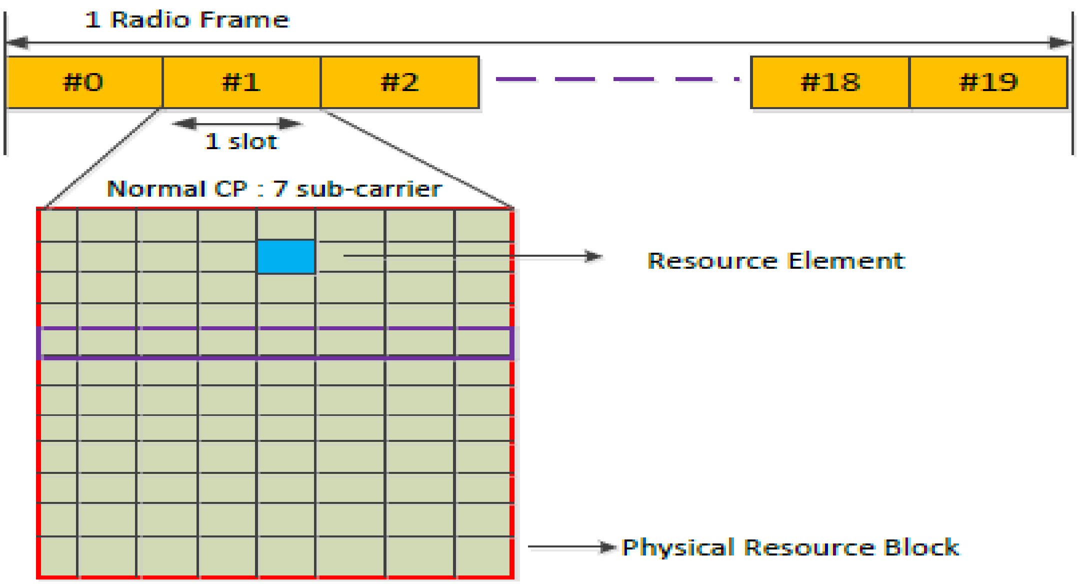


Figure 4. 1: **Physical Resource Block (PRB)**

## 4.2. Infrastructure based Relay Nodes (RNs)

The RNs cell is categorized into fixed and mobile RNs depending upon the infrastructure. RNs are used in distinct scenarios to improve data rates, coverage and to facilitate UEs' indoor and outdoor movements. The RNs can provision UEs' movements from indoor to outdoor. Besides, UEs experience satisfactory coverage through mounted RNs such as at the top of a bus or a train. The further classifications of the infrastructure based RNs are given below [40]:

### 4.2.1. Fixed RNs

Fixed RNs cell is mainly used to advance the coverage for those UEs, which are not close to the regular Donor eNB (DeNB) or base station which usually exits at the corner of the cells. Furthermore, the coverage holes due to shadowing are also improved. Fixed RNs can extend the cell coverage for the users outside the coverage of the regular base stations as shown in Figure.4.2. The functionalities of fixed RNs contain comparatively small antennas as compared to the antennas at the base stations. The RNs antennas are normally positioned at the top of a building, tower, poles, etc[90].

### 4.2.2. Mobile RNs

Mobile RNs according to 3GPP has considered providing satisfactory services to the users in fast-moving trains. However, in the recent literature, it has been shown that mobile RNs can also professionally improve the services in public vehicles for instance buses and trams. The purpose of mobile RNs is to offer coverage within a moving environment. The mobile RNs are positioned on the vehicle, train, etc. and create a communication path between the mobile UEs and the Base station (DeNB). The RNs communicate with the DeNB through the mobile relay link (backhaul), whereas using access link with the mobile UEs. Due to the vehicle restrictions and other safety measures, the antenna size of the mobile RNs is kept small and the functionalities of mobile RNs [90]

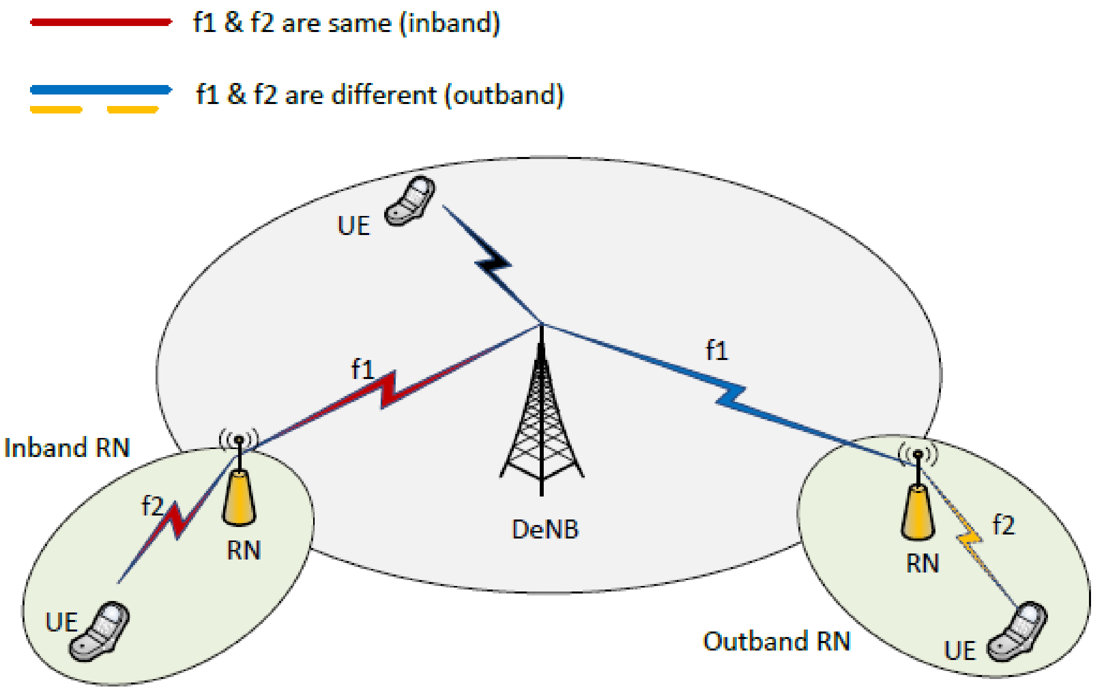


Figure 4. 2: **Relaying in 5G**

### 4.2.3. 5G Network Slicing

5G as a new generation of the mobile network is being actively discussed in the world of technology; network slicing surely is one of the most deliberated technologies nowadays. Mobile network operators such as China Mobile, and SK Telecom, and merchants such as Nokia, and Ericsson all know it as a model network architecture for the coming 5G period [43]. This novel technology allows operators to slice one physical network among numerous, virtual, E2E networks, each rationally isolated counting device, access, transport and core networks such as separating a Hard Disk Drive (HDD) into C and D drives and devoted for diverse kinds of services with different features and QoS requirements. Every network slice, committed resources for example resources within Network Functions Virtualization (NFV), Software Defined Networking (SDN), cloud computing, network bandwidth, QoS, and so on are certain as seen in Figure.4.3 [44][97][90].

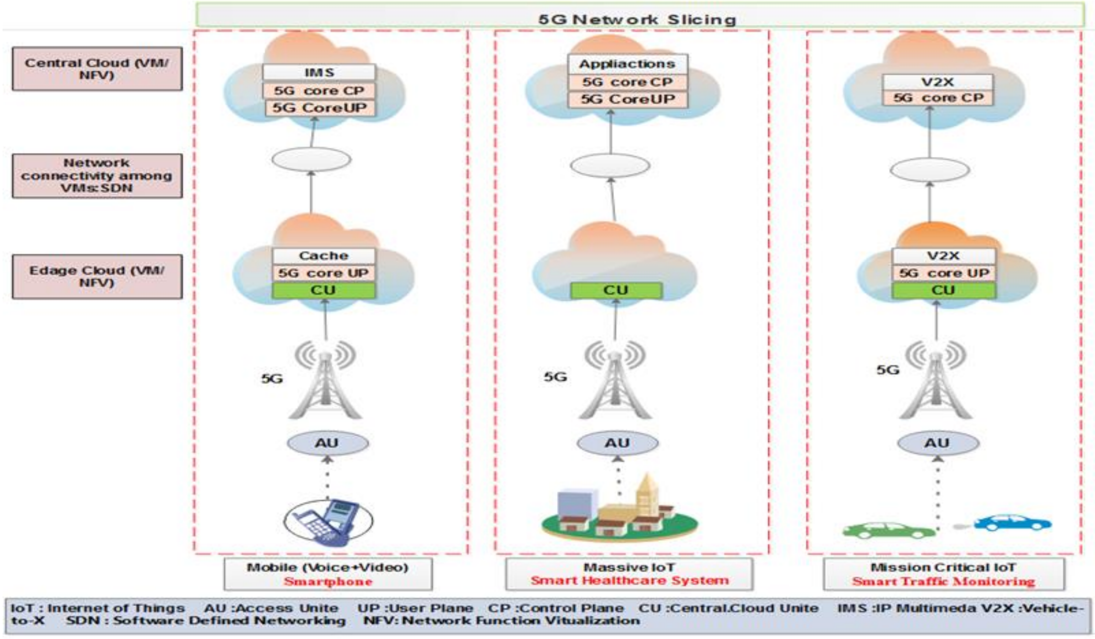


Figure 4. 3: **5G Network Slicing**

## 4.3. Data Traffic Aggregation Model

The proposed DTA model is relying on aggregating data from several IoT devices at the Packet Data Convergence Protocol (PDCP) layer of the RN and DeNB cells. The PDCP layer performs header compression, retransmission, and delivery of PDCP Session Data Units (SDUs), duplicate detection, etc. In the proposed model, the PDCP layer is used for the aggregation of the IoT devices' data in the uplink of 5G mobile networks. The main reason for selecting PDCP for aggregation in the uplink is to aggregate data packets with a minimum number of the additional headers as shown in Figure.4.4.

The individual data packets from several IoT devices approach the PHY layer of aggregation device with various intact headers such as Medium Access Control (MAC), Radio Link Control (RLC), and PDCP. The headers are removed as the received data is transported to the upper layers. On arriving at the PDCP, all the headers are removed and only the payloads from the individual devices are available, which are aggregated [95][96][64].

One single aggregation buffer *B* at the RNs cell is considered to aggregate IoT devices' data packets. This buffer aggregates data packets from different IoT devices ensuring and improving QoS for IoT device nodes as well as RNs and DeNB cells in 5G mobile network. In this implementation, the RN cell is worked as a user for the DeNB cell for

5G mobile network data traffic. In order to reach the maximum performance improvements in spectral efficiency, packet propagation delay, and cell throughput, I consider scenarios in which all the IoT devices communicate with the DeNB cell through RNs. The IoT devices' data aggregation Model is shown in Figure 4.5, and Table 4.1 as described as follows:

- Data from *K* IoT devices are considered for aggregation.
- The essential parameter for IoT devices' data aggregation is the maximum delay time *Tmax* for the packet at the RN cells and DeNB cell.

The maximum delay time *Tmax* is an essential parameter for IoT devices data packet and is calculated according to the various traffic classes of the IoT device priority. IoT devices data packets have different priorities according to their applications. For example, data packets received from the IoT devices deployed in Vehicular Communication System (VCS) scenario for the accident alerting or two vehicles communication have high priority over the packets from IoT devices, which are deployed in smart sensors. The data packets from these IoT devices having the highest priority along with a lower level of latency. Therefore, I initiate the *Tmax* value as the inter-send time of the IoT devices data with the highest priority. For example, in the simulation setup, these IoT device applications are separate in terms of customizing the low level of latency, the inter-send time of the IoT devices traffic model is 5 ms, which is the maximum time a packet is delayed at the RNs. The reasons behind determined the 5ms as the maximum time for the packets to go through the cell, as I have tried scale time between 10 until 0 ms as delay time to made buffering for incoming IP packets from different smart systems and the simulation showed that 5 ms was a perfect time that it can allow varieties of priorities level worked smoothly and exceptional to serve all of my smart systems used cases.

Thus, the value of the *Tmax* is initiated as 5 ms, which means that the data packets received from the distinct IoT devices are delayed for 5 ms at the RN [95] [96].

The value of *Tmax* is adaptive, i.e., the Model updates the value of *Tmax* if RN receives packets from a device, which has higher priority than the priorities of all the other devices in the queue of the RN. The data from all the IoT devices are buffered at the RN cell and DeNB cell. The individual IP headers of all the IoT devices are kept intact. The data packets are buffered until time delay approaches *Tmax*.

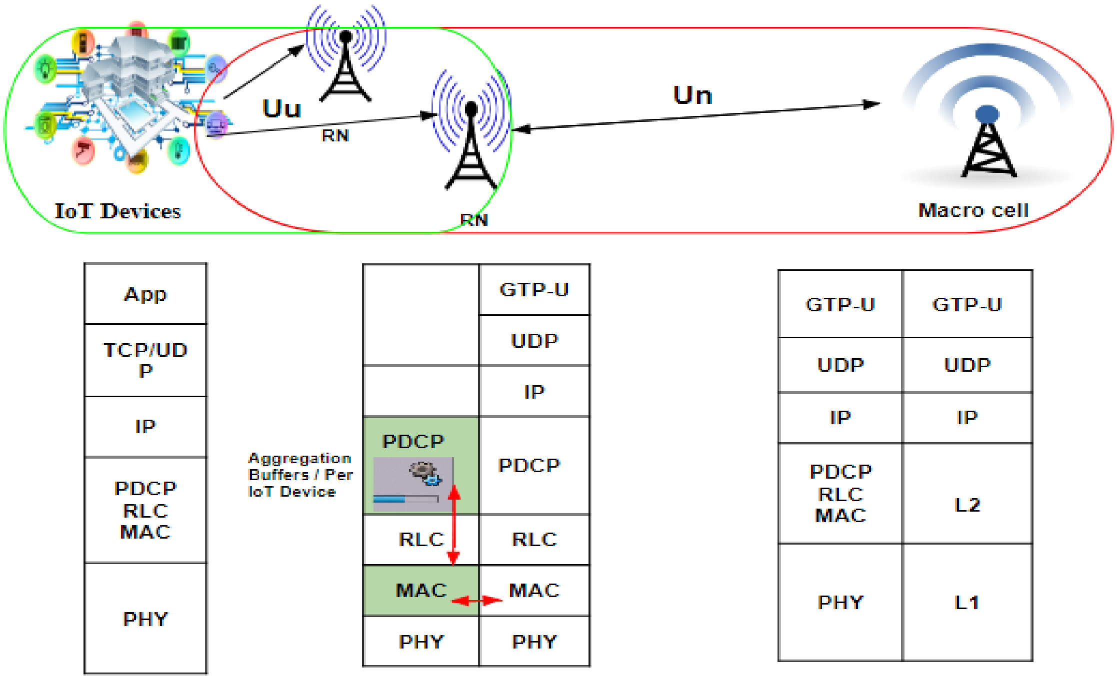


Figure 4. 4: **IoT Devices Data packets flow diagram**

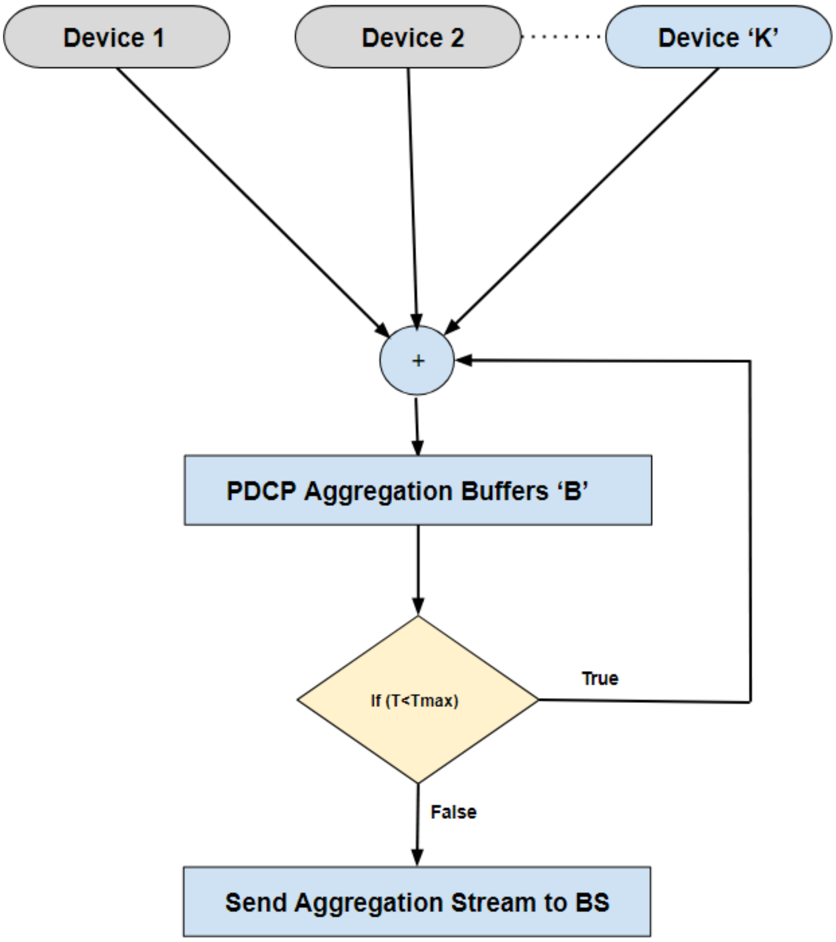


Figure 4. 5 : **Data Traffic Aggregation Model Flow chart**

Table 4. 1**: Data Traffic Aggregation model**

| Model 1: An overview of the proposed data aggregation Model in the RN PDCP. |
|---|
| Aim: Efficient utilization of PRBs among IoT devices initialization; |

***set expiry timer* T*max* == 5 *ms*;**<br>
***set Bmax* == (available TBS – RN Un overhead);**<br>
***set timer T* == 0;**<br>
***set multiplexing buffer B == 0;***<br>
**schedule RN and allocate PRBs (e.g., 5 PRBS are set for RN to analyse multiplexing process);**<br>
**schedule IoT devices within the coverage of RN for uplink transmission;**<br>
**while *packet arrival* == *TRUE* do**<br>
**start multiplexing process based on the value of timer and the size of the multiplexing buffer;**<br>
    **if *T < Tmax && B < Bmax* then accumulate incoming packet into buffer *B*;**<br>
        **increment timer T;**<br>
    **else**<br>
        **re-assemble aggregated packet of size *available TBS – overhead from buffer B*;**<br>
        **send large multiplexed packet to RN PHY via RN Un protocols;**<br>
        **add RN Un protocols overhead; route multiplexed packet to BS in next TTI;**<br>
        **reset timer *T*;**<br>
    **end if**<br>
**end while**

## 4.4. Resource Allocation Scheme (RAS)

To support the DTA model which I have developed, accompanied by the previous model Resource Allocation Scheme (RAS), this scheme mainly works on the application layers in the 5G mobile networks to offer exceptional QoS requirements over priorities level of IoT devices. In this case, I designed a form of virtualization radio resources in order to perform the RAS for application network slices. Certainly, the main aspect to be considered is the way PRBs are allocated to dissimilar slices to achieve the requirements of such slices [96][98]. The duty relevant to RAS becomes more challenging with network slicing, as it introduces a two-tier priority in the mobile network system. The first tier refers to the priority of different slices, i.e., inter-slice priority, as each slice has its priority defined according to the agreements between the network provider and the slice owner (tenants). The second tier refers to the priority between the users of the same slice, i.e., an intra-slice priority. Once looking at the solutions exploited over the existing 4G mobile network system to cope with PRBs, it emerges that 4G networks can maximize the QoS of the served users, however, they are not capable of performing the PRB allocation in slicing environments. This limitation is since RAS in a 4G mobile network system is performed by assigning the priorities to the requested services via the UE. This method thus fails when considering that in the 5G mobile network system different UEs may belong to different slices with different priorities, and thus such UEs should be managed by considering the priority of the slice they belong to plus the priority of the service they need [96][98].

This scheme is shown in Figure 4.6, it exploits a two-tier priority level relying on the idea that network slices communicate to a RAS entity with the desired QoS level. The RAS based on the priority of the slice decides about serving the slice. However, according to the inter- and intra-slice priority, the virtual network allocates the PRBs to the UEs or IoT devices of the admitted slices. Therefore, the RAS the resource allocation mission is performed with the purpose to maximize the Quality of Experience (QoE) of the users inside each slice, by considering the inter-slice priority. The QoE is measured by considering the effective throughput experienced by the users, normalized according to their maximum demanded data rate. With this target, the resources allocated to a slice with low priority could be reduced, if needed, down to the minimum amount capable of meeting the basic QoS requirements to admit new slice(s) with higher priority. Therefore, My proposal dynamically changes several network PRBs allocated to network slices[96][98],

according to the packets load without affecting the QoE of the users and while improving the network utilization.

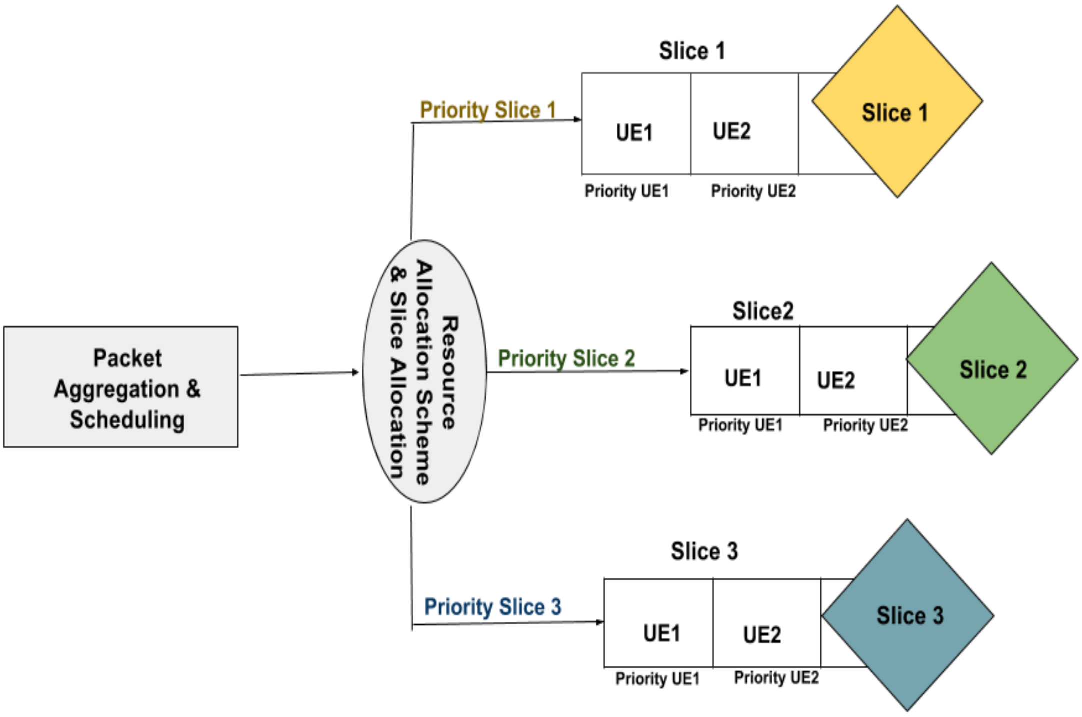


Figure 4. 6**: RAS with inter-slice and intra-slice priority**

## 4.5. Resource Allocation Scheme Environment

According to the 5G slicing technology concept, I have focused on classifying and measuring QoS requirement and data traffic of IoT device applications in the smart city use cases such as smartphones, smart healthcare system, and vehicular communication systems as results of IoT device applications (services) have data traffic characteristics in 5G slicing technology are relying on the content type of data, amounts typed of flow data, priority of data transmission and data transmission mode. As shown in Figure.4.7. These IoT devices Content type contains voice and video streaming; amounts type consists of different size, large size refers to the number of packets more than 1K bytes, small size refers to the number of packets less than 1K bytes [98]. Transmission method contains the periodic transmission, continuous transmission, burst transmission and time-response transmission; priority of transmitting consists of low, medium and high. Depending on the IoT device applications slicing, my research would have classified them into three main slices based on QoS requirements and data traffic types.

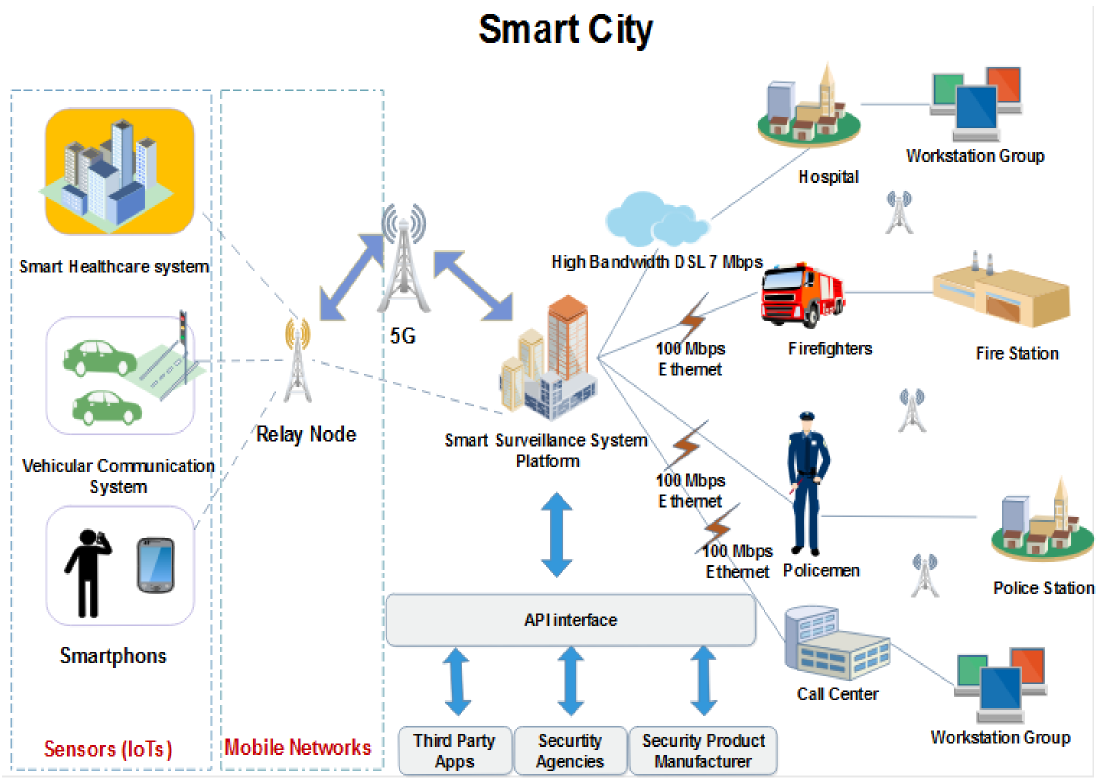


Figure 4. 7: **Smart systems in a smart city use case.**

## 4.6. Resources Allocation Scheme Strategy

As depicted in Figure. 4.8 My scheme consists of four main elements:, the physical resources (PRBs), the virtual network layer, the service slice layer and the RAS[96] [90], [98].

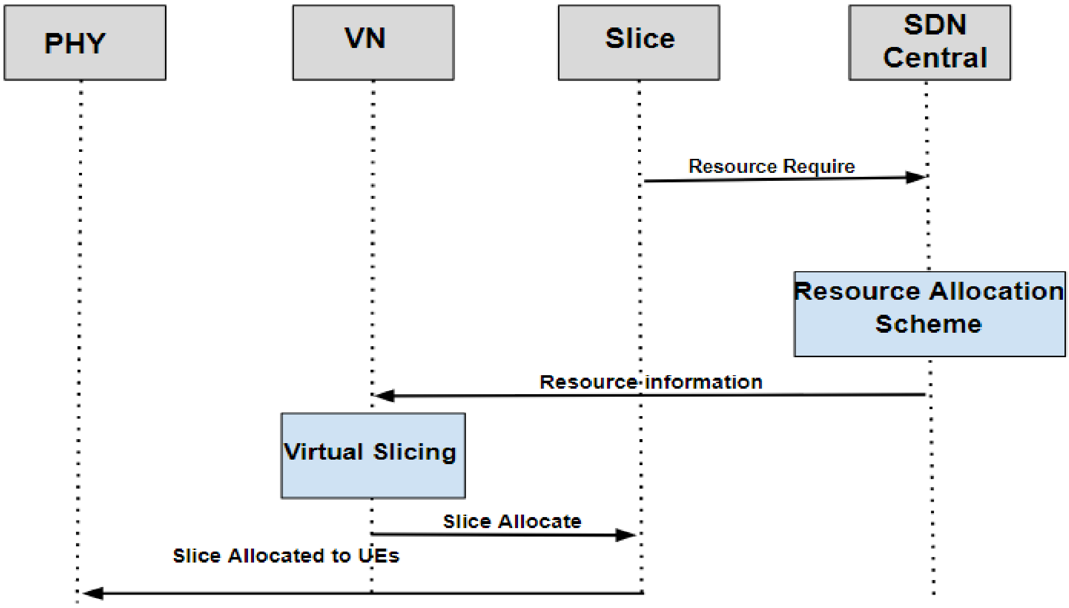


Figure 4. 8: **Flow chart of RAS**

### 4.6.1. Service Slices

The service slices offer different services (e.g., smartphones, vehicular communication system and smart healthcare system) which need resources to be served. I designate with $\mathbb{S} = \{1, 2, 3... S\}$ the set of slices in the virtual network [66][90], [98]. Each slice $s$ has a set of UEs, such a set is symbolized by $\mathbb{U}s = \{1, 2... Us\}$. Each slice $S$ performs a request to the RAS in terms of QoS restraints. In this section, I modelled such a request with $R_S mins$ and $R_S max$, which denotes the minimum and maximum data rates associated with the slice $s$, respectively. Each slice $s$ is characterized by a priority, $\mathbb{p}s$, where such priorities are defined with the constraint that $\sum s \in S\ \mathbb{p}s =$ 1.Similarly, each user $u$ belonging to the slice $s$, i.e.,$Us$ is characterized by a priority, where $\sum us \in Us\ \mu us = 1$

### 4.6.2. Virtual network

The virtual network layer delivers an abstraction of the physical network resources (PRBs), According to the decisions of the RAS, the virtual network slices the resources of the network to accommodate different slices. The virtual network receives the requests of different slices in terms of UEs to be served for each slice, and executes the subsequent allocation of physical resources according to the inter- and intra-slice priority while considering the QoE of UEs [66].

### 4.6.3. **Physical resources**

The physical resources denote the PRBs available in the virtual network. For the purpose of simplicity, I refer to the uplink channel of one RN cell. The total available bandwidth is indicated by $B$ MHz. The set $\mathbb{M}$ = {1, 2...$M$} represents the available sub-channels,

where the bandwidth of the generic sub-channel $m$ is $bm = \frac{B}{M}$. The total transmits Power $PTOT$ is uniformly allocated to each sub-channel, i.e., $pm = \frac{B}{M}$.

When PM is assigning the physical resources, I consider the channel conditions of the UEs. I assume that channel condition is determined by transmission path loss and shadowing components. The path loss and the shadowing fading path loss are assumed to be a Gaussian random variable with zero mean and $\sigma$ standard deviation equal to $8dB$ [99]. So, the path loss is based on the distance value $dus$ between a generic UE and the RNs.

## 4.7. Services Allocation Uses Case

The 5G mobile network terminal offers exceptional QoS through a diversity of networks. Nowadays, the mobile Internet users choose manually the wireless port of different Internet Service Providers (ISP) without having the opportunity to exploit the QoS history to choose the suitable mobile network linking for a provided service. In the future, the 5G phones will offer a chance for QoS analysis and storage of measured data traffic in the mobile network terminal. There are diverse QoS parameters (e.g., bandwidth, delay, jitter, and reliability), which will support in future the 5G mobile running in the mobile terminal. System processes will offer the best appropriate wireless connection based on needed QoS automatically. Therefore, I will consider various types of use case priorities as service allocation as shown in Figure.4.9 [90], [98]. These priorities types are based on different QoS requirement by various users and services the same as the following:

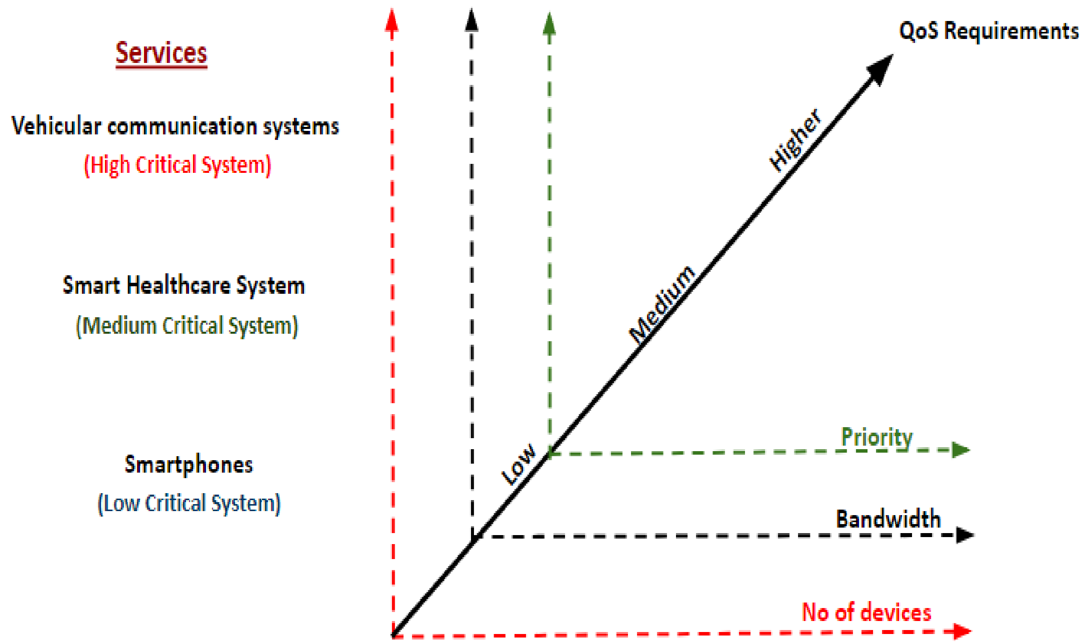


Figure 4. 9: **Services Allocation Use Cases**

- **Smartphones**

Smartphones and tablets are recent technologies that are represented as popular data traffic. Although smartphones are expected to continue as the key personal device and more development in terms of performance and ability, the number of personal devices growth was driven by such devices as wearable or sensors to reach millions in 2020. These devices the content type of mobile video streaming which is cause the total of the flow packets is regularly numerous megabytes or even tens of megabytes, it is many of packets; the transmission way is usually continual transmission; the priority is generally low due to the video requires broad bandwidth, and is likely to be blocked in congestion [90], [95].

