



ADDIS ABABA UNIVERSITY
ADDIS ABABA INSTITUTE OF TECHNOLOGY
SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING
TELECOMMUNICATION NETWORK ENGINEERING PROGRAM

Spatiotemporal Mobile Data Traffic Modeling: The Case of UMTS
Network in Addis Ababa, Ethiopia

By

Yosef Abera

Advisor

Dr. - Ing. Dereje Hailemariam

A Thesis Submitted to the School of Graduate Studies of Addis Ababa
University in Partial Fulfillment of the Requirements for the Degree of Master of Science
in Electrical Engineering

November, 2018

Addis Ababa, Ethiopia

ADDIS ABABA UNIVERSITY
ADDIS ABABA INSTITUTE OF TECHNOLOGY
SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING
TELECOMMUNICATION NETWORK ENGINEERING GRADUATE
PROGRAM

Spatiotemporal Mobile Data Traffic Modeling: The Case of UMTS
Network in Addis Ababa, Ethiopia

By: Yosef Abera

Approval by Board of Examiners

_____	_____
Chairman, School Graduate Committee	Signature
<u>Committee</u>	
<u>Dr. - Ing. Dereje Hailemariam</u>	_____
Advisor	Signature
_____	_____
Internal examiner	Signature
_____	_____
External examiner	Signature

Declaration

I declare that this thesis is my own work and does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or institute of higher learning. To the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to Addis Ababa Institute of Technology the non-exclusive right to reproduce and distribute my thesis, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Yosef Abera

Name

Signature:

Place: Addis Ababa

Date of Submission: _____

The above candidate has carried out research for the Master's thesis under my supervision.

Dr. -Ing. Dereje Hilemariam

Name

Signature

Abstract

The exponential growth in mobile data traffic is forcing telecom operators to invest on new infrastructures. But devising techniques for optimum network utilization, which can be provided by traffic modeling, has the potential to reduce investment costs. Modeling traffic variation in different service areas and time can also be applicable for energy efficient network planning, understanding customers' data traffic usage behavior and dynamic resource allocation.

In this thesis, based on the data collected from 734 Base Stations located in Addis Ababa, Ethiopia, the real traffic is modeled in space and time with a tunable accuracy. Firstly, a rectangle that can inscribe the geographical area of the city is selected and divided equally in to N by M smaller groups. Secondly, to understand the temporal behavior in the data, the time-series data traffic of each group is transformed to spectral domain by using Fast Fourier Transform where it is observed that all groups have the same four major frequency components but with different magnitude of coefficients and phases. Then, matrices corresponding to coefficient and phase values are transformed from spatial to spectral domain by applying Two Dimensional Discrete Cosine Transform. In spectral domain higher frequency components that contain less information are removed out and the remaining are used for the reverse transform that made the model to be complete in time and space.

The different application areas of the model for the operator require a varying level of accuracy which in turn is dependent on the truncation level of frequency components. As a result, a relation between model performance and truncation level is developed which formed the model as a tunable around accuracy. Finally, by defining parameters that use the model as an input, the network performance is analyzed and suggestions for the observed gaps are presented.

Index Terms: Spatio-temporal; modeling; UMTS Data Traffic; Fast Fourier Transform; Two Dimensional Discrete Cosine Transform.

Acknowledgment

All praise, honor and glory are to my Lord Jesus Christ for His richest grace and mercy for the accomplishment of this thesis.

I wish to express my deep gratitude to my advisor Dr. - Ing. Dereje Hailemariam for his supports beyond the expectations. He has been there providing his heartfelt guidance at all times and has given me invaluable provision, inspiration and suggestions in my quest for knowledge.

I would like to thank Ethio telecom for launching such a successful program and for providing the required data in conducting this thesis.

Finally, I would like to express my deepest gratitude to my dearest parents for their constant prayers and encouragement to me towards the completion of this thesis.

Publication

The outcome of this thesis is accepted for publication in ICTC2018 conference.

- Yosef Abera and Dereje Hailemariam, "*Spatio-temporal Mobile Data Traffic Modeling Using Fourier Transform Techniques*", The 9th International Conference on ICT Convergence (ICTC 2018), October 17 -19, 2018, Jeju Island, Korea.

