1-1-2012

Application-based Network Selection Algorithm in Integrated LTE-WLAN Systems

Leila Reyhani
Ryerson University

Recommended Citation

Reyhani, Leila, "Application-based Network Selection Algorithm in Integrated LTE-WLAN Systems" (2012). *Theses and dissertations.* Paper 1665.

APPLICATION-BASED NETWORK SELECTION ALGORITHM IN INTEGRATED LTE-WLAN SYSTEMS

by

Leila Reyhani

B.Sc. in Computer Hardware Engineering

Ferdowsi University of Mashhad, Mashhad, Iran, 2008

A Thesis

Presented to the School of Graduate Studies at

Ryerson University

in partial fulfilment of the

requirements for the degree of

Master of Applied Science

in the Program of Computer Networks

Toronto, Ontario, Canada, September 2012

Author's Declaration

I hereby declare that I am the sole author of this thesis.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

Author's Signature: _____

I further authorize Ryerson University to reproduce this thesis by photocopying or other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

Author's Signature: _____

Abstract

Application-based Network Selection Algorithm in Integrated LTE-WLAN Systems

©Leila Reyhani, 2012

Master of Applied Science
Electrical and Computer Engineering
Ryerson University

This research focuses on an application-based network resource selection algorithm in integrated LTE-WLAN systems. First, we study the structure of the LTE/WLAN overlaid systems and then propose an approach to select the network on LTE and WLAN interfaces in a user equipment. In the study we will test network's behavior change with the change of number of new arrival nodes to our model with different quality of service (QoS) and type of service (ToS) settings when implemnting the network selection algorithm in OPNET. A part of this algorithm works with fuzzy logic controller block which calculates the threshold needed for comparing data usage with the remaining free data for uplink with a cost effective aim. The procedure of calculating the threshold is also explained. This network selection algorithm, gives a better result in terms of QoS for real time applications, less delay variation for cellular network for uplink data transmission and uses the pre allocated cellular data (total data amount of download plus upload) as much as it can, avoiding extra costs.

Acknowledgement

First I thank God for this opportunity in my life.

I would like to express my sincere gratitude to my supervisor, Prof. Alagan Anpalagan for his guidance, patience, motivation, strict deadlines and at the same time positive attitude which helped me to finish this research and also have a big picture of life, not only my thesis.

I want to thank Dr. Bobby Ma for his generous time to help me, for his guidance, ideas and also giving me the opportunity of teacher assistantships for the department. Whom I could never stay in Canada without his help.

I would like to acknowledge the Department of Electrical and Computer Engineering and the School of Graduate Studies at Ryerson University for their support in terms of financial aid, and work experience as a graduate assistant.

I would also like to thank my mom and dad for all their care and supports, especially my mom's prayers.

Thanks are also due to my colleagues in the WiNCORE lab for being nice, friendly, helpful and professional, whom are now my good friends.

Contents

List of Figures

List of Tables

List of Acronyms

$3G$	Third Generation
$4G$	Fourth Generation
CS	Circuit Switch
$E-UTRAN$	Evolved UMTS Terrestrial Radio Access Network
EPC	Evolved Packet Core
FDD	F requency Division Duplex
$GGSN$	Gateway GPRS Support Node
HSS	Home Subscriber Server
ITU	International Telecommunication Union
LTE	Long Term Evolution
$MANET$	Mobile Ad-hoc Network
MN	Mobile Node
$MGWs$	Media Gateways
MSC	Mobile Switching Centre
OSI	Open Systems Interconnection
PDF	Packet Delivery Fraction
PS	Packet Switch
QoE	Quality of Experience
QoS	Quality of Service
RNC	Radio Network Controller
SAE	System Architecture Evolution
$SGSN$	Serving GPRS Support Node

ToS	Type of Service
TDD	Time Division Duplex
UE	User Equipment
$UMTS$	Universal Mobile Telecommunication System
VPN	Virtual Private Network
$VANET$	Vehicular Ad-hoc Network
VPN	Virtual Private Network
WEP	Wired Equivalent Privacy

Chapter 1

Introduction

There has been a significant interest and progress in LTE networks over the last few years. Nowadays all networks are almost heterogeneous, it means that they have more than just one available radio resource for access. Therefore, the question that is raised, is to find a resource allocation algorithm to satisfy the customer QoS requirements. The motivation of our work in this thesis was the concept of short-term contractual users in future wireless networks which employ various applications. Let us say, me as a customer want to have a pre-paid plan which has the option of a good data usage, a reasonable QoS and no more additional charges. This means short term contracts which I still can run various applications like Oovoo, Skype, Viber, etc... with data usage on my equipment. From the user's point of view, the perceived benefits are analogous to application performance, which is dependant to QoS requirements for the transport of user data from the peer application end points and user quality of experience (QoE) perception. From an operator's perspective, these factors need to be translated to traffic flows with different QoS profiles, that can further be translated into different charges, depending on which access networks are employed for data transmission. Therefore, the aim in this project is to achieve a good performance level for communication, based on applications to maximize user salification. This goal needs to honor the QoS requirements for applications and traffic flows while minimizing the total

charges incurred. Obviously, for any given combination of application traffic flows and QoS level, user utility will decrease as the monetary cost of enjoying it decreases. Monetary cost may be dependant on arbitrary traffic flow properties and the lowest value of it probably will be achieved through a combined use of multiple wireless access networks such as LTE and WLAN. In this study, we will first study a heterogenous LTE/WLAN architecture, then discuss on a user-centric quality of service based algorithm which is cost effective and avoids extra data usage as already mentioned. This algorithm is working for a small WLAN network overlaid by a LTE network. In this network we also implemented IPv6 flow management to provide mobile data offload support for LTE networks allowing multimode terminals, to send certain flows to WiFi hotspots, if any available. The average delay for background traffic can be minimized by network selection, also average throughput of network can be maximized by selecting the network for the load on interfaces. We considered outgoing traffic in this case. The algorithm is user satisfaction based which has two jobs: first by maximizing the load and network selection helps the user not to go over the pre-paid amount of data usage to avoid extra costs, secondly, by providing a good quality of service to satisfy the user from both sides: cost and QoS. Thus, the main contribution of this study includes: i) A proposed framework for user salification based network resource allocation. ii) Implementing the network resource selection algorithm at intra-operator level based on IPv6 flow management. iii) And implementation of a QoS user satisfaction policy for uplink traffic of the mentioned network. To organize our work, next chapter discusses basic concepts of WiFi and cellular systems. As this project was first done on an integration of WLAN/UMTS network and then it was migrated to the integration of WLAN/LTE, we will discuss the basic concepts first and then a brief explanation of the proposed architecture of the new 4G integrated system. After that, there will be the problem description and the proposed fuzzy-based user interface selection and network-centric resource allocation algorithm solution in chapter 3. The performance of proposed solution is evaluated in chapter 4 which is followed by conclusion and future works.

Chapter 2

WiFi and LTE

In this chapter we will discuss the basic concepts of WiFi and cellular, UMTS and LTE networks, and then we study a proposed architecture for integrated UMTS/WLAN and LTE/WLAN network with diverse applications in mind.

2.1   WiFi

WiFi, also known as 802.11b, has become the preferred technology for wireless local area networks based on IEEE 802.11b standard. It was designed for private uses but nowadays it is being used for public places to create so-called hot spots, and those which Wi-Fi users can obtain broadband Internet access [1]. To have a better understanding of its concept there are some challenges in WiFi that should be considered: Ease of use, security,virtual private network (VPN), mobility. Its technology consists of service providers, the franchisor, the WiFi carrier and the aggregator. First, we will give a brief introduction to the concept of the WiFi and then discuss some of mentioned challenges [1]. To understand the concept of WiFi in this section, we will paint a big picture like a global landscape equipped with inexpensive WiFi hotspots offering easy and almost all the time free service for each user equipped with a WiFi-capable PDA or laptop. This big picture is a kind of an extension from office environment to wide-area. As an example, we can use WiFi in many cases and

to understand the WiFi growth, imagine a situation that somebody is working in her/his office, she/he is going to another place equipped with WiFi hotspot. When she/ he opens the computer and after authentication and being connected to the Internet, the computer automatically builds a secure tunnel through the Internet back to the office Intranet. In an airplane we can use the plane's internal WiFi connected via satellite back to the earth. If none of these connections are available she/he can use a cellular network interface card in the computer. So in the scenario there was no any inconvenience in disconnection from the Internet [1]. Figure 2.1 is a very simple figure that shows the mentioned scenario.
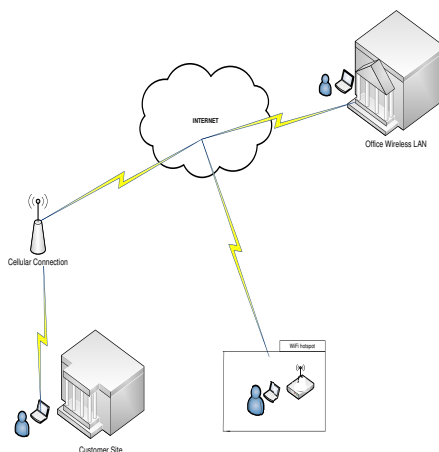
Figure 2.1: Non-stop connection.

Ease of use and access in WiFi are probably the most important challenges. One of the features made it much more easier to use is universal plug and play which makes it so easy to interconnect computing devices, in home, office and private environment. Security is another WiFi's challenge studied recently. As we know, one of the most important demands for WiFi was giving free public access. One important issue that can appear is security issues and existence of hackers. A very common scenario reported involves a hacker sitting somewhere near the WiFi hotspot and listening to the communications of the nearby personal communications. If the Wired Equivalent Privacy (WEP) is not enabled it is very easy to listen to the data transmitted wirelessly. Even if the WEP is enabled it is still not secure as

reported in [2], because the encryption key can be recovered by a hacker with a little effort. To have a better security, there is a VPN suggested on top of the unsecured WiFi network.

2.2   Introduction to LTE

Here, we will discuss a historical context for LTE systems and migration from 3G to 4G systems. First, we will give a brief information about UMTS and GSM systems and then our migration from a UMTS model to a LTE system will be described plus the issus for this evolution. UMTS first evolved into LTE and then into an enhanced version known as LTE-Advanced [3].

2.2.1   Architectural Review of UMTS and GSM

Global System for Mobile communications (GSM) has been one of the the largest technologies ever from the second-generation telecommunication systems. These generations enabled voice traffic to go wireless through the networks between nodes [4]. Because of the challenges of multimedia handing on-the-move, second generation systems were soon proven to be limited. So the mobile operators moved to a 3G system. Universal Mobile Telecommunication Systems, UMTS, are designed to globalize worldwide process of mobile communications and giving a basis for new services with high rate and high quality of data. After 3G was introduced, next step is to create a generation that can support up to 1 Gbps of data rate, combination of several available, evolving and emerging access technologies into a common platform. This high data rate of 4G systems is not feasible with the conventional 2G and 3G architecture. The reasons are first the transmission data rate for 4G systems is as high as four magnitude orders than what is in 3G systems. On the other hand, the spectrum that is being released for 4G systems was almost above 2-GHz band used by the 3G systems [4]. UMTS is a 3G wireless system that is part of the International Telecommunications Union's IMT-2000 vision of a global family of 3G mobile communications systems. UMTS was ex-

pected to deliver low-cost, high-capacity mobile communications, offering data rates up to 2 Mbps. The 3G cellular networks are designed to provide voice and data for mobile users. The sustainable per user data rate is hundreds of kbps limited by the total cell capacity of up to 2-3 Mbps [4]. Collaboration of national and regional telecommunication standard bodies designed LTE, in a process of a project named 3GPP, third Generation Partnership Project. An earlier version of 3GPP was a system named UMTS which formed LTE afterward. UMTS itself was an evolution from the GSM systems.

A mobile phone network is officially known as a public land mobile network (PLMN), a PLMN is run by a network operator such as Verizon or Vodafon [3]. GSM and UMTS have common network architecture, as we know that UMTS is an evolved version of GSM. There are 3 main components: a core network, a radio network and a user equipment like a mobile phone [3]. Each core network contains two domains, one is a circuit switched (CS) domain network which is responsible for phone calls at the covered area network. Our circuit switched network communicates with the public switched telephone network, PSTN, to let the users call beyond the covered area through other CS networks for other areas. Except the CS network for phone calls, there is another network which is packet switched and is being used for data transmission, such as email traffic and web page access. These data streams are between the user and external packet data networks, PDNs. An example for a PDN can be Internet [3].

For a better understanding of the concept, from the Figure 2.2 it can be seen, CS domain uses circuit switching method to transport the information. We can also see that there are two ways for transportation, in each way it dedicates a two-way connection for each of the individual phone call connection to provide a constant data rate and minimum delay for phone calls. This approach is inefficient, although the connection has enough capacity to handle the worst-case scenario while both users are speaking simultaneously. Usually this technique is being over dimensioned so it is not appropriate for data transfers as data rate can vary widely. In return, packet switched (PS) domain has another approach that a

data stream is divided into packets. Each packet has a different label with the destination address. Each router reads the address labeled for each packet and then forwards them to their destinations. As the network resources are shared between all users, this technique is more efficient in comparison with CS. On the other hand, if there are too many devices when trying to send data at the same time, delay may result.
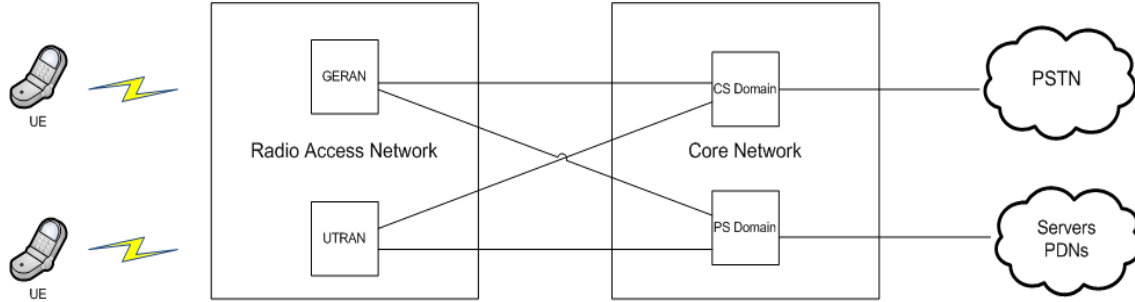

Figure 2.2: UMTS and GSM high level Architecture

The radio access network's job is to handle the radio communications between user and core network. In Figure 2.2, we can see there are actually two separate radio access networks, one is GSM EDGE Radio Access Network, GERAN, and the other one is UMTS Terrestrial Radio Access Network, UTRAN. These two different radio communication techniques are sharing a common core network with each other.