- **Smart Healthcare System (SHS)**

The smart healthcare system as sensitive data traffic is a promising model, which has currently achieved extensive attention in research and industry. A sensor Body-Area-Network (BAN) is generally positioned near the patient to gather information about the numerous health parameters, for instance, blood pressure, pulse rate, and temperature. Moreover, the patients are also monitored repeatedly by placing smart device sensors on the body of the patient when they are outside the hospitals or home. For handling critical situations, alarms are triggered to send messages to the related physicians for urgent treatment. In a smart healthcare system scenario, in order to monitor the patients frequently outside the medical centres (e.g., hospitals) the patients are equipped with smart devices that monitor various health parameters[90], [95].

- **Vehicular Communication System (VCS)**

A vehicular communication system allows the conversation of alerting information between a vehicle's infrastructure and the system applications over communication approaches and technologies. In this system, I will consider heavy data traffic. Vehicles connect with other Vehicles (V2V) or communicate with smart traffic monitoring servers, Vehicle to Infrastructure (V2I). This system application includes collision prevention and safety, parking time, Internet connectivity, transportation time, fuel consumption, video monitoring, etc. In the case of emergency, the information from devices positioned to monitor emergencies is transmitted to other networked vehicles within the communication range. To prevent any more accidents, the connection between the vehicles and the servers should be very fast for the detection of emergency messages and delivery of alerting messages. Since the reply time of the warning messages is very small, collision avoidance services request a high level of QoS (i.e., low latency), which can be supported by the 5G cellular networks. According to [87], the alerting messages are small size and must only be sent in serious circumstances for effective use of the communication network bandwidth. Traffic and infrastructure management play an important role in monitoring the issue of traffic congestion[58][63].

## 4.8. RAS Model

A RAS is based on the priority of new slices or users and provides a global optimization of the resources allocated to service slices at RNs or DeNB cell. The steps of My proposed RAS can be applied for admission control of new slices, by simply adjusting the parameters under consideration. When the new UE arrives on the network, by considering the QoE of the users in the same slice, I can derive an acceptance probability of the novel user in the virtual network by considering the constraints in terms of intra-slice priority as well as the QoE of served UEs. In My RAS, new UEs are accepted if the existing resources are enough to guarantee to satisfy at least the demand on the minimum data rate. The set of accepted users is thus offered as input to the resource allocation process [98].

Table 4. 2 : **Resources Allocation Scheme at RNs.**

| **Resources Allocation Scheme at RNs** |
|---|
| **Aim: Resources allocation based on services;** |

```
for s := 1 to S do
  for u := 1 to U do
    for P := 1 to p do
      Calculate ℚus∀Us ∈ 𝕌s;
      find UE xs with the max QoE;
      find UE js with the max QoE;
      while a new UE us ∈ Us enters
        the network do
        Calculate the new QoE value of us:qus;
        Then, find the neighbour QoE value of us:qus;
        if qus − qus >0 then
          if E [qus] < qus then
            Inject UE us;
            check priority order;
            if the priority order is the
               same then
              xs will be replaced by the
              new UE;
            else  js will be replaced by
                  the new UE;
            end
          end
          else
            Do not admit UE us;
          end
        end
        else
          generate accept probability
          p = −△QoE / T ;
          then, the new UE will be
          rejected based on the
          probability p;
        end
      end
    end
  end
end
end
```

## 4.9. Simulation Approach

OPNET simulation is used to assess the performance of the proposed DTA model and RAS scheme. Several protocols are simulated to evaluate the impact of IoT devices' data traffic on 5G mobile networks. The simulated proposed DTA model and RAS data traffic protocols include SMTP, VoIP, RTP and FTP. The main scenarios have been characterised into first scenario aggregation PRBs with RAS, second scenario aggregation PRBs without RAS and third scenario without both aggregation PRBs and RAS. The simulation has assessed the significant impact of IoT devices' data traffic on high priority data traffic. The E2E network performance has been improved by allocated data of several IoT devices, which is determined by simulating these scenarios. Considerable performance improvement is achieved in terms of average throughput per RNs cell, IoT nodes' average upload response time, nodes' average packet E2E delay and nodes' radio resource utilization. [100][90], [98].

### 4.9.1. **Simulation Setup**

I have developed and redesigned the IoT nodes along with LTE-A protocols, to work with the 5G mobile network. The remote server supports SMTP, VoIP, FTP, and RTP applications in the form of smart systems. The remote server and the Access Gateway (aGW) are interconnected with an Ethernet link with an average delay of 5 ms. The aGW node protocols include Internet Protocol (IP) and Ethernet. The aGW and eNB nodes (RN1, RN2, RN3 and DeNB) communicate through IP edge cloud (1, 2, 3 and 4). QoS parameters at the Transport Network (TN) guarantees QoS parameterization and traffic difference. The user mobility in a cell is matched by the mobility model by updating the location of the user at every sampling interval. The user mobility information is stored in the global server (Global-UE-Server). The channel model parameters for the air interface contain path loss, slow fading, and fast fading models [58]. Simulation modelling mostly focuses on the user plane to perform E2E performance evaluations. An inclusive explanation of the LTE-A simulation model and details about the protocol stacks. The different traffic QoS have been set according to the 3GPP standardization. The other simulation parameters are recorded in Table .4.3 [58][66].

Table 4. 3 **: Simulation parameters**

| **Parameters** | **Setting** |
|---|---|
| Simulation length | 1500 sec |
| Cell layout | 1 Enb |
| eNB coverage radius | 350 m |
| Min. Enb-UEs | 35 m |
| Max. terminal power | 23 dBm |
| **5G Parameters** | |
| 5G cell | 8*8 antennas |
| Cloud | Edge could |

| | |
|---|---|
| Capability | Enabled |
| RN Parameters | |
| PRBs for RN | 3 PRBs are allocated to RN by DeNB to evaluate PRB utilization. |
| Type of RN | Fixed |
| RN 1 | Supported by 4 antennas, 10 MHz TDD |
| RN 2 | Supported by 3 antennas, 5 MHz TDD |
| RN 3 | Supported by 2 antennas, 3 MHz TDD |
| TBS capacity | 1608 bits against MCS 16 and PRBs 5. Available service rate TBS-overhead (bits/TTI), 1608 (TBS)-352(overhead) =1256 bits/TTI. |
| Simulated scenarios | Aggregation PRBs with RAS, Aggregation PRBs Without RAS, Without both Aggregation PRBs and RAS. |
| **General Parameters** | |
| Terminal speed | 120 km/h |
| Mobility model | Random Way Point (RWP) |
| Frequency reuse factor | 1 |
| System Bandwidth | 5 MHz |
| Path loss | 128.1+37.6 *log* 10 *(R)*. R in km |
| Slow Fading | Log-normal shadowing, correlation 1, deviation 8 Db |
| Fast Fading | Jakes-like method |
| UE buffer size | ∞ |
| RN PDCP buffer size | 5 ms |
| Power control | Fractional PC, $\alpha$= 0.6, $Po$= -58 dBm |
| Applications | SMTP, VoIP, RTP and FTP. |

### 4.9.2. OPNET 5G Model Description

LTE-A simulation scenario is in the project editor with some of the most important entities of the simulation model. Whereas, the node's model of the DeNB and RNs cells implementation has been modified to 5G mobile network requirements, such as several antennas, edge cloud, small cells and high level of bandwidth as Figures 4.10 and 4.11 depict; more description of these entities is given below:

- **Applications:** Different applications such as VoIP, RTP, FTP, and SMTP are defined and configured in the applications.
- **Profile:** Various traffic models are defined in profiles. Moreover, the other operating parameters such as simulation length, start time, etc. are also defined in profiles to support applications requirement.
- **Mobility:** Mobility models of various users are defined. Moreover, channel conditions such as path loss, fading, etc. are also defined in mobility.
- **Global UE Server:** contains user’s data and transport functionalities.
- **Remote Server:** It is the application server.
- **IP Cloud:** In the form of edge clouds routes user data packets between DeNBs, RNs, and servers.

It also serves as a peer-to-peer connector between the transport network and servers.

- **Ethernet connectors (E1, E2, E3, and E4):** are connectors in the linked network.
- **DeNB:** DeNB models the functionalities of DeNB in E-UTRAN.
- **UE:** UEs represents different users in with various applications.

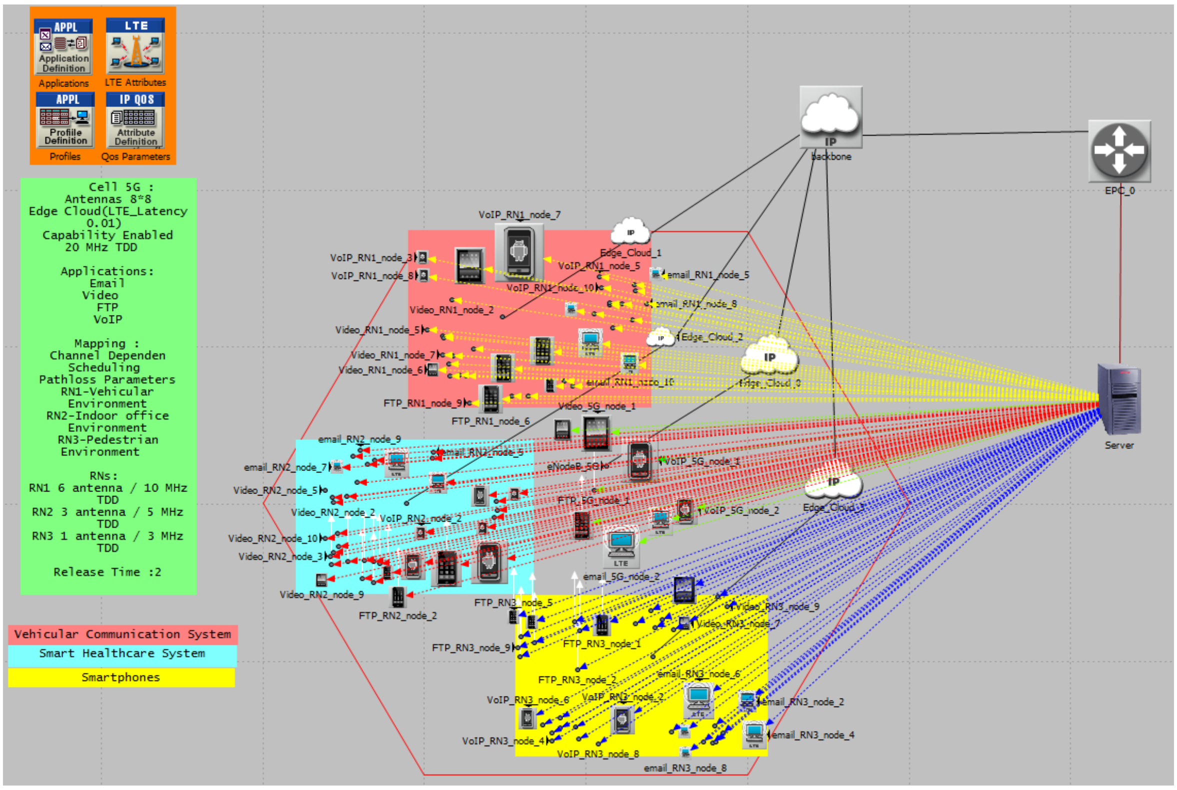


Figure 4. 10: **OPNET 5G Project**

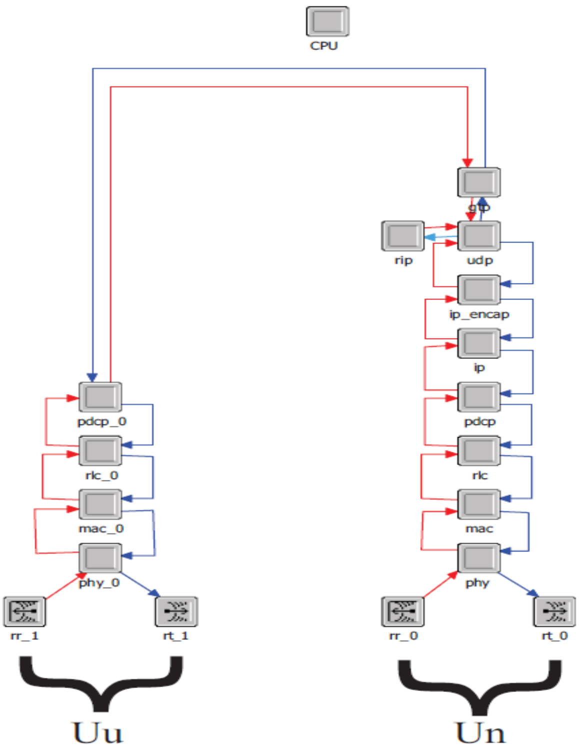


Figure 4. 11: **RN cell model**

### 4.9.3. **QoS of Radio Bearers**

The LTE QoS has gained considerable importance in the designing and planning of the networks. There are possibilities to use the LTE network for various operations. For example, some subscribers use the network services for emergency cases, while others use the services for entertainment purposes. QoS explains how a network serves subscribers due to the enclosed network architecture and protocols. In LTE, the term bearer can be defined as the flow of an IP packet between the UE and P-GW. Each bearer is linked with the QoS parameter. The network provides almost the same services to the packets which are linked to an individual or the same bearer. For establishing a communication path between UE and PDN, UE attempts to generate a bearer by default. Such bearers are called default bearers. The other bearers are named as dedicated bearers which are established to the PDNs. Establishing more than one bearer is possible. This is because one user demands several services and each service demands a specific bearer. For example, if a bearer is established, it is possible to generate more bearers in the presence of an existing bearer.

Moreover, the QoS value of an existing and newly created bearer is possible to vary. The bearer can be classified into Guaranteed Bit Rate (GBT) and Non-Guaranteed Bit Rate (Non-GBR).

- The GBR bearer has a minimum bandwidth which is allocated by the network for various services such as voice and video communication, regardless of whether they are used or not. Due to dedicated system bandwidth, the GBR bearer does not undergo any packet loss due to congestion and is free from latency[90], [98].
- The Non-GBR bearer is not allocated a specified bandwidth by the network. These bearers are used for best-effort services such as web browsing, SMTP, etc. These bearers might undergo packet loss due to congestion.
- Quality Control Identifier (QCI) describes how the network treats the received IP packets. The QCI value is differentiated according to the priority of the bearer, bearer delay budget and bearer packet loss rate. 3GPP has defined several QCI values in LTE which are summarized in Table 4.4.

Table 4. 4 **: LTE QCI values [6]**

| QCI | Resource | Delay | Priority | Error | Service type |
|---|---|---|---|---|---|
| | | | | | Conversational (VoIP) |
| 1 | | 100 ms | 2 | 10-2 | Conversational (Video) |
| 2 | | 150 ms | 4 | 10-3 | Real time gaming |
| 3 | | 50 ms | 3 | 10-3 | Non-conversational voice |
| 4 | GBR | 300 ms | 5 | 10-6 | IMS signalling |
| 5 | Non | 100 ms | 1 | 10-6 | Video Buffered streaming |
| 6 | GBR | 300 ms | 6 | 10-6 | TCP based (SMTP, HTTP, |
| 7 | | 100 ms | 7 | 10-3 | FTP) |
| 8 | | 300 ms | 8 | 10-6 | Voice, video and interactive |
| 9 | | 300 ms | 9 | 10-6 | gaming |
| | | | | | video buffering streaming |

### 4.9.4. **Radio Resourcing Aggregation and Allocation Models**

Packet scheduling is the distribution of radio resources between the radio bearers in RN and DeNB cells by the DTA and RAS Models. In 3GPP LTE standards, this task is performed by the PDCP and MAC schedulers in the RNs and DeNB cells. The allocation of the downlink and uplink PRBs by the RN cell to the UEs depends upon the data present in the buffers of the RNs cell and the UEs respectively. If the data for a UE is present in the buffer of the RN cell, then the RN cell allocates radio resources to the UE for downlink transmission if the RN cell has enough available PRBs and the QoS requirements of the other UEs located in the coverage area of the RNs cell are fulfilled. Similarly, in uplink transmission, the UEs transmit Buffer Status Report (BSR) information to the RN cell for granting PRBs if there is data present in the buffer of the UEs. UE BSR information also identifies the types of traffic in the UE buffer [90], [98]. The RN cell allocates PRBs for downlink and uplink according to the radio bearers QoS requirements of the UE. Time Domain-Maximum Throughput (TD-MT) scheduler provides the radio resources to the UEs close to RN and bears good channel conditions. The users at the cell-edge may not get radio resources. The TD-MT scheduler provides maximum throughput at the cost of fairness [90], [98], which can be expressed simply as in (Equation1):

$$PkT\ D = rk(t) \qquad (1)$$

### 4.9.5. **Simulation Scenarios**

The performance of the proposed model and scheme has been evaluated by three scenarios based on the algorithms, which have been applied along with RNs and DeNB cells. In the first scenario, an aggregation PRBs with RAS, in the second scenario an aggregation PRBs without RAS and the third scenario is without both aggregation PRBs and RAS as shown table 4.5. The data packets from all the active IoT devices, which are positioned in the vicinity of the RN and DeNB cells, are aggregated at the RN before being sent to the DeNB. Though, only the periodic per-hop control model is used in which the large aggregated data packets are served to guarantee full utilization of PRBs. The expiry timer is presented by 5 ms in order to limit the multiplexing delay particularly in the low loaded scenarios between RN and DeNB cells. In this situation, the aggregated packet is served after *T*max at the latest. All the above-stated scenarios are further sub-categorized into numerous sub-scenarios. In the first sub-scenario, VCS devices are placed in the vicinity of the RN1 cell, which is supported by 4 antennas and 10 MHz Time Domain Duplex (TDD) with a low level of priority 5 ms. The second sub-scenarios SHS devices are placed in the vicinity of the RN2 cell, which is supported by 3 antennas and 5 MHz TDD with a medium level of priority 10 ms. The Third sub-scenarios smartphones devices are placed in the vicinity of the RN3 cell, which is supported by 2 antennas and 3 MHz TDD with a low level of priority 15 ms.

Table 4. 5 : **Simulation Scenarios**

| **Scenarios** | **Smart Systems** | **Application Types** |
|---|---|---|
| (1) Aggregation PRBs with RAS | All | SMTP, VoIP,FTP &RTP |
| (2) Aggregation PRBs without RAS | All | SMTP, VoIP,FTP &RTP |
| (3) Without both Aggregation PRBs and RAS | All | SMTP, VoIP,FTP &RTP |

## 4.10. Simulation Results and Analysis

### 4.10.1. IoT Nodes Loading performance

All the scenarios for both the DTA model and the RAS scheme are simulated using OPNET 18.5 Modeler. The results for the average load packet for VoIP, RTP and SMTP nodes are shown in Figures 4.12, 4.13 and 4.14. The results display that the VoIP, RTP and SMTP IoT devices have the diverse load packets variation in all three scenarios even when allocated together with GBR bearers. The cause is the proportional varieties' distinguishing of priority, which is characterized by the RAS in scenario 1 comparing with scenarios 2 or 3. Besides, both VoIP, RTP and SMTP nodes have a significantly improved packet loading as a result of the PRBs DTA aggregation model, which allow the packets to use the maximum of PRBs in RNs cells.

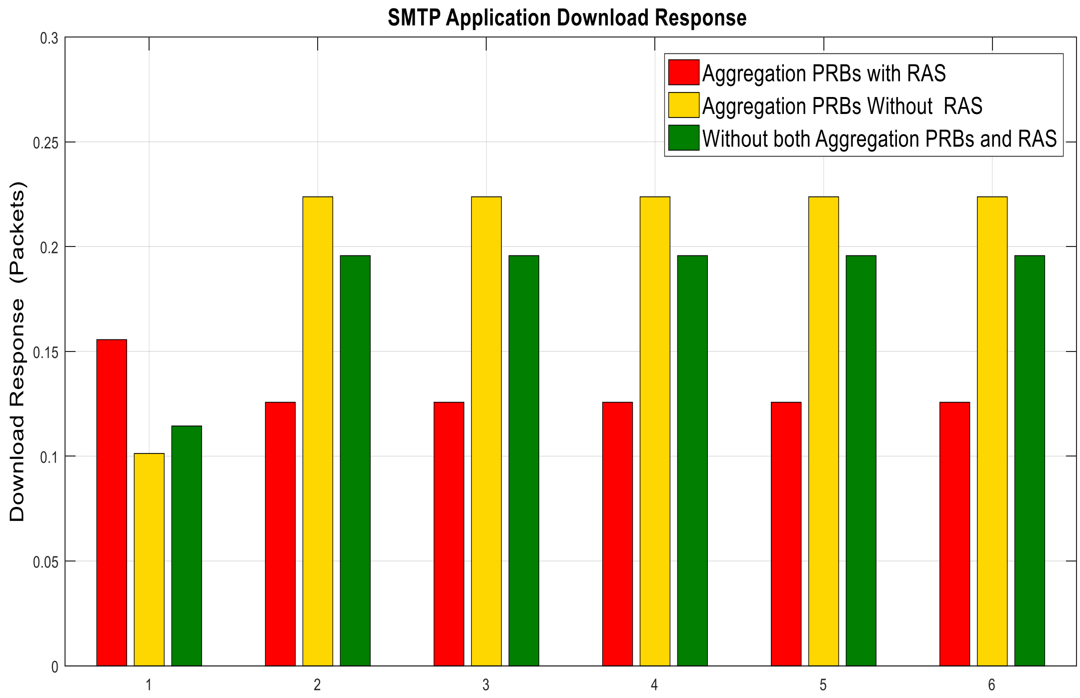


Figure 4. 12 : **SMTP Application Download Response**

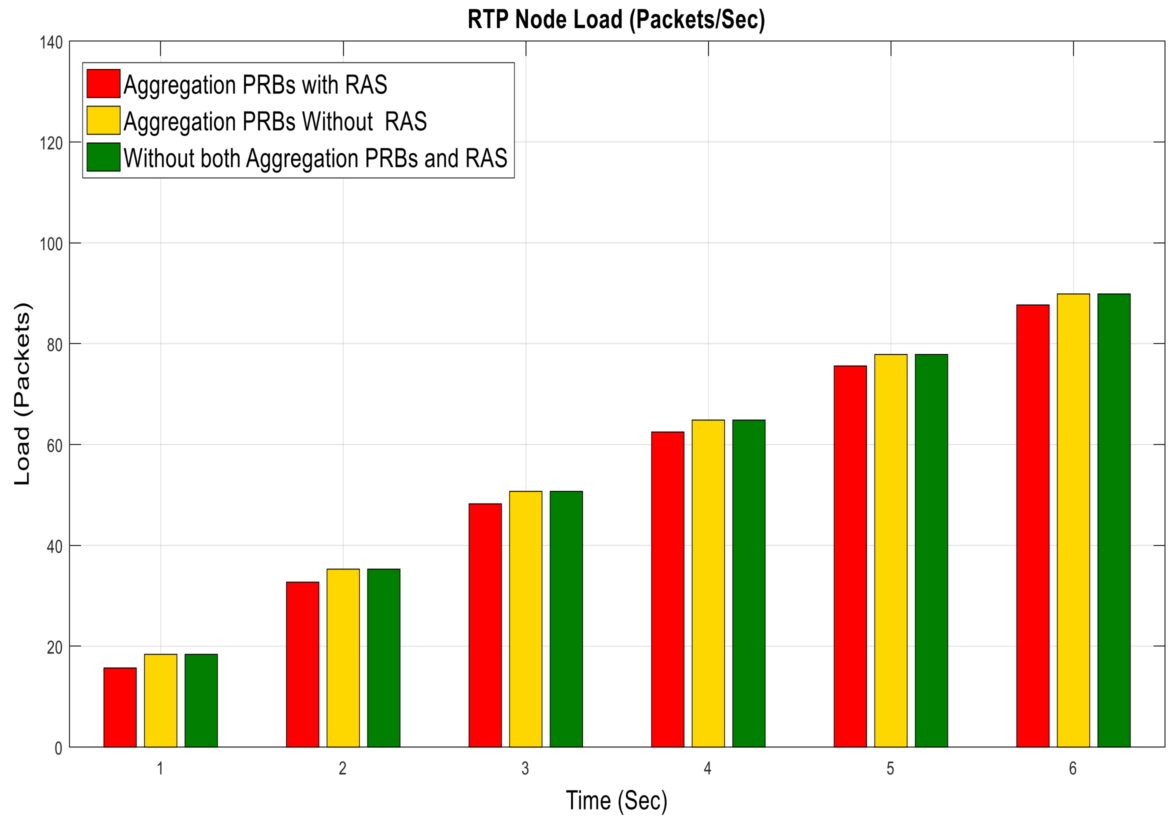


Figure 4. 13: **RTP Node Packets load**

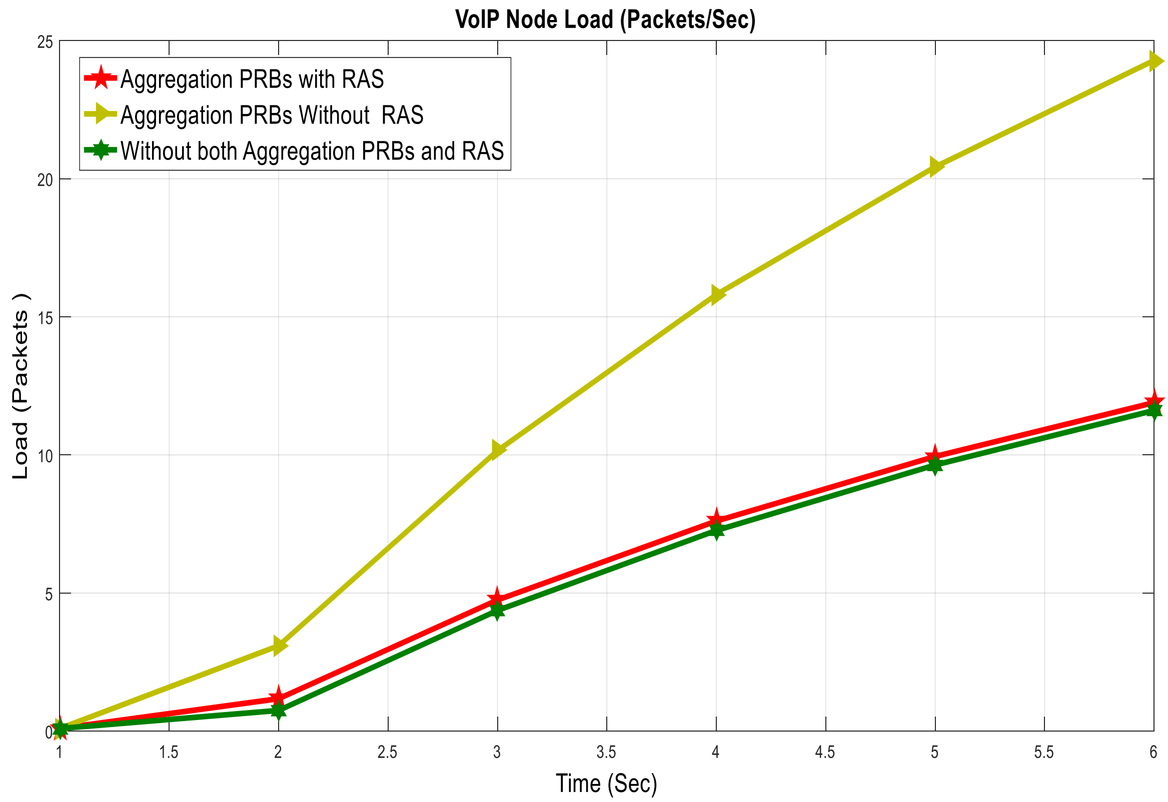


Figure 4. 14: **VoIP Node Packets load**

### 4.10.2. **IoT Nodes E2E Delay performance**

Meanwhile, in terms of the E2E delay performance, the VoIP nodes bearer has a relatively low level of packets E2E delay in scenario 1, in which the accrued data rate tends to get higher priority feature and will permanently be scheduled first based on RAS, however, in the scenario 2 as I have mentioned in the packets loading for VoIP node is really high as result of buffering the incoming packets into RNs cell without considering the priority of the traffic as same as scenario 1 by RAS. In addition, the VoIP node average E2E delay as shown in Figures 4.15, It can be seen that scenario 1 has somewhat better E2E delay compared to scenarios 2 and 3 ; this is because scenario 1 allocates the VoIP nodes to a higher MAC QoS class by allocating the this PRBs to VoIP users in this scenario[90].

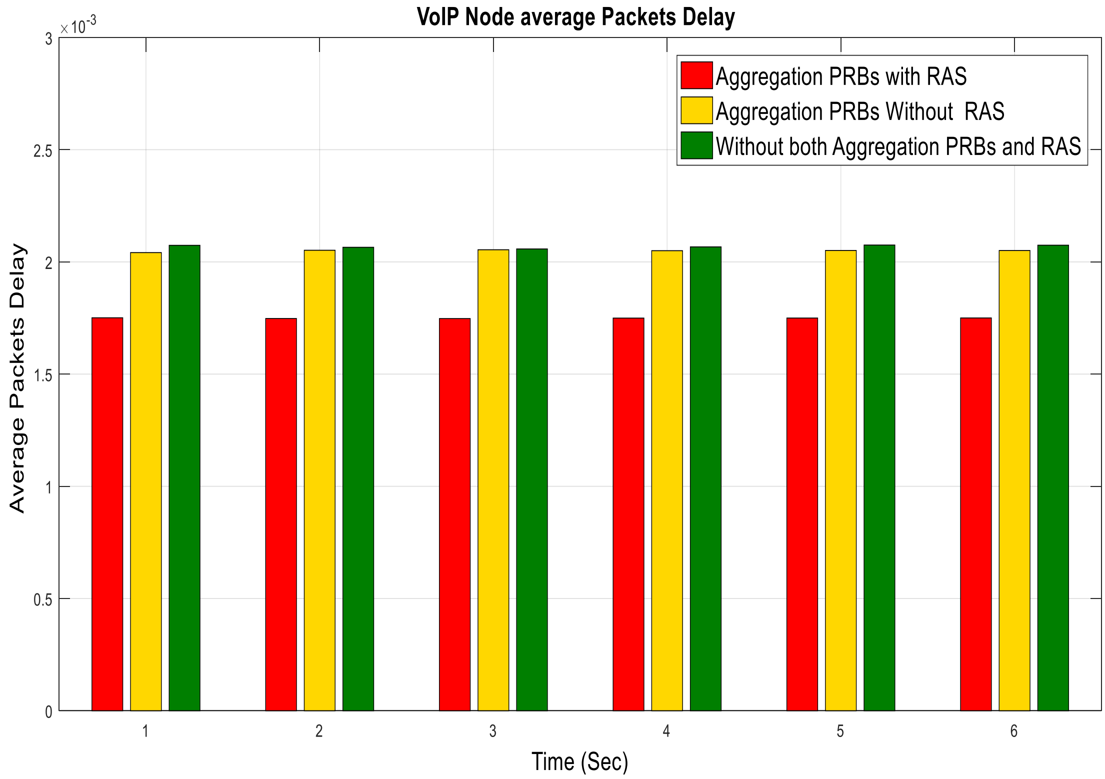


Figure 4. 13: VoIP Node E2E Packets Delay

Moreover, in Figure 4.16 the average Packets E2E delay variation for the RTP nodes. The result describes that the RTP nodes have worse performance in the scenario 2 compared to scenarios 1 and 3 where the RTP bearers are allocated by RAS into the GBR MAC classes by considering the priorities of RTP nodes. In scenarios 3, the RTP node shares the same non-GBR MAC QoS class with SMTP, FTP and VoIP nodes since the accumulated data rate of the RTP. Bearers are expressively high (~ 350 kbps), they do not become served all the time as a result of each node of them use only the assigned PRBs without aggregation incoming packets to RNs cell same as scenarios 1 with DTA model. The performance dropped down of the RTP nodes in scenario 2 as shown obviously with the average E2E delay. The RTP nodes suffer from significantly higher E2E delay performance compared to scenario1 where the RTP nodes have served with specific priority requirement by applying the RAS.