Table of Content

Abstract	i
Acknowledgment	ii
Publication	iii
Table of Content.....	iv
List of Figure	vi
List of Tables.....	viii
List of Acronyms.....	ix
1. Introduction.....	1
1.1 Statement of the Problem	3
1.2 Objective	3
1.2.1 General objective	3
1.2.2 Specific objectives.....	4
1.3 Literature Review	4
1.4 Methodology	6
1.5 Scope and limitation.....	6
1.5.1 Scope of the thesis	6
1.5.2 Limitation of the thesis	7
1.6 Contribution	7
1.7 Thesis Layout	7
2 Universal Mobile Telecommunications System	8
2.1 Introduction.....	8
2.2 UMTS Network Architecture.....	9
2.2.1 UMTS Core network.....	10
2.2.2 UMTS Terrestrial Radio Access Network.....	10
2.3 UMTS Network Functions and Protocols	12

2.3.1	UMTS Functions.....	12
2.3.2	UMTS Protocols.....	13
2.4	UMTS Quality of Services	16
3	Fundamentals of Fourier Analysis	19
3.1	General Introduction.....	19
3.1.1	Fourier Series	19
3.1.2	Fourier Transform.....	20
3.2	Fast Fourier Transform	21
3.3	Two dimensional Discrete Cosine Transform	24
4	UMTS Data Traffic Modeling Using Fourier Transform	28
4.1	Basics of Modeling.....	28
4.1.1	Time Series Analysis and Modeling.....	28
4.1.2	Spatial Modeling	29
4.2	Model Development procedure	30
4.2.1	Traffic Model in Time	31
4.2.2	Traffic Model in Space.....	33
4.2.3	Traffic Model in Space and Time	37
5	Results and Discussion	38
5.1	Network Utilization	46
6	Conclusions and Recommendations.....	52
6.1	Conclusions	52
6.2	Recommendations.....	53
	References.....	55
	Appendices.....	58

List of Figure

Figure 1.1 Mobile data traffic carried by serving RATs in Addis Ababa [4]	2
Figure 2.1 UMTS Network architecture [18].	9
Figure 2.2 UTRAN Logical Architecture [19].....	12
Figure 2.3 UMTS protocol stack [21].	14
Figure 2.4 UTRAN general protocol model [22].....	15
Figure 2.5 UMTS bearer service architecture [23].....	17
Figure 3.1 Flow graph of the decimation-in-time decomposition of an N-point DFT computation into two (N/2)-point DFT computations (N = 8) [25].	23
Figure 3.2 Flow graph of complete decimation-in-time decomposition of an 8-point DFT computation [25].	24
Figure 3.3 Basis functions of 2D-DCT with n and m equals to 8 [26]	26
Figure 4.1 Map of the city inscribed in square.....	31
Figure 4.2 Traffic of sample groups in spectral domain.....	32
Figure 4.3 Plot of real data vs. temporal model.....	33
Figure 4.4 Energy content of spatial parameters' spectrum.....	35
Figure 5.1 Model accuracy vs. Transformed coefficients' frequency truncation level.....	39
Figure 5.2 Model accuracy vs. Transformed phases' frequency truncation level.....	40
Figure 5.3 Performance of the model for different grouping	41
Figure 5.4 Model accuracy vs. spatial frequency truncation level	41
Figure 5.5 Real data traffic vs. model output for 9 pm	42
Figure 5.6 Real data traffic vs. model output for 1 pm	43
Figure 5.7 Real data traffic vs. model output for 4AM.....	43
Figure 5.8 Traffic distribution of the city	44
Figure 5.9 Probability density function of base stations peak hour traffic	45
Figure 5.10 Cumulative Distribution Function of NodeB's busy hour Traffic.....	45
Figure 5.11 Busy Hour Distribution of the groups.....	47
Figure 5.12 PHU (%) using average of traffic in space and distributed in time.....	48

Figure 5.13 PHU of the whole city (using average of traffic in both space and time).....	49
Figure 5.14 Plot of Ideal (optimal) network vs. Real network traffic.....	50
Figure 5.15 PHU of real and ideal network.....	51