Each UE shows the user devices in the mentioned figure, which communicates with the radio network through the air interface. Also there is an important concept which we used in this thesis: upload link (Uplink) or reverse link. UL means the direction from mobile to network and in return there is downlink or forward link [3] which is the direction from network to mobile node. Our work is on uplink interface. A roaming mode is a situation when a mobile node can work outside of the coverage area of its own network operator. It is happening when the node is using other available public resources from two land mobile networks: the visited network which the mobile node is in it now and operator's home network.

2.2.2 Architectural Review of Core Network

In Figure 2.3, there is a simple internal architecture of the core network. In the CS domains, media gateways (MGWs) are responsible to route phone calls and, mobile switching center (MSC) servers are handling the signaling messages for setting up, managing and tearing down the phone calls. In packet switched network, UMTS, a GGSN, gateway GPRS support node, is acting as an interface to the server and packet data network in the outside world. Also a serving GPRS support node, SGSN, is routing the data between the base station and GGSN for signaling messages. In Figure 2.3, Home Subscriber Server (HSS) is a central database that contains all the information about the network operator's subscribers and is shared between two parts of the network domains [3].
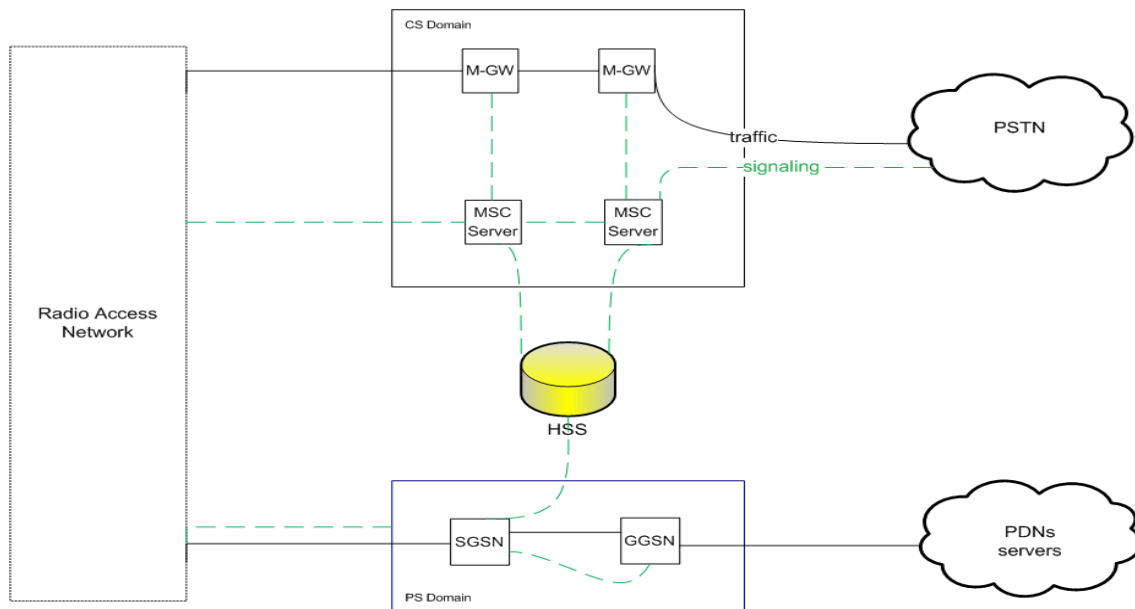

Figure 2.3: Core Network Architecture

2.2.3 Architectural Review of Radio Network

In this part, first we will review the radio network of UMTS and then we will make a comparison between LTE and UMTS systems. In Figure 2.4, the most important part of UMTS radio network is called NodeB which is the base station of the radio network that has

one or more sets of antennas. Each station is using the antenna to communicate with the UEs in each sector. A sector is being controlled by an antenna, a typical base station uses three antennas to control three sectors. In this study, the word "cell" is the same as sector. Each cell has a limited size which is defined by the maximum range at which the receiver can successfully hear the transmitter [3]. The maximum combined data rate of all the mobiles in the cell shows the limited capacity of each cell. These limitations are the difference between different types of cells. Macrocells in this concept are making wide areas with coverage of a few kilometers. Microcells have the coverage area of a few hundred meters with a greater capacity [3]. Picocells are large indoor environments with a few ten meters coverage like offices or shopping centers. And finally femtocells provide a few meters coverage area that people can buy the base stations for their homes. Each base station and user equipment transmits on a certain radio frequency named carrier frequency. Bandwidth is an occupation of a certain amount of frequency spectrum around each carrier. To understand the concept better let us assume that if a mobile is transmitting with a carrier frequency of 1000MHz and a bandwidth of 10MHz, it means that the frequency range is 995-1005MHz.

To segregate the mobiles' and base stations' transmissions from each other not to have any interference, UMTS and also LTE systems have two access methods: FDD and TDD. For frequency division duplex, the base station transmits on one carrier frequency and for mobile on another carrier frequency. On the other hand, for time division duplex, base station and mobiles transmit at the same carrier frequency but at different times. When a node is moving from one cell to another, there are two techniques to handle this situation: one is handover the other one is cell reselection. Soft handover happens when in UMTS, a node is communicating with more than one cell. In Figure 2.4 we can see that RNC is a part of network where all the base stations are being grouped together. Radio network controller is passing the voice information between radio and core network. Also it controls invisible signalling messages of mobile's radio communications. Each RNC supports a few hundreds of base stations, and a typical network has about a few tens of RNCs.
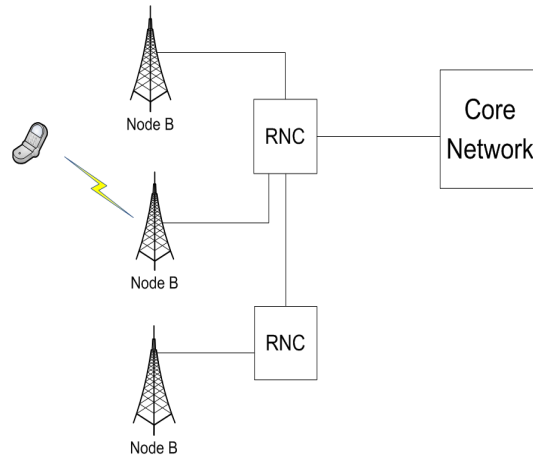
Figure 2.4: Radio Network Architecture

2.3 Migration from UMTS to LTE

The very first reason for this migration, coming to mind is the growth of mobile data. From 2010, the growth of data started to be more and more, which made the need of a system with higher capacity was mandatory. Also from 2007, Apple iPhone applications came out and from 2008 Android's, so the need of a system with higher data capacity was inevitable [3]. Also network operators encouraged the growth of mobile data by a flat rate-charging that permitted unlimited data downloads. To make more capacity for the system, there are three main approaches developed, the most important can be the use of smaller cells. The channel capacity is the maximum data rate a single cell can handle. If we have more base stations and reduced cell sizes, the capacity of network will be increased.

Second way which is kind of limited, is increasing the bandwidth. Radio spectrum is managed by the regional and national regulators. As the amount of radio spectrum is limited and there are a variety of applications such as military communication and radio astronomy that are using them, this approach is limited as mentioned before. Third approach is to improve the whole communication system and technology which is the main reason of creation of LTE [3].

In Figure 2.5 we can get the idea of immigration from UMTS to LTE and the components

replaced. In 2004, 3GPP began the study of the long term evolution of UMTS. As a result, 3GPP's communication systems were kept competitive over timescales of 10 years and beyond. From Figure 2.5, we can see in the new architecture for LTE, we have a component named EPC which is a direct replacement for the packet switched domain of UMTS and GSM. It is responsible to distribute all type of information to the user, voice and data, it also uses the same traditional packet switching technologies for data alone. For the circuit switched domain, there is no equivalent like this, but instead all voice calls are being transported as voice over IP (VoIP). The radio communications of EPC and mobile are handled by E-UTRAN. This is a direct replacement of UTRAN. This new architecture was designed as part of two 3GPP work items, one is system architecture evolution, system architecture evolution (SAE) to cover the core network, and the other one is LTE, long term evolution to cover the radio access network.
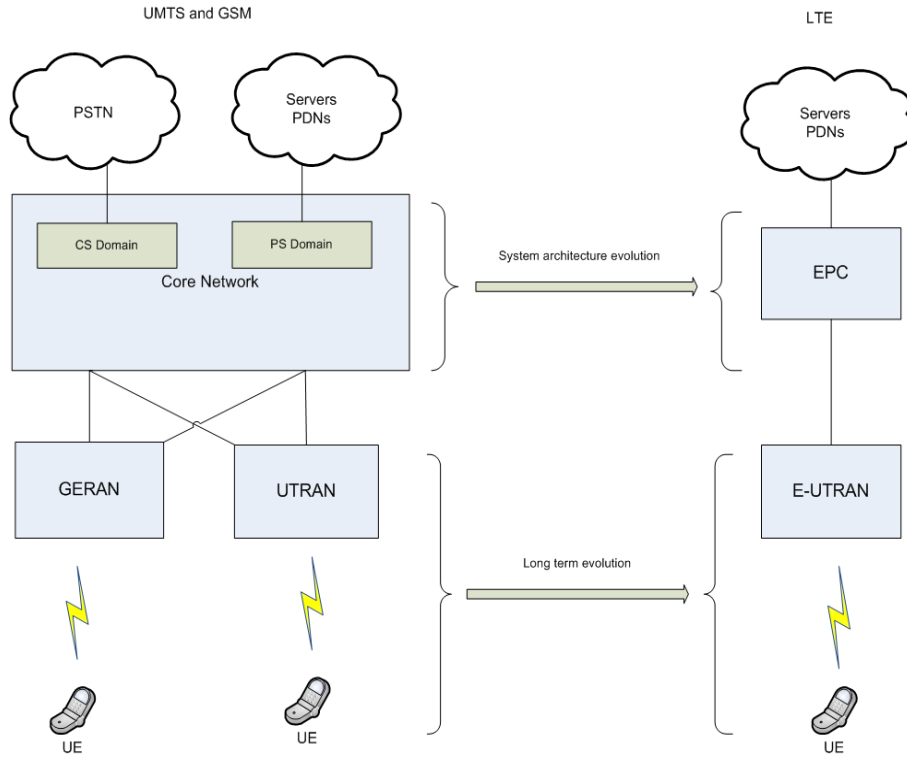

Figure 2.5: UMTS to LTE Migration

We exactly followed the same migration path in this project. Before applying any policy

with the algorithm, we needed a framework to simulate an integrated WLAN/cellular system. First model we simulated was the integrated WLAN/UMTS system based on a tight coupling architecture which will be explained in the next chapter, and then we migrated from UMTS to an LTE network. From Figures 4.4 and 4.3 we can see the explained concept and used components through this migration.

We can see the RNC, SGSN and GGSN all are gathered in one component named EPC. Smaller circles are showing the area covered by the access point which is connected to the cellular core network through the MANET gateway. This gateway is connected to EPC in LTE case and in the UMTS network it is connected to Internet through the GGSN. It is called tight coupling when the WLAN is connected to the Internet through the wireless core network. In the next chapter we will discuss the design of the network, problem description and fuzzy logic solution.

Chapter 3

Proposed Solution and Methodology

In this chapter we will give a brief description of the problem and the proposed solution. In the next chapter we will discuss the details of how to set up the simulated network parameters, how to model this network from the scratch and we will provide an appendix file of how to link 3rd party program codes in C++ and MATLAB to OPNET simulated files. In this study, we used OPNET modeler 17.1 to simulate a small LTE/WLAN heterogenous network. Our goal in this study is to test an application-based network selection algorithm run on a system in an integrated WLAN/LTE small network. We want to provide data flow for LTE mobile nodes and at the same time if the node switched to a WLAN hot spot, switching to WLAN should be available. Also we are running 4 applications, VoIP and video as conversational applications, web-browsing as interactive and email as a background traffic. At the end, we will separate the load for LTE and WLAN when the WLAN network is available and study network's behavior before and after running the algorithm. IPV6 is available for the nodes, so the mobile IP will be set for all our mobile nodes.

3.1   Design Overview

As mentioned before, nowadays wireless networks are becoming to be heterogeneous, as different networks with different Radio Access Technologies (RAT) are working together

cooperatively [5]. First goal of the project is to integrate a wireless LAN network into a cellular system. We assume our cellular network is covering a wide area. As in the background review in chapter 2, we will consider the WLAN network as a microcell to be limited in space and overlaid by the cellular network as a macro cell. The small WLAN network is a sample of our mobile ad hoc network. In this scenario, a mobile user is connected to WLAN to download the data and on the other hand through LTE network to obtain access to voice network. We first simulated the mentioned network for an integrated WLAN/UMTS small network and then migrated from UMTS to LTE. There will be more discussions about the UMTS and LTE, migration, and finally, results in upcoming sections and chapters.