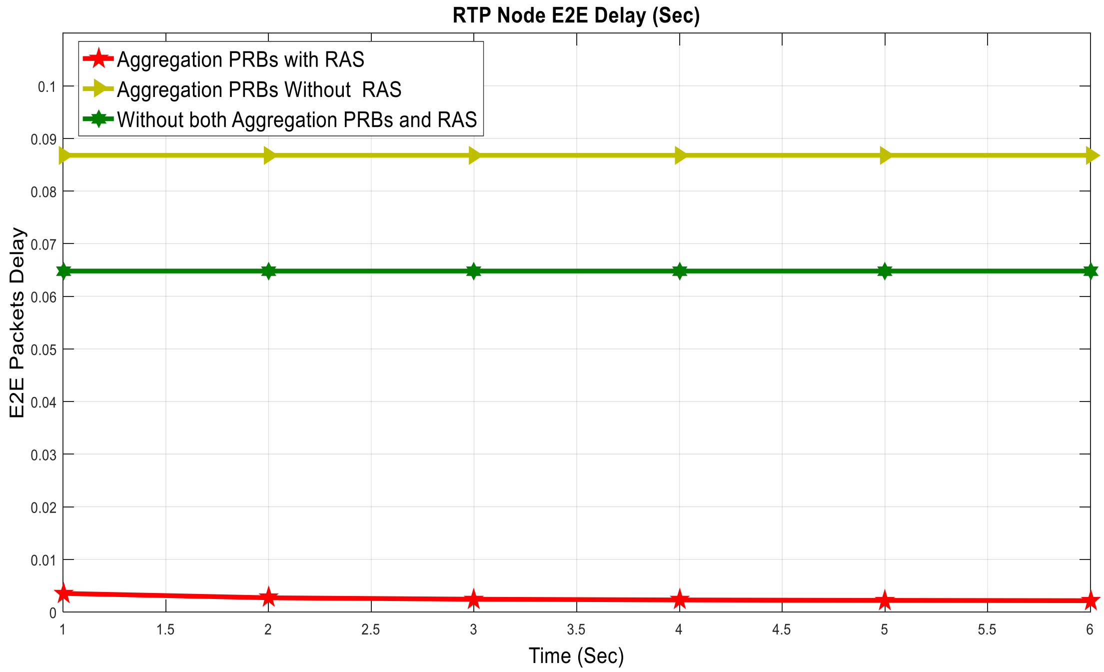


Figure 4. 14 : **RTP Node Average packets Delay**

### 4.10.3. **IoT Nodes Throughput performance**

On the other hand, there is an observation of the VoIP and FTP nodes throughput results as seen in Figures.4.17 and 4.18. It can be observed that the VoIP nodes have much better application performance scenario 2 compared with scenarios 1 and 3, when I have applied the DTA model to aggregate the packets that are going through the RNs cell, which improves the data rate performance, even if I have not allocated on a lower MAC QoS class as I can see in scenario 1. Also, mostly when it is not mixed with the FTP, SMTP and RTP nodes and is allocated to a lower MAC QoS class than FTP nodes. Also, I can observe that with FTP nodes which are slightly improved compared with scenario1, this is because of the QoS weight in scenario1, which is considered the priority in a different level based on applications and critical smart systems requirements compared to scenarios 2 or 3.

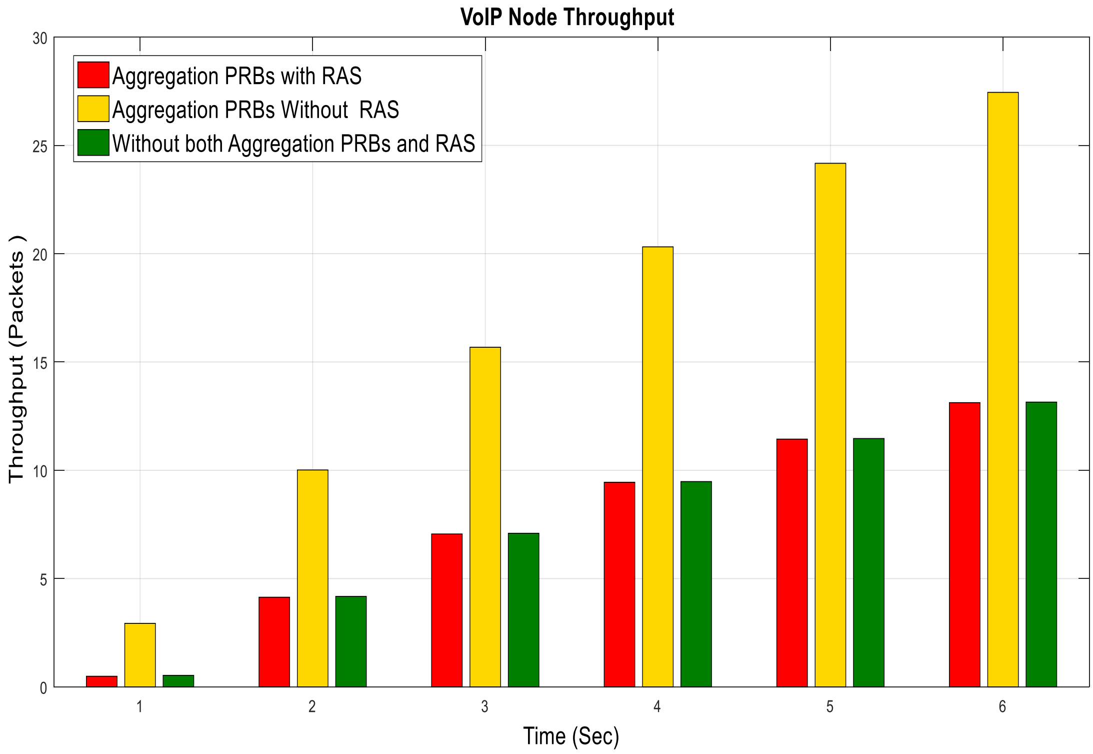


Figure 4. 15**: VoIP Node Throughput**

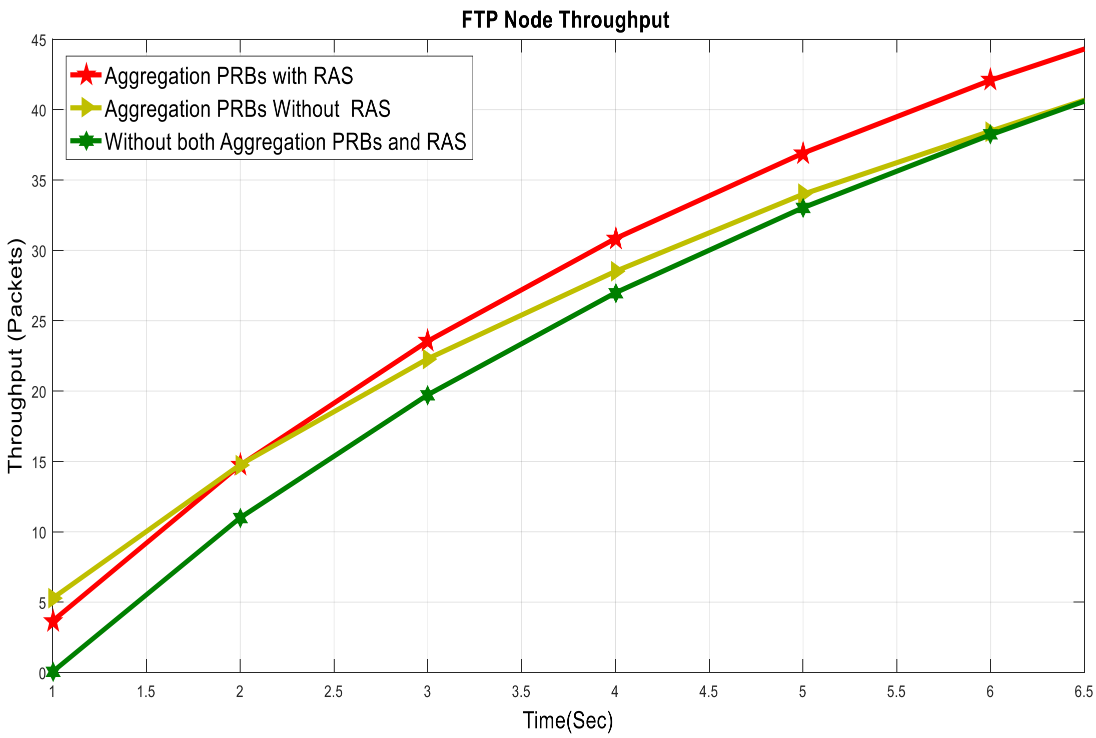


Figure 4. 16: **FTP Node Throughput**

### 4.10.4. **RNs and DeNB cells Performance**

Lastly, I have done comparisons among the three scenarios in terms of assessing the overall performance of the small cells (RNs) and macrocell (DeNB), as I can see in the Figures 4.19, 4.20 and 4.21, the RNs cell packets' loading is decreased in scenario 2 compared with scenarios 1 and 3 as a result of the DTA model, since the incoming packets from different IoT nodes are aggregated in scenario 2 without considering the priorities of these IoT nodes, however, it has been observed in scenario 1 with the RAS where I have allocated the packets with an appropriate priority. However, the results have shown a significant improvement in the E2E delay in the RNs cell in scenario1, which I can describe by enhancing the priorities of the IoT devices by allocating to the lowest MAC layers QoS class as mentioned in table 4.4. However, the results in Figure 4.21 scenario 2 and 3 have a high level of E2E delay, the reasons behind that are that I have not applied the RAS to allocating and customizing the radio resources among IoT devices.

Moreover, the DeNB cell PRBs utilization performance results are seen in Figure.4.19., it has shown that in the scenario 1 it has the maximum level of a radio resources fairness used among the IoT devices of different smart systems, as a result of applied DTA model in the PDCP layer of DeNB cell, in addition, it has been observed that these radio resources are allocated with high level by using a RAS in the MAC layer of DeNB cell.

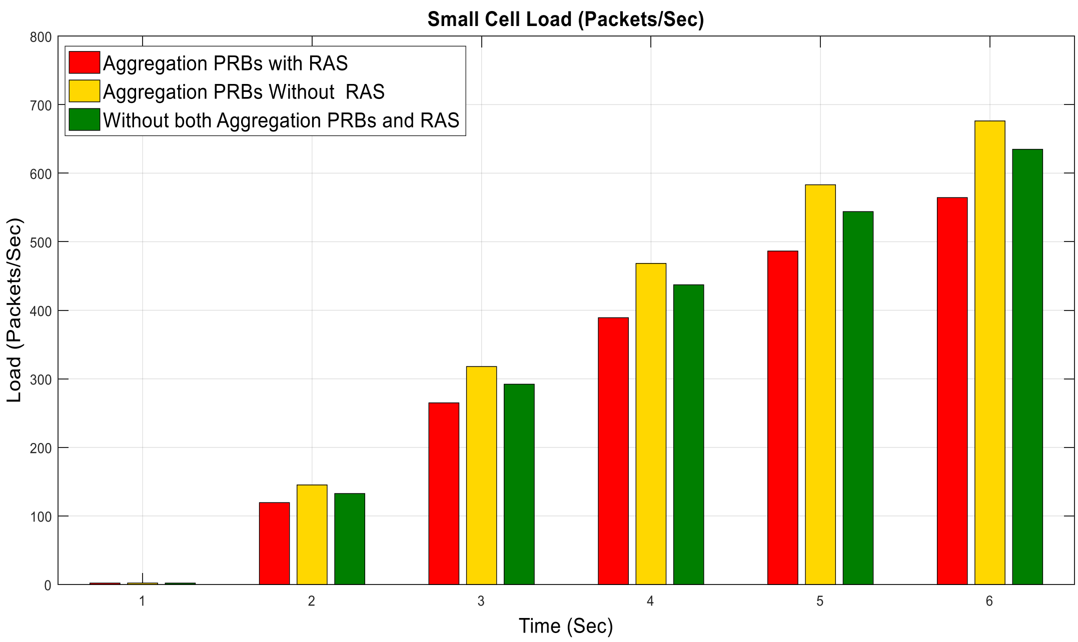


Figure 4. 17 **: RNs Cell Load per Packets**

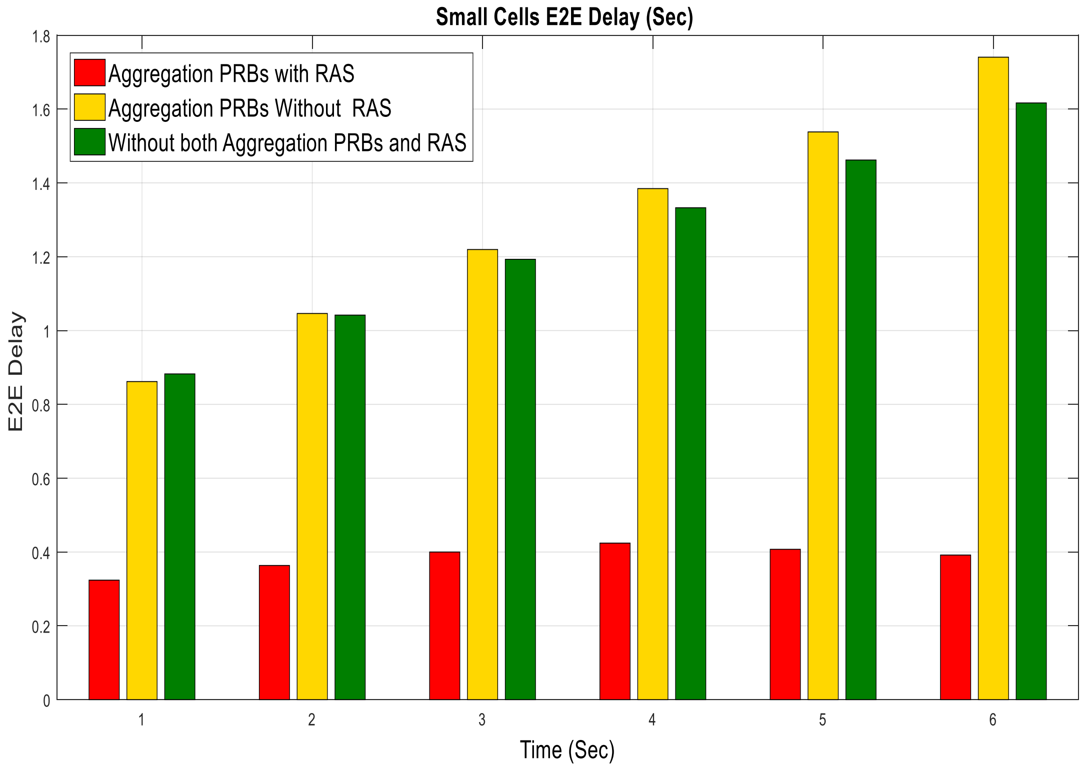


Figure 4. 18 **: RNs Cell E2E Delay**

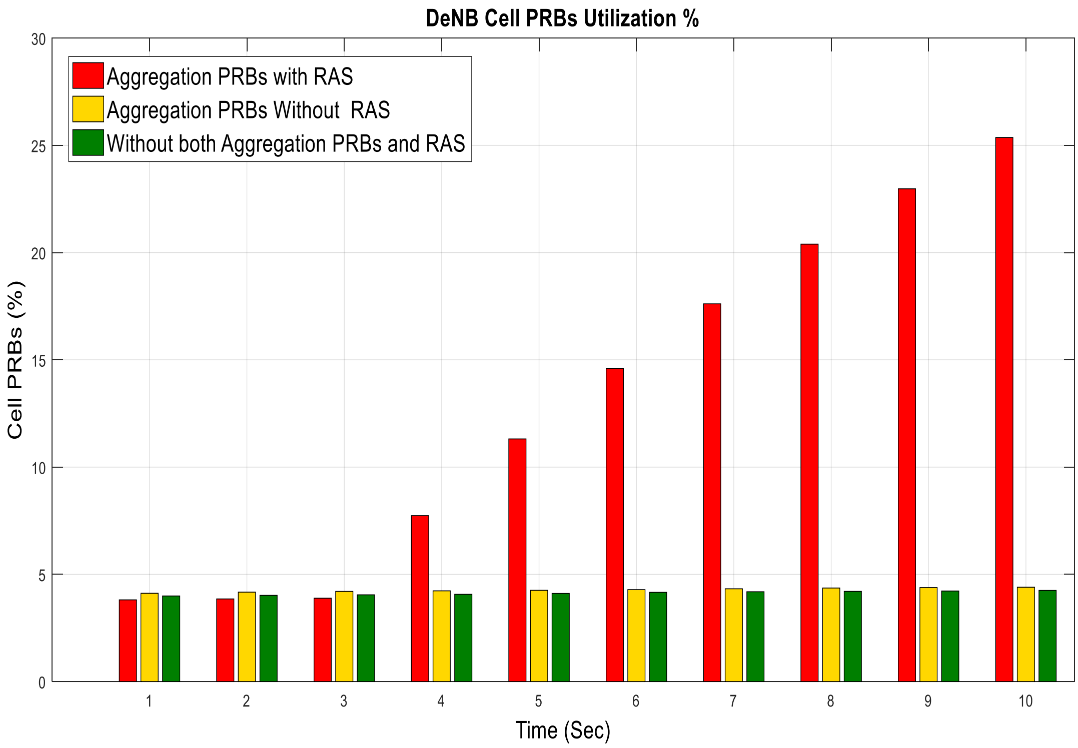


Figure 4. 19: **DeNB Cell PRBs utilization %**

## 4.11. Rational Discussion of the Simulations

This chapter proposed the DTA model and RAS scheme. I have developed the DTA model in fixed RNs cell for uplink in 5G mobile networks. It improves the radio resource utilization for smart systems over 5G mobile networks. It offers a maximum multiplexing gain in the PDCP layer for data packets from several IoT devices along with considering diverse priorities to solve packets E2E delay. Also, in this chapter, I have presented a novel RSA scheme for radio resource allocation in the 5G mobile networks with network slicing. This scheme is a heuristic-based prioritized resource allocation among IoT devices in terms of the scheduling mechanism in the MAC layer of RNs and DeNB cells, which takes into consideration both the inter-and the intra-slice priority and executes the resource allocation accordingly in order to meet the QoS requirements dictated by the service slice. My scheme increases the QoE experienced by mobile UEs as well as allowing better management of network resources. In the implementation, the RNs and DeNB cells are used to aggregate PRBs and allocate these radio resource in different priorities in the form of slicing for IoT devices. This has enhanced the performance in terms of cell throughput and E2E delay of IoT devices data traffic for different scenarios.

# CHAPTER 5: APPLYING SCHEDULING MECHANISMS OVER 5G MOBILE NETWORK

## 5.1. Introduction

In this chapter I have done investigation and comparison among packet scheduling mechanisms, this part of the thesis is to evaluate and compare the performance of three scheduling mechanisms designed for 4G and 5G mobile networks in terms of user's suitability to enhance the priorities order, which will reflect on IoT devices and smart systems QoS requirement such as throughput, load and delay as I have stated in the previous chapter. The results from My performance evaluation allowed us to draw conclusions on the performance of the three-packet scheduler mechanisms and point out the strengths and weakness that are common to schedulers under study. This would help design the scheme of the scheduler at RNs and DeNB cells appropriately.

As a result, choosing an appropriate packet traffic scheduling scheme is a critical element in a case of enhancing quality of service over 5G mobile networks, especially, in the near future mobile networks are expected to face challenges due to the massive number of IoT devices and their data traffic with various QoS requirements, specifically 5G mobile networks is anticipated to offer higher bandwidth with the best QoE among the IoT devices and users [101]. The main impact on QoS in the near future is that there will be billions of IoT devices such as smartphones, SHS and VCS. They require different priorities in terms of using 5G mobile networks, therefore, present a comparison of the packet traffic scheduling mechanisms applied based on data traffic slicing model through RNs and DeNB cells. In this chapter, I present the executed data traffic slicing model along with the most outstanding packet traffic scheduling mechanisms such as WFQ, PQ and FIFO in the 5G mobile network. Consequently, there is strong challenge and inspiration beyond packet scheduling mechanisms and radio resource provision in order to recover smart systems performance by supporting the spectral productivity of the wireless interface and hence upgrade inclusive network capacity [101].

### 5.1.1. Packet Queueing Challenges

Real-time applications, such as video monitoring in critical smart systems are essential to gain the advantage of the QoS adaptation by any network., It is highly critical to offer QoS for priority application such as Vehicular Communication System (VCS) and Smart Healthcare System (SHS) [73]. The highest priority of such packets must be efficiently handled by packet scheduling mechanisms. In contrast, fewer priority packets would also be managed and conveyed fairly. So, it is the serious task of developing scheduling mechanisms to balance among all these standards [74]. Since radio resources are essential assets and rarely obtainable in mobile networks, therefore, efficient utilization is required. The novel communication technologies, such as LTE, LTE-A, and 5G mobile networks, make use of multiple carriers' schemes to offer better data rates and to ensure high QoS demands. As I have mentioned in the previous chapter the PRB unit is allocable in the 5G mobile networks to a single IoT device. Under favourable channel conditions, the PRB can transmit several kilobytes of data. These multiple carriers' schemes are able of transmitting a large amount of data. However, in the case of IoT communication, both narrowband and broadband applications must be considered to enhance QoS requirements. Especially, these applications have different sizes of data traffic, which need QoS specifications such as real-time, accuracy and priority [73]. If one PRB is allocated to a single IoT device for data transmission of just a few bytes, then it might cause severe wastage of radio resources. Also, the different types of data traffic should be considered in the 5G slices approach in terms of network operators providing an appropriate QoS for different types of data traffic. Therefore, utilization of the full radio resources and classifying data traffic should be an ideal solution for data traffic explosion and the fairness of priorities.

### 5.1.2. **Data Traffic Priority Types**

The 5G mobile networks terminal offers exceptional QoS through a diversity of networks. Nowadays, the mobile Internet users choose manually the wireless port of different Internet Service Providers (ISP) without having the opportunity to exploit the QoS history to choose the suitable mobile network linking for a provided service. In the upcoming years, the 5G phone will offer an occasion for QoS analysis and storage of measured data traffic in the mobile network terminal. There are diverse QoS parameters (e.g., bandwidth, delay, jitter, and reliability) which will support in future 5G mobile running in the mobile terminal. System processes will offer the best appropriate wireless connection based on needed QoS automatically[90][101]. Therefore, I will consider dissimilar types of data traffic (e.g., sensitive, popular, and heavy traffic in term of the fairness of priorities) as shown in Figure 5.1 Data Traffic Queuing model [102].

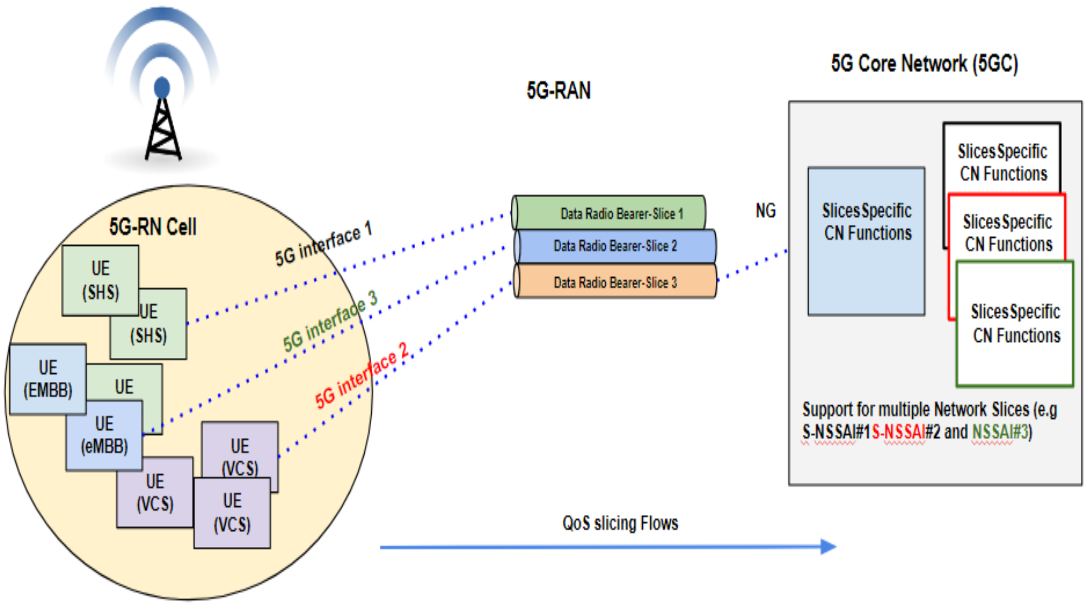


Figure 5. 1**: Data Traffic Queueing Model.**

## 5.2. Comparison of Packet Scheduling Mechanisms

In this part of the chapter, I have proposed a novel comparison among packet scheduling mechanisms based on the 5G mobile network slicing model, that relies on smart systems in the smart city case study, which I stated in the previous chapter. I have considered the concept of network slices to differentiate smart systems such as services priority. Therefore, the delay of the packets which happens at the output buffer of an RNs cell and DeNB cells is called queuing priority. Such a priority is dealt with efficiently and equally by several packet scheduling mechanisms. QoS and Fairness demands are the most significant features offered by any scheduler. Priority Queuing (PQ), First-In-First-Out (FIFO) and Weighted Fair Queuing (WFQ) are a few of the most frequently utilized packet scheduling mechanisms. In this chapter of the thesis is applied packet scheduling mechanisms in order to the classified fairness services priority for QoS evaluation in protocols include FTP, VoIP, SMTP and RTP. These protocols have been applied in OPNET simulation into IoT devices relying on the priority of the data traffic, I have considered some QoS requirements such as delay, load and throughput for diverse packet scheduling mechanisms in the 5G mobile networks.

### 5.2.1. Packet Scheduling Mechanisms

The three core packet scheduling mechanisms used in this part of the chapter are FIFO, PQ, and WFQ. This section shows these three packet scheduling mechanisms in further details.

### 5.2.2. First-in-First-out (FIFO)

The simplest method to schedule a packet in any network is FIFO. Here the first packet in the queue is served first in a specific time slot, regardless of any prioritization, protection or even fairness[103]. Therefore, it is very simple to execute. However, it fails to reach all other scheduling customize excluding complexity. FIFO suffers from the Head of Line (HOL) issue, which means that if the first packet in the queue is blocked for any cause, the rest is blocked even though the link is idle [103].

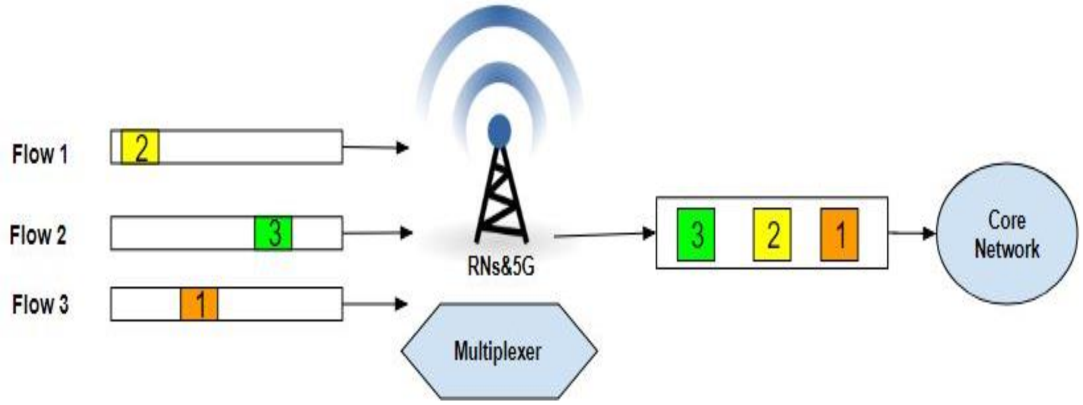


Figure 5. 2**: First-in-First-out Mechanism**

### 5.2.3. Priority Queuing (PQ)

PQ is designed to cope with the issue of FIFO, which does not offer any priority to any data traffic or any class. PQ normally confirms the fastest service of high priority data at each point where it is applied [103]. It provides firm priority to the traffic, which is very essential. The location of each packet in one of four queues known as high, medium, normal or low is achieved relying on the allocated priority of each packet[103]. The possible disadvantage of this scheduling mechanism is that the lower level traffic cannot be assisted for a long time if the high priority is usually there [103].

Consequently, the lower class will be affected by a starving issue that leads to a major discard of the packets.

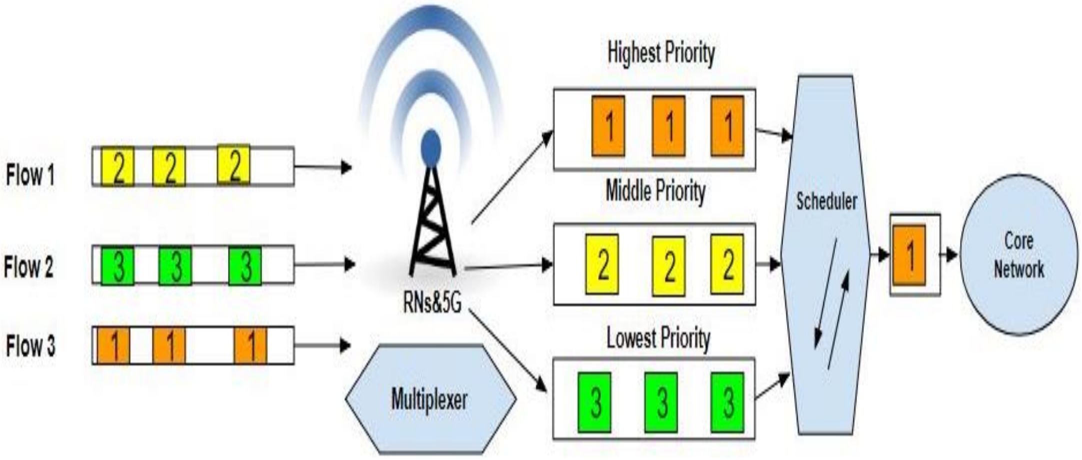


Figure 5. 3**: Priority Queuing Mechanism**

## 5.2.4. **Weighted Fair Queuing (WFQ)**

WFQ is a queuing mechanism relying on data packet flow and the applied realization of Generalized Processor Sharing (GPS) structure that is a theoretical theory and sustains decent fairness [100]. Two things are done seamlessly in WFQ, first reactive traffic is scheduled to the front of the queue for the decrease of response time, and secondly, it shares the residual bandwidth between high-bandwidth flows in a fair mode [100]. WFQ commonly looks into the matter that queues do not starve for bandwidth and all packets must gain the anticipated services. WFQ can notice the superiority bit marked in the IP packet header of each packet and in order with that marking, it classifies the priority levels of packets, with the growth of the superiority value, WFQ assigns more bandwidth to that exact packet to avert congestion [100]

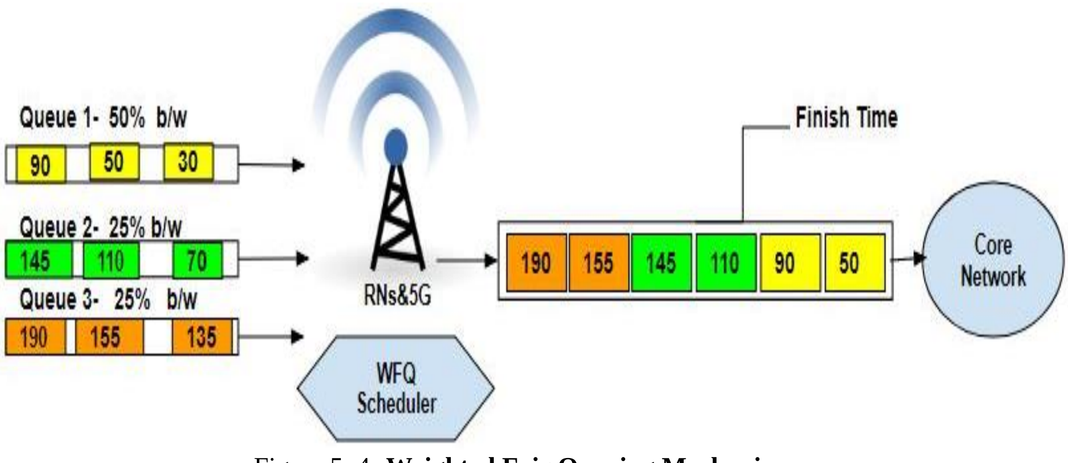


Figure 5. 4**: Weighted Fair Queuing Mechanism**

## 5.3. Comparison Transmission Packet Scheduling Mechanisms

In this part, I presented the advantages and disadvantages of packet scheduling mechanisms in terms of transmitting the packets from the source of IoT devices via uplink and downlink of RNs and DeNB cells until they reach the core network of the 5G mobile networks. Besides, I have done comparisons of the QoS requirement parameters of the IoT devices in the uplink of the 5G mobile networks as I can see in tables 5.1 and 5.2:

Table 5. 1**: Scheduling Mechanisms Transmission Comparison**

| Queuing Mechanisms | Advantages | Disadvantages |
|---|---|---|
| **FIFO** | • Simple and fast (one single queue with a simple scheduling mechanism)<br>• Supported on all platforms<br>• Supported in all switching paths<br>• Supported in all OS versions (above 10.0) | • Unfair allocation of bandwidth among multiple flows<br>• Causes starvation (aggressive flows can monopolize links)<br>• Causes jitter (bursts or packet trains temporarily fill the queue) |
| **PQ** | • Provides low-delay propagation to high-priority packets<br>• Supported on most platforms<br>• Supported in all OS versions (above 10.0) | • Starvation of lower-priority classes when higher-priority<br>• classes are congested<br>• Manual configuration of classification on every hop |
| **WFQ** | • Simple configuration (classification does not have to be configured)<br>• Guarantees throughput to all flows<br>• Drops packets of most aggressive flows<br>• Supported on most platforms | • Multiple flows can end up in one queue<br>• Does not support the configuration of classification<br>• Performance limitations due to complex classification and scheduling mechanisms |

Table 5. 2 **: Scheduling Mechanisms QoS parameters**

| QoS variables | FIFO | PQ | WFQ |
|---|---|---|---|
| Default on interfaces | >2 Mbps | No | <=2 Mbps |
| No of Queues | 1 | 4 | Dynamic |
| Configurable Classes | No | Yes | No |
| Bandwidth Allocation | Automatic | Automatic | Automatic |
| Provides for Minimal Delay | No | Yes | No |
| Modern Implementation | Yes | No | No |

### 5.3.1. Data Traffic Priority Types and PQ Mechanism

In the data traffic slices model, I have considered enhancing QoS by efficient utilization of the 5G radio resources PRBs for IoT and the principal idea of the PQ Mechanism. I have found that the PQ mechanism is the appropriate packet scheduling mechanism in terms of supporting the various queuing priorities, which is based on the priority of the packets, the highest priority is transferred to the output port first and then the packets with lower priority and so on as illustrated in data traffic slices model in Figure 5.6. Therefore, I design My smart systems environment in three levels of priorities high *(slice1)*, medium *(slice2)*, and low *(slice3)*, relying on the data traffic types as follows:

- Vehicular Communication System (VCS) as sensitive data with high priority (5 ms)
- Smart Healthcare System (SHS) as heavy data with medium priority (10 ms)
- smartphones as popular data with low priority (15 ms)

This data traffic will work in a form of slicing over the 5G mobile network in the uplink path between RN and DeNB cells based on user plane interface as shown in Figure 5.5.