List of Tables

Table 2.1 Releases of UTRAN air interface technologies.....	11
Table 2.2 UMTS QoS Parameters	17
Table 2.3 UMTS Traffic classes' properties and application.	18
Table 3.1 The horizontal and vertical frequency components of a block.....	25
Table 3.2 Quantization matrix for quality level of 50.....	27
Table 4.1 Parameters variation in space domain	33
Table 4.2 Matrix for average traffic (coefficient a_0) in MB of each group.....	34
Table 4.3 Matrix for α_1 (in rad) of each group.....	34
Table 4.4 Matrix of coefficient values for a_0 in spectral domain	36
Table 4.5 Matrix of coefficient values for α_1 (rad) in spectral domain.....	36
Table 5.1 Truncation level of spatial parameters.....	42
Table 5.2 PHU (%) using average of traffic in time and distributed in space	48
Table 5.3 Amplitude value of major frequency components.....	50

List of Acronyms

2D – DCT	Two Dimensional Discrete Cosine Transform
2G	2 nd Generation
3G	3 rd Generation
4G	4 th Generation
AR	Auto Regression
AS	Access Stratum
AuC	Authentication Center
CC	Call Control
CN	Core Network
CS	Circuit Switched
DCT	Discrete Cosine Transform
DFT	Discrete Fourier Transform
DIT	Decimation In Time
EIR	Equipment Identity Register
FFT	Fast Fourier Transform
FDD	Frequency Division Duplex
GMSC	Gateway Mobile Switching Center
GGSN	Gateway GPRS Support Node
GPRS	General Packet Radio System
GSM	Global System for Mobile communication
HLR	Home Location Register
HSPA	High Speed Packet Access
IP	Internet Protocol
IMS	IP Multimedia Subsystem
LTE	Long Term Evolution

MA	Moving Average
MAC	Medium Access Control
MM	Mobility Management
MSC	Mobile Switching Center
NAS	Non Access Stratum
NSS	Network Subsystem
PS	Packet Switched
PRS	Performance Reporting System
PSTN	Public Switched Telephone Network
QoS	Quality of Service
RAT	Radio Access Technology
RLC	Radio Link Control
RNC	Radio Network Controller
RRC	Radio Resource Control
SM	Session Management
SGSN	Serving GPRS Support Node
SPPP	Spatial Poison Point Process
TD-SCDMA	Time Division Synchronous Code Division Multiple Access
TDD	Time Division Duplex
UE	User Equipment
UMTS	Universal Mobile Telecommunication System
UTRAN	UMTS Terrestrial Radio Access Network
VLR	Visitor Location Register
WCDMA	Wideband Code Division Multiple Access

1. Introduction

The advent of new technologies, such as smartphones, has led data traffic carried by cellular networks to grow exponentially. According to the forecast study on global Internet Protocol (IP) traffic growth from Cisco, mobile data is a leading area with an estimated Compound Annual Growth Rate of 46 percent from 2016 to 2021 [1]. In order to fulfill such demand growth, in addition to the common trends such as network expansion and upgrade, there is a need to know how the network is being utilized, so that an optimal operation may be achieved. One way of getting the network utilization is through traffic modeling which can provide the traffic usage amount and peak hour distribution of the different coverage areas for different times of the day. These outputs of the modeling give an insight to manage the network resources effectively and meet the growing demand from other strategies such as dynamic resource allocation and energy efficient network planning. Creating an adaptive network that circulates its resources according to the demand makes the network more efficient and developing a marketing strategy that diverts customer usage behavior from peak to other hours helps network utilization to be uniformly distributed throughout the day by avoiding congestions. Additionally, in networks where congestion and utilization is not the issue, traffic modeling can be applied for energy efficient network planning and also for the routine operation and maintenance activities [2].

In cellular network planning, getting the right figure about the customer demand is a major factor to obtain an effective output. Modeling the data traffic and performing an appropriate prediction of the customer demand can definitely reduce the possible occurrence of network challenges related with Quality of Service (QoS) in the future. The first and most decisive point in the forecasting process is to find the data traffic model that reflects the customer demand in the real network [3].

Introduction

Currently in Ethio telecom, the sole telecom service provider in Ethiopia, cellular network data services are delivered through all (i.e., 2nd, 3rd and 4th Generations) Radio Access Technologies (RATs). As depicted in Figure 1.1, around eighty three percent of the data traffic in the capital, Addis Ababa, is carried by the 3rd Generation (3G) Universal Mobile Telecommunications System (UMTS) technology. This is the main reason for this research to target on the 3G network's data traffic.