First, we will give a brief information about our UMTS/WLAN network. This design needs an architecture that can support to work in parallel. For integrating UMTS and WLAN, there are two points of integration: one possible case is WLAN as a standalone network, which means it is not connected to UMTS network. A user may subscribe to both UMTS and WLAN services through the same service provider. This scenario is called roaming mode in [6]. The other case that has been implemented in real world was the "Hot-Spot Scenario", which occurs when an operator is offering a high-speed data connection in some hot-spot areas, for example in the school building. In this case, WLAN is used for data connection only and terminals have two interfaces, one is to connect to cellular network and the other one for WLAN. As WLANs are in minority and they are not everywhere, they form micro cells, and UMTS networks that are covering wider areas form macro cells. Figure 3.1 shows the concept.

An overview and details of the design and simulation setup will be discussed in chapter 4.
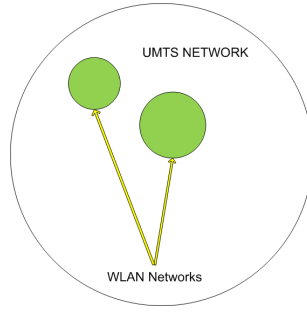
Figure 3.1: Macro and Micro cells

3.2 Problem Description

Our goal on such an integrated system for WLAN and UMTS network is to develop an application-based network selection system which can give a reasonable QoS to the user and is cost effective in terms of cellular data usage of the network and then we will show how to migrate from this system to an LTE/WLAN integrated system. We considered it from the customer point of view. With the following assumptions, we propose an algorithm to solve bottleneck created by too much traffic on just one interface.

Let us assume that the user has an iPad which has both interfaces for 3G/4G and WLAN. Also the data plan for iPad is limited to 5GB per month. User is at school and can use both services, just to show what kind of problems the user may face, imagine she/he is through a job interview with skype and suddenly the connection with WLAN loses or the speed is slowed down that can not support that much transmission of data. As we know applications such as video conferencing or voice over IP are categorized in conversational group, their type is real-time. Therefore, the quality of service (QoS) is a very important concept which has to be taken care of, as the customer needs low delay and jitter and, clear voice and video. This approach is bringing the lowest cost and the best QoS for the customer. QoS is a very important feature which can be combined with the network selection and resource allocation concept to reach the best result.

To have a better understanding, this approach is a key technique in heterogenous systems

to maximize the usage of radio resources and make the whole system efficient [5]. Root of the concept is from CRRM, Common Radio Resource Management [7]. Network selection algorithm has been used to solve problems just for only a single factor, some only on one application.

In this study, we use the algorithm for two different interfaces and different types of applications run on our system as already mentioned. It makes the algorithm more complicated. If we want to talk about it generally, a good network selection algorithm considers many things, and should be able to exchange the context information. This algorithm can be considered as a controlling tool. Because the network is heterogeneous, it consists of UMTS or LTE and WLAN, except our radio resources, some important parameters, such as QoS, UE speed of movement and cost of usage of resources have to be taken into account. If we want to consider everything, complexity of the problem will make it impossible to implement and solve. In this study, we will consider the QoS based on the radio resource usage which can make the performance of the network better. Imagine a situation where a user equipment, using a WLAN resource, is inside a hotspot which is overloaded, so the probability of blocking is getting higher and QoS is under the impact of it. Therefore, to avoid congestion, the UE has to switch to the other radio resource, UMTS or LTE, if it is available. Therefore, the very goal of this study is to use of the pre-paid data as much as we can, to avoid congestion, and provide a better QoS for applications with the help of network selection policies. In the next section, we will discuss how to develop an algorithm to solve the problem.

3.3   Suggested Solution and Algorithm Proposal

In the proposed algorithm, we used QoS approaches from OPNET modeler's library plus a fuzzy logic to estimate a threshold for choosing the appropriate interface, which will be explained in detail. This algorithm is working on network/IP layer as the separation of

interfaces is happening in IP layer in each node in OPNET modeler process block. We can see that in Figure 3.2.

From ip_encap block to ip block, in the Figure 3.2, if the incoming traffic to the interfaces is cellular voice, it straightly goes to cellular interface. In Figure 3.3 we can see that if the traffic is data, there is an approach to choose the proper interface to avoid congestion and achieve the best usage for the data in a cost effective approach. This approach first examines the data. In this study, the data has been categorized on real and non-real time. Real-time data is the one which needs interactive type of service, and as mentioned in the previous section, non real-time applications get the best effort ToS from the network.

In the flow chart in Figure 3.3, we assumed that there is 5GB of pre-paid data from cellular network per month. Also algorithm's goal is to reach the most usage of pre-paid data, avoiding extra charges for user. Therefore, there is a concept like threshold needed not go over it and at the same time reach a reasonable QoS which can satisfy the user. After recognizing the data as real/non-real time (categorizing based on ToS) next decision is based on how much data remains from the data quota of the day. This data usage is based on the estimation that the algorithm calculates. Calculation of data usage will be explained in the next section. The idea of data usage prediction comes from [8]. We consider applications VoIP and video conferencing as real-time and FTP, email and web browsing as non-real time. Although web browsing is a kind of interactive traffic, the delay is not that important and can satisfy the user. Both real time and non-real time sides follow the same procedure; first, the prediction of the application usage. The amount of data usage is changed during the execution of the simulation. As the applications running inside the network are sending more and more data, the available pre-defined quota is decreasing.

Decision making of using which interface is based on this amount of remaining data. There is a threshold defined for each real and non-real side which the estimated amount of data is compared to it. This threshold is calculated with a fuzzy algorithm and there are two functions to calculate the threshold for each side, because the amount of data transmitted

17

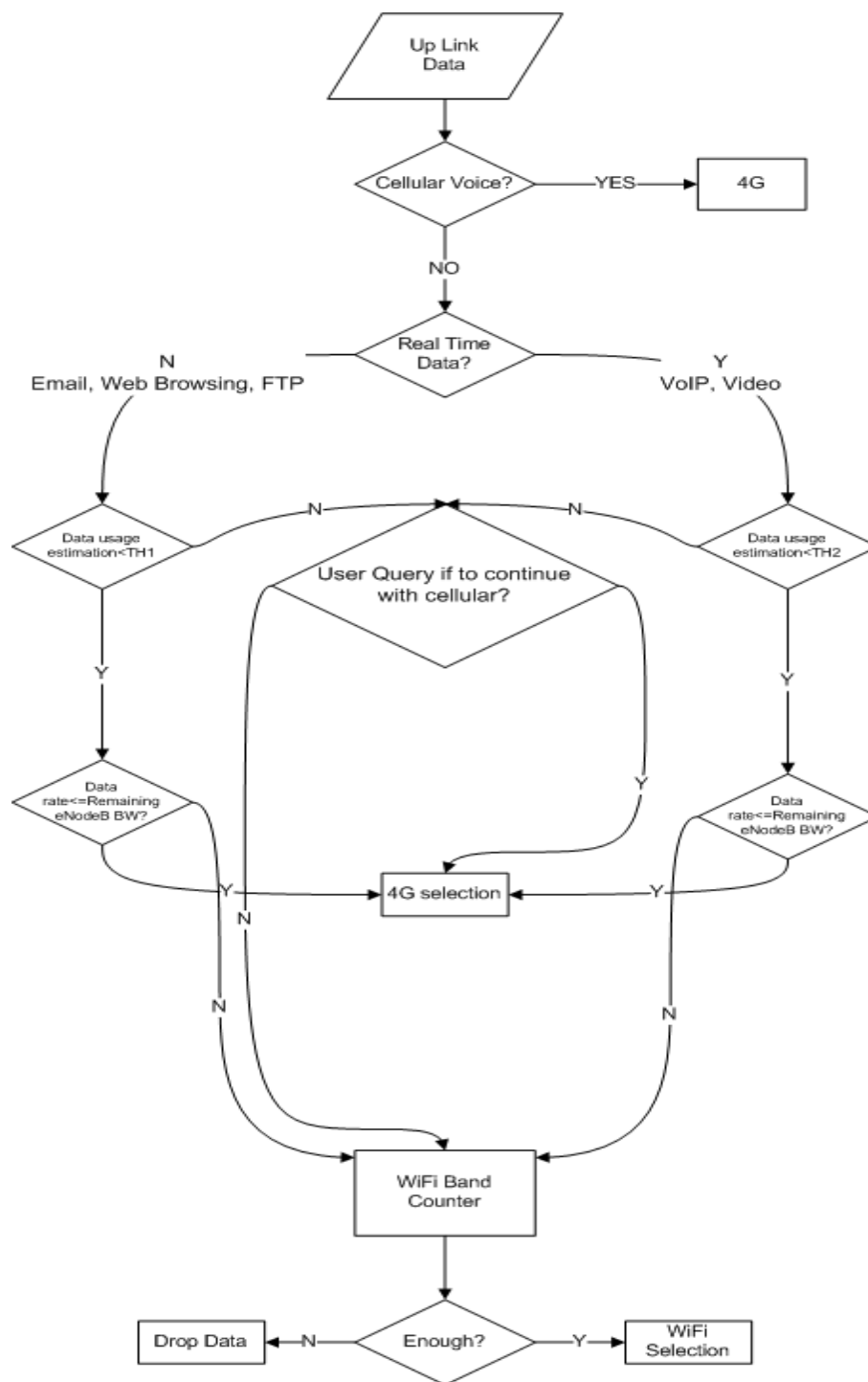Figure 3.2: Node Editor of a LTE Node Model in OPNET

Figure 3.3: Network Selection Algorithm Flowchart

for non-real and real time applications are different. If the mentioned estimated data usage was over the threshold, it sends the data to the WiFi side. Otherwise, it goes to the next step and considers the amount of data rate for the application data transmission and compares it with the remaining bandwidth of the transmitter. If it has enough bandwidth, then data is accepted on the 3G/4G side, otherwise WiFi will take care of it. Also if it was over the threshold it will ask user if she/he wants to continue with cellular network, if yes there will be no change of interface, if not it sends the data to WiFi side. In the next section, we will have a quick overview of fuzzy logic and will explain how we can fit this logic to calculate the threshold.

3.3.1 An introduction to fuzzy solutions

As we already discussed, fuzzy logic was the approach to measure the amount of the mentioned threshold. In this section, we will give a brief introduction to fuzzy concept and then show how to solve a part of our problem with this logic. The threshold part of the flow chart in Figure 3.3, where the estimated data is compared with, is calculated with fuzzy logic. The concept is that the remaining data quota is the data quota of the day minus the already used data is the current amount of data we have. If the current remaining amount of data is less than 25% of the day quota, it is considered to be small. If it is between 20% to 80% is it considered to be medium size and if it is over 75% of the day quota, it is large. It will be explained completely in two next sections. As an introduction, we can say fuzzy system was first proposed by professor Lotfi A. Zadeh, in 1965 when he presented his seminal paper on "fuzzy sets" [9]. Dr. Lotfi Zadeh showed that fuzzy logic unlike classical logic can realize values between false (0) and true (1). Basically, he transformed the crisp set into the continuous set. Fuzzy sets thus have movable boundaries, i.e., the elements of such sets not only represent true or false values but also represent the degree of truth or degree of falseness for each input. Fuzzy logic is the part of artificial intelligence or machine learning which interprets human actions. Computers can interpret only true or false values but a

human being can reason the degree of truth or degree of falseness. Fuzzy models interpret the human actions and are also called intelligent systems. The Japanese were successful in applying fuzzy logic in commercial goods such as automobiles and washing machines and their current fuzzy logic market is estimated around 2 billion dollars. Fuzzy logic has mostly been applied to control systems. Fuzzy control systems interpret the expert human and replace them for performing certain tasks such as control of a power plant. Fuzzy controllers apply decision rules (if-then rules) by making use of critical variables to interpolate the output between the crisp boundaries. Some typical examples where fuzzy logic has been implemented are railway (Sendai Railways in Japan), automobile industries (transmission and braking), heating and cooling systems, copy machines, washing machines.

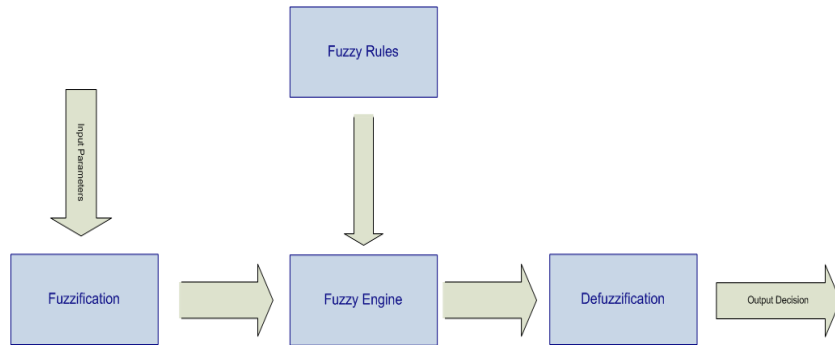Process of this logic's approach for problem solving can be found in Figure3.4

Figure 3.4: Fuzzy Function Block

In the following sections we will have a brief overview on fuzzification, fuzzy rules, and defuzzification techniques, and the way we implemented them in our case.

3.3.2 Fuzzy Logic Applied to Thresholding

Fuzzification is the process of changing a real scalar value into a fuzzy value. This is achieved with the different types of fuzzifiers. There are generally three types of fuzzifiers, which are used for the fuzzification process; they are Singleton fuzzifier, Gaussian fuzzifier, and Trapezoidal or Triangular fuzzifier. In our work, we used triangular Mamdani approach [10].

3.3.3　Fuzzy Rules

Fuzzy systems are built to replace the human expert with a machine using the logic a human would use to perform the tasks. Suppose we ask someone how hot it is today. He may tell us that it is hot, moderately hot or cold. He cannot tell us the exact temperature. Unlike classical logic which can only interpret the crisp set such as hot or cold, fuzzy logic has the capability to interpret the natural language. Thus, fuzzy logic can make human-like interpretations and is a very useful tool in artificial intelligence, machine learning and automation. Fuzzy logic operates on the basis of rules which are expressed in the form of "If-Then" constructs, also know as horn clauses. A fuzzy proposition can be an atomic or compound sentence. For example: "Remaining quota is small" is an atomic fuzzy proposition. "Remaining quota is small and degree of data acceptance is low" is a compound fuzzy proposition. Compound fuzzy relations are expressed with fuzzy connectives such as "and", or "and complement".