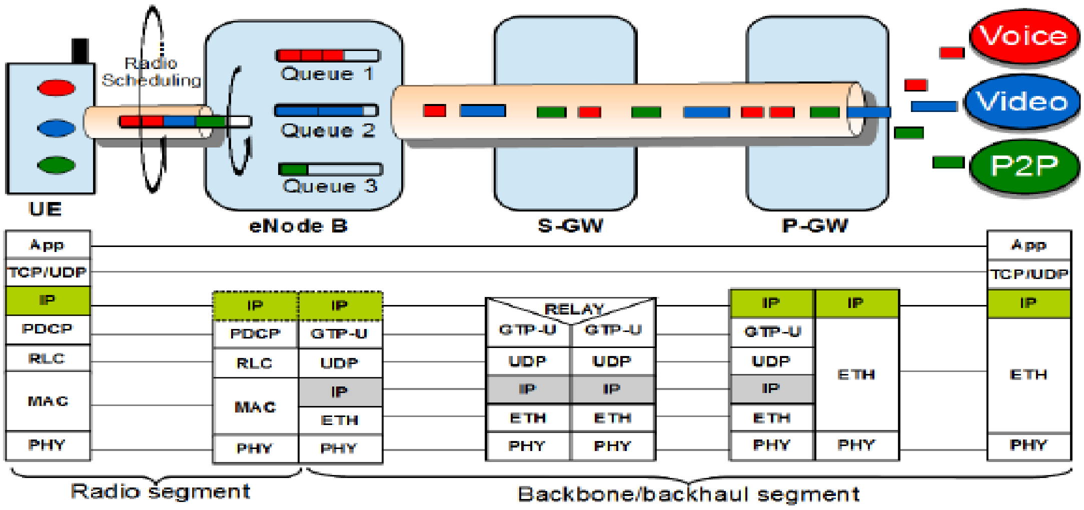


Figure 5. 5**: proposed model for data traffic slices**

## 5.4. Simulation Approach

OPNET simulation is an extensive and powerful simulation software tool with a wide variety of capabilities. It enables the possibility to simulate entire heterogeneous networks with various protocols. The simulated communication network design as shown in Figure 5.7, consists of 3 RNs cells, DeNB cell, 32 users and 4 different application servers.

### 5.4.1. **Simulation setting**

In this part of the chapter, I have considered the LTE-A nodes with LTE-A protocols that I have modified to be suitable along with 5G mobile networks features. The remote server includes SMTP, FTP, VoIP and RTP applications among all smart systems. The remote server and the aGW are connected with an Ethernet link with an average delay of 20 ms. the aGW node protocols contain Internet Protocol (IP) and Ethernet. The aGW and Enb nodes (DeNB1, and RN1, RN2, RN3) connect over IP edge cloud (1, 2, 3 and 4). QoS Parameters at the Transport Network (TN) guarantees QoS parameterization and traffic difference as seen in Figure 5.8 Table 5.3. The user mobility in a cell is coordinated by the mobility model by updating the location of the users at every single interval. The user's mobility data is saved on the Global Server (Global-UE-Server). The channel model parameters for the air interface cover slow fading, fast fading models, and path loss. The simulation demonstrating generally emphasises on the user plane to execute E2E performance assessments[58]. The several traffic QoS have been established withthe 3GPP standardization. Which I have considered the bearers can be categorized addicted to (GBT) and Non-GBR) in the above chapter.

Table 5. 3 : **Simulation parameters**

| Parameters | Setting |
|---|---|
| Simulation length | 1000 sec |
| Cell layout | 1 Enb |
| eNB coverage radius | 350 m |
| Min. Enb-UEs | 35 m |
| Max. terminal power | 23 dBm |
| **5G Parameters** | |
| 5G cell | 8*8 antennas |
| Cloud | Edge could |
| Capability | Enabled |
| **RN Parameters** | |
| PRBs for RN | 3 PRBs are allocated to RN by DeNB to evaluate PRB utilization. |
| Type of RN | Fixed |
| RN 1 | Supported by 4 antennas, 10 MHz TDD |
| RN 2 | Supported by 3 antennas, 5 MHz TDD |
| RN 3 | Supported by 2 antennas, 3 MHz TDD |
| TBS capacity | 1608 bits against MCS 16 and PRBs 5. Available service rate TBS-overhead (bits/TTI), 1608 (TBS)-352(overhead) =1256 bits/TTI. |
| Simulated scenarios | FIFO Model, PQ Model, WFQ Model |
| **General Parameters** | |
| Terminal speed | 120 km/h |
| Mobility model | Random Way Point (RWP) |
| Frequency reuse factor | 1 |
| System Bandwidth | 5 MHz |
| Path loss | 128.1+37.6 *log* 10 *(R)*. R in km |
| Slow Fading | Log-normal shadowing, correlation 1, deviation 8 Db |
| Fast Fading | Jakes-like method |
| UE buffer size | ∞ |
| RN PDCP buffer size | 5 ms |
| Power control | Fractional PC, α= 0.6, *Po*= -58 dBm |
| Applications | SMTP, VoIP, RTP and FTP. |

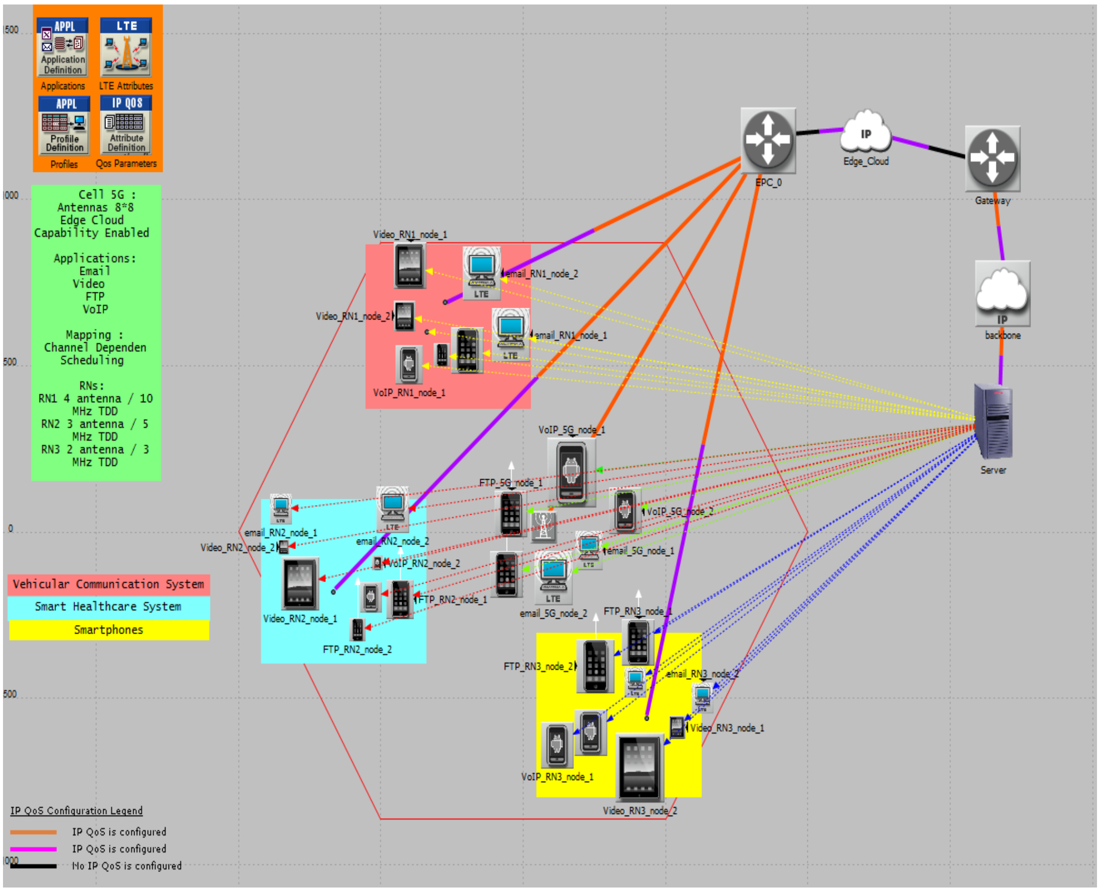


Figure 5. 6 : **OPNETS Scheduling Mechanisms Models**

### 5.4.2. **Scenarios**

Three scenarios are proposed in this paper, the initial scenario used (FIFO). The second scenario used (PQ) and the third scenario proposed the (WFQ) as queuing Mechanisms in 5G heterogeneous networks environment.

Table 5. 4 : **Packet scheduling Mechanism Comparisons**

| Scenarios | Cells | Application Types |
|---|---|---|
| **(1) FIFO Model** | RNs and DeNB | SMTP, VoIP,FTP &RTP |
| **(2) PQ Model** | RNs and DeNB | SMTP, VoIP,FTP &RTP |
| **(3) WFQ Model** | RNs and DeNB | SMTP, VoIP,FTP &RTP |

## 5.5. **Simulation Results and Analysis**

In this section, I review the performance of the proposed comparison of three queuing packet scheduling mechanisms in 5G heterogeneous networks, the scenarios are simulated according to PQ, FIFO and WFQ Models, which are applied on DeNB cell and RNs cells to link the users (smart systems) by the main server. They were considered to compare the performance of both the highest uplink load, throughput and a lower E2E delay between users using different applications among three smart systems, I have designed these smart systems with four applications, SMTP, FTP, VoIP and RTP also, I have stated three levels of packet priorities for each smart system, as I have mentioned above.

### 5.5.1. IoT Nodes performance

The results shown in Figure 5.7 in the case of SMTP node in three smart systems, the FIFO and WFQ models have higher SMTP upload response time compared with the PQ model which has a lower SMTP upload response time for all smart systems, which is a given indication that FIFO and WFQ scheduling mechanisms have supported packets loading from the sources until the final destination of 5G mobile networks. Besides the VoIP node in three smart systems as shown in Figure 5.8 that PQ model has a slightly lower delay in the E2E VoIP packets than FIFO and WFQ models, which is supported the data traffic convey from smart systems to reach main server especially, for the vehicular communication system.

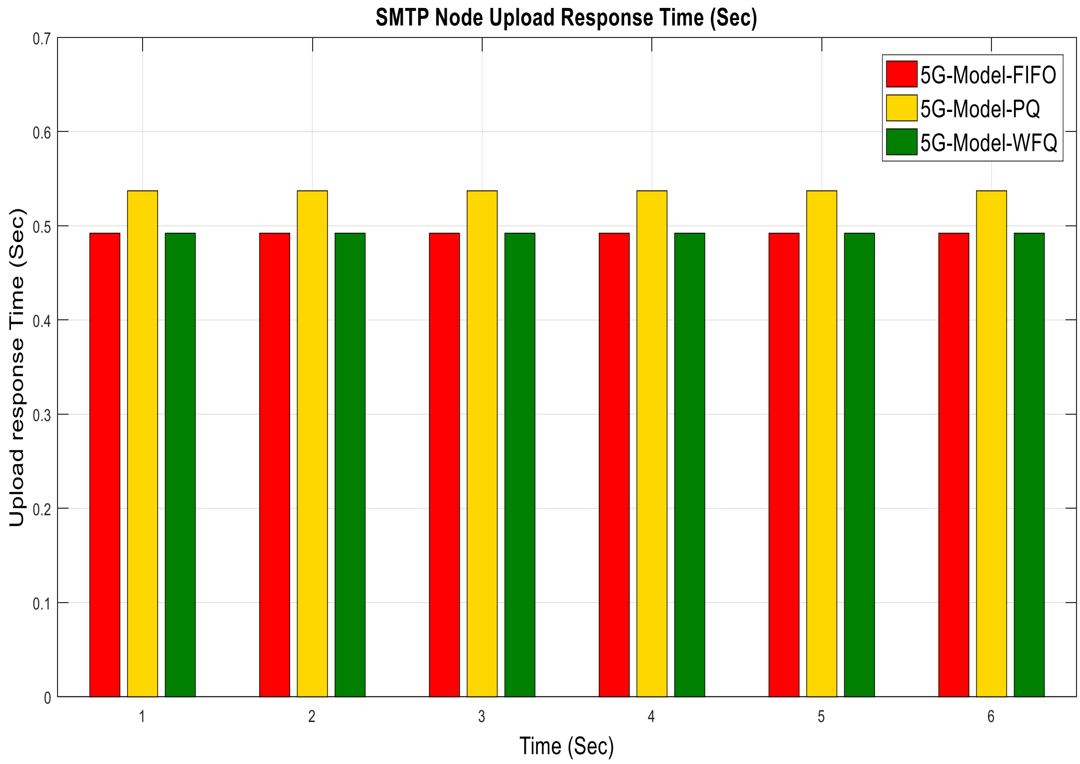


Figure 5. 7: **SMTP Node Upload Response Time**

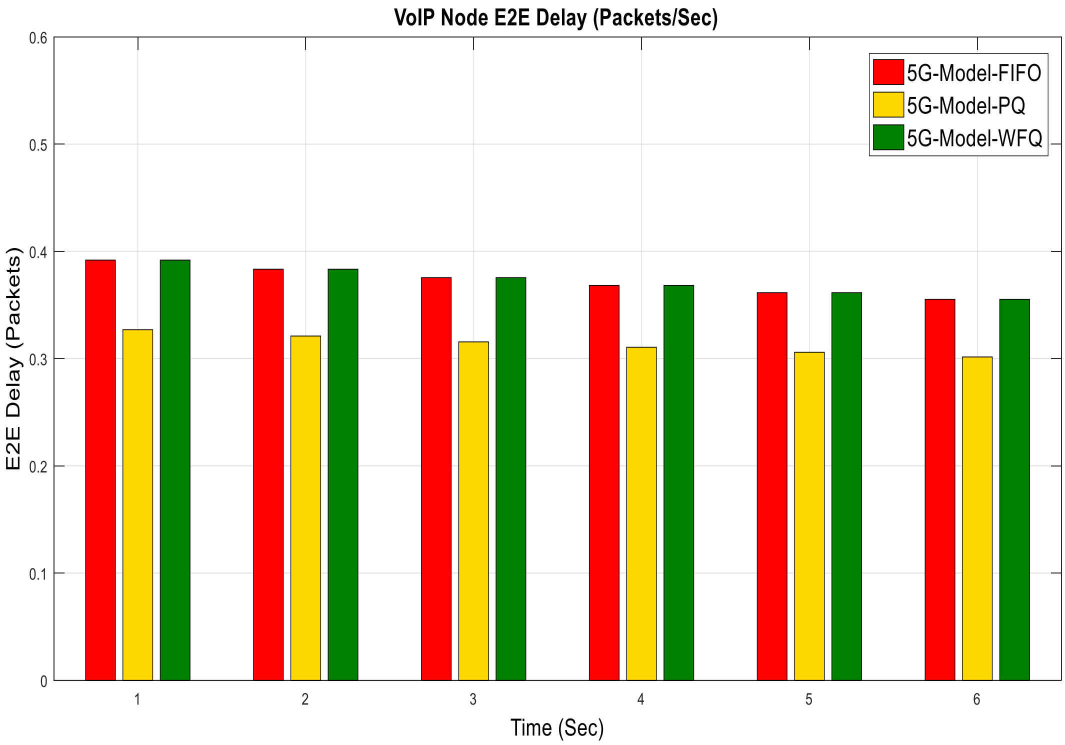


Figure 5. 8 : **VoIP Node E2E Delay**

### 5.5.2. **Cells Performance**

On the other hand, as I have compared the three scenarios on DeNB and RNs cells as shown in Figures 5.9, 5.10, and 5.11, for instance, in Figure.5.10 I have utilized the three queuing scheduling mechanisms over DeNB cell, I found PQ model has slightly lower processing delay in uplink packets than FIFO and WFQ models, which have a similar delay in transferring packets from users, which is offered that PQ model can support packets processing delay lower level and can solve the issue of critical smart systems. Moreover, Figure.5.10 illustrated RNs cell in terms of throughput the PQ model has the higher value than FIFO and WFQ models, which can be used to enhance the heavy data traffic in both vehicular communication systems and smart health care systems. In addition, in Figure 5.11 and 5.12 I have used these three queuing scheduling mechanisms over RNs cells, it has illustrated FIFO and WFQ models have the lower packets loading compared with the PQ model, which has the highest rate of loading that can be useful to support different smart systems loading requirements such as smart healthcare systems. Overall, after analysis and simulation of the three models in the 5G mobile networks, I found that PQ is the best option of packet scheduling mechanisms to enhance both highest uplink E2E delay and throughput, which can support a low level of latency in data traffic demand such as vehicular communication systems. However, the FIFO and WFQ models have been observed to enhance the packets loading, which can be used for heavy data traffic such as smart health care systems among other smart systems.

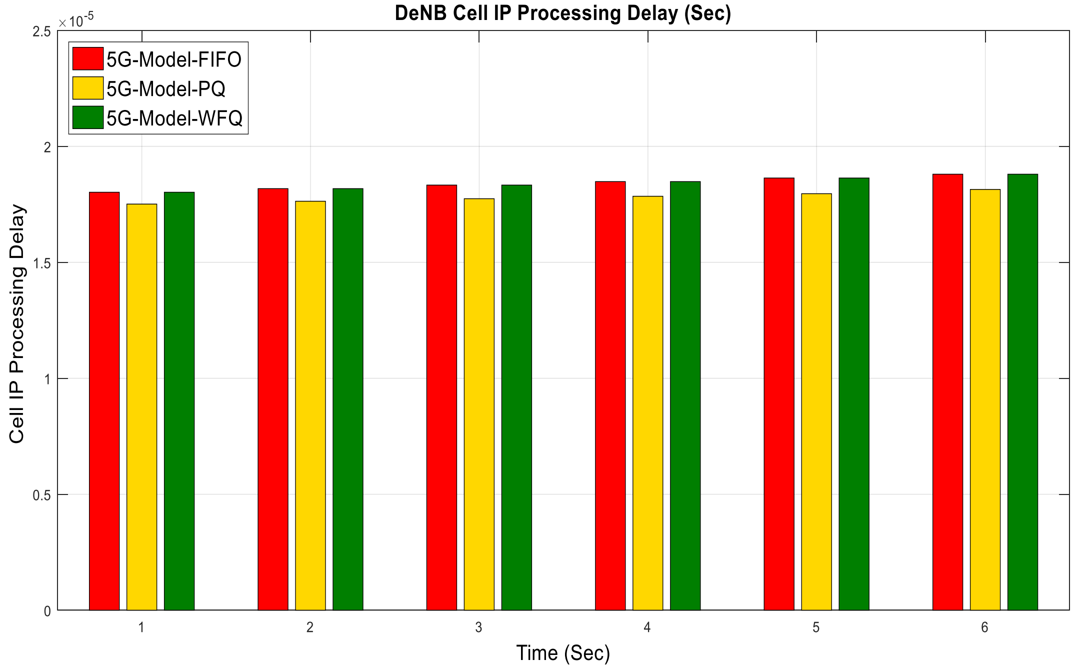


Figure 5. 9: **DeNB Cell IP Processing Delay**

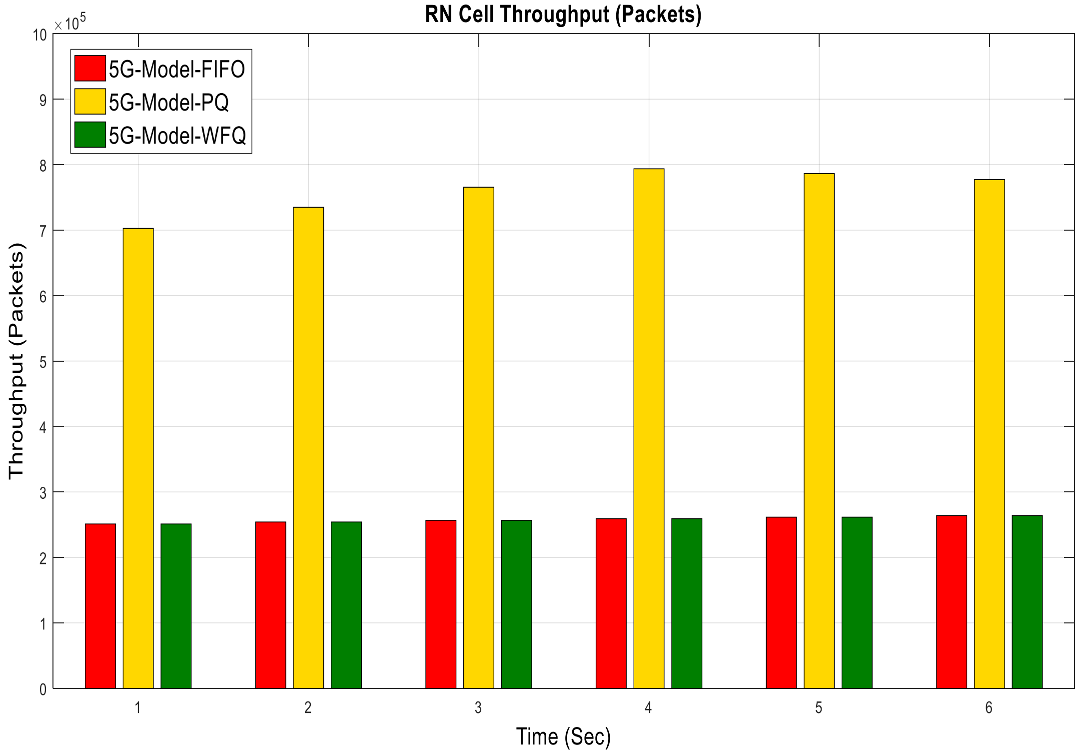


Figure 5. 10 : **RNs Cell Throughput**

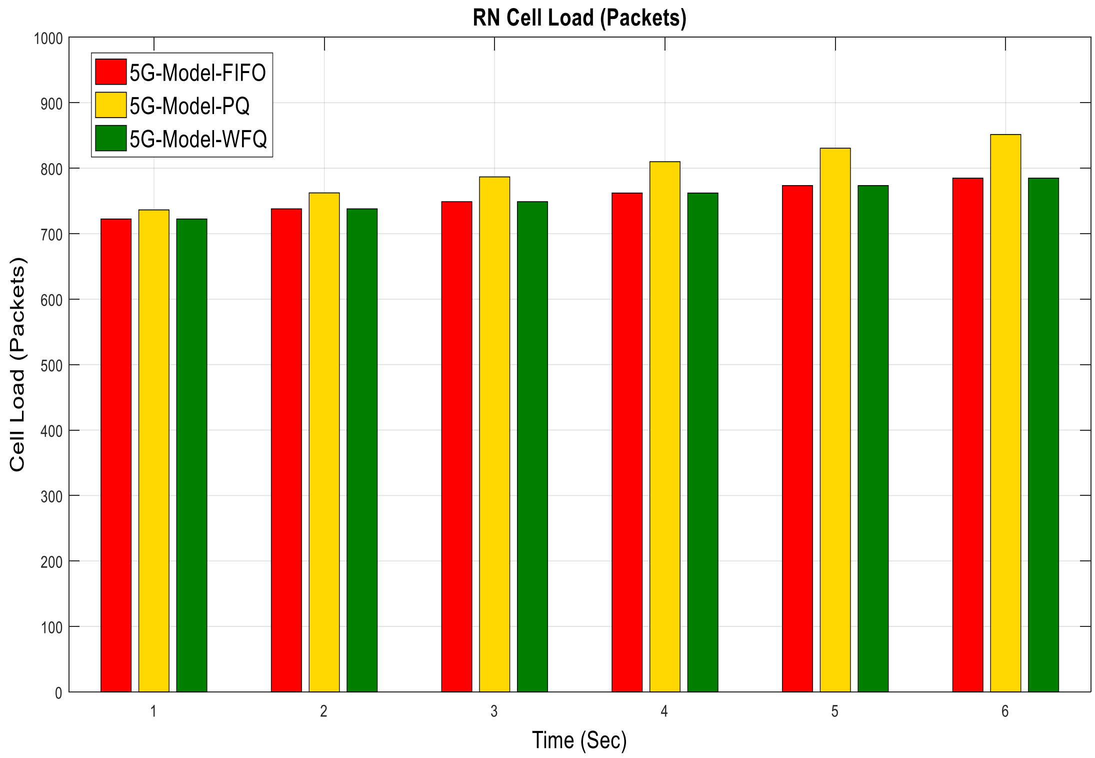

Figure 5. 11: **RN Cells Load**

## 5.6. Rational Discussion of the Simulations

In this chapter, I have proposed the data traffic slicing model over the small cell RNs and the macrocell DeNB within the uplink in 5G mobile networks. It has investigated the radio resource utilization for smart systems over scheduling mechanisms include FIFO, PQ, and WFQ, this has provided an overview of comparative advantages and disadvantages of scheduling mechanisms in terms of E2E packets transmission also QoS parameters. Moreover, I have applied and assessed the FIFO, PQ, and WFQ models over the RNs and DeNB cells of OPNET simulation, I found PQ model as the appropriate scheduling mechanism in the case of supporting various priorities queuing for data traffic, which demand a lower level of latency and throughput such as vehicular communication system. Also, the simulation has shown the FIFO and WFQ models have improved the packet loading particularly in the heavy data traffic such as smart healthcare system. Overall these comparisons have been based on smart systems QoS need in a smart city case study to support and assist the operations of diverse systems (e.g., vehicular communication system, smart healthcare system, and smartphones). It can present opportunities for more research in terms of resolving data traffic explosion and fairness of services area.

# CHAPTER 6: SLICING ALLOCATION MANAGEMENT MODEL

## 6.1. Introduction

In the next few years with the fast growth of IoT with billions of devices, the 5G mobile networks will be required to offer massive connectivity of IoT devices and meet the demand of QoS such as low latency, these QoS requirements are really important for Vehicular Communication System (VCS), where the communication system of connected vehicles: Vehicle-to-Vehicle (V2V), Vehicle-to-Roadside (V2Rs), Vehicle-to-Infrastructure (V2I), and Vehicle-to-everything (V2X) are the basis of intelligent and connected transportation systems where all vehicles and infrastructure systems are interconnected with each other. [104][105]. On the other hand, this massive number of IoT devices and services are connected through various Heterogeneous Networks (HetNets) such as RNs and DeNB cells. Although many of these will only be sending and receiving relatively small amounts of data, they will make new demands, in terms of managing the total accumulation of data and number of physical connections. However, current 3GPP networks cannot connect a higher number of users and establish and maintain a healthy transmission at the same time due to inherited control-plane limitations and scheduling limitations respectively [106][105].

Hence, new scheduling and access control mechanisms are required to reduce the amount of control plane signalling for IoT users. Nonetheless, emergency service providers like Smart Healthcare System (SHS) need real-time data availability and a higher level of reliability in order to deal with critical situations more effectively. Police, fire, and ambulance services must have highly reliable voice links without having issues like call dropping and unresponsive networks. Today, some of these are provided using dedicated networks, but they have limited data capacity of throughput and require high investment just to provide reasonable coverage [107]. Moreover, these systems do not guarantee to fulfill futuristic service requirements in terms of high data rates and real-time interactions. So new network technologies and innovations are required for "ultra-reliable" scenarios, where the ability to connect and operate in situations of severely degraded or complete lack of infrastructure must be assured. This is based on using a device to device direct communication, ad-hoc backhaul and networking, and flexible reconfiguration of networks[106][108].

Furthermore, mobile networks are estimated to face challenges as a result of demanding QoS requirements of futuristic IoT services such as a provision of radio resources to a massive number of IoT devices, prioritization, and inter-device communication [44][108].. The existing mobile systems such as 4G and upcoming 5G might run out of capacity due to the increasing IoT devices traffic, resulting in the performance degradation of regular mobile data traffic [109]. On the other hand, IoT devices demand various types of QoS levels to facilitate different services. For example, SHS devices convey big sized data that are sensitive to delay [110][111]. The RAN is the smallest radio resource unit call PRB, which is allocated to a single device for data transferred in 5G mobile networks. In smart systems, different devices are transmitting numerous sizes of data; where some transmit small size of data traffic, therefore, the capacity of the PRB is not fully used and without radio allocation in the shaping of slicing, as a result of this significant effect of performing the smart systems.

To tackle these issues the contribution of this chapter has presented a Slice Allocation Management (SAM) model for efficient utilization PRBs via HetNets cells such as RNs and DeNB and customize dedicated slices to specific IoT devices and smart systems. Therefore, based on 5G SAM functional architecture radio resources will be efficiently exploited by managing the data of different IoT devices for each slice individually, also the Mobile Edge Computing (MEC) for each slice will reduce the latency and improve the QoS [111].

### 6.1.1. **Heterogeneous IoT Service Challenges**

The main challenge of the IoT devices in the near future with high densification and heterogeneity of mobile networks, particularly, when these IoT devices increase to reach more than twenty billion in 2020. Besides, the world of mobile network the fast growth such as 4G and 5G mobile networks and ever-increasing demand for services with high QoS request, they need to are managing of network resources as becomes a permanently more challenging step that requires to be correctly designed to advance network performance. Nevertheless, network slicing is born as an emerging business to operators by allowing them to sell the customized slices to various tenants at different prices and QoS demands, in this situation, network slicing is getting always-increasing importance as an effective approach to introduce flexibility in the management of network resources. A slice is a gathering of network resources, selected in order to satisfy the demands of such things as QoS of the service(s) to be delivered by the slice. Slicing aims to introduce flexibility and higher utilization of network resources by offering only the network resources necessary to fulfil the requirements of the slices enabled in the smart systems. My approach here is designed to exploit and manage the RAN capacity of RNs and DeNB cells in slicing form, which allows us to customize and reduce the PRBs wastage in each slice in terms of technical requirements such as mobility management and priorities, also QoS requirements such as latency, throughput and loading. In addition, I have designed the MEC Model to reduce the latency of IoT devices based on finding the closed edge cloud to these IoT devices, which also set up the appropriate priority for each IoT node. Therefore, I have used the smart city use case as the smart system, which forms the network slices by differentiating the data traffic smartly in terms of QoS requirements of each slice such as in VCS, SHS and smartphones.

## 6.2. System Model

In this section, I have proposed the SAM Model relying on SAM functional architecture to support specific slicing network, which will concentrate on categorizing and dedicated QoS demand of IoT devices such as Smartphone, VCSs, and SHS.