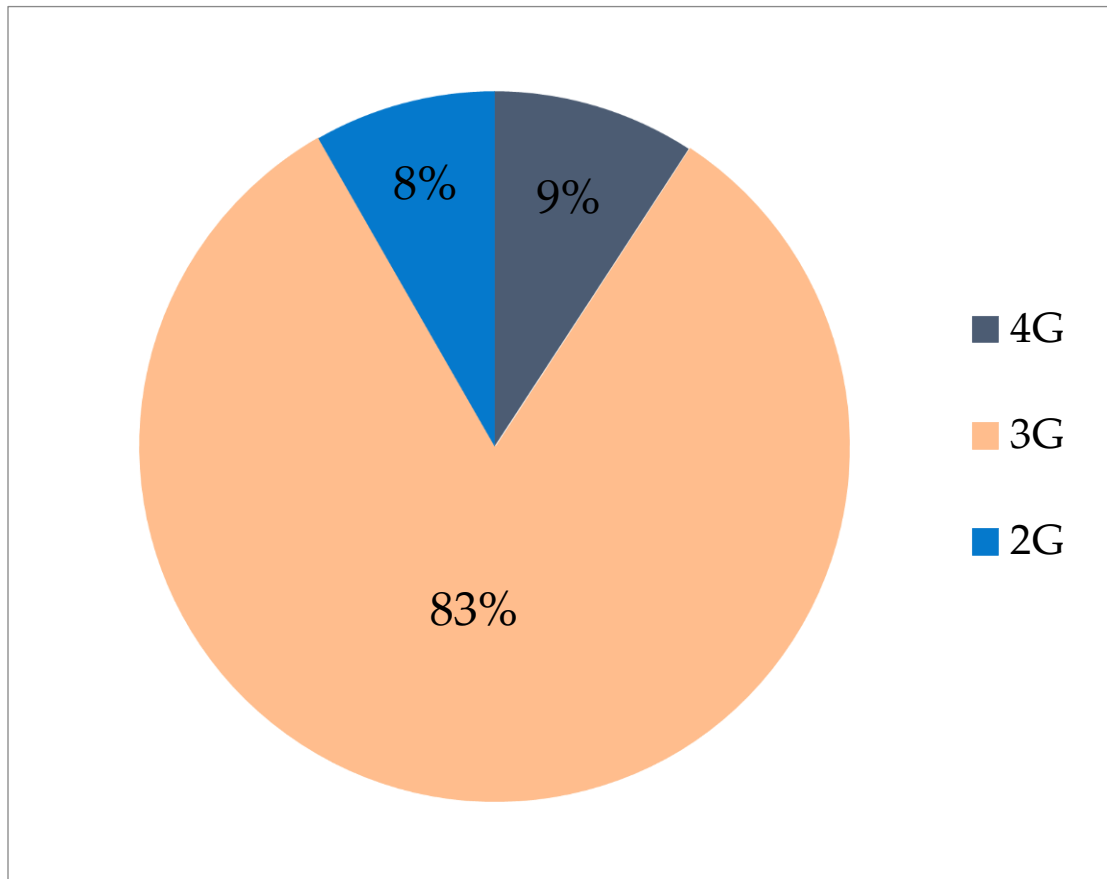


Figure 1.1 Mobile data traffic carried by serving RATs in Addis Ababa [4]

In this thesis, the geographical and temporal mobile data traffic volume distribution in Addis Ababa is analyzed and a model that can capture the traffic variation in both (i.e., geographical and temporal) domains is developed. As indicated above the vast majority of cellular network data traffic in the city is carried by the UMTS technology. So, the dataset related with UMTS network is used for the trend analysis and model development works of this research.

1.1 Statement of the Problem

One of the main factors that drive Ethio telecom towards network expansion and new project execution is the congestion level in the network. Even if it is a useful parameter in the optimization process, it doesn't show the actual network utilization level. Knowing the utilization level rather than the congestion invites other optional measures which can avoid unnecessary expenditure costs. Capturing traffic variation for different time of the service areas enable to directly acquire network utilization level.