Mamdani proposed a fuzzy implication rule for fuzzy control in 1977. It is a simplified version of Zadeh implication operator. The Mamdani fuzzy logic operator is given as:

$$\varphi\left[\mu_A\left(x\right),\mu_B\left(y\right)\right] \equiv \mu_A\left(x\right) \wedge \mu_B\left(y\right) \ , \tag{3.1}$$

where μ_A and μ_B are fuzzy proposition. In our case μ_A is maximum amount of remaining data, and μ_B is degree of acceptance, and φ is minimum of these two propositions which apply the rules.

Here, the propositions are "Remaining quota is large or remaining quota is medium" and "Data acceptance". Then calculate the threshold. The linguistic variable is "remaining quota (RQ)" and rule is "acceptance". The linguistic values for RQ are large, medium and small. The membership function for RQ in the universe of discourse, μ, is given below.

for small remaining quota, which means between 0 and 10% of the whole data is too small, and from 10% to 25% it decreases linearly and after 25% it is not small anymore.

Therefore, in our case:

$$\mu_{Small}(x) = \begin{cases} 1, & if & x < 10 \\ -\left((x-10)/15\right), & if & 10 < x < 25 \\ 0, & if & x > 25 \end{cases} \tag{3.2}$$

for medium between 20% and 80% it's medium:

$$\mu_{medium}(x) = \begin{cases} (x-20)/30, & if & 20 < x < 50 \\ -\left((x-50)/30\right), & if & 50 < x < 80 \\ 0, & else \end{cases} \tag{3.3}$$

for large from 75% and 90% it's moderately large and from 90 to 100 % it is too large:

$$\mu_{large}(x) = \begin{cases} 0, & if & x < 75 \\ (x-75)/15, & if & 75 < x < 90 \\ 1, & if & x > 90 \end{cases} \tag{3.4}$$

Also, acceptance and non acceptance rule formulation are as follows: These two rules have to be exactly defined opposite of each other, although we will use only the accept rule.

$$\mu_{notAccept}(x) = \begin{cases} 0, & if & x > 75 \\ -\left((x-20)/55\right), & if & 20 < x < 75 \\ 1, & if & x < 20 \end{cases} \tag{3.5}$$

$$\mu_{Accept}(x) = \begin{cases} 0, & if & x < 20 \\ (x-20)/55, & if & 20 < x < 75 \\ 1, & if & x > 75 \end{cases} \tag{3.6}$$

The plots for the membership functions of remaining quota (fuzzification of the remaining quota) and the rule, acceptance and not acceptance of data are given in Figures 3.5 and 3.6
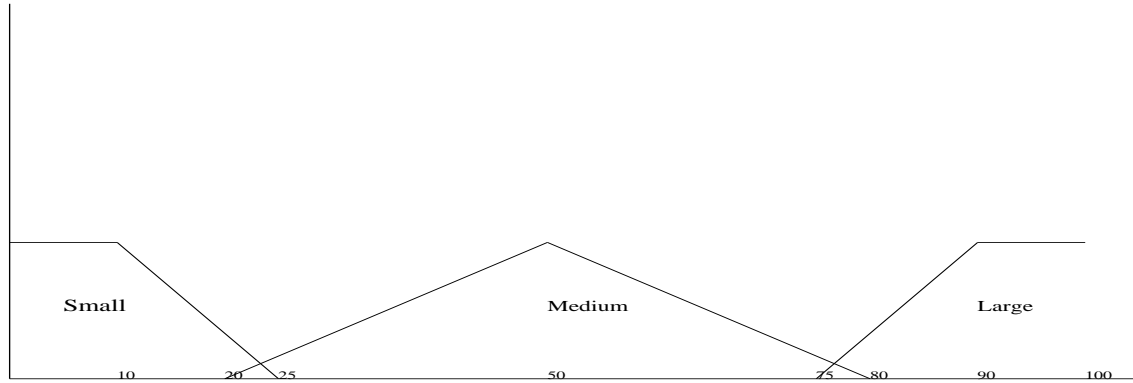
respectively.

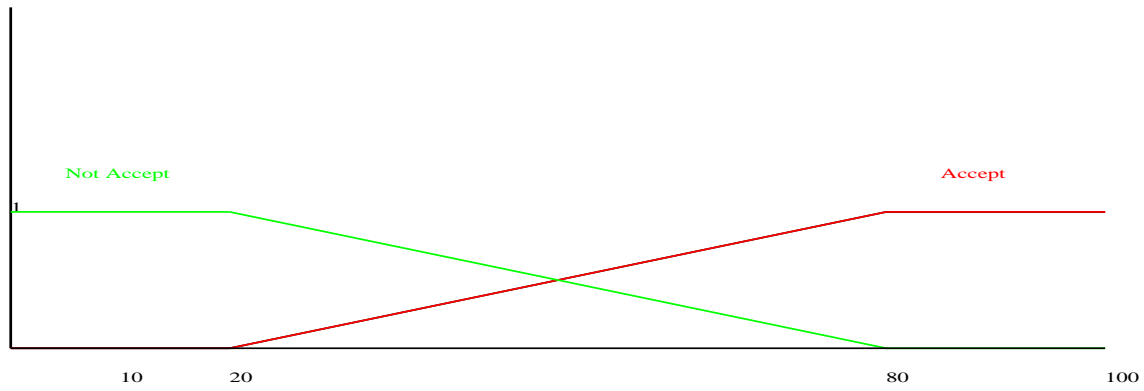Figure 3.5: Fuzzification Membership

Figure 3.6: Fuzzification Rules which we used them as Input

To apply the Mamdani implication rule for the proposed problem, the following rule is applied: The "or" connective is replaced with the max (union) operator. The maximum of the two membership functions is evaluated for the antecedent part of the fuzzy rules. The Mamdani implication operator (i.e., min operator) is applied between the resulting antecedent membership function and the consequent membership function. Therefore, if the data remaining is small and rule is not acceptance, the threshold is small, in which case the estimated data will be definitely over it. If it is medium and the acceptance factor is rising linearly, the function should make a medium threshold which may lead to choosing 3G/4G interface based on the amount of estimation. And if the remaining data is large it should

24

be accepted, so the amount of threshold is very large. As it is supposed to, as much as the remaining quota is decreasing, the threshold amount decreases too.

3.3.4 Defuzzification Technique

Fuzzy logic is a rule-based system written in the form of horn clauses (i.e., if-then rules). These rules are stored in the knowledge base of the system. The input to the fuzzy system is a scalar value that is fuzzified. The set of rules is applied to the fuzzified input. The output of each rule is fuzzy. These fuzzy outputs need to be converted into a scalar output quantity so that the nature of the action to be performed can be determined by the system. The process of converting the fuzzy output is called defuzzification. Before an output is defuzzified, all the fuzzy outputs of the system are aggregated with an union operator. The union is the max of the set of given membership functions and can be expressed as the following equation.

$$\mu_A = \bigcup_i \left(\mu_i(x) \right), \tag{3.7}$$

There are many defuzzification techniques, to name some, we can refer to maximum defuzzification, centroid technique and weighted average defuzzification technique. Between all, centroid approach is the most popular one which we used in this work and we will give a brief explanation on it. This centroid defuzzification method is also known as center of gravity or center of area defuzzification. This technique was developed by Sugeno in 1985 [10]. This is the most commonly used technique and is very accurate. The centroid defuzzification technique can be expressed as:

$$x\star = \frac{\int \mu_i(x) x d_x}{\int \mu_i(x) d_x}. \tag{3.8}$$

Therefore, from 3.8, $x\star$ is the defuzzified output. The threshold we want to reach, $\mu_i(x)$, is the aggregated membership function and $x\star$ is the output variable. The only disadvantage

of this method is that it is computationally difficult for complex membership functions [10]. In Figure 3.7 we can see how defuzzification is calculating the mentioned threshold. For a better understanding, Figures 3.7 and 3.8 show the threshold calculation results. As we can see in Figure 3.8, when the remaining data is small and the acceptance rule is applied, which is "not accept" for this amount, result of the threshold is approximately between 0% and 20%.

Rule View results in Figure 3.7 shows how Mamdani's approach calculates the threshold in this case. For each amount of remaining data when it applies the rules, less than 10% threshold is 0, for a medium amont of data remaining threshold is arounf 65% and for large it is around 85%, graph shows the defuzzified output from Equation 3.8
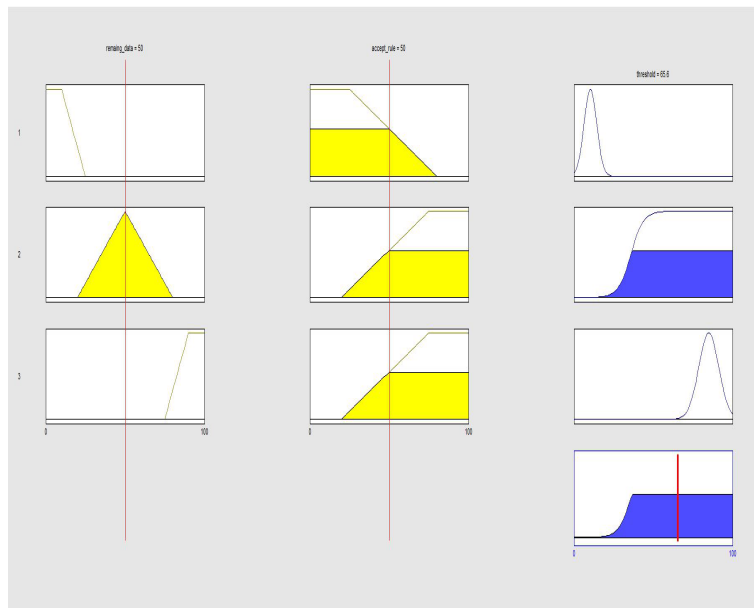

Figure 3.7: Rule Result View

Figure 3.8 is the surface view, we can see when the remaining data is medium and large (over 20% of daily quota), the data should be accepted, as the result, calculated threshold is over 40% to 50% of the remaining quota.
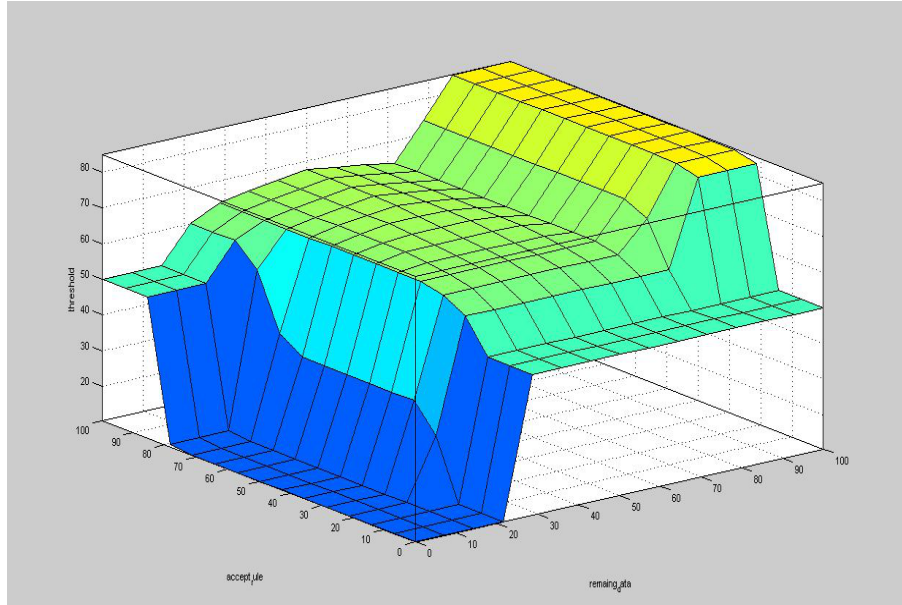

Figure 3.8: Threshold Calculation Result Surface Mode

Chapter 4

Implementation and Simulation Results

In this chapter, we review the implementation of the project, our simulation, related works and finally a comparison between our results and related works. First we will give an overview of related works on UMTS/WLAN and LTE/WLAN simulations. Then there will be a discussion on our work and the differences between it and related works, and a good indicator that our work performs well.

4.1   Related Work

In this section, we will discuss what has been done before, their results and and a brief explanation and analysis of the results.

Karl and Christer [11] proposed and evaluated a new access method for network selection in a combined LTE/WLAN network. Their proposed approach takes all available UE parameters, network resource in use and available candidate chosen into the account. Also their algorithm performs in an optimized manner for network selection. Their model is based on a well-known approximate solution to bin packaging problem with OPNET 16.1 simulator.

Probably our work is the closest to what Sun, Wu and Sha did in [5]. Their work was on

load balancing on a 3G/WLAN integrated network. Their approach is based on an algorithm for radio resource allocation to get a better QoS. One part was done with fuzzy logic. Our study is different from their work in many ways. To name a few, their approach was to separate the voice and data at first, but we classified our traffic based on the ToS. Secondly, we always keep 10-15 percentage of the data quota for the 4G voice because the telephone calls are the first priority on a cellphone device, but in their algorithm it first uses the 3G quota. Also our approach considers one threshold for each ToS. Here we can see that there is one fuzzy controller before the network selection procedure to calculate the probability of 3G or WLAN selection with fuzzy approach. We also calculated the threshold with fuzzy logic. Our approach can work better as when UE exceeds using the remaining quota instead of dropping data, it asks the user to continue with 4G or switch to WLAN. If user does not care about extra fees and cares about handover of switching between cellular and WLAN, she/he would select the 4G and stays with it, if not it will be switched to WLAN.