### 6.2.1. SAM Functional Architecture

From a functional perception, the 5G-RAN consists of two types of network functions (NFs)1, each distributing the full radio access functionality to interact with the UE over the radio interface: gNBs, using the 5G New Radio (NR) interface; and ng-eNBs, using an evolution of the LTE interface. Focusing on 5G NR access, gNBs are linked to the 5G Core network (5GC) utilizing NG interfaces and may be interconnected with other gNBs and ng-eNBs over Xn interfaces. To present modularity and support different deployment options, 3GPP has also standardised the F1 interface that functionally splits a gNB into a gNB Central Unit (gNB-CU) for upper protocol layer processing and a gNB Distributed Unit (gNB-DU) for lower protocol layer processing[112]. A single gNB, regardless of whether it is divided into gNB-CU/ gNB-DU or not, handles the operation of one or more 5G (RN) cells. Each 5G cell, individually identified by a *cell ID*, is allocated with specific radio resources such as

RF carriers which are operated under a common set of control channels (e.g. synchronisation, broadcast). The 5G (RN) cell interface is being designed with high flexible OFDM based waveforms with different numerologies (e.g. different subcarrier spacing and cyclic prefix lengths and adaptable time-frequency frame structures such as selectable slot durations and dynamic assignment of downlink/uplink transmission direction) [112]. Furthermore, the 5G (RN) cell interface allows for UEs served via the same 5G (RN) cell to be instructed to receive or transmit using only a subset of the cell resource grid. Eventually, this flexibility of the 5G cell interface allows UEs with diverse access types such as enhanced Mobile Broadband [eMBB], massive Machine Type Communications [mMTC], and Ultra-Reliable Low Latency Communications [URLLC] to be concurrently multiplexed over the same 5G cell, as shown in Figure 6.1. From a service perception, the overall 5G network 5G RAN and 5GC is considered to support a PDU Connectivity Service, such as a service that provides an exchange of Protocol Data Units (PDUs) such as IPv4, IPv6, Ethernet or Unstructured data packets between a UE and an external data network reachable from the 5GC. The PDU Connectivity Service is realized via the establishment of one or multiple PDU sessions, which are the logical associations created between the 5GC and the UE to handle the data packet exchanges [113]. On this basis, the understanding of network slicing relies on the principle that each PDU session is associated with a specific *Network Slice*.

In fact, a *Network Slice,* which is defined as a logical network that delivers specific network abilities and network characteristics, allows providing a differentiated network behaviour to UEs that are attached to the same 5G network such as to the same 5GC, uniquely identified by a Public Land Mobile Network [PLMN] identity, but have PDU sessions connected with different delivered *Network Slices*. Moreover, to differentiated traffic treatment, *Network Slices* can also be used to serve diverse customers separately as per an agreed Service Level Agreement (SLA). Therefore, a *Network Slice* is officially identified in 3GPPspecifications [113] by a Single Network Slice Selection Assistance Information (S-NSSAI) identifier, which is unique within a PLMN and is comprised of a Slice/Service type (SST), denoting the expected network behaviour, and a Slice Differentiator (SD), differentiating amongst multiple network slices of the same Slice/Service type. Therefore, each PDU session activated between a UE and a 5GC/PLMN network is associated with one and only one S-NSSAI so that the corresponding traffic flows. Denoted as QoS flows in the 5G network are handled according to whatever behaviour is pre-established for the allocated S-NSSAI. The allocation of the serving S-NSSAI is decided between the UE and 5GC based on such as subscription rights and communicated to the 5G-RAN through signalling. Within the 5G-RAN, the pre-established behaviour associated with the S-NSSAI can be then enforced by the proper handling of Data Radio Bearers (DRBs), which are the delivery services provided by the 5G-RAN over the radio interface such as specific scheduling rules and/or radio protocol stack configuration for the corresponding DRBs. Furthermore, to supporting multiple S-NSSAIs of a particular 5GC/PLMN, the 5G-RAN could also serve multiple 5GC/PLMN networks by leveraging the sort of RAN sharing solutions introduced for legacy technologies such as 3GPP Multi-Operator Core Network (MOCN). Hence, gNBs could be linked to several 5GCs and the shared 5G cells could broadcast information about the reachable 5GC/PLMN networks as well as support flexible access control mechanisms per PLMN/S-NSSAI such as 5G Unified Access Control mechanisms.

### 6.2.2. Slicing Allocation Management Model

I propose a novel SAM model by operating and managing the slices by reducing the loss of multiplexing. If the number of slices is low, the wastage for each slice is also low. In the proposed SAM model, virtualized links and nodes are managed in by the Service-Based Slice Allocator (SBSA), including, an Operation Support System (OSS) and Business Support System (BSS) are expanded to implement a SAM Model as shown in Figure 6.2 [28]. The SAM Model starts when the service operator needs to execute a service to the user demand (e.g., smart healthcare system, smartphone and VCS, (1) the service operator sends an inquiry to SAM requesting service release admittance accompanied by a list of the technical service required (e.g., priority management, access area range, etc.) and their QoS requirements (e.g., lower latency limit, high bandwidth limit, service specification protocol). (2) SAM calculates the required resources and chooses one of three options: (A) allocate current slice, (B) allocate current slice after expansion, and (C) create service-dedicated slice on basis of service requirements and current slice's utilization [28].

If option (B) or (C) is selected, (3) SAM instructs the Network Functions Virtualization Orchestrator (NFVO) to expand the current slice or create a new slice. Then, (4) SAM sends the SBSA the service-related information and the forwarding destination of the service. Then, (5) SAM notifies the service operator of the results (which option was selected) and the access point. (6) The service operator embeds the access point into the background application of the service or its IoT devices (7) so that when someone uses the services, the service information is provided to the user. Finally, (8) the service user accesses the SBSA with the information, and (9) the SBSA transmits the service traffic to the assigned slice along with various technical and QoS requirements based on equations (1) and (2) below, which consist of  the priority or latency of the packets, the highest priority and lower latency need is conveyed on the output port first and then the packets with lower priority with high latency and so on as illustrated in the service slices Model as showed Figure 6.4. Therefore, I design My smart systems environment in three technical and QoS requirements high (slice1), medium (slice2), and low (slice3), rely on the data traffic types as follow:

- Vehicular Communication Systems (VCS)as sensitive data traffics
- Smart Healthcare System (SHS) as heavy data traffics
- Smartphone as popular data traffics

This data traffic will work in slicing over the 5G mobile network in the uplink path between RNs and DeNB based on user plane interface. This Model and slice operating method reduce the number of slices to the minimum, thus improving multiplexing gain by accommodating more service traffic in a slice. Compared to current monolithic EPC architecture, since service traffic's time-varying resource demand patterns with bursts of high demand periods and low- utilization services are complementary, the more multiplexing service traffic in a slice, the less total capacity required to satisfy the demand of all the services.

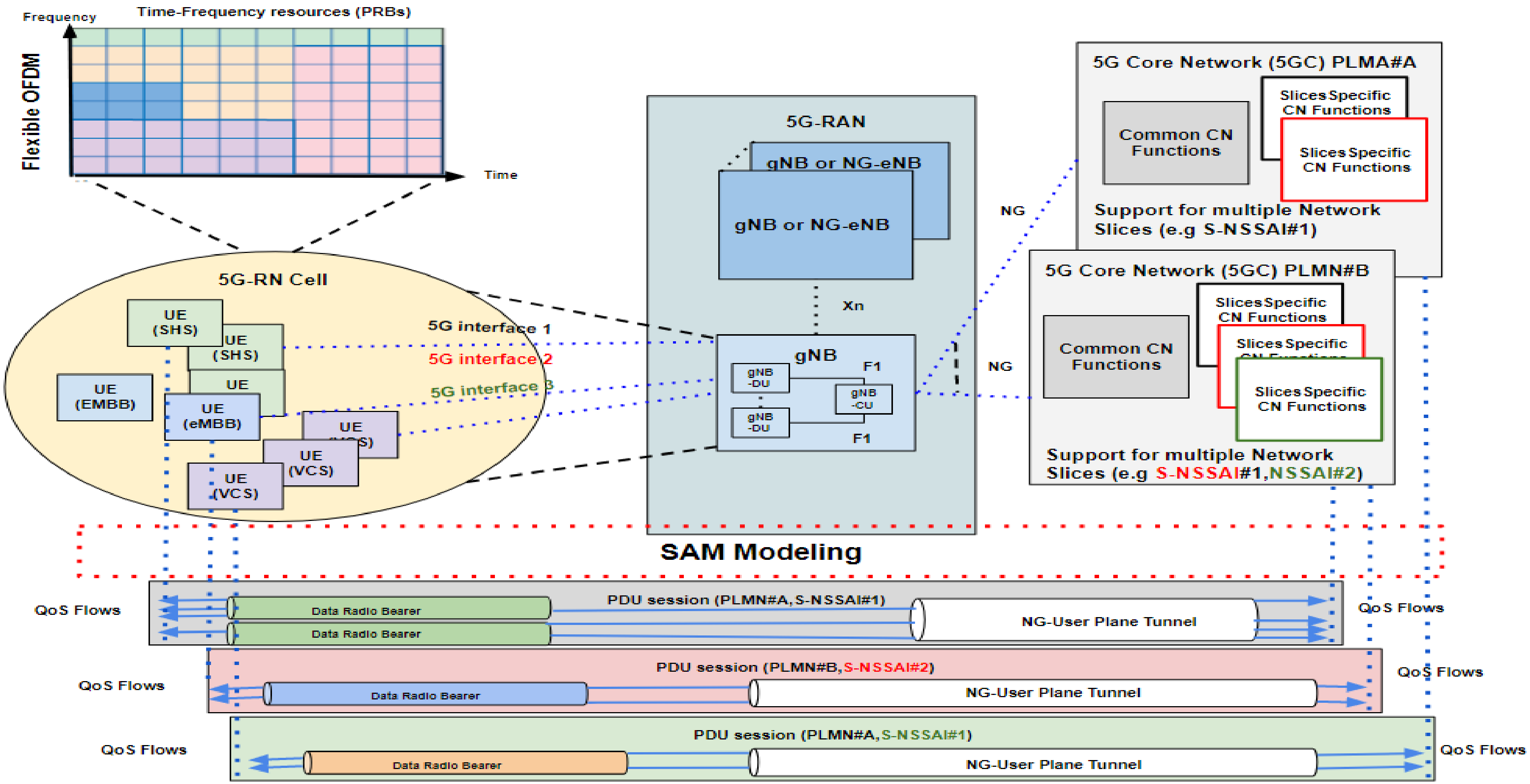


Figure 6. 1: **SAM Functional Architecture**

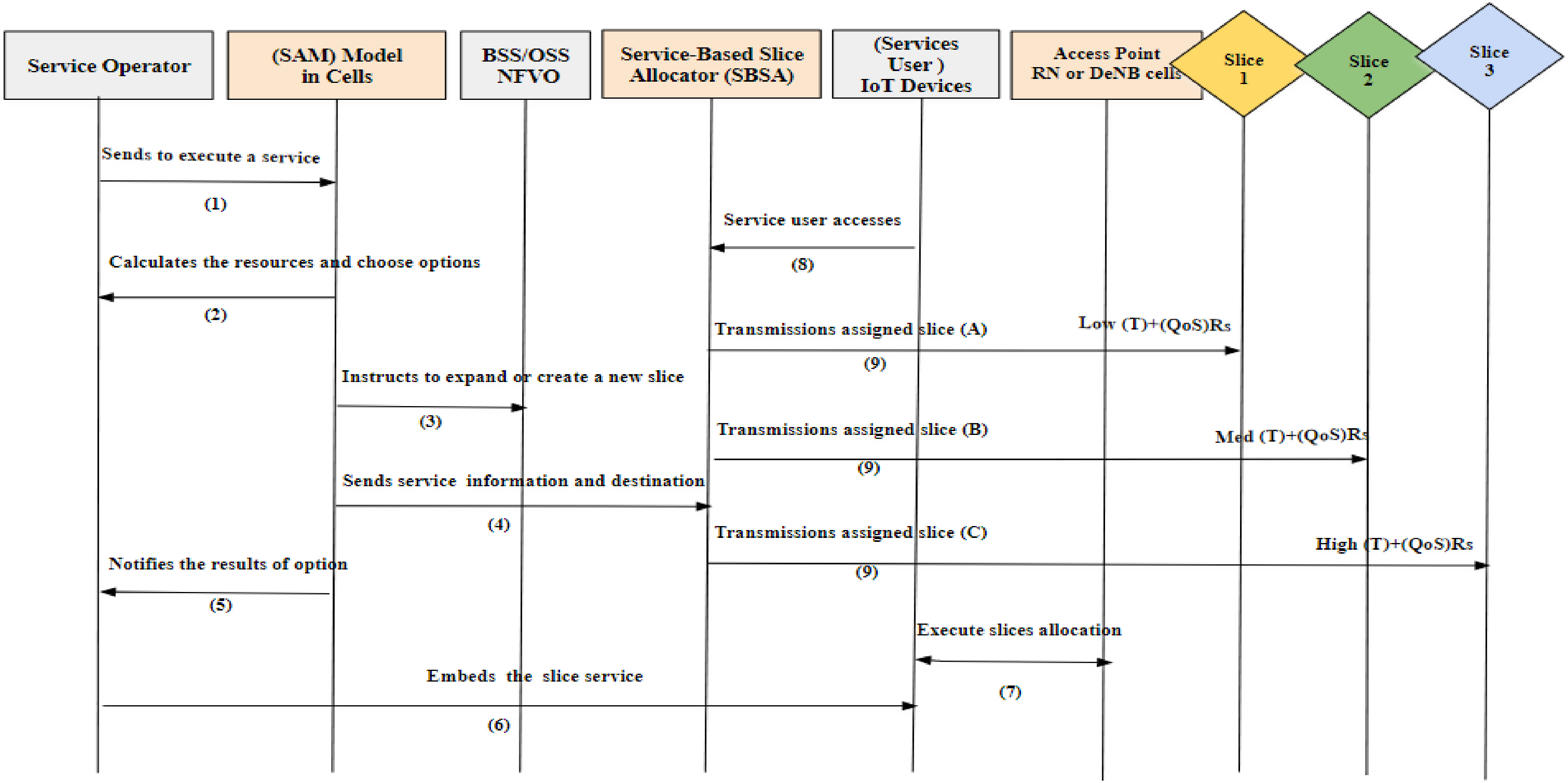


Figure 6. 2 **: Slicing Allocation Management Model**

### 6.2.3. **Vehicular Communication System (VCS)**

In this, as I have taken the VCS as the main use case, which enables the exchange of information between vehicles' infrastructure and VCS applications through communication methods and technologies. In VCS, vehicles communicate with other vehicles (V2V) or communicate with the VCS server's infrastructure (V2I). The VCS applications include collision avoidance and safety, parking time, the Internet connectivity, transportation time, fuel consumption etc. [114]. Several research efforts were made to investigate the support of IoT communication in VCS [115]. To explain the use of IoT in VCS, few VCS applications are discussed in previous chapters. Collision avoidance and onboard security are the most significant applications of the VCS. When driving a vehicle which is not networked with the VCS server, the decisions are made depending upon the information within the Line-of-Sight (LoS) of the vehicle. According to [115], the main purpose of using wireless IoT communication is to deal with such LoS limitations in order to avoid an accident. In the case of emergency, the information from devices positioned to monitor emergencies is transmitted to other networked vehicles within the communication range. To avoid any further accidents, the communication between the server and vehicles must be very fast for the detection of emergency messages and delivering of warning messages. Since the response time against the warning messages is very small, so collision avoidance services demand high QoS services and low latency. According to[114], the warning messages are small in size and should only be sent in critical situations for efficient utilization of the communication network bandwidth. Traffic and infrastructure management plays a significant role in controlling the problem of road congestion. All over the world, every day the drivers face the problem of road congestion that not only increases fuel consumption which leads to more emissions and causes an increase in pollution but also causes high tension for the drivers [114]. A better-managed infrastructure improves productivity and reduces the factors of costs and pollution in society. VCS tackles the problem by providing a bidirectional IoT communication. Such applications do not demand high data rates as few parameters like time, speed and vehicle identification are required. As a result of the low latency which will be delivered by the fundamental access network, 5G network emergency services which depend on massive IoT and device-to-device connections will be categorised by higher throughput, QoE, higher QoS and low buffer demands for the IoT devices [116] [117].

Table 6. 1: **Use cases performance metrics** [95]

| Services | Traffic types | No of Devices | Loading | Priority | Latency | Mobility | Throughput |
|---|---|---|---|---|---|---|---|
| **VCS** | Sensitive | Thousands | High | 5ms | Low | Very High | Very High |
| **SHS** | Heavy | Thousands | Very High | 10ms | Low | Medium | High |
| **Smartphones** | Popular | Billions | Medium | 15ms | High | High | Medium |

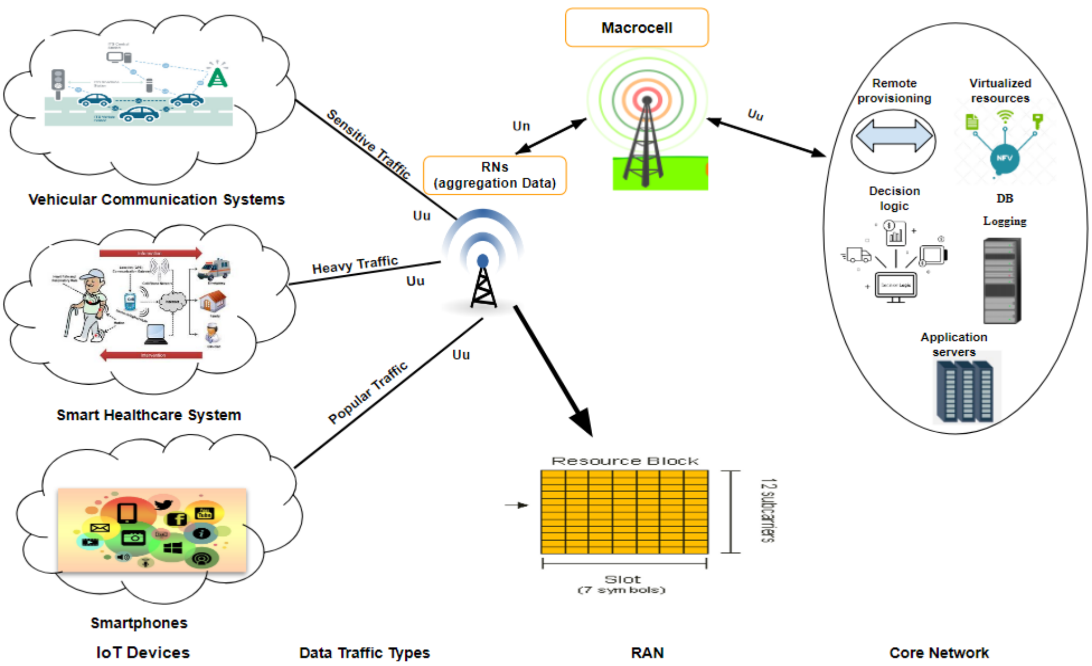


Figure 6. 3 **: SAM model Environment**

### 6.2.4. **Service Slices Model**

I consider $t_n$ and $q_n$for each slice $n$ (n = 1, 2, 3.) where $t_n$ represents technical requirements such as mobility management, tunnelling while $q_n$ represents QoS criteria such as maximum bandwidth, minimum latency stated in Table 6.1. The algorithm used by SAM model to decide whether to (A) allocate existing slice, (B) allocate existing slice after expansion, or (C) create service-dedicated slice on basis of service requirements and current slice's utilization is diagrammed in Figure 6.3. When the service operator $S$ requires a slice, which is characterized by $t_s$and$q_s$. It will send $t_s$and $q_s$ to SAM model and slice allocation will be calculated at SAM model according to equation (1) and (2). $d_{t\ (n)}$ and $d_{q\ (n)}$represent the difference between the required ($t_s$, $q_s$ ) and current slice's ($t_n$ ,$q_{\ n}$).

$$d_{t\ (n)}= t_n - t_s \tag{2}$$

$$d_{q\ (n)}=q_{n-}q_s \tag{3}$$

First, it determines whether $S$ can be accommodated in a slice by calculating the parameters $d_t\ (n)$ and $d_q\ (n)$ . For every slice $n$, if one or both parameters are always negative, accommodating S by using a current slice is impossible because no slice meets S's technical requirements or/and QoS criteria. Then, for every slice $n$, SAM model calculates $C_{en}$ , which is the cost of expanding the current slice, $C_{oen}$ , which is the cost of operating the current slice after expansion, $C_C$ , which is the cost of creating a service-dedicated slice, $C_O$ , which is the cost of operating a created service-dedicated slice, and $I$, which is the loss of multiplexing gain. Next, SAM model calculates $d_{ctech}$ , which is the difference between cost of expanding a slice and the cost of creating a slice.

$$d_{ctech\ (n)}= C_{en+}\ C_{oen-}\ (C_{C+}\ \ C_o\ )\text{+I} \tag{4}$$

If some $d_{ctech}\ (n)$ is negative, SAM model decides that the expansion cost is lower than the slice creation cost and expands the slice with the lowest $d_{ctech}\ (n)$. The service example, in this case, would fall within non-emergency communication such as smartphones as shown in Table 6.1. It could be high-resolution RTP streaming, which is feasible only if some priority parameters are expanded. Compared to a physical architecture, the proposed architecture expands virtually, so it is easier and more cost-effective to scale-up or scale-down.

If all $d_{ctech}$ $(n)$ is positive, SAM model decides that the lowest slice expansion cost is higher than the slice creation cost and creates a new slice of the service. The service example, in this case, would fall within emergency communication systems such as SHS and VCS as shown in Table 6.1. It could be remote surgery service, which has service requirements (e.g., throughput, latency, and topology) that are very stringent. With the proposed architecture, the network provisioning cost for a service with a small number of users and tough requirements would be reduced greatly because of the reduced hardware cost and higher multiplexing gain.

If $d_t$ $(n)$ and $d_q$ $(n)$ are both positive, SAM model decides that there is no technical problem with accommodating the service in a current slice. However, for a service that uses much more than enough node functions such as priority management at the slice, creating a slice with minimum functionality would reduce operation cost. Therefore, SAM model decides whether to create a slice from the commercial phase. It calculates the cost of operating the service $C_n$ for each slice with positive $d_t$ $(n)$ and $d_q$ $(n)$. Then, it calculates $d_{Ccomm}$ $(n)$, which is the difference between the slice expansion cost and the slice creation cost.

$$d_{Ccomm}(n) = C_n - (C_c + C_o + I) \quad (5)$$

If $d_{Ccomm}$ $(n)$, is negative, SAM model decides to accommodate the service by using an existing slice. If some slices have a negative $d_{Ccomm}$ $(n)$, SAM model accommodates the service in the slice with the lowest $d_{Ccomm}$ $(n)$. The service example, in this case, would fall within middle emergency communication systems such as VCS as shown in Table 6.1. It could be voice communication or browsing with requirements that fit within the current slice's capacity. If $d_{Ccomm}$ $(n)$ is positive, the cost of operating the current slice's excessive node functions is higher than the slice creation cost. SAM model thus creates a service dedicated slice for the service. The service example, in this case, would fall within non-emergency communication such as smartphones. It could be a smartphones service, which has a massive number of devices and low functional requirements

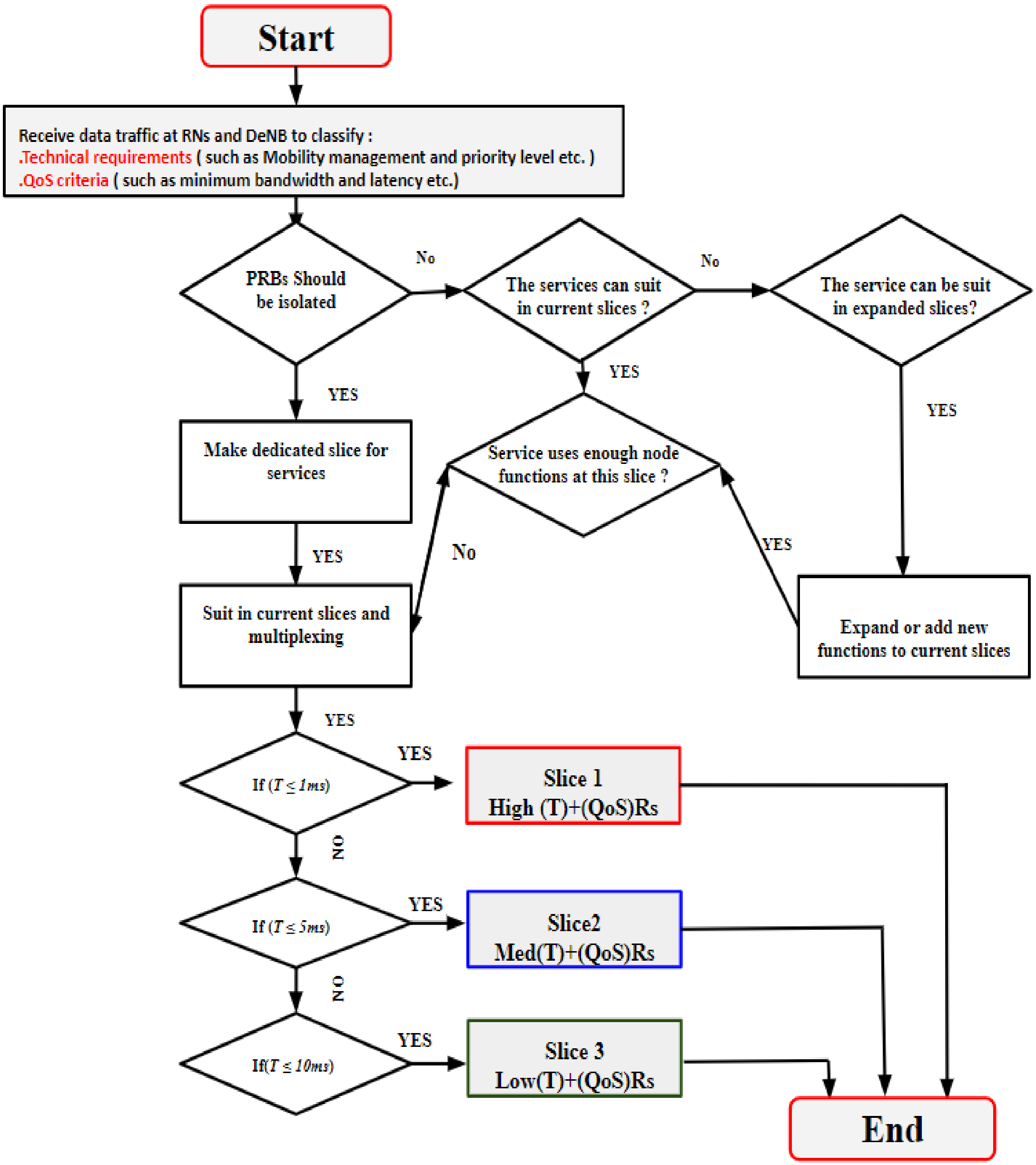


Figure 6. 4 : **SAM Model**

## 6.3. SAM Edge Cloud Model

In this section, I have presented novel Models in which can place Mobile Edge Computing (MEC) direct the IoT devices data traffic as the nearest edge cloud with assigned priorities that should reduce bandwidth consumption by placing caches at the edge of targeted IoT nodes in the smart systems.  Edge or fog cloud facilities will also participate in the management of network latency, routing, and load balancing, becoming the ingress points for the data coming from multiple heterogeneous sources and deciding if it must be analyzed locally or conveyed through a specific path to the cloud for further processing. Such a complex ecosystem is referred to as edge-fog cloud computing. Its scalability, flexibility, and performance characteristics represent a driving force for a new type of applications that involve effective and efficient data management and analytics, such as VCS or SHS applications needing reliable low latency and data traffic management connected  [117] as illustrated in Figure. 6.5.

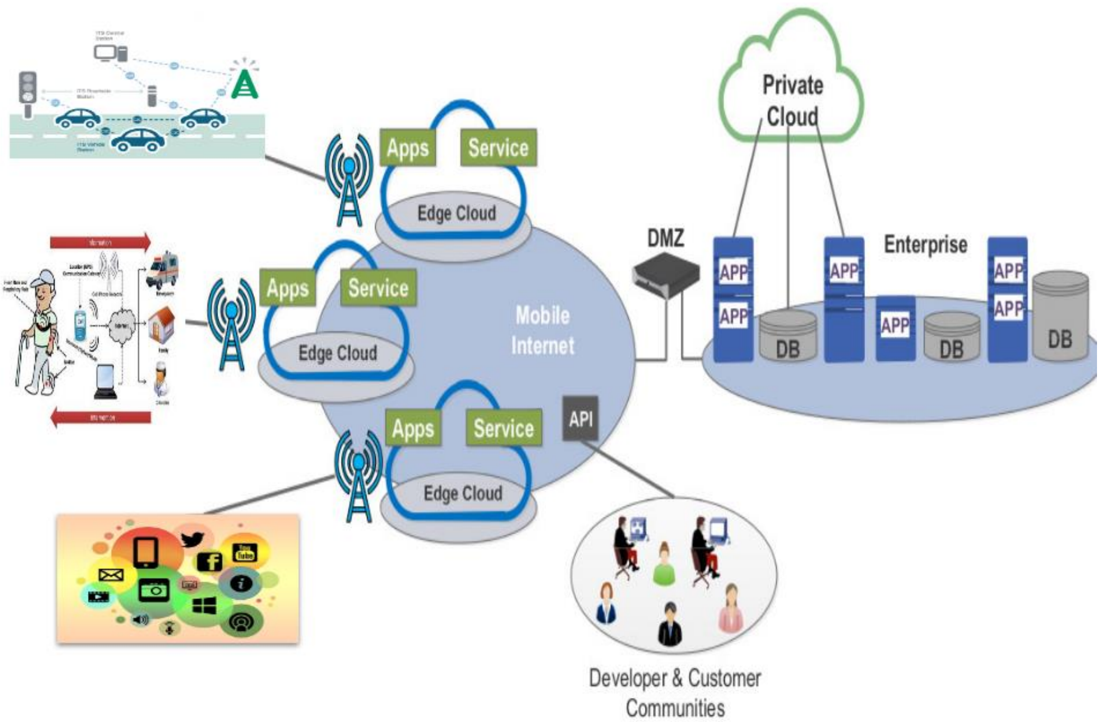


Figure 6. 5: **SAM Edge Cloud Model**

### 6.3.1. **MEC Placement**

Several applications and systems, such as VCS, SHS, RTP streaming, and machine control can benefit from placing MECs close to the user. MECs' running application layer services such as control processes, data pre-processing, and caching may create significant performance improvements when being run at the network edge. The benefits of edge processing include low latency, caching at the edge, and reduction of data transfer to the core and local significance of data (including device-to-device communication).

With the term edge node, I mainly refer to RNs/DeNB cells, but the network may contain several hierarchical levels of edge nodes between the cells and the central data centre, such as regional nodes. The goal of the edge processing is to select those MECs that most benefit from being close to the IoT device and, based on their priority, place them on the available nodes. In the case of all the capacity of the edge nodes being already allocated, the MECs with the lowest need for being at the edge should be moved farther away from the edge. MECs implementing common functions, i.e. functions applicable to most of the users in the slice, should be placed on every edge node that is part of a slice. Thus, when the slice is created, MECs are created at all identified edge nodes. When a new edge node becomes part of a slice, an IoT device belonging to the slice is joining, the MECs are installed on the node.

Each MEC is assigned a priority for being located at the edge. The priority is increased by factors such as the need to minimize the latency or reduce bandwidth consumption by placing caches at the edge. On the other hand, each MEC has a cost in terms of the consumed resources. If MECs are placed at the edge, they have to be replicated to more nodes than in the centralized case. Although the needed capacity of the MEC is lower in the edge case, there is still overhead associated with each new MEC. Therefore, the priority also needs to consider the cost. Thus, if the cost is high in comparison to the benefits offered by placing at the edge, the priority should be low, as I have specified each slice with different priorities.

Besides, in terms of placing IoT devices closed to MEC a configuration is required. The network must be configured to assign the IoT devices to use the appropriate (closest) MECs. The choice of MEC software to install also depends on the location within the network. A MEC at the edge typically does not need to handle the same amount of traffic as the corresponding MEC more centrally located. Therefore, the MEC size, the image to be deployed or the scaling must be considered. This also affects the placement algorithm in that the capacity required by the MEC is a function of the location or the number of users served by the MEC. An MEC flow chart for this MEC service is provided in Figure 6.6. Precisely, radio conditions are monitored through a Radio Network Information Service (RNIS) specified per user (or per network slice). Different actions might be activated: the MEC controller may directly adjust network resources allocated to different network slices to efficiently handle slice SLA violations.

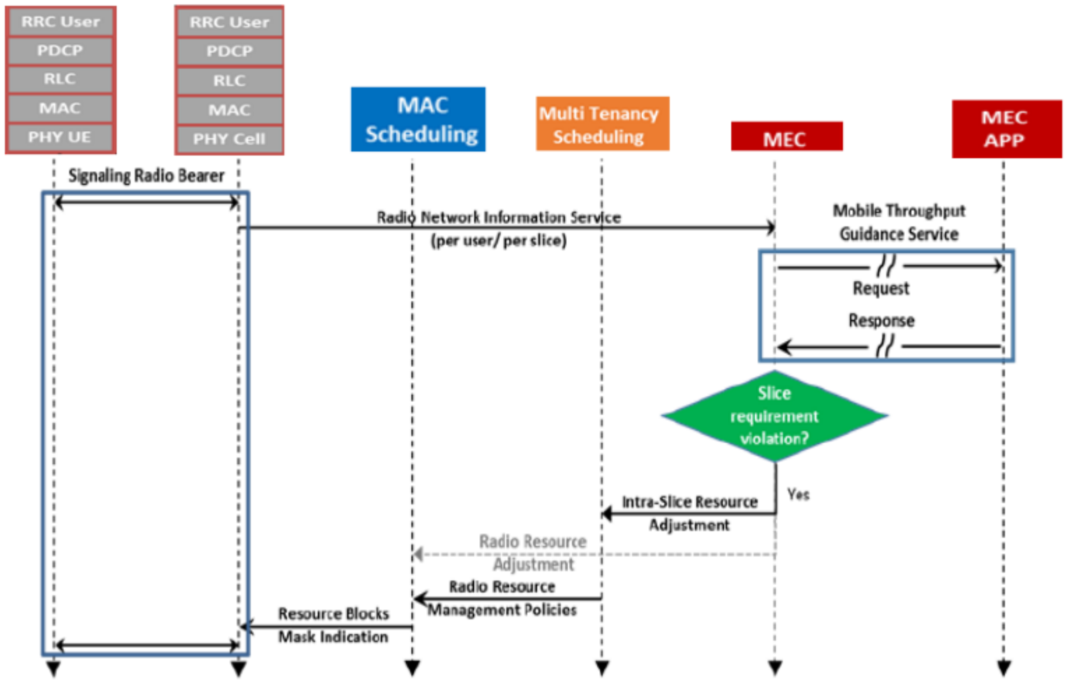


Figure 6. 6: **SAM Flow chart placement close to the edge**

### 6.3.2. MEC Placement algorithms

In the placement MECs algorithm, I start from the cells and users traverse the network toward the core data centre while looking for an available location for the edge MECs. I place the MEC near to the first cell with enough capacity for the user or IoT device. If they encounter the IoT device or user with the same MEC already installed, I stop. This may happen at the edge of the cell already if there is another user part of the slice. Along the path, I may need to move any of the existing MECs having a lower priority far away from the edge placement. These are moved toward the core using the same algorithm. Thus, when I encounter a MEC with a lower preference of the priority user for being at the edge, I remove the lower preference user and place the higher preference on the current MEC, and finally continue the algorithm for placing the lower priority user or IoT device. The algorithm is presented in Table 6.2.

Table 6. 2**: Model for Placing MECs at Edge**

Model **1: aim, Steps**

**Aim: Placing MECs at Edge**

**Steps:**

1. $\boldsymbol{n :=}$ **starting node**
2. $\boldsymbol{m :=}$ **MEC instance to place**
3. **if** $\boldsymbol{T_m \in \{T\omega \mid \forall\omega \in m_n\}}$ **then stop**
4. **if** $\boldsymbol{C_n \geq C_m}$ **then** $\boldsymbol{M_n := M_n \cup \{m\}}$; **stop**
5. $\boldsymbol{\omega := arg_\omega\ min\{P\omega \mid \forall\omega \in Mn\}}$
6. **if** $\boldsymbol{P\omega < Pm\ and\ C_n + C_\omega - C_m > 0}$ **then** $\boldsymbol{M_n := M_n \setminus \{\omega\} \cup \{m\};\ M = \omega}$
7. **if** $\boldsymbol{P_n = \emptyset}$ **then stop with failure**
8. $\boldsymbol{n := P_n}$
9. **Go to 3**

In this algorithm, I denote $C_n$ as the available capacity of node (users) $n$ and $C_m$ as the capacity required by MEC $m$. In practical applications, $C_n$ and $C_m$ are vectors consisting of multiple properties such as CPU power, memory, disk space, etc. each represented as an element. To consider the dependency of the location (e.g. the number of users served) $C_m$ can be replaced by a function $C_m$ (*d*, *u*), where *d* denotes the distance from the edge and *u* denotes the estimated number of users (or bandwidth) served. I further denote $\boldsymbol{M_n}$ as the set of MECs currently allocated to node *n*, $\boldsymbol{T_m}$ as the type or class of the MEC *m*, $\boldsymbol{P_n}$as the parent node of node *n* in the hierarchical network topology, and $\boldsymbol{Pm}$ as the edge priority of MEC *m*. The edge priority indicates how important it is for the MEC to run at the edge.

I run the algorithm once per cell included in the slice. Thus, when the algorithm starts, *n* is the cell to which the user is connected. If later a new cell is added to the slice, the algorithm is run again. The above algorithm, in its simplicity, can only replace a single existing MEC with a lower priority MEC. The algorithm can be extended into a version that moves several of the existing MECs as needed, shown in Table 6.3.

Table 6. 3 **: Model for Moving Lower Priority**

| **Model 2: Aim, Steps** |
|---|
| **Aim: Moving Several Lower Priority MECs at Edge** |
| Steps: |
| 1. $\boldsymbol{n} :=$ **starting node** |
| 2. $\boldsymbol{m} :=$ **MEC instance to place** |
| 3. **if** $\boldsymbol{T_m} \in \{\boldsymbol{T_\omega} \mid \forall_\omega \in \boldsymbol{M_n}\}$ **then stop** |
| 4. $\boldsymbol{M_{low}} := \{\boldsymbol{\omega} \mid \forall_\omega \in \boldsymbol{V_n}, \boldsymbol{P_\omega} < \boldsymbol{P_m}\}$ |
| 5. $\boldsymbol{C_{low}} := \sum_\omega \in \boldsymbol{M_{low}}\ \boldsymbol{C_\omega}$ |
| 6. **if** $\boldsymbol{C_n} + \boldsymbol{C_{low}} - \boldsymbol{C_m} < 0$ **then** $\boldsymbol{n} := \boldsymbol{P_n}$ **; go to 3** |
| 7. $\boldsymbol{M_n} := \boldsymbol{M_n} \setminus \boldsymbol{M_{low}} \cup \{\boldsymbol{m}\}$ |
| 8. **For each** $\boldsymbol{m} \in \boldsymbol{M_{low}}$ **run this algorthem with** $\boldsymbol{n} := \boldsymbol{P_n}$ |

Depending on the policy, moving MECs can consider priorities across network slices. Thus, a higher priority MEC in one slice might cause a lower priority MEC in another slice to be relocated. This requires that priorities are specified in a uniform way. Using such a policy may improve the use of processing resources and the overall QoE across slices, but on the other hand, causes undesired dependency between slices where the slice performance may be impacted by events in other slices. Mobility also affects deciding which MECs are located at the edge. MECs that are affected negatively by mobility need to be farther away from the edge. These are the MECs that are specific to a given user or groups of users. To consider for these cases, a lower priority (possibly even negative) can be assigned to these MECs. In a multi-level hierarchical network, an alternative approach would be to start the algorithm from a higher layer starting node, e.g. $\boldsymbol{P}_{(\boldsymbol{n})}$, for these MECs. For removed MECs or the case of the last user of a slice leaving the cell, a similar process is started in reverse. The aim is to optimize the use of edge nodes in this new situation.

Table 6. 4: **Major Symbols Use**

| Slice Allocation Models Symbols | | | |
|---|---|---|---|
| Symb. | Meaning | Symb. | Meaning |
| $\boldsymbol{n}$ | Each Slice | ***Coen*** | The cost of operating the current slice after expansion |
| $\boldsymbol{t}$ | Technical requirements | ***Cc*** | cost of creating a service-dedicated slice |
| $\boldsymbol{q}$ | QoS criteria | ***Co*** | The cost of operating a newly created service-dedicated slice. |
| $\boldsymbol{S}$ | Service operator | ***dctech*** | The difference between cost of expanding a slice and the cost of creating a slice |
| $\boldsymbol{d}$ | Difference between the required and current slice's | ***l*** | The loss of multiplexing gain |
| **Edge Cloud Allocation Models Symbols** | | | |
| $\boldsymbol{n}$ | Node (user) | $\boldsymbol{d}$ | Distance from the edge |
| $\boldsymbol{m}$ | MEC (edge cloud) | $\boldsymbol{u}$ | Estimated number of users (or bandwidth) served |
| $\boldsymbol{C_n}$ | Capacity of node | $\boldsymbol{T_m}$ | The type or class of the MEC |
| $\boldsymbol{C_m}$ | Capacity required by MEC | $\boldsymbol{P_n}$ | The parent node of node |
| $\boldsymbol{M_n}$ | The set of MECs currently allocated to node | $\boldsymbol{Pm}$ | The edge priority of MEC |
| ***Cen*** | The cost of expanding the current slice | ***dCcomm*** | The difference between the slice expansion cost and the slice creation cost |

## 6.4. Simulation Approach

In this section, I proposed the simulation tool was OPNET version 18.5, I have considered the LTE-A nodes with LTE-A protocols I have modified to be suitable along with 5G features. The remote server includes SMTP, FTP, VoIP and RTP applications among all smart systems. The remote server and the aGW are connected with an Ethernet link with an average delay of 20 ms. the aGW node protocols contain Internet Protocol (IP) and Ethernet. The aGW and Enb nodes (DeNB1, and RNs1, 2, 3 cells) connect over IP edge cloud (1, 2, 3 and 4). QoS Parameters at the Transport Network (TN) guarantees QoS parameterization and traffic difference as seen in Figure 6.7 and Table 6.5. The user mobility in a cell is coordinated by the mobility model by updating the location of the users at every single interval. The user's mobility data is saved on the Global Server (Global-UE-Server). The channel framework parameters for the air interface cover slow fading, fast fading models, and path loss. The simulation emphasises on the user plane to execute E2E performance assessments [90]. The several traffic QoS have been established with the 3GPP standardization.

### 6.4.1. **Simulation Setup**

OPNET is a simulation used to assess the performance of the proposed model. Several scenarios are simulated to evaluate the impact of smart devices data traffic on regular 4G and 5G mobile networks data traffic. The simulated 4G and 5G data traffic applications include FTP, VoIP SMTP and RTP. The scenarios are categorized into four scenarios, the first one was designed for 4G mobile networks without density connection of devices and the other three were designed for supporting 5G mobile networks the first two 5G mobile networks in density connection of IoT devices without small cells and edge clouds, the last one 5G mobile networks with density connection of IoT devices in the form of slicing based on My smart systems use case requirement and it was supporting of RN cells and edge clouds. The results show the significant impact of IoT devices data traffic on high priority data traffic. The E2E network performance has been improved by allocating data of several IoT devices, which is determined by simulating several scenarios. Considerable performance improvement is achieved in terms of average IoT device and cell throughput, average upload response time, average packet E2E delay and radio resource utilization in the SMTP, VoIP, FTP and RTP applications. [29].

Table 6. 5 **: Simulation parameters**

| **Parameters** | **Setting** |
|---|---|
| Simulation length | 2000 sec |
| Cell layout | 1 Enb |
| eNB coverage radius | 350 m |
| Min. Enb-UEs | 35 m |
| Max. terminal power | 23 dBm |
| **5G Parameters** | |
| 5G cell | 8*8 antennas |
| Cloud | 4 Edge clouds with difference latencies and support different priorities. |
| Capability | Enabled |
| **RN Parameters** | |
| PRBs are allocated to Cells | RN 1 = 50 PRBs, RN 2= 25 PRBs, RN3 = 15 PRBs and DeNB 50 PRBs to evaluate PRB utilization. |
| Type of RN | Fixed |
| RN 1 | Supported by 6 antennas, 10 MHz TDD |
| RN 2 | Supported by 3 antennas, 5 MHz TDD |
| RN 3 | Supported by 1 antenna, 3 MHz TDD |
| TBS capacity | 1608 bits against MCS 16 and PRBs 5. Available service rate TBS-overhead (bits/TTI), 1608 (TBS)-352(overhead) =1256 bits/TTI. |
| Simulated scenarios | 4G Mobile Broadband No Density, 5G Density No Edge Clouds, 5G Density No Small Cells and SAM Density with Small Cells Edge Clouds |
| **General Parameters** | |
| Terminal speed | 120 km/h |
| Mobility model | Random Way Point (RWP) |
| Frequency reuse factor | 1 |
| System Bandwidth | 25 MHz |
| Path loss | 128.1+37.6 log 10 (R). R in km |
| Slow Fading | Log-normal shadowing, correlation 1, deviation 8 Db |
| Fast Fading | Jakes-like method |
| UE buffer size | ∞ |
| RN PDCP buffer size | ∞ |
| Power control | Fractional PC, α= 0.6, Po= -58 dBm |
| Applications | SMTP, VoIP, RTP and FTP. |

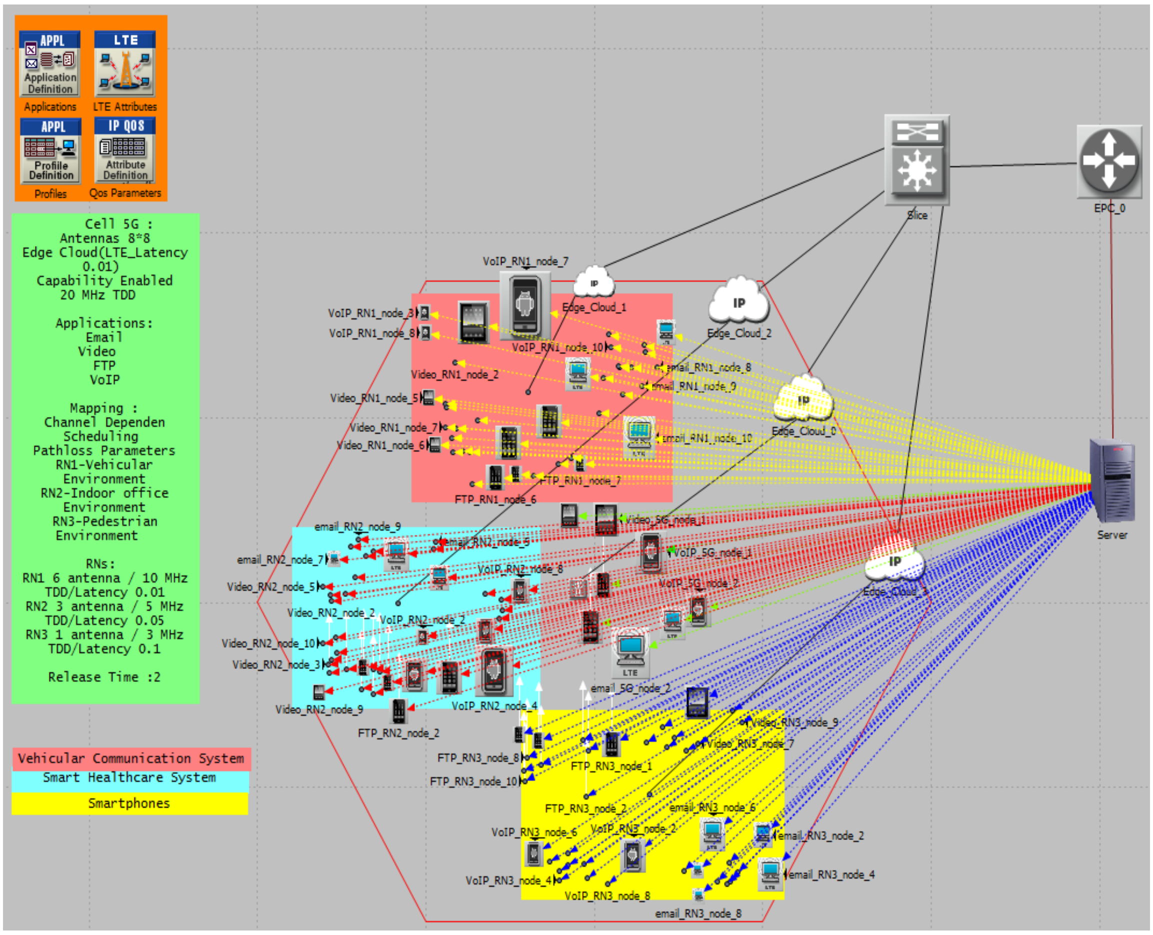

Figure 6. 7 **: OPNET 5G Project**

## 6.4.2. **Simulation Scenarios**

In this section, I compare the performance of My proposed SAM model, it will be evaluated by four scenarios relying on RNs and DeNB cells to assess the three slices of smart systems based on throughput, load and latency requirements within density environment of IoT devices. In the first scenario is 4G mobile networks with traditional 4G Smartphone without density environment of IoT devices connection, the second scenario is 5G mobile networks density environment of IoT devices connection without edge clouds, the third scenario is 5G mobile networks density environment of IoT devices connection without supporting by small cells and the fourth scenario is My SAM model in 5G mobile networks density environment of IoT devices connection with edge clouds and small cells as shown Table 6.6. The data packets from all the active smart devices, which are positioned in the nearness of the RNs and DeNB cells, are allocated slice services at the RN cell before being sent to the DeNB[118].

However, only the periodic per-hop control model is used in which the large allocation data packets are served to guarantee full utilization of RAN. The expiry timer is presented in order to limit the multiplexing delay particularly in the low loaded scenarios between RN and DeNB cells. In this situation, the allocated packet is served after *T*max at the latest. All the overhead stated scenarios are further sub-categorized into three sub-scenarios. In the first sub-scenario VCS the IoT devices are placed near to RN1 cell, which is supported by 6 antennas and 10 MHz TDD with a low level of priority 1 ms in both the small cell and edge cloud. In the second sub-scenario SHS the IoT devices are placed near to RN2 cell, which is supported by 3 antennas and 5 MHz TDD with a medium level of priority 5 ms in both the small cell and edge cloud. In the third sub-scenario Smartphone devices are placed near to RN3 cell, which is supported by 1 antenna and 3 MHz TDD with a medium level of priority 10 ms in both the small cell and edge cloud.

Table 6. 6 **: Simulation Scenarios**

| Scenarios | DeNB Cell | Small Cells | (MEC) | SDN (Slicer) | Smart Systems | Application Types |
|---|---|---|---|---|---|---|
| **(1) 4G Mobile Broadband No Density** | **4G** | **Yes** | **No** | **No** | **Mobile Broadband** | **SMTP, VoIP,FTP & RTP** |
| **(2) 5G Density No Edge Clouds** | **5G + MIMO** | **Yes** | **No** | **Yes** | **All** | **SMTP, VoIP,FTP & RTP** |
| **(3) 5G Density No Small Cells** | **5G + MIMO** | **No** | **Yes** | **Yes** | **All** | **SMTP, VoIP,FTP & RTP** |
| **(4) SAM Density with Small Cells Edge Clouds** | **5G + MIMO** | **Yes** | **Yes** | **Yes** | **All** | **SMTP, VoIP,FTP & RTP** |

## 6.5. Simulation Results and Analysis

### 6.5.1. IoT Nodes E2E delay

In this subsection, I have compared between the four scenarios and the three sub-scenarios as mentioned above: the simulation results showed the average air interface packet E2E delay for FTP and SMTP nodes are shown in Figures 6.8 and 6.9. The results display that the FTP and SMTP nodes have the diverse E2E delay variation in all four scenarios even when allocated together with GBR bearers. The cause is the proportional varieties distinguishing of priorities, which is characterized by SAM Model in “scenario2, 3 and 4”. Meanwhile, the VoIP bearer has a relatively low level of packets E2E delay in the “scenario 4” compared to “scenario1, 2, and 3”. As a result of the support by SAM and MEC placing Models within the small cells and edge clouds with RAN allocation in the shape of slicing, it tends to get higher priority feature and will permanently be scheduled first. Then, the average packets E2E delay of VoIP and RTP node are shown in Figures 6.10 and 6.11. It can be seen that “scenario 4" has somewhat better packets E2E delay compared to “scenario1, 2 and 3". However, I can see an example "scenario1", in RTP node has better performance compared to "scenario 2, and 3", as a result of there being no density of IoT devices connection that causes a low level of packets overload and cognition, also, the "scenario 4” is a result of SAM and MEC placing Models allocating RAN to VoIP and RTP node bearers in

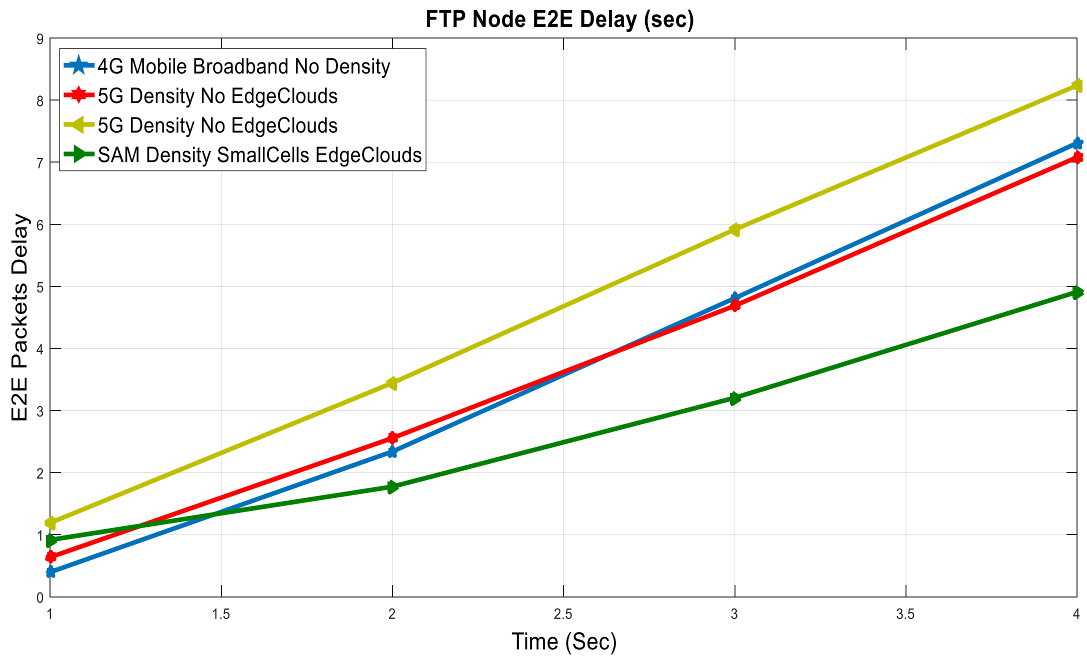


Figure 6. 8 : **FTP Node E2E Delay**

the RN1 to customize the slice 2 to support SHS with a higher MAC QoS class for VoIP SMTP IoT devices in this sub-scenario.

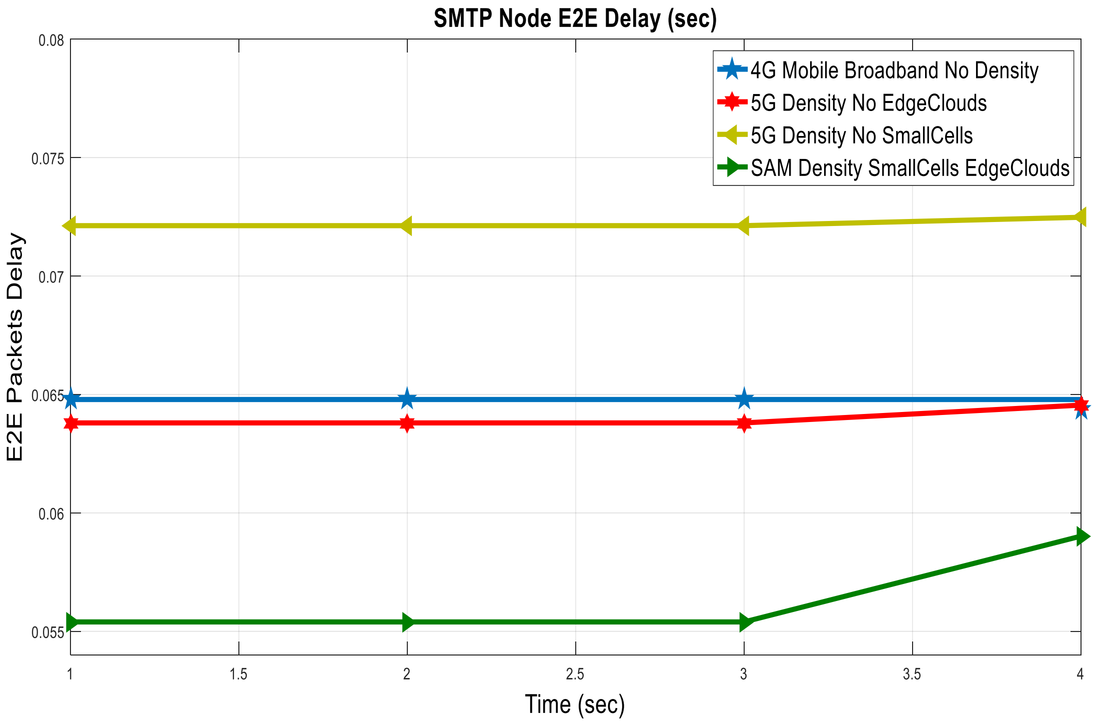


Figure 6. 9 : **SMTP Node E2E Delay**

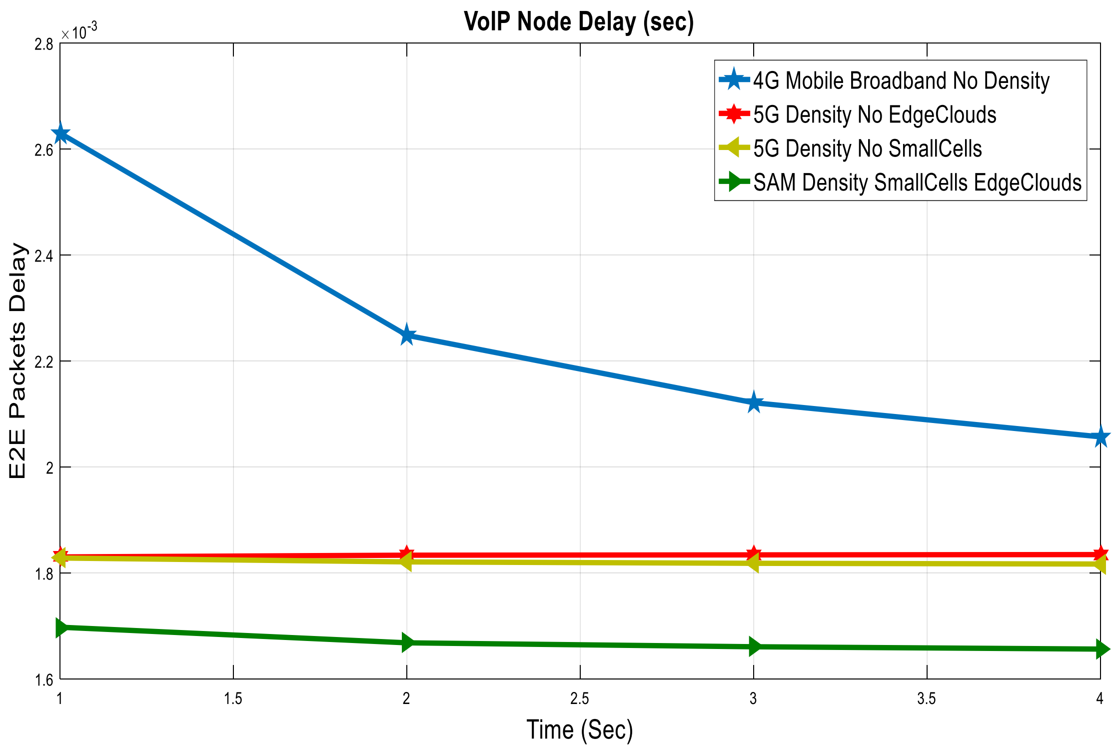


Figure 6. 10: **VoIP Node Delay**

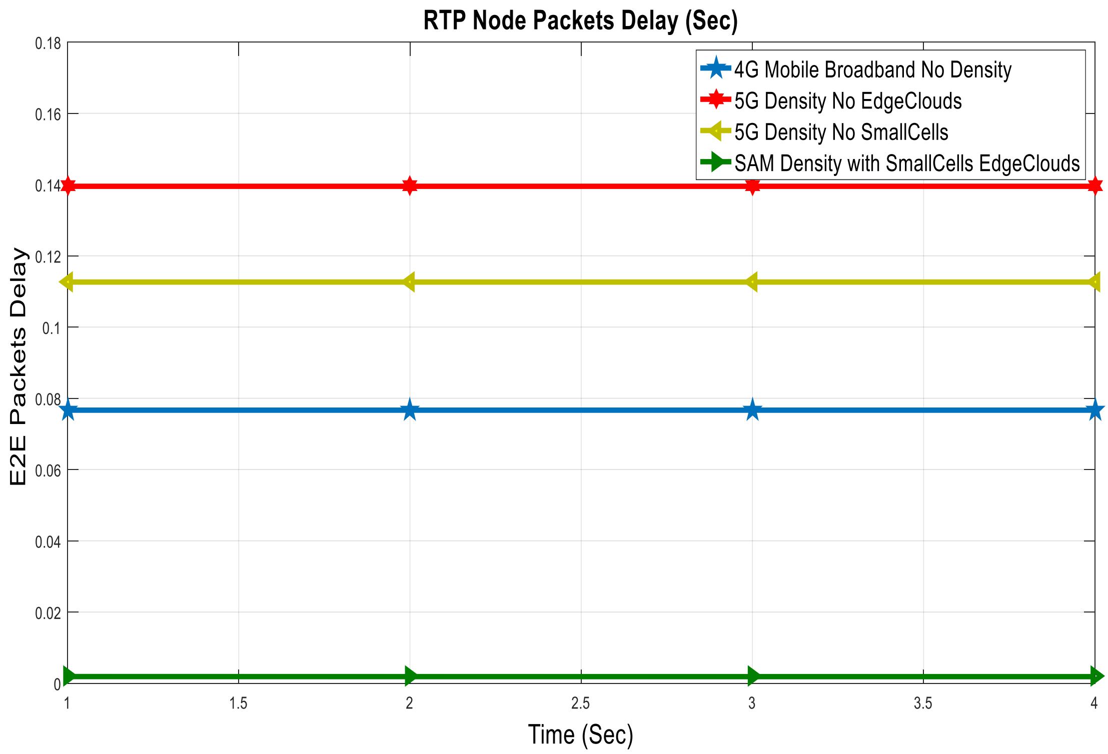


Figure 6. 11 : **RTP Node Packets Delay**

### 6.5.2. **IoT Nodes Loading**

The performance packet load of the VoIP node in "scenario 1 and 4" is shown in Figure 6.12. The VoIP node has a significantly higher ratio of loading performance compared to "scenario 2, and 3" as a result "scenario1" has no density and massive connectivity of IoT devices and "scenario 4" has customising slice by SAM and MEC placing Models. On the other hand, the SMTP node in Figure.6.13, in "scenario 4" has best loading compared with the other scenarios, where the SMTP node has served with specific priority requirement in "scenario 4" in the form of slicing based on SAM and MEC placing Models with existing small cells and edge clouds or without, mostly when it is not mixed with the FTP and RTP nodes and is allocated to a lower MAC QoS class than FTP. This is since the QoS is customized in form slices the same as in "scenario 4" and not customized in the "scenario1".