In [12], Yousun and Aesoon studied vertical handover which enables users to move around various wireless networks continuing with their on-going services. This matter requires fast and efficient handover control. They developed an open platform concept to support various wireless technologies. Their proposed vertical handover platform was verified for their 3G LTE test-bed. This platform recognizes the individual network such as WLAN and LTE, and is dynamically used in multimode reconstructing devices or resources of the terminal according to the standard of network selection. The proposed platform in that paper provides the transparency about the multiple network interface for upper service elements of mobile terminals and guarantees the reuse of the terminal protocol and resource element.

In [13], authors proposed a dynamic QoS negotiation scheme that allows users dynamically negotiate the service levels required for the traffic and reach them through one or more interfaces. This way for bandwidth aggregation leads to transmission of data belonging to a single application via multiple paths with different characteristics. This may result in an out-of-order delivery of data packets to the receiver and cause additional delay. The

proposed QoS negotiation system in that paper considers movement between different APs and fairness of network resource usage as already mentioned. The results show that their approach is better in minimizing packet loss and delay, and also enhancing the scalability of the system.

Umar et al. studied a user-centric quality of experience (QoE) based resource allocation with different ToS, several operators and users in [14]. They proposed a framework for future short-term user operator contractual vision with a game-theoretic approach. They also proposed a formulation for the utility function for users and operators. They studied their approach with OPNET for different types of traffics such as FTP, VoIP and video conferencing over LTE and WLAN with the implementation of IPV6 as we did in this study.

Bazzi in [15] focused on the gain in terms of capacity that deployment of WLAN hot spots provide in a fully covered area by a UMTS network. The focus was on the possible advantage without advanced centralized multi-radio resource management in an urban scenario, with attention to the voice service. The results show that even with consideration of a deployment covering less than 3 percentage of the total area, it's still possible to get an acceptable quality with the increase of number of users. Also it was highlighted that the results require smart WLAN call admission control and, only if the terminals are equipped with multimode interfaces they take advantage of any available WLAN.

In [16] authors proposed a design for integrating UMTS and ad hoc networks, which the UMTS part of our study was inpired by. Their proposed design was based on the inter-system handover modeling as the inter-SGSN (Serving GPRS Support Node) handover and the connecting the gateway in the ad hoc network with the GGSN (Gateway GPRS Support Node) in the UMTS core network. The main reason to integrate these two networks was to provide mobile nodes with high speed connectivity yet with small infrastructure support. Their results show that handover time is managed in 80-120ms. For the MANET part of their project, it resulted in significant improvement for total throughput and low packet loss when the load of the network is heavy.

In [17], authors envision heterogeneous wireless network architecture by integrating IEEE 802.11p VANETs with 3GPP LTE for a seamless data connectivity. Issues on cluster-head based multicasting and QoS are explored in their study. A proposed approach for election mechanism of an adaptive multi-metric cluster head to manage the VANET sub-clusters to provide multimedia communication services over a spatially-apart vehicular groups. Except the effective cluster-head selection mechanism, there is a suggested approach for a virtual 2-hop overlay mesh-based shared tree for lower level VANET multicasting. Their results show that the integrated system has acceptable results in terms of throughput and end-end delay for LTE network.

In [18], Shankar and Danajayan investigated QoS for different services. They analyzed the performance of wideband power base (WPB) and throughput based call admission control with load balancing in hybrid coupled and Radio over Fiber (RoF) based WLAN/UMTS networks to enhance QoS. WPB scheme works as follows: Any time a new user seeks admission into the system it definitely adds certain amount of interference to the system. The criterion for uplink admission of connection is based on comparison of interference the new user would add to the system. If ΔI is the interference a new user adds to system and I_s is the existing interference in the system, I_T is the interface threshold, they follow this rule:

$$I_s + \Delta I \leq I_T. \tag{4.1}$$

We used the same network parameters as in [18] for our simulation and also tight coupled integration method. They did the load balancing in the midway routers for uplink and downlink traffic and we did on user interfaces.

In [19] as other papers, authors developed a utility-based access selection algorithm targeting to achieve the network selection between UMTS and WLAN networks. The performance of the algorithm is evaluated using OPNET. They considered the problem in the second layer. The load factors for uplink and downlink are calculated in UTRAN (RNC and

Node-B) based on the signal energy, noise spectrum, chip rate and station bit rate.

4.2   OPNET Setup

In this section, we will give a brief overview of the model setting in OPNET Modeler. For UMTS case, in Figure 4.3 we used the same components, except for EPC which was replaced insted of RNC, SGSN and GGSN. The application which were run in the network. So I will explain the settings on LTE case. For the access point in Figure 4.4 we used a wirless LAN router, wlan-ethernet-slip4-router fixed node, to connect it to the ISP, we used a T1 link. For the ISP's router we used an ethernet2-slip8-gtwy fixed node. As a MANET gateway to connect to EPC part, manet-gtwy-wlan-ethernet-slip4 was the choice. To connect these 2 components together we used a T1 link. To connect the MANET gateway and EPC part a PPP-DS3 was used. EPC was the lte-epc-atm8-ethernet8-slip8-adv EPC model which was connected with a PPP-D3 link to both IP cloud and eNodeB. We ran this model for 2, 4, 6, 8 and 10 nodes. Nodes are fixed with no trajectory and all had access to both wireless networks. The duration of the run was 1400 seconds. For each UE, all application services were available, also the same setting for application services was done on the server. The flows were from all UEs to the server. Supported profile which runs the applications on the UEs was set on all 5 mentioned applications we have in the network.

We have run 5 different applications in 700 seconds, FTP and email can be considered as a background traffic, web surfing is interaction traffic and voice and video are real-time. All the traffics are low quality to avoid bottleneck, because we used only one server for all applications as you can see on Figure 4.3 and 4.4. Table 4.1 shows the setting for the mentioned applications. Also there are 4 IP-GSM-Voice flows from four UEs to the server with 12 packets/s with the average packet size of 72.5 Bytes and the traffic start time of 80 and traffic duration of 350 seconds. And the same flow rom the server to a LTE iPad.

Table 4.1: Application Setup

Characteristic	Video	VoIP	HTTP	FTP	Email
ToS	ToS(5) and ToS(6)	ToS(6)	ToS(0)	ToS(0)	ToS(0)
Packet Interval	Constant(0.25) and Constant(2)	Exp(0.352)	Exp(720) page interval	Exp(3600)	Exp(3600)
Size	Frame Size 128*120 pixel and Const(10000)	1 Voice frame/packet	Object Size 500Kb with 1 Object/page	File Size 1000KB	Email Size 500Kb
Start Time	1 sec after simulation start time	5 sec after simulation start time	5 sec after simulation start time	5 sec after simulation start time	5 sec after simulation start time

There are some more considerations for some applications which we did not put in the table, for Voice over IP (PCM quality): Incoming/Outgoing Talk spurt length in Seconds was set on exponential (0.352), For De/Compression Delay=0.02 sec. Also we set the conversation environment as a quiet room with a land phone.

Just as a note for a better understanding, an unbroken sequence of on-intervals is called a spurt. The sequence of off-intervals is called a gap. A talk spurt is a period of time in which the listener does not detect a pause. This can be the same as a spurt, but it may also contain several gaps, which are not audible to a human or not interpreted as such. Analogue to this is the definition of pause, which is the interval in which the listener does not detect any speech or does not interpret the spurts as speech. Again, a pause could be the same as a gap, but some spurts might be present, due to for instance noise on the transmission line. These will not be interpreted as speech bursts and therefore they are part of the pause period. For the simulated LTE system, we will have the same types of traffic.

4.3 Simulation and Results

In this section we will overview of our design and network parameters and then discuss the results.

4.3.1 Network Components and Design Parameters

In this project, our aim is to reach a data delivery cost effectively with consideration of bottleneck and QoS for different types of services. The scenario consists of one LTE and wireless LAN access point technologies. In the LTE cell, UEs are moving around freely (consisting of iPads and iPhones), and inside the LTE coverage there's just one hot spot to cover for WLAN. Our heterogeneous system was tightly coupled with WALN and LTE. There are basically two types of interworking in WLAN/cellular heterogeneous systems: tight coupling and loose coupling. For loose coupling design, the WLAN gateway does not have any direct connection to the cellular system elements and WLAN traffic would not go through the cellular core network [20]. In a tight coupling interworking architecture approach, which we used for this project [19] a WLAN is connected to a LTE core network as other LTE radio access networks, where all the e-NodeBs are separated. It can be implemented on both 3G and 3GPP networks. So when it comes to the tight coupling, the WLAN gateway implements all the cellular system protocols such as authentication, mobility management, etc. In this approach, we used tight coupling as we wanted to implement IPv6 on the mobile nodes as we wanted their addresses to be fixed. Figure 4.1 and 4.2 show the tight and loose interworking respectively. One of the advantages of tight coupling is the easy maintenance. But loose coupling offers several advantages as deployment and traffic engineering of cellular/WLAN systems [21].

Our UMTS model as we can see in Figure 4.3 is made up of 3 main components of UMTS systems: RNC, SGSN and GGSN which the function of each one has been described in previous chapter. Figure 4.3 shows the topology of our network for UMTS/WLAN before
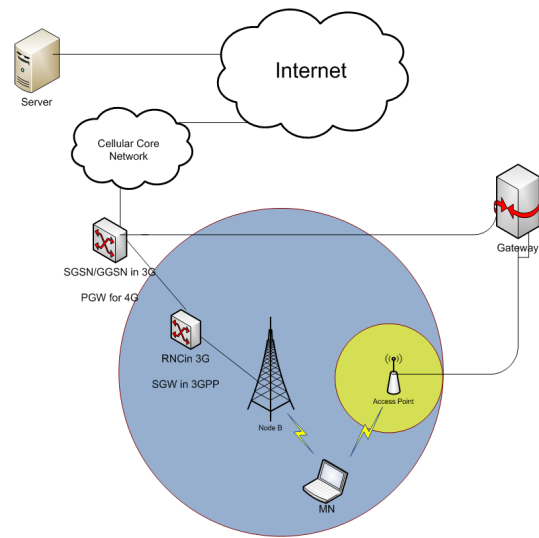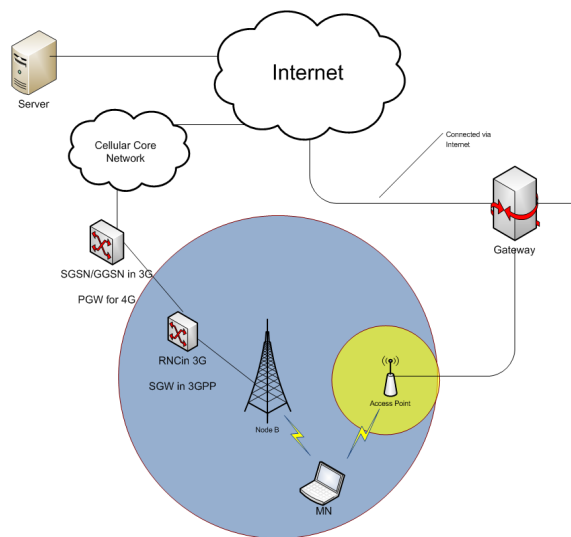
Figure 4.1: Tight Coupling

Figure 4.2: Loose Coupling

migration. Inside the subnet we have put the nodes with the wireless LAN coverage is provided by the access point whose coverage was highlighted with the lighter circle area.
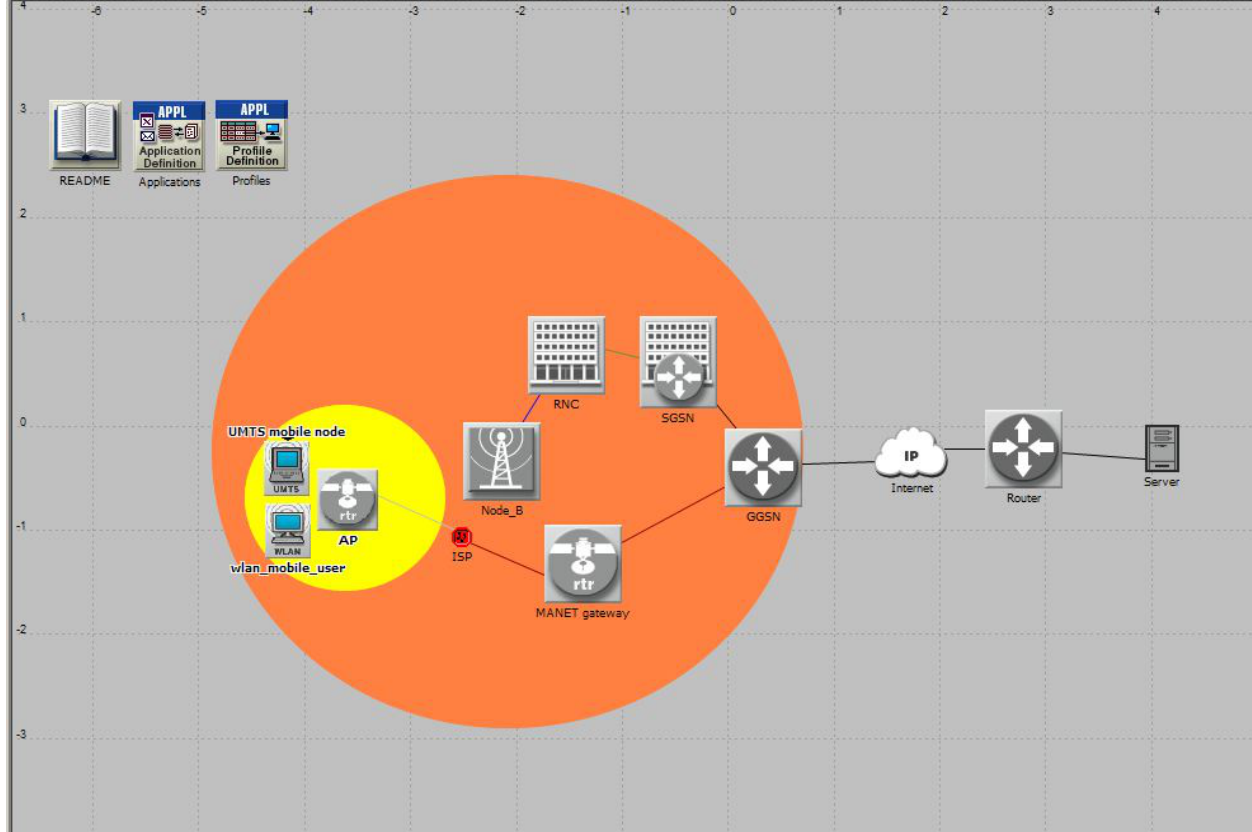

Figure 4.3: Our UMTS/WLAN Network simulated in opnet

Here, we should note that our algorithms's goal is to reach the maximum data quota of the prepaid plan for the LTE device and best QoS for the running applications. Figure 4.4 shows the network's design with 6 nodes. WLAN's data rate is 11Mbps, and for LTE with 10MHz bandwidth, UL and DL base frequencies are 1.92GHz and 2.11GHz respectively.

In next sections, we will disuss on some of the results.
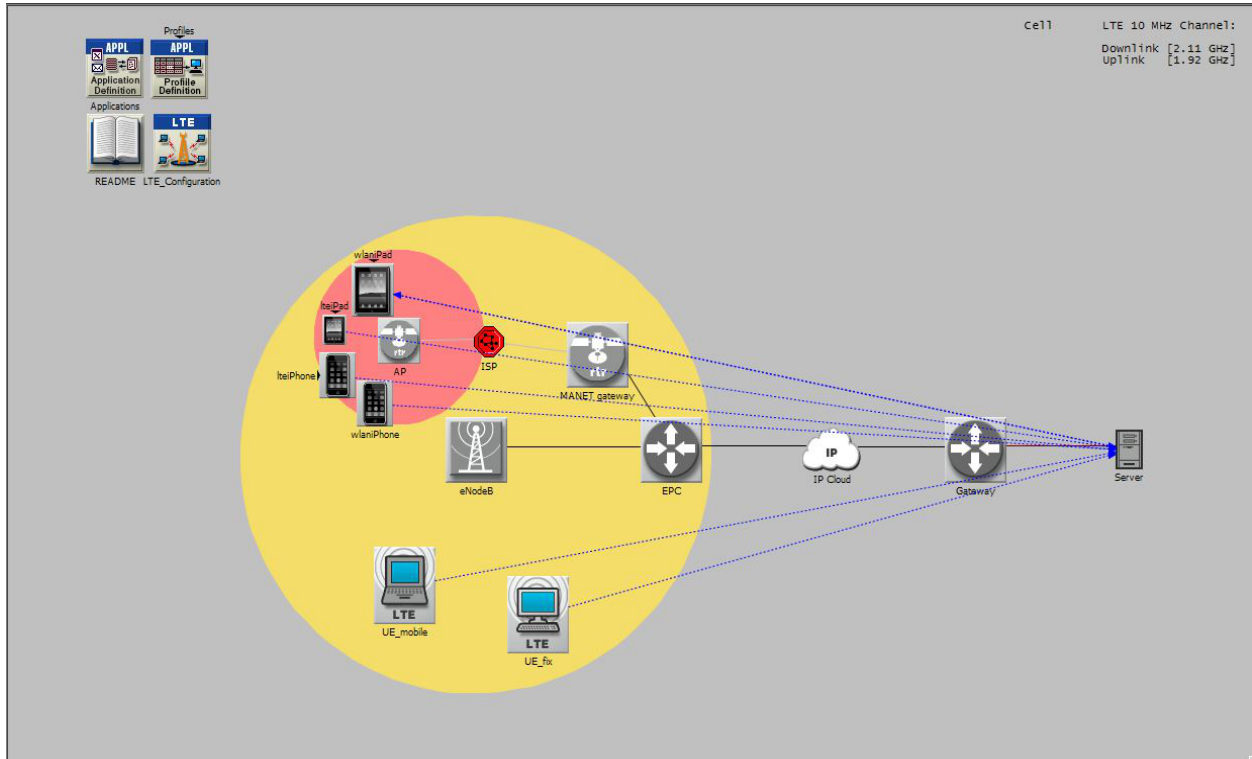
Figure 4.4: Our LTE/WLAN Network simulated in opnet

4.3.2   Application Results

The results of the simulation, before and after applying the algorithm, are discussed in this section. We also give the analysis of the network's reaction to the changes. First we will explain QoS results, and then data usage for both applications and networks. All the following studied results are for about 50 minutes of simulation run duration.

Email Application

For email application which has the ToS (0), best effort, before and after applying the algorithm through the policies, we expect the network's performance for Email in terms of upload delay gets better. And at the same time we want a reasonable response time for each application while we expect network to use the most of the prepaid data quota through the cellular network avoiding extra costs for data (to deliver data cost effectively) through

cellular as mentioned before. So, In Figure 4.5 and 4.6 we can see the average amount of data sent and received is increased by adding more nodes to the network and we get a higher probability of congestion. After applying the policies, because of the goal of maximizing the cellular data usage and at the same time maintain a reasonable QoS for each application, we can see as the load has been sent and traffic does not go through from just one single interface. Therefore, the system can deliver more data for email application and less data drops.
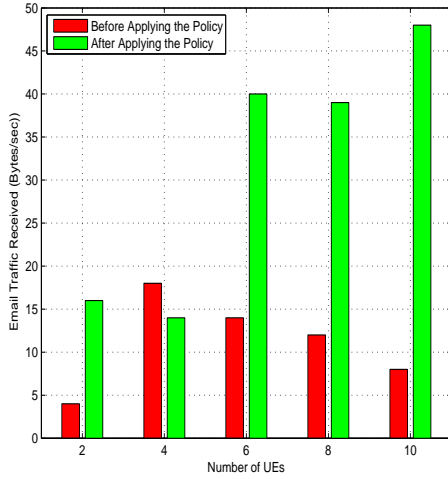
Figure 4.5: Email Traffic received
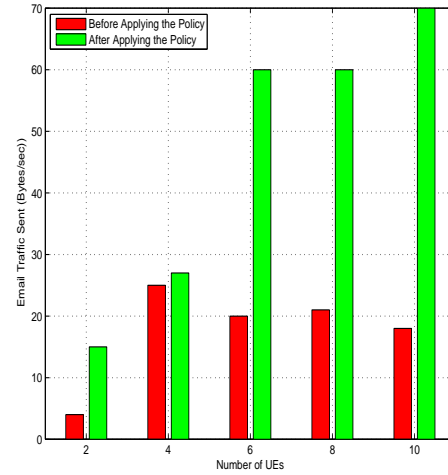
Figure 4.6: Email Traffic sent

FTP Application Results

For file transfer application, as we can see from Figure 4.7 and 4.8, in terms of traffic sent and received, after network selection, the amount of data transmitted will increase as the algorithm reduced the probability of data drop.

FTP download response time is a parameter that shows the time elapsed between sending a request and receiving the response packet. It is measured from the time a client application sends a request to server to the time it receives the response packet. As we can see in Figure 4.9 in terms of FTP download response time, the algorithm does not work as well as
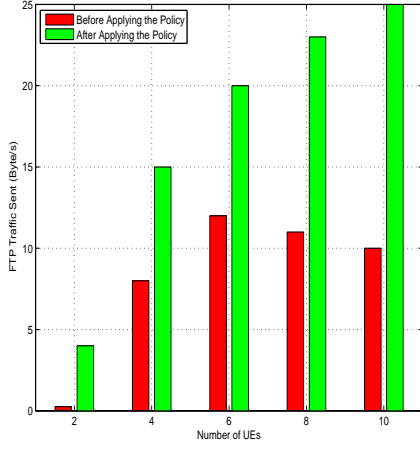
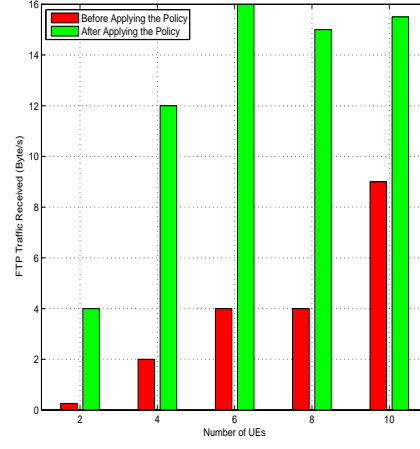Figure 4.7: FTP traffice sent comparison

Figure 4.8: FTP traffic received comparison

before. It is because the calculating the threshold and processing of the algorithm are time consuming, therefore, the application data transmission has to wait when the congestion probability increases with number of nodes. We can see the same trade off as email application. Mentioned trade off is between FTP download delay and data transmission which leads to more delay and more data transmission.
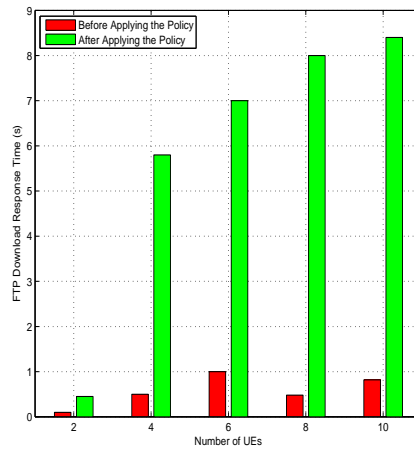

Figure 4.9: FTP download delay

Voice Results

Here we discuss the behavior of the network by adding more nodes, to have more probability of congestion, before and after applying the algorithm for VoIP application. We also study how the algorithm impacts on this application. First we study VoIP average traffic sent plot to have a better understanding of misbehavior of the network in terms of jitter and end to end delay. According to Figure 4.10, applying the algorithm can help to reach the goal of maximizing usage of free data quota from the cellular system. The data drop decreases as the result of selecting the network. Before applying the algorithm to interfaces, data sent for VoIP was less which shows the amount of data drop for VoIP application was more.
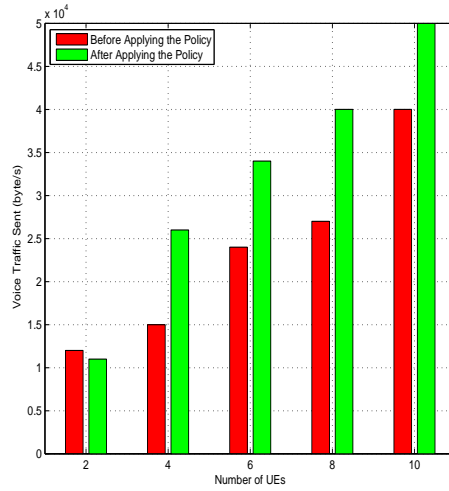

Figure 4.10: Average sent traffic comparison for VoIP application

End to end delay is the total voice packet delay, called "analog-analog" or "mouth-to-ear" delay. This parameter is equal to network delay plus encoding delay plus decoding delay plus compression delay plus decompression delay. Here we can see from the plot that there is a trade off between usage of all free data and end to end delay. As already mentioned, our algorithm considers calculation time overhead on the system. As the 4G voice has to go through LTE network and it is different with VoIP applications, it will suffer more in terms of delay and it is more under the influence of the mentioned overhead. For the case of more

than 4 nodes added to the model and more congestion probability, system's tolerance for end to end delay decreases. It means it has more end to end delay. Therefore, end to end delay increases. But still there is not an outstanding change for end to end delay and we still can get a reasonable response for VoIP. Figure 4.11 shows this concept. As we know, all these values are gathered from all nodes, so when we claim 2 seconds of end to end delay for voip it means approximately each node experienced 0.2 second of delay for transmitting average of 50KB/s of GSM traffic flow.
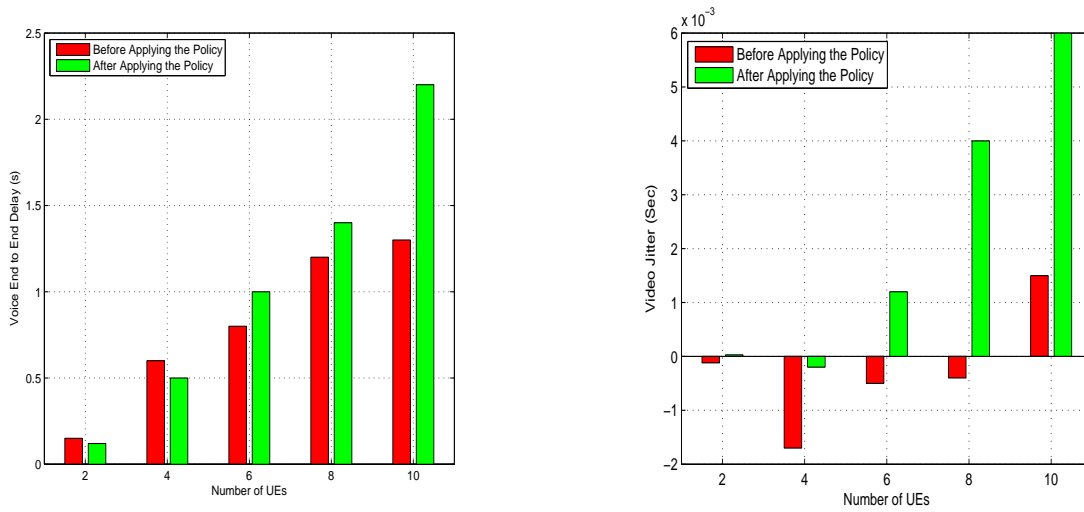

Figure 4.11: VoIP end to end average delay

Figure 4.12: Average jitter comparison for VoIP application

Jitter is delay variance for VoIP packets. It means that if two consequetive packets leave the source with time stamp of t_1 and t_2, and then played back at the destination with time stamps of t_3 and t_4, then jitter is : $(t_4-t_3)-(t_2-t_1)$. For better understanding, if the packets are sent with 10ms interval, positive jitter means more than 10ms interval in reception and negative jitter means less than 10ms when the packets get to the destination. This parameter is under the influence of end to end delay. Comparing the results in Figure 4.12, shows that average jitter for voice application increases by adding more traffic to model.