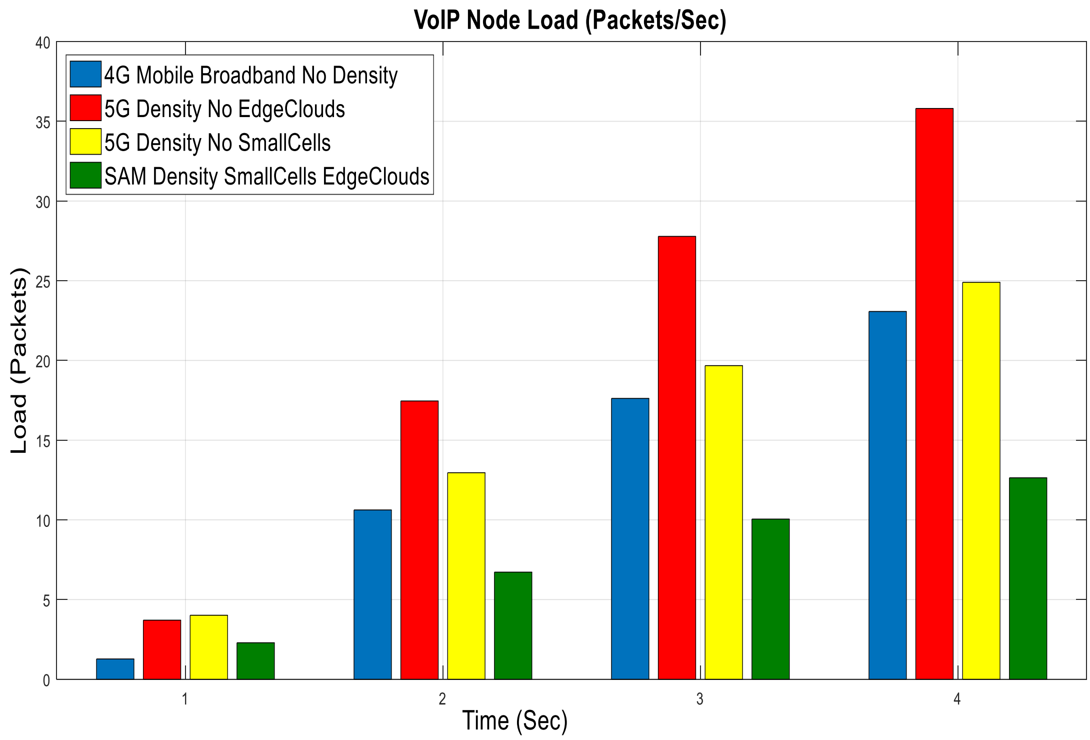


Figure 6. 12: **VoIP Node Load per Packets**

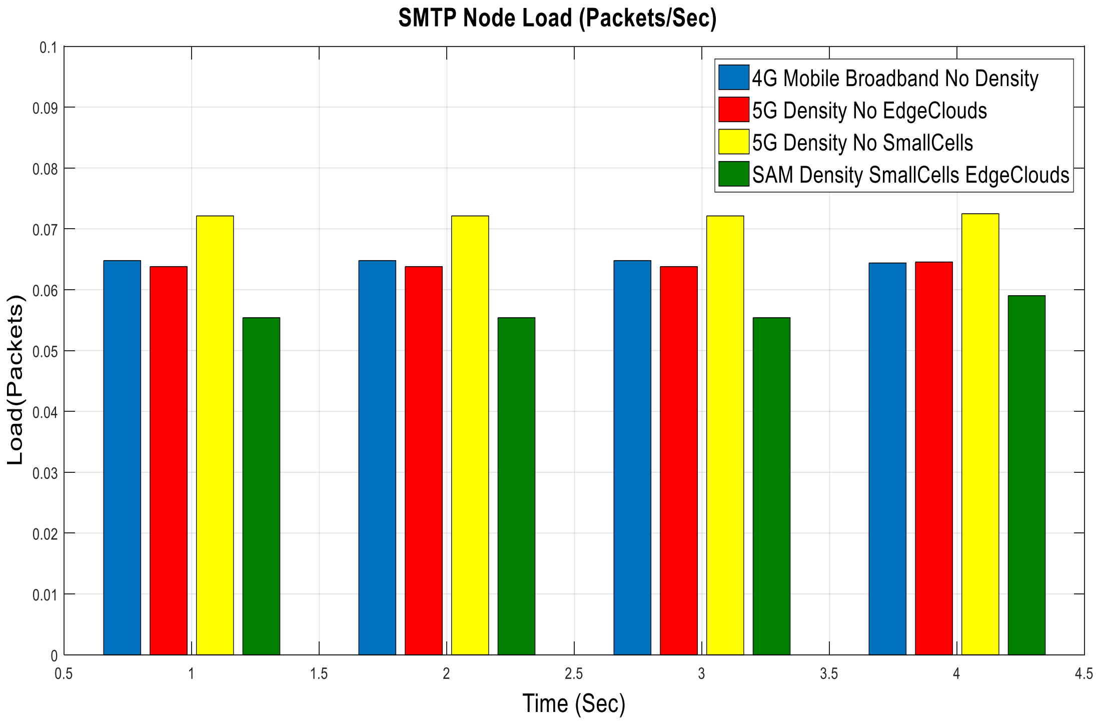


Figure 6. 13: **SMTP Node Load per Packets**

In addition, as I have considered the priority in a different level based on applications and smart systems need in the RNs, DeNB cells and edge clouds in “scenario2, 3, and 4” based on SAM and MEC placing Models. As I can see in FTP and RTP nodes packets loading as shown the results in Figures 6.14 and 6.15 as already predictable, the RTP node performance is decreased when going from fully mixed scenario to fully separate one. Where the FTP node packets’ loading time becomes improved in the “scenario 4”. This is due to the FTP node being allocated to the lower level of MAC QoS class and is supported with low priority in the RNs, DeNB cells and edge clouds as compared to the other applications. However, offering the FTP node lower priority is realistic since FTP is not the real-time application and in real life, it is acceptable for the FTP IoT devices to wait a couple more seconds for their files to be sent, while the same cannot be accepted when it comes to real-time applications such as RTP or VoIP.

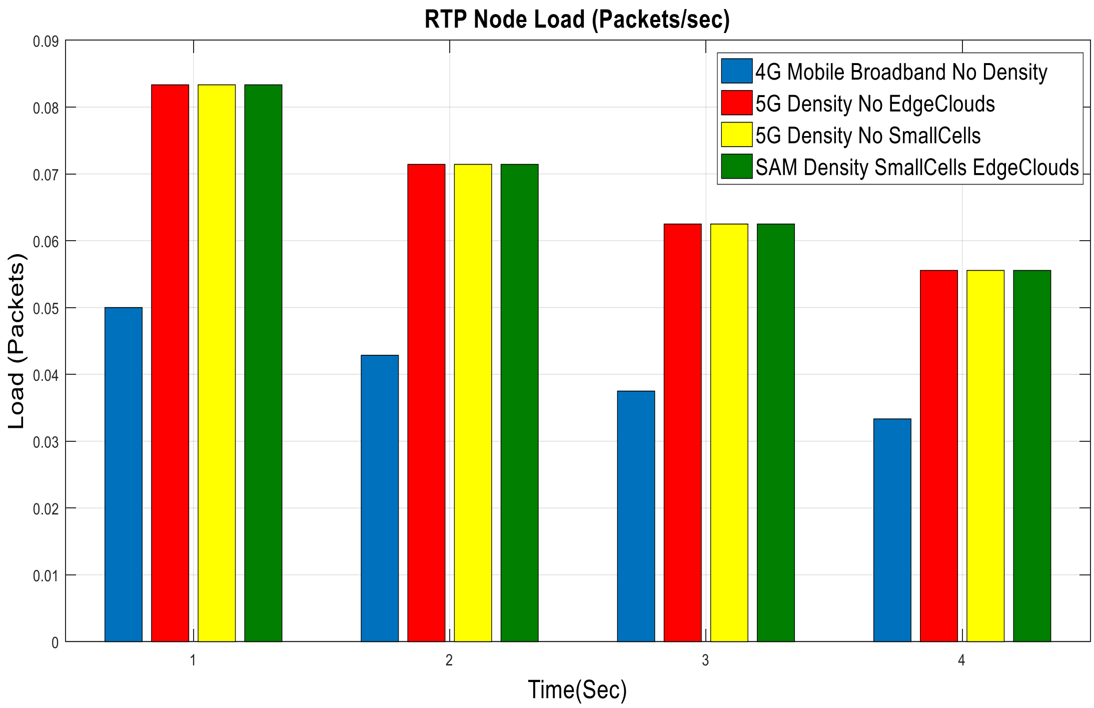


Figure 6. 12 **: RTP Node Load per Packets**

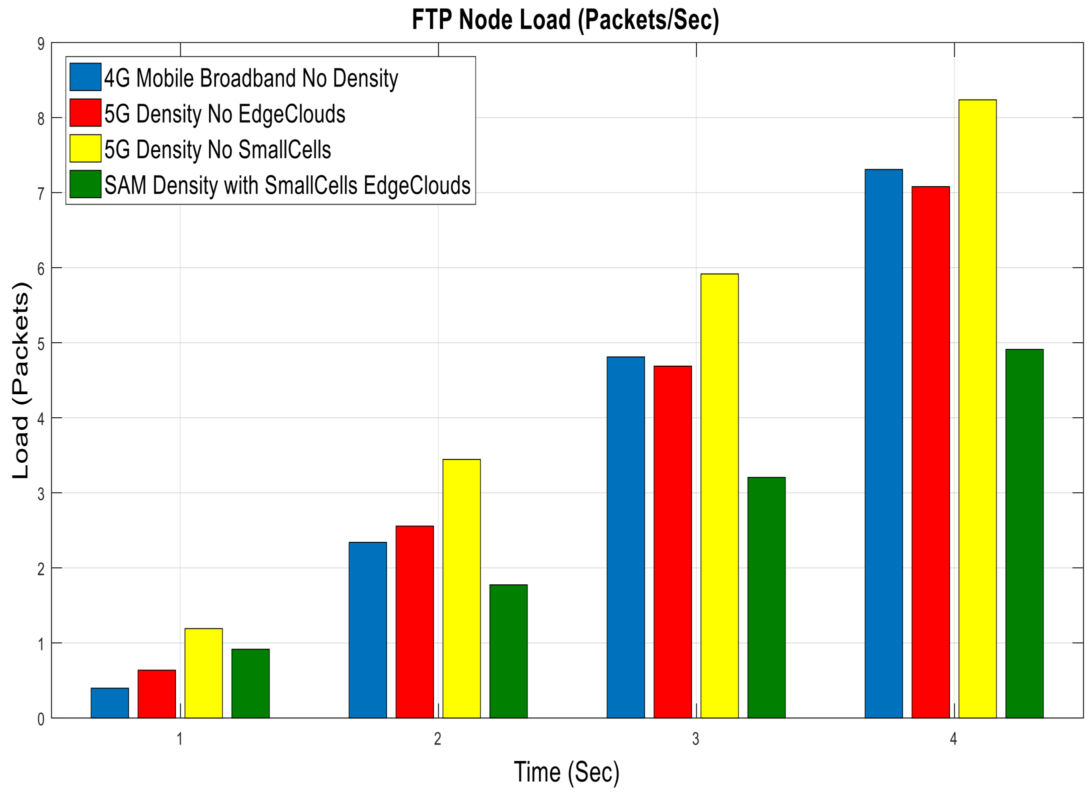


Figure 6. 13 : **FTP Node Load per Packets**

### 6.5.3. **IoT Nodes Throughput**

Moreover, in the case of throughput FTP and VoIP nodes as shown in Figures 6.16 and 6.17 the throughput for the FTP and VoIP IoT devices, the result describes that the FTP and VoIP IoT devices have worse performance in the "scenario1" compared to "scenario2,3 or 4" where the FTP nodes are allocated into the GBR MAC classes. In the "scenario 4", the FTP and VoIP nodes do not share the same non-GBR MAC QoS class with SMTP, RTP nodes in the same slice, which has different priority in the RNs, DeNB cells and edge clouds, since I customize the slices and IoT device such as the data rate of the FTP node. Besides, I have observed in all nodes downlink and uplink throughput has a significantly higher ratio in the "scenario 4" compared to "scenario1, 2 or 3" as results seen in Figures 6.18 and 6.19 demonstrate. This is due to the SAM and MEC placing Models allocating the packets to the different levels of MAC QoS class and being supported with low priority in the RNs, DeNB cells and edge clouds.

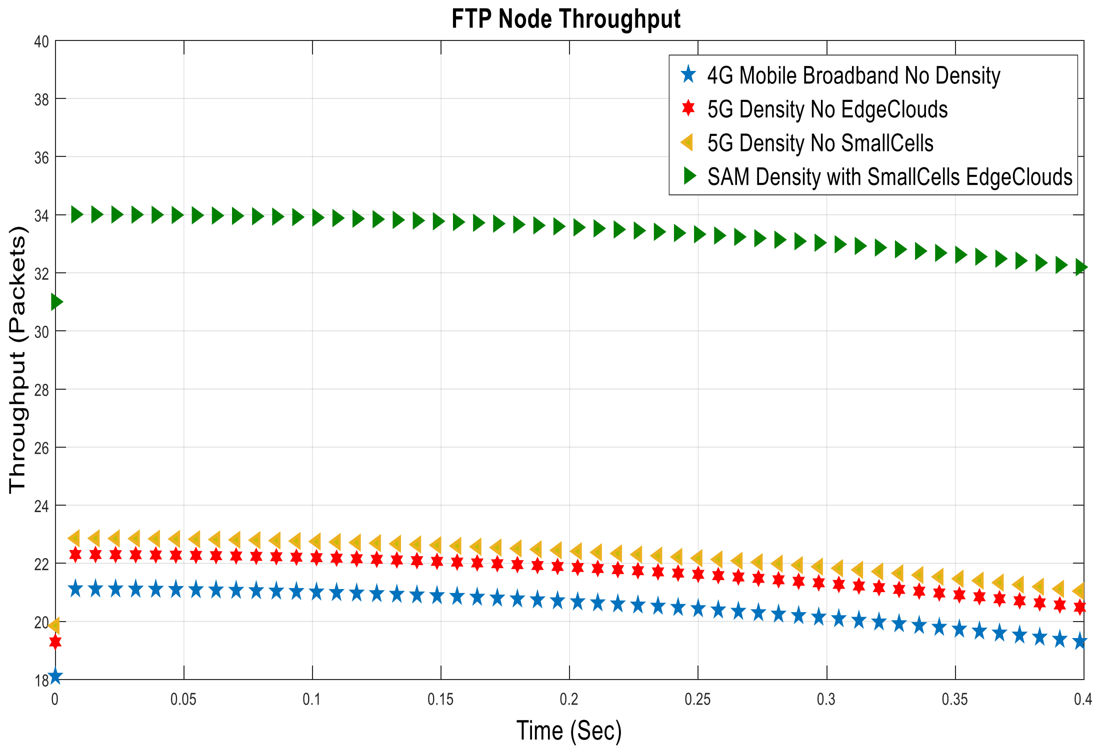


Figure 6. 14 **: FTP Node Throughput**

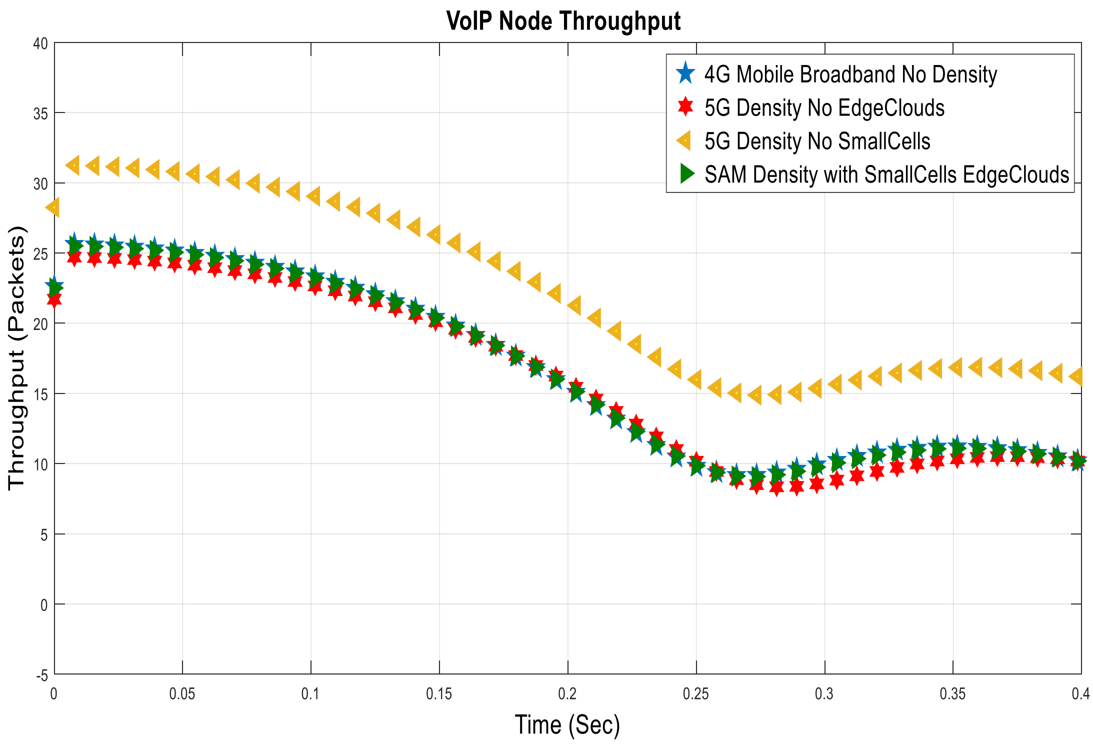


Figure 6. 15 **: VoIP Node Throughput**

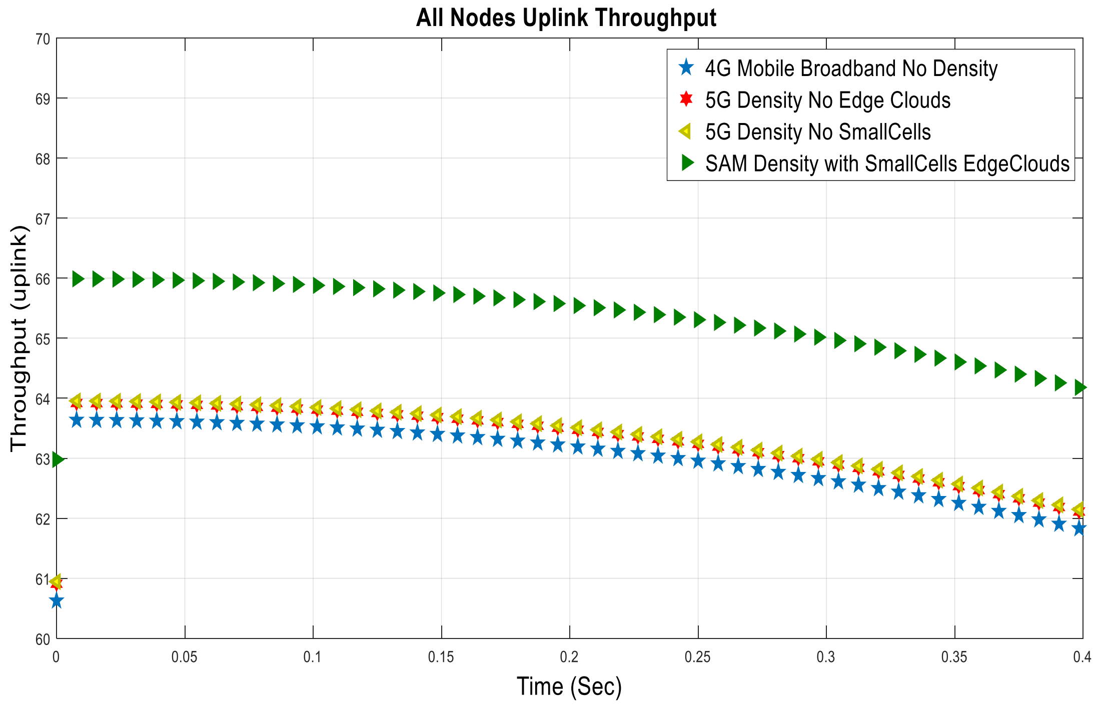


Figure 6. 16 : **All Nodes Uplink Throughput**

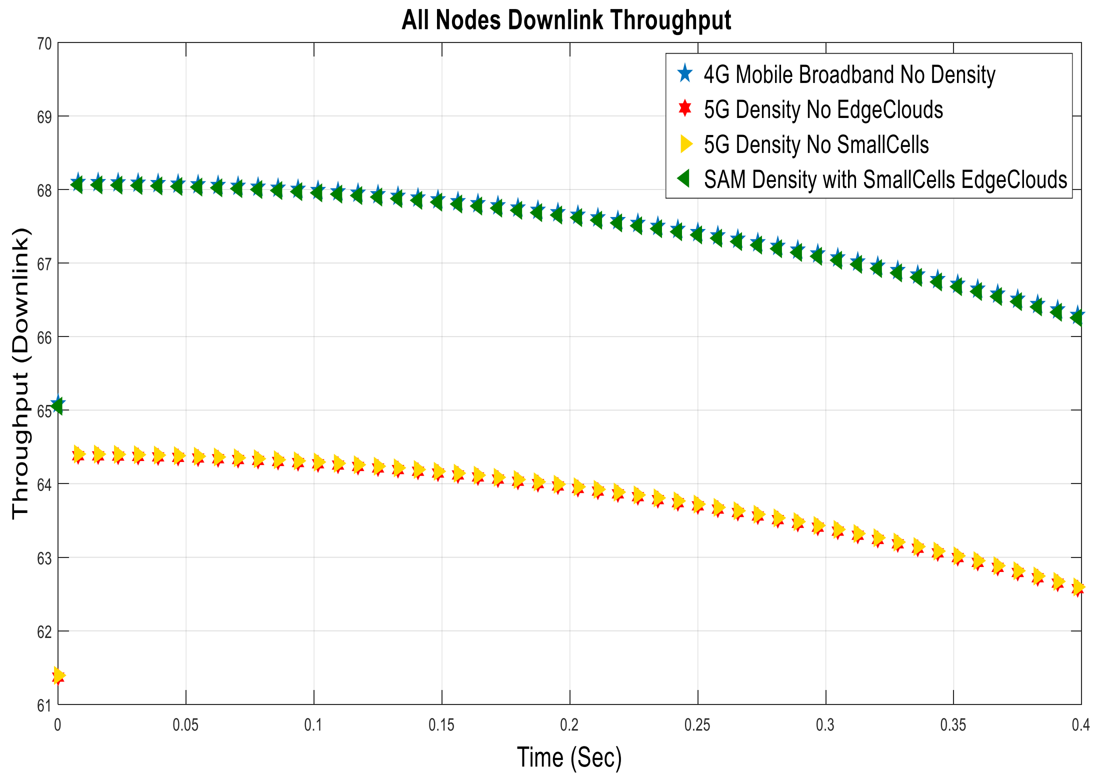


Figure 6. 17 : **All Nodes Downlink Throughput**

### 6.5.4. RNs and DeNB cells Performance

Finally, the RNs and DeNB cells load, delay and throughput are shown in Figures 6.20, 6.21, and 6.22. It can be seen that overall load, delay and throughput of cells are evaluated in just one scenario, which is “scenario 4” (SAM Density with Small Cells and Edge Clouds), where the IoT devices are connected directly to DeNB while some other IoT devices placed close to cell edge communicate with DeNB through the RNs, therefore, this part is focused on the individual cell to evaluate the overall performance and assess the SAM Model slicing process, when the single IoT devices are communicating individually via RNs or DeNB cells with the core networks. SAM Model has a positive impact on the network by freeing the network resources, which ultimately increases the DeNB cell load, throughput and improving the networks E2E packets delay. The overall DeNB cell load and throughput are lower levels compare with small cells as a result of enhancing slicing of the smart systems and separating the load and throughput among RN1, RN2 and RN3 cells with the result that, for example, the load and throughput in the RN1 has a significantly higher performance for VCS compare with RN2 SHS and RN3 (Smartphone).

Therefore, in the density environment of IoT devices in which I have considered both the VCS and SHS devices in the RN1 and RN2 cells, along with normal Smartphone users in RN3 cell, the performance evaluation showed there are RAN allocated in each slice based on smart systems QoS requirements, the single IoT devices are communicating individually via RNs or DeNB cells with the core networks. Moreover, the performance evaluation showed there are RAN allocated in each slice based on smart systems QoS requirements, the cells have different levels of E2E delay based on the slices close to RNs cells, the RN1 has lower level of E2E delay can offer to the IoT devices closed to this cell such as VCS, compared with the RN3 has the high level of E2E delay can serve the IoT devices closed to this cell such as Smartphone.

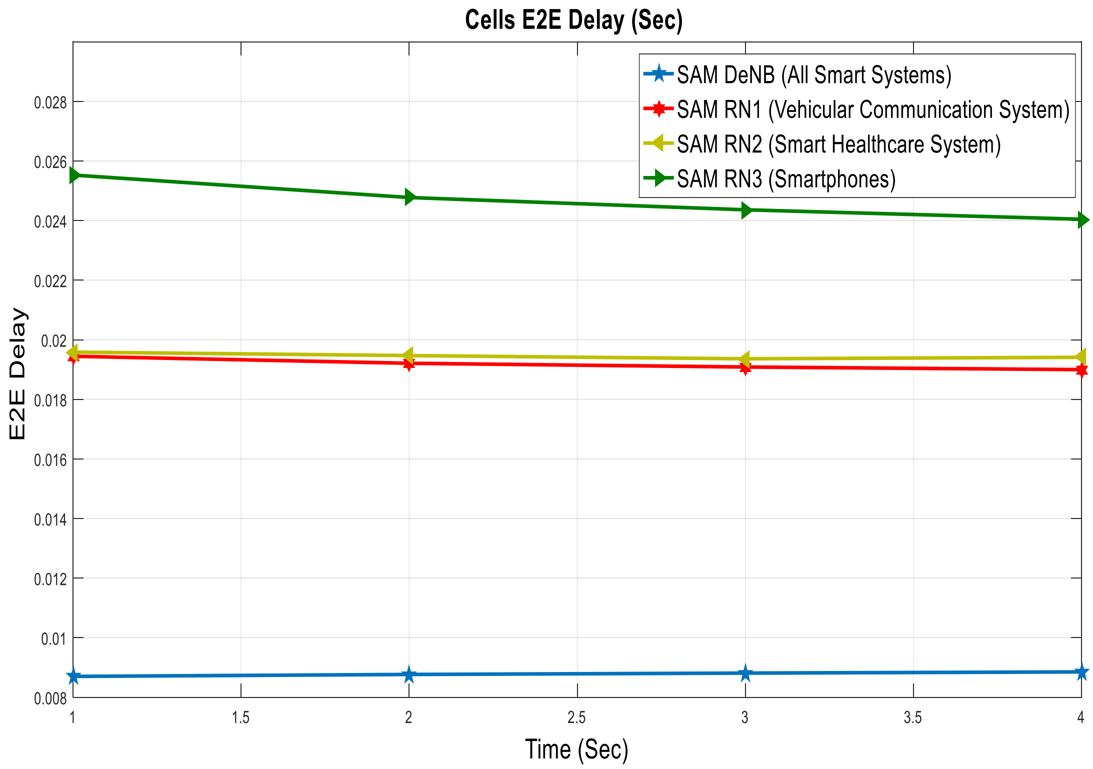


Figure 6. 18: **Cells E2E Delay**

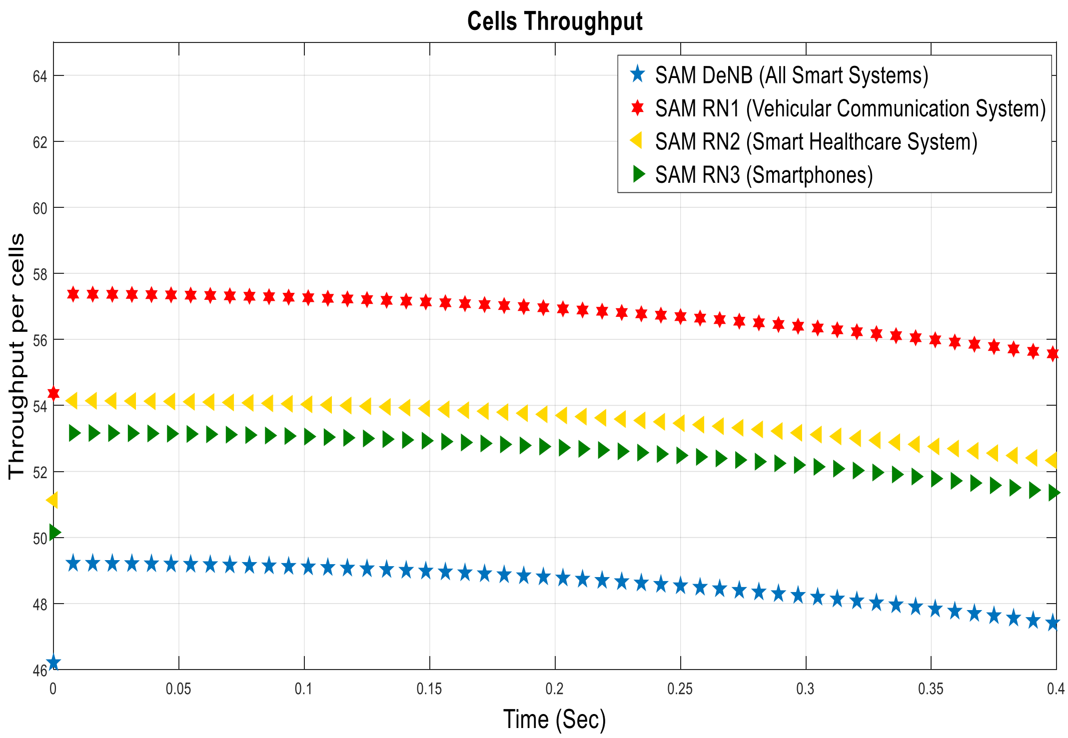


Figure 6. 19: **Cell Load per Packets**

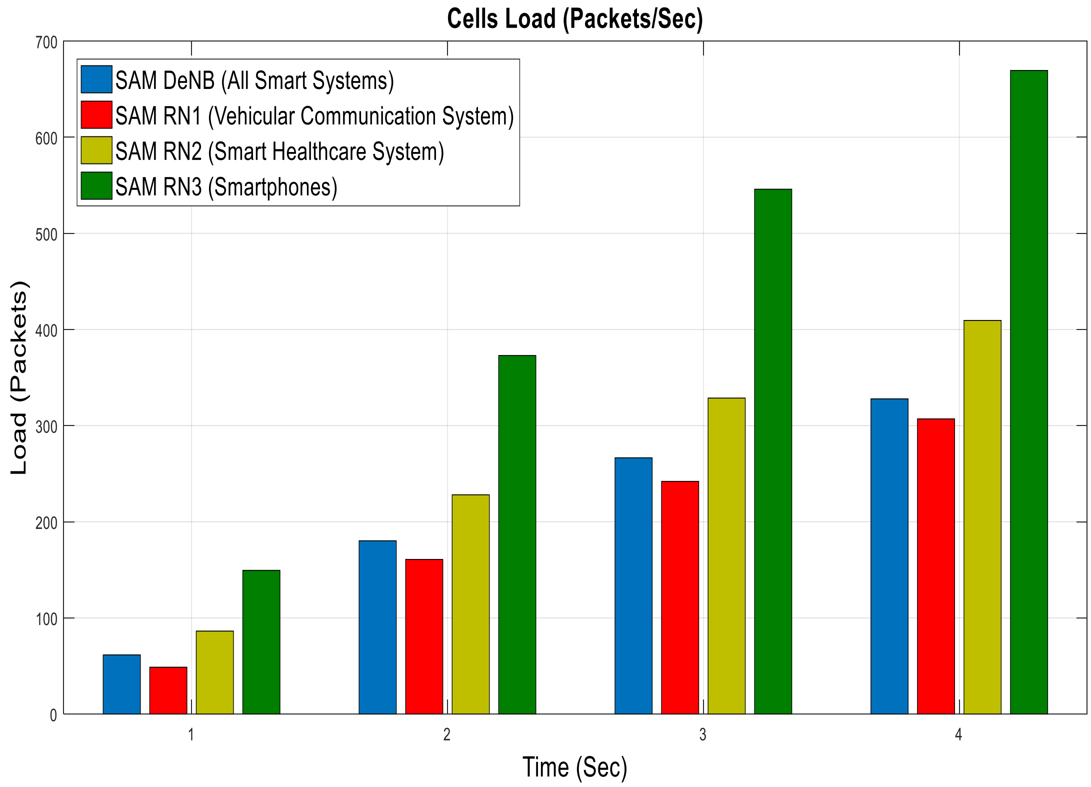


Figure 6. 20 : **Cells Throughput**

## 6.6. Rational Discussion of the Simulations

In this chapter, I have proposed SAM and MEC models, I have allocated unique slices for each smart system application such as VCS, SHS, and smartphone in the 5G mobile networks, based on the technical and QoS requirements. It uses well-recognized and promising network slicing technology, enabling cost-effective service deployment and an effective operational model.SAM and MEC models were working in the density environment of IoT devices in which I have considered both the VCS and SHS devices in the RN1 and RN2 cells, along with normal Smartphone users in RN3 cell, the performance evaluation showed there are RAN allocated in each slice based on smart systems QoS requirements, the single IoT devices are communicating individually via RNs or DeNB cells with the core networks. Moreover, the performance evaluation showed there are RAN allocated in each slice based on smart systems QoS requirements, the cells have different levels of E2E delay based on the slices close to RNs cells, the RN1 has lower level of E2E delay can offer to the IoT devices closed to this cell such as VCS, compared with the RN3 has the high level of E2E delay can serve the IoT devices closed to this cell such as Smartphone.

I have reduced the loss of multiplexing gain that occurs with conventional slicing, which leads to resource wastage. My proposed models use dynamic slice-based separation, inter-slice resource optimization, and service multiplexing in a single network. This chapter provides flexibility for all proven network operation techniques while enabling IoT heterogeneous service requirements effectively including overall technical and QoS requirements. These models enable the realization of customizing the dedicated slice method to solve the issue of high densification and heterogeneity of wireless networks in 5G mobile network and demand for high speed such as VCS including the low latency, low loading, high throughput, and massive scalability.

# CHAPTER 7: CONCLUSION AND FUTURE WORK

## 7.1. CONCLUSION

The main findings of this thesis were, firstly, I have developed the SAM and MEC models were working in the density environment of IoT devices in which I have considered both the VCS and SHS devices in the RN1 and RN2 cells, along with normal Smartphone users in RN3 cell, the performance evaluation showed there are RAN allocated in each slice based on smart systems QoS requirements, the single IoT devices are communicating individually via RNs or DeNB cells with the core networks. Moreover, the performance evaluation showed there are RAN allocated in each slice based on smart systems QoS requirements, the cells have different levels of E2E delay based on the slices close to RNs cells, the RN1 has lower level of E2E delay can offer to the IoT devices closed to this cell such as VCS, compared with the RN3 has the high level of E2E delay can serve the IoT devices closed to this cell such as Smartphone.

Secondly, I have applied and assessed the FIFO, PQ, and WFQ models over the RNs and DeNB cells of OPNET simulation, I found PQ model as the appropriate scheduling mechanism in the case of supporting various priorities queuing for data traffic, which demand a lower level of latency and throughput such as vehicular communication system. Also, the simulation has shown the FIFO and WFQ models have improved the packet loading particularly in the heavy data traffic such as smart healthcare system.

Thirdly, I have developed the DTA model is fixed RNs cell for uplink in 5G mobile networks. It improves the radio resource utilization for smart systems over 5G mobile networks. It offers a maximum multiplexing gain in the PDCP layer for data packets from several IoT devices along with considering diverse priorities to solve packets E2E delay. Also, in this chapter, I have presented a novel RSA scheme for radio resource allocation in the 5G mobile networks with network slicing. This scheme is a heuristic-based prioritized resource allocation among IoT devices in terms of the scheduling mechanism in the MAC layer of RNs and DeNB cells, which takes into consideration both the inter-and the intra-slice priority and executes the resource allocation accordingly in order to meet the QoS requirements dictated by the service slice.

The focus of this thesis was on providing three relevant solutions in the context of the RAN, considering the challenges that arise by the massive data traffic of heterogeneous types of IoT services in the density area, this part of the mobile network architecture compared to the core. While each of the individual solutions focused on addressing different problems, the resulting systems are highly complementary and contribute in the bigger picture of creating a next generation RAN that is flexible and adaptable to a wide range of use cases. Big attention was given to creating realizable system designs, something that is evident both from the concrete prototype implementations that complement them and from the extensive evaluation results that indicate the feasibility and performance of the proposed models and approaches. The high-level conclusions for each of these solutions are presented in the following subsections.

The contribution of this thesis attempted to address the problem of bringing virtualization and slicing capabilities to the RAN in a way that ensures the efficiency in the allocation of the PRBs and the fully functional and performance isolation of the co-existing slices. Towards this direction, the customized RAN slicing system was proposed, which enables the dynamic virtualization of RNs and DeNB cells and the flexible customization of slices to meet their respective service needs. Customized is based on a packet aggregation and allocation components that is the key enabler of the RAN virtualization. Among others, the aggregation builds on a set of abstractions that are specifically introduced for ensuring the isolated distribution of the PRBs to slices. It also allocates data traffic by controlling and extending the priorities of critical smart systems such as VCS and SHS to allow the introduction of multiple virtual control planes, each corresponding to a different slice. To complement the basic design of customized RAN slicing, an additional extension was also considered to provide the needs of 5G slicing service providers. The evaluation results and the case studies that were performed using a concrete prototype implementation revealed both the feasibility of the proposed approach and the benefits that it can bring compared to the state-of-the-art.

These thesis approaches are assessed over the OPNET simulator to measure the performance of the SAM and DTA models along with an assessment of the packet scheduling mechanism. The simulation considers IoT devices in several smart systems such as VCS, SHS and smartphones also, diverse protocols contain SMTP, FTP, VoIP and RTP. Simulation results displayed an important upgrading in IoT nodes packets transmission through RNs and DeNB cells, in My SAM model scenario comparing with other scenarios with or without density area. The model has enhanced such as E2E delay in FTP node by reaching 1ms, loading in VoIP node by 80% and throughput of all nodes in the uplink side of networks by 66%.

In addition, the results showed significant impact of IoT data traffic with several priority, networks E2E performance is improved by aggregating data traffic of several IoT devices with DTA model, which is determined by simulating several scenarios, considerable performance improvement is achieved in terms of average cell throughput, upload response time, packet E2E delay and radio resource utilization. Finally, the result found the PQ packet scheduling mechanism as the appropriate scheduling mechanism in case of supporting several priorities queuing for data traffic. I have improved the uplink network infrastructure of the 5G heterogeneous network by applying RNs as small cells and MEC as Edge cloud. These improvements can be useful for 5G network slicing operators and tenants in utilizing the appropriate slices in both QoS and technical support.

In summary, the novel contributions presented in this thesis are outlined below:

- A comprehensive state-of-the-art in three main topics are covered:
  - Data traffic aggregation models
  - Slices resource management models
  - Packet scheduling mechanisms
- Developed Data Traffic Aggregation (DTA) model is relying on aggregating packet data.
- Investigation and comparison among packet scheduling mechanisms.
- Designed and investigated a Slice Allocation Management (SAM) model based on critical services.
- Assess and investigate the above contribution works by using OPNET Simulation:
  - Design and simulate data traffic of IoT heterogeneous services in RNs and DeNB cells.

### 7.1.1. Limitations and Future work

This section summarizes the limitations and future work opportunities concerning the contributions made in this thesis:

**Security concerns** – the security issue that needs to be further considered is related to the security concerns raised by the Virtual File System (VSF) updating Model. Depending on the environment in which RAN slicing operates, the VSF updating mechanism could be potentially exploited in cases when the development and deployment of third-party applications are allowed. Developing a sandbox environment with controlled permissions for the execution of the VSFs is a very important topic for further research [119].

**Scale and scope** – Improving the scalability of RAN slicing for wide-area settings by introducing another layer of control and broadening its scope to go beyond the control and management of the PRBs in the 5G mobile network RAN by considering other domains like the core network and multi-RAT settings would provide a more holistic SDN solution for future mobile networks [120].

**Slicing the UE** – The current design of the SAM model of UEs to slices with the additional capability of supporting Over The Top (OTT) services over the extensions considered. However, in some use cases, multi-slice abilities might still be required, like for example allowing an IoT device to be connected to two different slices at the same time for work and personal use correspondingly [121]. Perhaps the most important challenge when trying to present such multi-slice abilities is associated with the need to provide guarantees regarding the conflict-free operation of slices. For instance, dealing with mobility and power management in such settings is not straightforward, as the roles and the responsibilities of the different slices are not clear. Presenting an approach to resolve such conflicts and regulate the operation of multi-slice environments is an interesting topic for future research [121].

**Multi-RAT settings** – The focus of SAM model was in allowing multi-tenancy in indoor RNs and DeNB cells environment with shared PRBs. The practicality of the network could be further enhanced by also considering unlicensed PRBs in the context of both 3GPP-based on RATs such as LTE-U and Multifibre, but also technologies such as WiFi, as this could significantly enlarge the capacity of the network[111][122]. However, the opportunistic nature of the spectrum presents extra limitations in terms of its allocation to the tenants, since its availability can change at an even higher time separately compared to the scenario of the shared spectrum such as every few ms. Investigating the applicability of the proposed dynamic pricing mechanism in such cases and classifying ways to improve its performance is another interesting topic for future work [111][122].

# References


[1] Cisco Systems 2017, “Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021 White Paper,” *Cisco*, 2017.

[2] CPI, “Market report 2017,” 2017.

[3] ITU, “R15-SG05-C-0040,” in *IMT-2020.TECH PERF REQ*, 2017.

[4] NGMN Alliance, “NGMN 5G White Paper,” *ngmn*, 2015.

[5] J.-H. Choi and D.-J. Shin, “Generalized RACH-Less Handover for Seamless Mobility in 5G and Beyond Mobile Networks,” *IEEE Wirel. Commun. Lett.*, vol. 8, no. 4, pp. 1264–1267, 2019.

[6] Y. Ji, J. Zhang, Y. Xiao, and Z. Liu, “5G flexible optical transport networks with large-capacity, low-latency and high-efficiency,” *China Commun.*, vol. 16, no. 5, pp. 19–32, 2019.

[7] N. Baranasuriya, V. Navda, V. N. Padmanabhan, and S. Gilbert, “QProbe: locating the bottleneck in cellular communication,” in *Proceedings of the 11th ACM Conference on Emerging Networking Experiments and Technologies - CoNEXT ’15*, 2015.

[8] NEC, “Making 5G a Reality,” *White Pap.*, 2016.

[9] N. Abbas, Y. Zhang, A. Taherkordi, and T. Skeie, “Mobile Edge Computing: A Survey,” *IEEE Internet of Things Journal*. 2018.

[10] B. Keogh and A. Zhu, “Wideband Self-Interference Cancellation for 5G Full-Duplex Radio Using a Near-Field Sensor Array,” *2018 IEEE MTT-S Int. Microw. Work. Ser. 5G Hardw. Syst. Technol. IMWS-5G 2018*, no. 1, pp. 1–3, 2018.

[11] T. Tojo, S. Okada, Y. Hirata, and S. Yasukawa, “An Integrated Control of IP and Optical Network for Multi-grade Virtualized Networks,” *2018 4th IEEE Conf. Netw. Softwarization Work. NetSoft 2018*, no. NetSoft, pp. 1–9, 2018.

[12] I. Da Silva *et al.*, “Impact of network slicing on 5G Radio Access Networks,” *EUCNC 2016 - Eur. Conf. Networks Commun.*, pp. 153–157, 2016.

[13] S. Y. Lien and K. C. Chen, “Massive access management for QoS guarantees in 3GPP machine-to-machine communications,” *IEEE Commun. Lett.*, vol. 15, no. 3, pp. 311–313, 2011.

[14] W. H. Chin, Z. Fan, and R. Haines, “Emerging technologies and research challenges for 5G wireless networks,” *IEEE Wirel. Commun.*, vol. 21, no. 2, pp. 106–112, 2014.

[15] M. Iftikhar, B. Landfeldt, S. Zeadally, and A. Zomaya, “Service level agreements (SLAs) parameter negotiation between heterogeneous 4G wireless network operators,” *Pervasive Mob. Comput.*, vol. 7, no. 5, pp. 525–544, 2011.

[16] R. Montero, F. Agraz, A. Pages, and S. Spadaro, "End-to-End 5G Service Deployment and Orchestration in Optical Networks with QoE Guarantees," *Int. Conf. Transparent Opt. Networks*, vol. 2018-July, pp. 1–4, 2018.

[17] J. Nightingale, P. Salva-Garcia, J. M. A. Calero, and Q. Wang, "5G-QoE: QoE modelling for ultra-HD video streaming in 5G networks," *IEEE Trans. Broadcast.*, vol. 64, no. 2, pp. 621–634, 2018.

[18] G. Liu, Y. Huang, F. Wang, J. Liu, and Q. Wang, "5G features from operation perspective and fundamental performance validation by field trial," *China Commun.*, vol. 15, no. 11, pp. 51–61, 2018.

[19] E. Coronado, S. N. Khan, and R. Riggio, " 5G-EmPOWER : A Software-Defined Networking Platform for 5G Radio Access Networks ," *IEEE Trans. Netw. Serv. Manag.*, vol. 16, no. 2, pp. 715–728, 2019.

[20] S. Ajah, A. Al-Sherbaz, S. Turner, and P. Picton, "Machine-to-machine Communications Energy Efficiencies: The Implications of Different M2M Communications Specifications," *Int. J. Wire. Mob. Comput.*, vol. 8, no. 1, pp. 15–26, 2015.

[21] M. Chen, J. Wan, and F. Li, "Machine-to-machine communications: Architectures, standards and applications," *KSII Transactions on Internet and Information Systems*, vol. 6, no. 2. pp. 480–497, 2012.

[22] K. C. Chen and S. Y. Lien, "Machine-to-machine communications: Technologies and challenges," *Ad Hoc Networks*, vol. 18, pp. 3–23, 2014.

[23] A. K. Jain, R. Acharya, S. Jakhar, and T. Mishra, "Fifth Generation (5G) Wireless Technology 'Revolution in Telecommunication,'" *Proc. Int. Conf. Inven. Commun. Comput. Technol. ICICCT 2018*, pp. 1867–1872, 2018.

[24] I. Abdalla and S. Venkatesan, "Remote subscription management of M2M terminals in 4G cellular wireless networks," in *Proceedings - Conference on Local Computer Networks, LCN*, 2012, pp. 877–885.

[25] A. Farrel, "Recent Developments in Service Function Chaining (SFC) and Network Slicing in Backhaul and Metro Networks in Support of 5G," *Int. Conf. Transparent Opt. Networks*, vol. 2018-July, pp. 1–4, 2018.

[26] Y. Huo, X. Dong, W. Xu, and M. Yuen, "Cellular and WiFi Co-design for 5G User Equipment," *IEEE 5G World Forum, 5GWF 2018 - Conf. Proc.*, pp. 256–261, 2018.

[27] B. C. Kuszmaul, "The RACE network architecture," in *Proceedings of 9th International Parallel Processing Symposium*, 2009, no. Wiley, pp. 508–513.

[28] H. Holma and A. Toskala, *LTE for UMTS - OFDMA and SC-FDMA based radio access*. Chichester, UK: John Wiley & Sons, Ltd, 2009.

[29] E. Dahlman, S. Parkvall, and J. Skold, *4G: LTE/LTE-Advanced for Mobile Broadband*. 2013.

[30] I. B. Khriplovich and A. A. Pomeransky, "Equations of Motion of Spinning Relativistic Particle in Electromagnetic and Gravitational Fields," *J. Can. Dent. Assoc.*, vol. 70, no. 3, pp. 156–7, Mar. 1998.

[31] 3GPP TR 36.814 V9.0.0, "Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects (Release 9)," *3rd Gener. Partnersh. Proj. Tech. Rep,* 2010.

[32] K. Kondepu, F. Giannone, S. Vural, B. Riemer, P. Castoldi, and L. Valcarenghi, "Experimental demonstration of 5G Virtual EPC recovery in federated testbeds," *2019 IFIP/IEEE Symp. Integr. Netw. Serv. Manag. IM 2019*, pp. 712–713, 2019.

[33] R. M. Rao, M. Fontaine, and R. Veisllari, "A Reconfigurable Architecture for Packet Based 5G Transport Networks," *IEEE 5G World Forum, 5GWF 2018 - Conf. Proc.*, pp. 474–477, 2018.

[34] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation," *ETSI TR 136 311*, no. 36.211, pp. 1–85, 2008.

[35] R. Fronthaul, G. Beam-switching, and M. Ip, "Multi RAT ( WiFi / LTE / 5G ) Mobile Network featuring," *Eur. Conf. Opt. Commun. 2018*, no. 1, pp. 7–9, 2018.

[36] T. Specification, "TS 136 101 - V12.7.0 - LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio transmission and reception (3GPP TS 36.101 version 12.7.0 Release 12)," *Etsi*, vol. 0, 2015.

[37] F. Boccardi *et al.*, "Multiple-antenna techniques in LTE-advanced," *IEEE Commun. Mag.*, 2012.

[38] Y. Ni, J. Liang, X. Shi, and D. Ban, "Research on Key Technology in 5G Mobile Communication Network," *Proc. - 2019 Int. Conf. Intell. Transp. Big Data Smart City, ICITBS 2019*, pp. 199–201, 2019.

[39] P. Antonio, F. Grimaccia, and M. Mussetta, "Architecture and methods for innovative heterogeneous wireless sensor network applications," *Remote Sens.*, 2012.

[40] D. Niyato, E. Hossain, D. I. K. D. I. Kim, and Z. H. Z. Han, "Relay-centric radio resource management and network planning in IEEE 802.16 j mobile multihop relay networks," *... Commun. IEEE ...*, vol. 8, no. 12, pp. 6115–6125, 2009.

[41] H. Wang, P. Zhang, J. Li, and X. You, "Radio propagation and wireless coverage of LSAA-based 5G millimeter-wave mobile communication systems," *China Commun.*, vol. 16, no. 5, pp. 1–18, 2019.

[42] Y. Sui, J. Vihriala, A. Papadogiannis, M. Sternad, W. Yang, and T. Svensson, "Moving cells: A promising solution to boost performance for vehicular users," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 62–68, 2013.

[43] A. Annunziato, "5G vision: NGMN - 5G initiative," in *IEEE Vehicular Technology Conference*, 2015, vol. 2015.

[44] M. Iwamura, "NGMN view on 5G architecture," in *IEEE Vehicular Technology Conference*, 2015, vol. 2015.

[45] W. Lee, T. Na, and J. Kim, "How to Create a Network Slice? - A 5G Core Network Perspective," *Int. Conf. Adv. Commun. Technol. ICACT*, vol. 2019-

Febru, pp. 616–619, 2019.

[46] L. Ma, X. Wen, L. Wang, Z. Lu, and R. Knopp, "An SDN/NFV Based Framework for Management and Deployment of Service Based 5G Core Network," no. October, pp. 86–98, 2018.

[47] David S. Watson, Mary Ann Piette, Osman Sezgen, and Naoya Motegi, "Machine to Machine (M2M) Technology in Demand Responsive Commercial Buildings ," *2004 ACEEE Summer Study Energy Effic. Build. Pacific Grove*, 2004.

[48] L. Atzori, A. Iera, and G. Morabito, "The Internet of Things: A survey," *Comput. Networks*, 2010.

[49] A. Ivanov, "The internet of things," *IEEE Design and Test*, 2014.

[50] C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos, "Context aware computing for the internet of things: A survey," *IEEE Commun. Surv. Tutorials*, 2014.

[51] R. Ricart-Sanchez, P. Malagon, J. M. Alcaraz-Calero, and Q. Wang, "Hardware-Accelerated Firewall for 5G Mobile Networks," *Proc. - Int. Conf. Netw. Protoc. ICNP*, vol. 2018-Septe, pp. 446–447, 2018.

[52] M. Chen, J. Wan, and F. Li, "Machine-to-machine communications: Architectures, standards and applications," *KSII Transactions on Internet and Information Systems*. 2012.

[53] R. Xie, R. Wang, T. Huang, and J. Wu, "A game theoretic approach for hierarchical caching resource sharing in 5G networks with virtualization," *China Commun.*, vol. 16, no. 7, pp. 32–48, 2019.

[54] A. Varga, "A Quick Overview Of The OMNeT ++ IDE," *Omnet++*, 2017. .

[55] S. Siraj, A. K. Gupta, and Badgujar-Rinku, "Network Simulation Tools Survey," *Int. J. Adv. Res. Comput. Commun. Eng. Vol. 1, Issue 4, June 2012*, 2012.

[56] S. M. Woolley, D. Posada, and K. A. Crandall, "A comparison of phylogenetic network methods using computer simulation," *PLoS One*, 2008.

[57] I. S. Hammoodi, B. G. Stewart, A. Kocian, and S. G. McMeekin, "A comprehensive performance study of OPNET modeler for ZigBee wireless sensor networks," in *NGMAST 2009 - 3rd International Conference on Next Generation Mobile Applications, Services and Technologies*, 2009, pp. 357–362.

[58] Z. Lu and H. Yang, *Unlocking the power of OPNET modeler*. 2012.

[59] V. Pandey, "A review on data aggregation techniques in wireless sensor network," *J. Electron. Electr. Eng.*, vol. 1, no. 2, pp. 1–8, 2010.

[60] T. Salam, W. U. Rehman, and X. Tao, "Cooperative Data Aggregation and Dynamic Resource Allocation for Massive Machine Type Communication," *IEEE Access*, vol. 6, pp. 4145–4158, 2018.

[61] Y. E. Massad, M. Goyeneche, J. J. Astrain, and J. Villadangos, "Data

Aggregation in Wireless Sensor Networks," *2008 3rd Int. Conf. Inf. Commun. Technol. From Theory to Appl.*, vol. 2, no. June, pp. 1040–1052, 2008.

[62] F. Balteanu, H. Modi, Y. Zhu, S. Drogi, and S. Khesbak, "Envelope tracking system for high power applications in uplink 4G/5G LTE advanced," *Asia-Pacific Microw. Conf. Proceedings, APMC*, vol. 2018-November, pp. 863–865, 2019.

[63] S.-Y. T. S.-I. S. M.-H. Tsai, "Effect of Data Aggregation in M2M Networks," *Wirel. Pers. Multimed. Commun. (WPMC), 2012 15th Int. Symp.*, pp. 95–99, 2012.

[64] T. Salam, W. U. Rehman, and X. Tao, "Data Aggregation in Massive Machine Type Communication: Challenges and Solutions," *IEEE Access*, vol. 7, pp. 41921–41946, 2019.

[65] F. Ahmad, S. N. K. Marwat, Y. Zaki, and C. Goerg, "Tailoring LTE-Advanced for M2M Communication using Wireless Inband Relay Node," *WTC 2014; World Telecommunications Congress 2014*. 2014.

[66] X. Li, S. Liu, F. Wu, S. Kumari, and J. J. P. C. Rodrigues, "Privacy preserving data aggregation scheme for mobile edge computing assisted IoT applications," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4755–4763, 2019.

[67] D. Gong and Y. Yang, "Low-latency SINR-based data gathering in wireless sensor networks," *IEEE Trans. Wirel. Commun.*, vol. 13, no. 6, pp. 3207–3221, 2014.

[68] P. M. Egidius, A. M. Abu-Mahfouz, M. Ndiaye, and G. P. Hancke, "Data aggregation in software-defined wireless sensor networks: A review," *Proc. IEEE Int. Conf. Ind. Technol.*, vol. 2019-February, pp. 1749–1754, 2019.

[69] M. Hasan, E. Hossain, and D. I. Kim, "Resource allocation under channel uncertainties for relay-aided device-to-device communication underlaying LTE-A cellular networks," *IEEE Trans. Wirel. Commun.*, vol. 13, no. 4, pp. 2322–2338, 2014.

[70] M. Shan, G. Chen, D. Luo, X. Zhu, and X. Wu, "Building Maximum Lifetime Shortest Path Data Aggregation Trees in Wireless Sensor Networks," *ACM Trans. Sens. Networks*, vol. 11, no. 1, pp. 1–24, 2014.

[71] M. Dagar and S. Mahajan, "Data Aggregation in Wireless Sensor Network : A Survey," *Int. J. Inf. Comput. Technol.*, vol. 3, no. 3, pp. 167–174, 2013.

[72] M. K. Sroya, "Comparison of Various Scheduling Schemes for Voice over LTE Networks," pp. 28–33, 2013.

[73] T. Velmurugan, H. Chandra, and S. Balaji, "Comparison of queuing disciplines for differentiated services using OPNET," in *ARTCom 2009 - International Conference on Advances in Recent Technologies in Communication and Computing*, 2009.

[74] R. K. Pradhan and M. A. Gregory, "Comparison of queuing disciplines for fiber to the home networks," *2012 Int. Conf. Comput. Inf. Sci. ICCIS 2012 - A Conf. World Eng. Sci. Technol. Congr. ESTCON 2012 - Conf. Proc.*, vol. 2, pp. 751–754, 2012.

[75] M. Jiang, M. Condoluci, and T. Mahmoodi, “Network slicing in 5G: An auction-based model,” in *IEEE International Conference on Communications*, 2017.

[76] M. A. Hasabelnaby, H. A. I. Selmy, and M. I. Dessouky, “Joint optimal transceiver placement and resource allocation schemes for redirected cooperative hybrid FSO/mmW 5G fronthaul networks,” *J. Opt. Commun. Netw.*, vol. 10, no. 12, pp. 975–990, 2018.

[77] H. Zhang, N. Liu, X. Chu, K. Long, A. H. Aghvami, and V. C. M. Leung, “Network Slicing Based 5G and Future Mobile Networks: Mobility, Resource Management, and Challenges,” *IEEE Commun. Mag.*, vol. 55, no. 8, pp. 138–145, 2017.

[78] M. Bagaa, T. Taleb, A. Laghrissi, A. Ksentini, and H. Flinck, “Coalitional game for the creation of efficient virtual core network slices in 5G mobile systems,” *IEEE J. Sel. Areas Commun.*, vol. 36, no. 3, pp. 1–18, 2018.

[79] I. Afolabi, M. Bagaa, T. Taleb, and H. Flinck, “End-To-end network slicing enabled through network function virtualization,” in *2017 IEEE Conference on Standards for Communications and Networking, CSCN 2017*, 2017.

[80] D. Sattar and A. Matrawy, “Optimal Slice Allocation in 5G Core Networks,” *IEEE Netw. Lett.*, vol. 1, no. 2, pp. 48–51, 2019.

[81] A. Kammoun, N. Tabbane, G. Diaz, A. Dandoush, and N. Achir, “End-to-end efficient heuristic algorithm for 5G network slicing,” in *Proceedings - International Conference on Advanced Information Networking and Applications, AINA*, 2018.

[82] J. Costa-Requena, A. Poutanen, S. Vural, G. Kamel, C. Clark, and S. K. Roy, “SDN-Based UPF for Mobile Backhaul Network Slicing,” *2018 Eur. Conf. Networks Commun. EuCNC 2018*, pp. 48–53, 2018.

[83] M. I. Kamel, L. B. Le, and A. Girard, “LTE wireless network virtualization: Dynamic slicing via flexible scheduling,” in *IEEE Vehicular Technology Conference*, 2014.

[84] K. Nguyen, L. Zhe-Tao, and H. Sekiya, “Virtualization for Flexibility and Network-Aware on 5G Mobile Devices,” *2019 IEEE 43rd Annu. Comput. Softw. Appl. Conf.*, vol. 1, pp. 928–929, 2019.

[85] N. Nikaein *et al.*, “Network Store: Exploring Slicing in Future 5G Networks,” in *Proceedings of the 10th International Workshop on Mobility in the Evolving Internet Architecture*, 2015.

[86] R. Martinez, R. Vilalta, M. Requena, R. Casellas, R. Munoz, and J. Mangues, “Experimental SDN control solutions for automatic operations and management of 5G services in a fixed mobile converged packet-optical network,” *22nd Conf. Opt. Netw. Des. Model. ONDM 2018 - Proc.*, pp. 214–219, 2018.

[87] J. G. Andrews *et al.*, “What will 5G be?,” *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[88] H. J. Einsiedler, A. Gavras, P. Sellstedt, R. Aguiar, R. Trivisonno, and D.

Lavaux, “System design for 5G converged networks,” in *2015 European Conference on Networks and Communications, EuCNC 2015*, 2015, pp. 391–396.

[89] N. Panwar, S. Sharma, and A. K. Singh, “A survey on 5G: The next generation of mobile communication,” *Phys. Commun.*, vol. 18, pp. 64–84, 2016.

[90] M. Dighriri, G. M. Lee, T. Baker, and L. J. Moores, “Measuring and Classification of Smart Systems Data Traffic Over 5G Mobile Networks,” in *Technology for Smart Futures*, A. B. Dastbaz M., Arabnia H., Ed. Springer, Cham, 2015.

[91] Y. Zaki, L. Zhao, C. Goerg, and A. Timm-Giel, “LTE wireless virtualization and spectrum management,” in *2010 3rd Joint IFIP Wireless and Mobile Networking Conference, WMNC 2010*, 2010, pp. 1–6.

[92] C. Liang, F. R. Yu, and X. Zhang, “Information-centric network function virtualization over 5g mobile wireless networks,” *IEEE Netw.*, vol. 29, no. 3, pp. 68–74, May 2015.

[93] K. Zhu and E. Hossain, “Virtualization of 5G Cellular Networks as a Hierarchical Combinatorial Auction,” *IEEE Trans. Mob. Comput.*, vol. 15, no. 10, pp. 2640–2654, Oct. 2016.

[94] M. M. Rahman, C. Despins, and S. Affes, “HetNet Cloud: Leveraging SDN & Cloud Computing for Wireless Access Virtualization,” in *2015 IEEE International Conference on Ubiquitous Wireless Broadband, ICUWB 2015*, 2015, pp. 1–5.

[95] M. Dighriri, A. S. D. Alfoudi, G. M. Lee, and T. Baker, “Data Traffic Model in Machine to Machine Communications over 5G Network Slicing,” in *Proceedings - 2016 9th International Conference on Developments in eSystems Engineering, DeSE 2016*, 2017, pp. 239–244.

[96] M. Dighriri, G. M. Lee, and T. Baker, “Measurement and classification of smart systems data traffic over 5g mobile networks,” in *Technology for Smart Futures*, 2017.

[97] A. S. D. Alfoudi, G. M. Lee, and M. Dighriri, “Seamless LTE-WiFi Architecture for Offloading the Overloaded LTE with Efficient UE Authentication,” in *Proceedings - 2016 9th International Conference on Developments in eSystems Engineering, DeSE 2016*, 2017, pp. 118–122.

[98] M. Dighriri, G. M. Lee, and T. Baker, “Resource Allocation Scheme in 5G Network Slices,” 2015.

[99] M. Hasan, E. Hossain, and D. I. Kim, “Resource allocation under channel uncertainties for relay-aided device-to-device communication underlaying LTE-A cellular networks,” *IEEE Trans. Wirel. Commun.*, vol. 13, no. 4, pp. 2322–2338, Apr. 2014.

[100] G. Zirong and Z. Huaxin, “Simulation and analysis of weighted fair queueing algorithms in OPNET,” in *Proceedings - 2009 International Conference on Computer Modeling and Simulation, ICCMS 2009*, 2009, pp. 114–118.

[101] M. Dighriri, A. S. D. Alfoudi, G. M. Lee, T. Baker, and R. Pereira, "Comparison Data Traffic Scheduling Techniques for Classifying QoS over 5G Mobile Networks," in *2017 31st International Conference on Advanced Information Networking and Applications Workshops (WAINA)*, 2017, pp. 492–497.

[102] C. S. Lee, G. M. Lee, and W. S. Rhee, "Smart Ubiquitous Networks for future telecommunication environments," *Comput. Stand. Interfaces*, vol. 36, no. 2, pp. 412–422, 2014.

[103] Y. Miaji and S. Hassan, "Comparative simulation of scheduling mechanism in packet switching network," in *Proceedings - 2nd International Conference on Network Applications, Protocols and Services, NETAPPS 2010*, 2010, pp. 141–147.

[104] NGMN Alliance, "Description of Network Slicing Concept," *ngmn*, 2016.

[105] A. Tzanakaki, M. P. Anastasopoulos, and D. Simeonidou, "Converged optical, wireless, and data center network infrastructures for 5G services," *J. Opt. Commun. Netw.*, vol. 11, no. 2, pp. A111–A122, 2019.

[106] Y. Shoji, M. Ito, K. Nakauchi, L. Zhong, Y. Kitatsuji, and H. Yokota, "Bring your own network - A network management technique to mitigate the impact of signaling traffic on network resource utilization," in *2014 IEEE 11th Consumer Communications and Networking Conference, CCNC 2014*, 2014, pp. 182–187.

[107] NGMN, "Description of Network Slicing Concept by NGMN Alliance," 2016.

[108] F. Gabriel, G. T. Nguyen, R. S. Schmoll, J. A. Cabrera, M. Muehleisen, and F. H. P. Fitzek, "Practical deployment of network coding for real-time applications in 5G networks," *CCNC 2018 - 2018 15th IEEE Annu. Consum. Commun. Netw. Conf.*, vol. 2018-Janua, pp. 1–2, 2018.

[109] A. Gupta and R. K. Jha, "A Survey of 5G Network: Architecture and Emerging Technologies," *IEEE Access*. 2015.

[110] 5GPPP Architecture Group, "5G PPP Architecture Working Group: View on 5G Architecture," 2016.

[111] T. X. Do and Y. Kim, "Latency-aware placement for state management functions in service-based 5G mobile core network," *2018 IEEE 7th Int. Conf. Commun. Electron. ICCE 2018*, no. 978, pp. 102–106, 2018.

[112] 3GPP, "NR; NR and NG-RAN Overall Description; Stage 2," 2018.

[113] 3GPP SA2 Chairman Frank Mademann, "System architecture milestone of 5G Phase 1 is achieved," *3GPP SA2 Chairman*, 2017. .

[114] V. Kapsalis, C. Fidas, L. Hadellis, C. Karavasilis, M. Galetakis, and C. Katsenos, "A networking platform for real-time monitoring and rule-based control of transport fleets and transferred goods," in *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 2010.

[115] I. Cha, Y. Shah, A. Schmidt, A. Leicher, and M. Meyerstein, "Trust in M2M communication," *IEEE Veh. Technol. Mag.*, 2009.

[116] J. He and W. Song, “AppRAN: Application-oriented radio access network sharing in mobile networks,” in *IEEE International Conference on Communications*, 2015.

[117] J. Bartelt *et al.*, “5G transport network requirements for the next generation fronthaul interface,” *Eurasip J. Wirel. Commun. Netw.*, vol. 2017, no. 1, 2017.

[118] “Handbook of simulation: Principles, methodology, advances, applications, and practice,” *J. Manuf. Syst.*, 2008.

[119] S. P. Bendale and J. Rajesh Prasad, “Security Threats and Challenges in Future Mobile Wireless Networks,” *Proc. - 2018 IEEE Glob. Conf. Wirel. Comput. Networking, GCWCN 2018*, pp. 146–150, 2019.

[120] S.-H. Lee, T.-U. Park, S.-H. Cho, H.-K. Cho, and K.-C. Cho, “A Study on Simplified Test Bench Evaluation Model of Application Layer Protocols for Pre-5G Service,” *2019 Elev. Int. Conf. Ubiquitous Futur. Networks*, pp. 688–691, 2019.

[121] C. Song *et al.*, “Hierarchical edge cloud enabling network slicing for 5G optical fronthaul,” *J. Opt. Commun. Netw.*, vol. 11, no. 4, pp. B60–B70, 2019.

[122] T. Varum, A. Ramos, and J. N. Matos, “Planar microstrip series-fed array for 5G applications with beamforming capabilities,” *2018 IEEE MTT-S Int. Microw. Work. Ser. 5G Hardw. Syst. Technol. IMWS-5G 2018*, 2018.