Video Conferencing Results

In this section we study the results for video conferencing application. Packet delay variation is a parameter to measure the variance among end to end delays for video packets. Packet end to end delay, is the time taken for the video application packet to a get to a destination node application layer. This statistic recorded data from all the nodes in the network. Comparing the results for this parameter, in Figure 4.13, shows that our model did not have a good response to higher congestion probability. As a result of adding more nodes and traffic to the network, network drops the video data for more than 6 nodes case and does not transmit video data anymore as the result of traffic jam and drops them completely. Results show that after applying the policy and balancing the video load on both interfaces, model's tolerance increased in terms of higher rate of traffic in the network. Considering to give a higher QoS requirement to video application with ToS(5) and usage of free data on cellular side by the nodes, end to end delay decreases. The reason is because the traffic sent and received for this application was increased with less data drops, this case will be studied next.
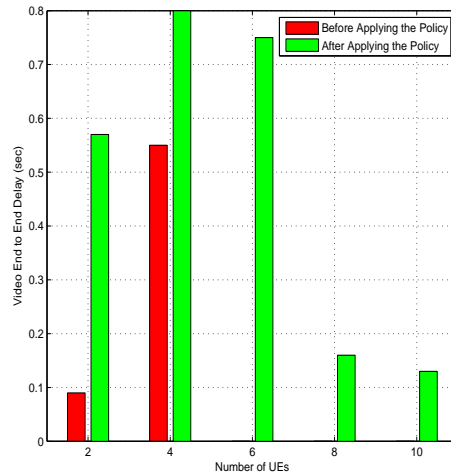

Figure 4.13: Average video end to end delay

After applying the rules to the interfaces, in Figure 4.14 there is a considerable difference

between previous result in terms of stability for packet delay variation of video traffic. This parameter is influenced by end to end delay.
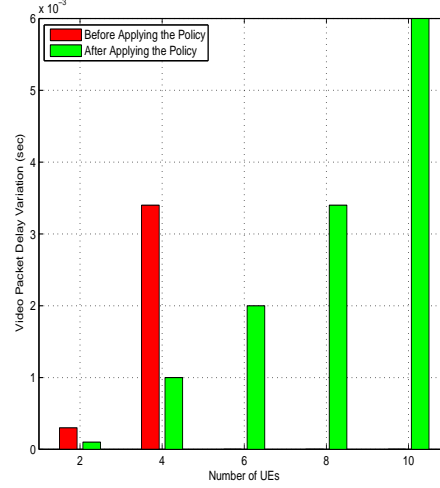

Figure 4.14: Packet delay variation average for video traffic

From Figure 4.15, we can see for uplink (sending) traffic, before applying the policy the model can tolerate it for less than 8 nodes added to network. In case of more than 8, when it comes to 10 nodes, because of the traffic jam on just one interface to transmit the data, it drops sending traffic and stops the application. After running the algorithm, system's tolerance to higher data transmission works better and goes higher. Because of splitting the traffic on both interfaces, data drops reduce and sending data increases.

In terms of receiving data, before applying the policy the network did not have tolerance to receive it with traffic jam. As our algorithm considers total sent and received data of cellular network while calculating the mentioned threshold, we can see because of more data sent on uplink, the downlink load decreases for more than 4 additional nodes in the network as it was increasing for traffic sent case. Before applying the policy for the same reason of traffic jam and tolerance of network, which was explained for previous plot, data is dropped for more than 4 nodes. Therefore, we can see network's data drop for video application decreases and the QoS is better for this application with having less video conferencing call
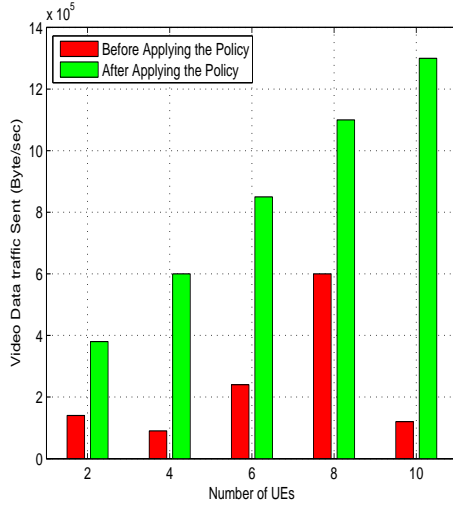
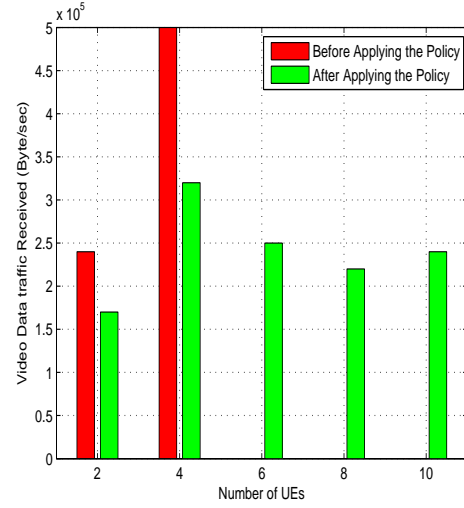Figure 4.15: Average Video Conferencing traffic sent

Figure 4.16: Average Video Conferencing traffic received

drops.

4.3.3 LTE Results

In this section we study the results for LTE network and see how the algorithm worked on the LTE network in terms of how much uplink load transmitted and how it affected the QoS. First in Figure 4.17 uplink load represents the total uplink (from UE to eNodeB) higher layer data traffic arriving to the LTE layer in packets/sec. This statistic was collected by all the UEs in the network. As we can see the amount of data uploaded from the user side to the network was decreased as the total traffic transmitted becomes higher during the run of simulation, the threshold gets smaller as a result of usage of the data, the policy forces the outgoing traffic goes down to the other interface unless the user wants to go on with more cost, to save the remaining quota for future. But it is not very stable on LTE uplink data transmission. After applying the policy, it gets more stable as the system has to meet the good usage of LTE data quota criteria. The result we took for this part was the proposed algorithm did not work completely stable in terms of LTE uplink transmitted data as it was

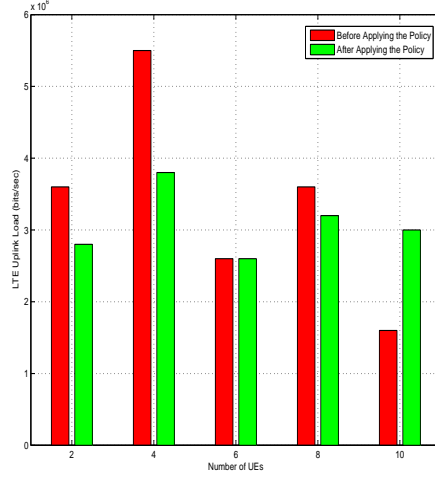before, but it could transmit more data after applying the policy.

Figure 4.17: LTE uplink traffic comparison

For downlink load we did not put any policy for downlink, but as the algorithms considers total amount of data usage for download and upload, as the data upload decreases in previous plot, it can has the decreasing effect on the downloaded. The amount of free data (up and down links together) is limited. We can see as the result of increasing in threshold due to generating more traffic inside the network, the policy made the data drop for future quota saving, so the downlink load drops a little bit, we can see this in Figure 4.18.

As a result of the network selection policy, we can see the downlink delay is increasing with more probability of congestion. Therefore, our algorithm did not work better in terms of downlink delay for cellular network. We can see the results for downlink average delay in Figure 4.19.

4.3.4 WLAN Results

Here, we analyze the results for WLAN part of this tight coupling heterogeneous network.

In terms of delay, there is a parameter representative of end to end delay of all packets received by the wireless LAN MACs of all nodes in the network and forwarded to the higher
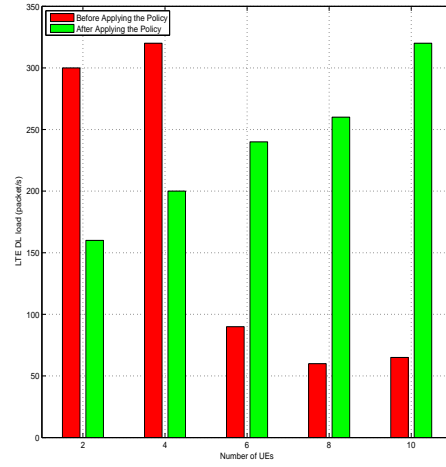
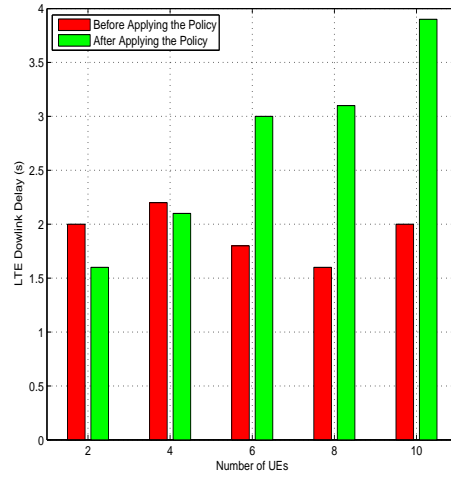Figure 4.18: LTE downlink traffic comparison

Figure 4.19: LTE downlink average delay comparison

46

layer. Therefore, even if the policy was implemented just on uplink interface, after applying it, the load on WLAN interface decreases as the device tries to make a better usage of cellular prepaid data, so cellular system takes some of the load. As the result average delay on WLAN inerface decreases and gets more stable. This comparison is illustrated in Figure 4.20.
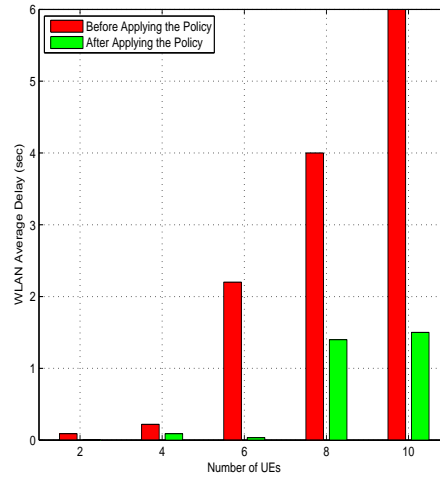

Figure 4.20: Average WLAN delay comparison

The following statistic dimensioned in order to measure the network load in bits/s submitted to wireless LAN layers by all higher layers in all WLAN nodes of the network. As we can see from the results, the load of the WLAN network increased as the result of less data drops and network selection algorithm which made some of the load goes through the WLAN interface.

In terms of retransmission attempt, this statistic demonstrates the total number of retransmission attempts by all WLAN MACs in the network until either packet is successfully transmitted or it is discarded as a result of reaching short or long retry limit. Retransmission attempt also includes retry count increments due to internal collisions. Figure 4.22 shows that, internal collisions after applying the policy and higher probability of collision, increases as the data traffic increased and the process of threshold calculation is time consuming, the
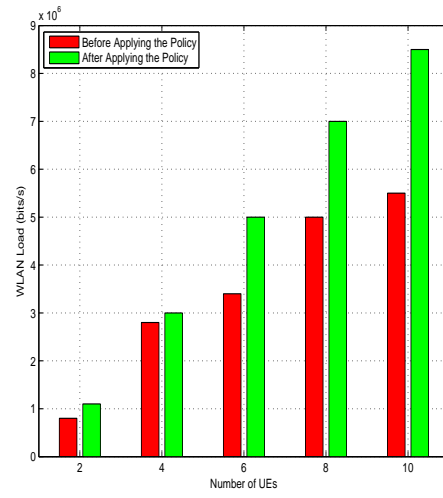
Figure 4.21: Average WLAN load comparison

data buffers get full faster than before, so the collision probability gets higher.
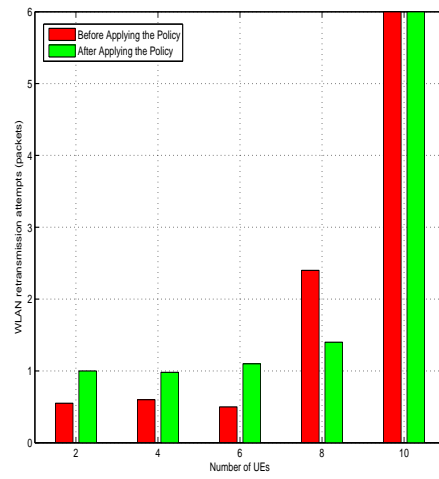

Figure 4.22: Average WLAN Retransmission attempt comparison

4.4 Results for Total Amount of Sent Data Through Each Interface

In this section we will discuss about how much traffic was sent through each interface before and after applying the policy. Then we study if the algorithm helped not to go over the data usage threshold of each month. It is to prove that our approach met the criteria. As we can see from Figure 4.24 that total amount of data that has been sent, through the WLAN interfaces of UEs, for cases of 2 to 10 nodes, was less than after applying the policy. This amount of data shows how much data was uploaded successfully from the IP layer in one day. As we expected, for 5GB of data per month, we have approximately total 170.66MB of data quota per day. The graph shows that from the LTE interface after applying the policy, 130.5MB data was sent in one day in case of 10 nodes. As we studied earlier in Figure 4.17 the amount of uplink data sent from LTE side decreased as the result of network selection algorithm. This algorithm makes the data drop from 190.94MB/day to 100.48MB/day in case of 10 nodes. So it meets the criteria not to go over the threshold of 170.66MB/day data usage. Also it still has kept enough quota to download. From WLAN side, we can see from the Figure 4.23 that amount of data transmitted increased as the result of less drop as the result of running the algorithm and more data transmitted on this side. Here, also as we already studied in Figure 4.22 the average WLAN load increased, the same as on the interface.

The amount of mentioned data, successfully sent data from each interface, is calculated from following equation.

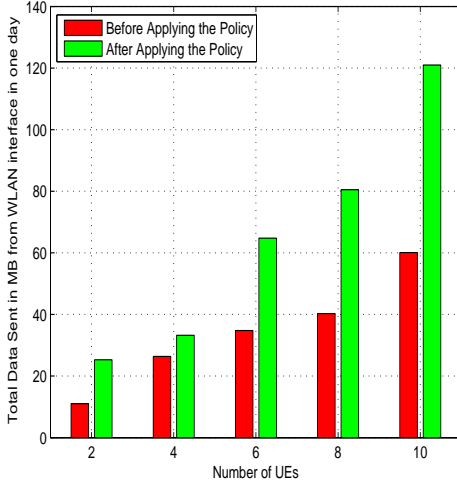$$SentData = CreatedData - DroppedData. \tag{4.2}$$

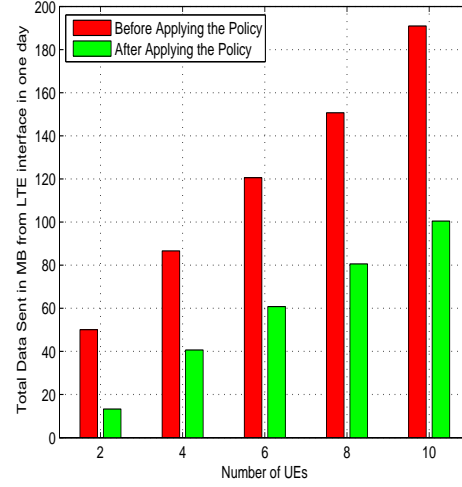Figure 4.23: Data Sent from WLAN interface in one day

Figure 4.24: Data Sent from LTE interface in one day

4.5 Results Summary

To summarize, the network selection algorithm on the uplink system was given and implemented considering QoS for each application. Network selection algorithm which was explained in chapter 3 was introduced to select the network for the system data for uplink interface considering a cost effective communication and giving a reasonable response time to each application. Cost effectiveness of this approach is achieved by the way each UE tries to make the usage of cellular prepaid data as much as possible and not go over this amount of data at the same time it has to take care of giving a reasonable QoS to each application. After applying the policy on uplink interfaces, for each application and cellular/WLAN channel we discuss the networks's performance.

For FTP application, results show that traffic sent and received are increased as the result of less data drop, but trade off was more delay in terms of download as the data transmission rate was higher. For email application, we got the same results as FTP. More transmitted data and as a result more delay on upload response time, still we can see the delay is reasonable in terms of QoS to the email application. The other important application

which needed a good QoS was video conferencing. From the results, we can see the algorithm worked well on ToS (5), for video application. In general, this application got a better service in our model with less data drop. Although the end to end delay was more than before for the case of adding more than 4 nodes in the network, as the data transmission for this application is higher but with a reasonable response time.

For VoIP, our algorithm could work better in terms of seding data, but in terms of delay it could not make the delay better as the cellular channel was dedicated more than before to transmit the data. Sending more data has the trade off of a worst delay as the result of calculation overhead. So far we can see as the algorithm was based on the applications' type of services, with the goal of network selection considering QoS and optimized usage of free data. As we could allocate the network resource to outgoing data traffic with more usage of cellular free data, for the applications with higher ToS like video or voice, we could send more data but not any improvement in terms of delay except for video. It was the same with email and FTP. On each channel, WLAN and LTE, comparing the results shows that, the proposed policy works better in terms of average delay on WLAN as trying to maximize usage of cellular data and network selection at the same time. We can see for LTE delay got bit more stable with more download data transmission. At the end, we studied the amount of uplink on each interface to see if the algorithm could meet the criteria, as the results show, it dropped the data in case of going over the threshold which was expected.

Chapter 5

Conclusion and Future Work

In this chapter there will be a brief review of what we did in the study, results and constraints of this approach and possible future works.

5.1 Conclusion

This research proposed a new algorithm to allocate the network resource for the load on uplink interfaces of user equipments. First there is a proposed tight coupling architecture to study this algorithm which the wireless LAN network is overlaid by a cellular, LTE, eNodeB coverage. The proposed framework was based on user salification and network resource allocation on uplink with the help of network selection policy. Implementation of this algorithm was inter level based in IPv6 flow management on each interface. Also QoS user salification policy was set by the help of the algorithm for uplink traffic of the mentioned network. As discussed in summary part of previous chapter, our algorithm worked generally better on sent and received data, but not the best in terms of delay. Although there was not a big change in terms of delay, it could not make any improvement. Instead it gave a better service for each application and met the criteria of best usage of pre-paid LTE data.

5.2　Future Work

For future work, there are a number of open problems worth pursuing. One of the concerns is considering the incoming traffic on the interfaces. There is another possibility to work on balancing the load on incoming (download data) ports. A second problem is an idea of limiting the WLAN download and upload data. This limitation that we could assume is for the WLAN: up to 1G it's free and then after the usage of 1G, it will be like "pay as you go". The same is the case as happening in some public WLAN hotspots like airports. It will make the problem more complicated and we can consider the implementation as a future work.

Chapter 6

References

Bibliography

[1] P. S. Henry and H. Luo, "Wifi: What's next?," *IEEE Communication Magazine*, December 2002.

[2] I. M. S. Fluhrer and A.Shamir, "Weaknesses in the key scheduling algorithm of rc4," *8th Annual Workshop. Sel. Areas on Cryptography*, Aug 2001.

[3] C. Cox, *An Introduction to LTE*. Wiley, 2012.

[4] G. Aggelou, *Mobile Ad Hoc Networks from wireless LANs to 4G networks*. Mc Graw-Hill, 2005.

[5] S. Jin and W. Xuanli, "Load balancing algorithm with multi-service in heterogeneous wireless networks," *IEEE 6th International ICST Conference on Communication and Networking in China (CHINACOM)*, vol. 6, 2011.

[6] M. Jaseemuddin, "An architecture for integrating umts and 802.11 wlan networks," *Eighth IEEE International Symposium on Computers and Communication, ISCC'03*, 2003.

[7] L. Giupponi and R. Agusti, "A novel joint radio resource management approach with reinforcement learning mechanism," *IEEE Press*, pp. 621–626, 2005.

[8] Q. Song and A. Jamalipour, "Network selection in an integrated wireless lan and umts environment using mathematical modeling and computing techniques," *IEEE Wireless Communication Magazine*, vol. 2, pp. 42–48, June 2005.

[9] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, pp. 335–353, 1965.

[10] T. J. Ross, *Fuzzy Logic with Engineering Applications*. McGraw-Hill, Inc, 1995.

[11] K. Anderson and C. hlund, "Optimized access network selection in a combined wlan/lte environment," *Springer, Wireless press communications*, October 2011.

[12] Y. Hwang and A. Park, "Vertical handover platform over applying the open api for wlan and 3g lte systems," *IEEE*, 2008.

[13] J. C. F. andTarik Taleb, "Bandwidth aggregation-aware dynamic qos negotiation for real-time video streaming in next-generation wireless networks," *IEEE Transaction on Multimedia*, vol. 11, October 2009.

[14] U. Toseef and M. A. Khan, "User satisfaction based resource allocation in future heterogeneous wireless networks," *IEEE, 2011 nineth annual Communication Networks and Services Research Conference*, 2011.

[15] A. Bazzi, "Wlan hot spots to increase umts capacity," *21st annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communication*, 2010.

[16] J. Wu and M. Jaseemuddin, "Integrating umts and mobile ad hoc networks," *IEEE WiMob*, 2006.

[17] R. Sivaraj and A. K. Gopalakrishna, "Qos-enabled group communication in integrated vanet-lte heterogeneous wireless networks," *IEEE 7th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, 2011.

[18] R. Shankar and P. DanaJayan, "An adaptive call admission control scheme with load balancing for qos enhancement in rof based converged umts/wlan," *International Journal of Computers and Communications*, vol. 5, August 30 2011.

[19] Y. Zhou and Y. Rong, "Utility-based load balancing in umts/wlan internetworking systems," *Radio and Wireless Symposium*, 2008.

[20] A. Munir and V. S.Wong, "Interworking architectures for ip multimedia subsystems," *Springer, Mobile Network Applications*, April 16th 2008.

[21] L. Ma and F. Yu, "A new method to support umts/wlan vertical handover using sctp," *IEEE Wireless Communication*, pp. 44–53, August 2004.

Appendix A

OPNET Process Models

In this part of the document we will give a brief overview of how the work was done and the process models we used for this work to be done. Next chapter has the information of the codes inside of each process models. Node Editor, is a user interface tool which a user can create and edit the internal structure of a device or node. A node can represent a computer, a switch, eNodeB, iPad, a network cloud or any network component we use in designing a model. In Figure 3.2 we can see inside of a LTE node. It is consist of modules which are used to transmit packets, process data, store data, route packets, etc. As we can see in this figure, LTE node has all seven layers of the OSI model. Process editor is where actually we can write codes to implement algorithms and protocols. Process editor allows you to visually depict the logic process via the state transition diagrams, STDs. In STDs a logic is composed of several states. States can transit between each other if certain conditions are triggered. So when we call a function in one of the state transitions it can transit one state to another. Figure A.1 shows inside the ip layer of the process editor of the LTE node which connects two interfaces together.

Here we should apply the QoS policy in IP layer for outgoing data (uplink transmission). We need to co-simulate the model with a C third-party program. To do so, we have to interact with an external system during the simulation. We need and External System
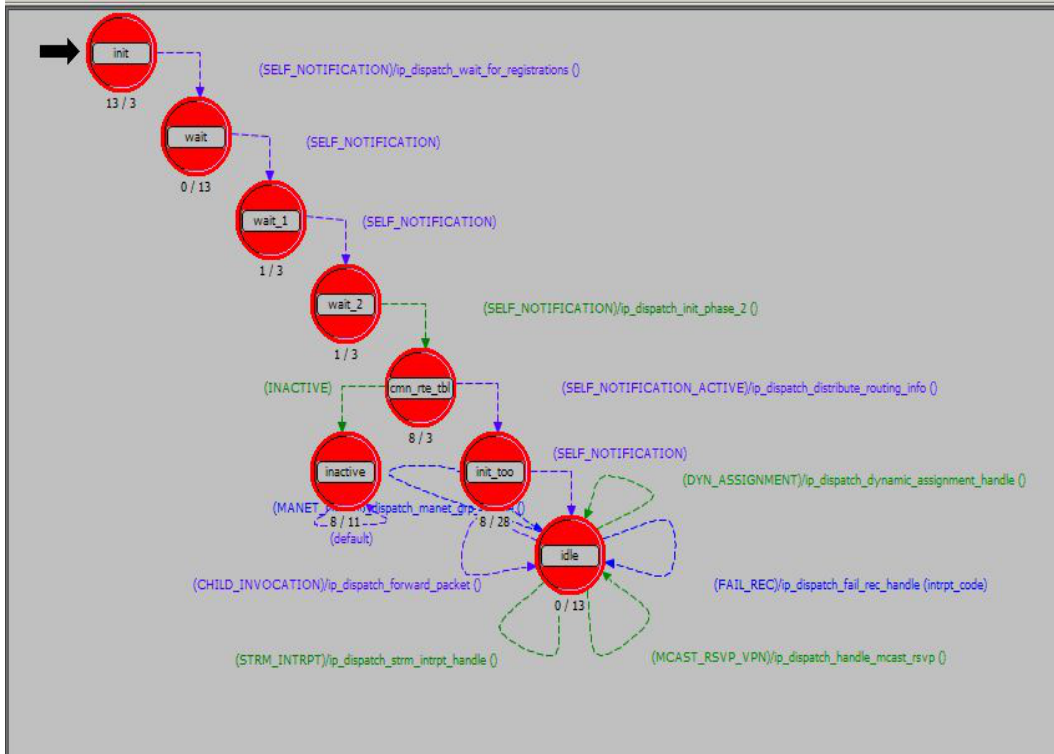
Figure A.1: Process editor of IP layer of LTE node

Definition (ESD) model, which defines a set of interfaces that allows process models in OPNET modeler to communicate with external programs. These interfaces can be read by both external programs and OPNET process models. We also need Esys (External System) module. The Esys module is a node domain that can be placed into a node domain and processed as are other domain modules. This module allows user's simulation model to talk to external systems. Simulator description file is a plain-text file containing statements that specify how to built co-simulation. This file has to be placed in OPNET directory with the ".sd" extension. External Simulation Access (ESA) API package contains functions that can read and write interfaces' values defined in ESD model from external code during co-simulation. This external code refers to the node of an external system or a link to another external program. This package also contains functions that are able to control the simulation flow process, read/write text, and issue debugging commands from the external codes. External system is the last thing we need, it represents any external program, device

or code that is in co-simulation with OPNET modeler. All the files are included to the final version of thesis.

Appendix B

Pseudo Code

Here we present the pseudo code of the algorithm.

1: For all outgoing traffic check the traffic type

2: if (traffic is voip or voice)

3: Estimated usage for ToS(6 and 5)=average amount of 2 last runs;

4: Calculate the remaining data: data quota minus previous data usage;

5: Fuzzify the remaining data from equations 3.2, 3.3 and 3.4;

6: Define the acceptance degree based on amount of data remaining;

7: Calculate Threshold2 from equation 3.8;

8: If (estimated data less than threshold2)

9: send data through Cellular;

else

10: send data through WLAN;

11: else if (traffic is FTP HTTP or Email)

12: Estimated usage for ToS(0)=average amount of 2 last runs;

13: Calculate the remaining data: data quota minus previous data usage;

14: Fuzzify the remaining data from equations 3.2, 3.3 and 3.4;

15: Define the acceptance degree based on amount of data remaining;

16: Calculate Threshold1 from equation 3.8;

17: If (estimated data is less than threshold1)

18: send data through Cellular

else

19: send data through WLAN;