Additionally, once expansion or new project execution is decided for the cellular network, the planning process in the company uses the marketing department's analysis results as a major input. The main drawback with this approach is the right figure about customer demand cannot be obtained and also it considers the data traffic usage distribution to be homogeneous throughout the coverage area.

So this thesis is intended to develop a model that can capture traffic variation for determining the network utilization level and to enable a heterogeneous traffic distribution assumption in planning.

1.2 Objective

1.2.1 General objective

The main objective of this research is to study and model the geographical and time variation of the UMTS network data traffic in Addis Ababa. For that, Fast Fourier Transform (FFT) is used for modeling the traffic in time domain and Two Dimensional Discrete Cosine Transform (2D-DCT) and its inverse in space domain.

1.2.2 Specific objectives

The specific aims of the research are:

- ❖ To analyze the dataset from the UMTS network in Addis Ababa and extract geographical and time behavior of customer data traffic usage.
- ❖ To study the application of Fourier Transform techniques such as 2D-DCT and FFT for traffic modeling.
- ❖ Developing a mathematical model that determines the UMTS network data traffic volume of any location in the capital and at any time.
- ❖ To evaluate the network utilization level, so that it may be used as an input for different application areas and in quantifying the customer data traffic usage behavior.
- ❖ Based on the network utilization result, forwarding recommendations for the operator to have an efficient network operation.
- ❖ To simulate the mathematical model of the real traffic by using Matlab.

1.3 Literature Review

For modeling data traffic in mobile cellular network, literatures have used different approaches. In [5] the traffic modeling is performed in space and time dimension where the traffic variation pattern for a single base station in real cellular network is modeled. The paper proposes sinusoid superposition model to represent the temporal traffic variation and lognormal distribution is used for spatial modeling. Based on this, a third combined spatial-temporal traffic modeling approach is developed. In this research a similar approach as the temporal modeling proposed in the paper is followed. But the difference is that they selected three types of areas and generalized the model with two major frequency components whereas in this research case the whole city's traffic is

Introduction

considered and the model is generalized for all areas of the city by taking four major frequency components.

Some other papers make the modeling either in time or space domain separately. Temporal modeling like sinusoid superposition as in [6] is performed where theoretical analysis for traffic profiling is done. To make this model more practical [7] and [8] added a poisson distributed random process over it. In [9] traffic distribution is modeled spatially by using Two-Dimensional Discrete Cosine Transform (2D-DCT). This approach is adopted in this research but the difference is that they performed it for determining the average traffic of a day for three selected regions in a city where as in this research the approach is applied to capture the spatial distribution of parameters from the temporal model that correspond to the different areas of the whole city. In [10], different statistical distributions are tested for their fitness to the spatial distribution of base stations' traffic density and found that log-normal mixtures fit in a better way than the others. Additionally, different modeling methods are proposed for the spatial domain such as Homogeneous spatial poisson point process (SPPP) [11], K-tier SPPP [12], lognormal distribution [13], exponential distribution [14] and 2-dimensional Gaussian model [15].

In regard with the forecasting model, there are previous works which are directly related with Addis Ababa UMTS network traffic demand forecasting and capacity planning. The paper in [16] has made the comparison between Seasonal Auto-Regression Integrated Moving Average (SARIMA) and Auto-Regression Integrated Moving Average (ARIMA) forecasting model in order to determine the one that best fits to the collected dataset. The result has shown that SARIMA is a better model for the city's UMTS network data traffic demand growth. Finally the result of SARIMA based demand forecasting is used for planning the network and a better UMTS data traffic model is proposed. But the main critique in the paper is that it directly used the raw

Introduction

dataset from core network which makes the usage behavior assumption to be homogeneous across the whole coverage area. In [17], the paper has done the forecast for voice and data traffic of UMTS network in Addis Ababa using artificial neural network. The dataset was collected from the core network element for 14 consecutive months. But this paper also doesn't consider the heterogeneity of data usage distribution in space.

1.4 Methodology

In order to conduct this research, initially the dataset for Addis Ababa's UMTS network PS traffic is extracted from Ethio telecom Performance Reporting System (PRS). Then literatures in the field of cellular network data traffic modeling and related area were reviewed. With an insight from the studies reviewed, the raw dataset was analyzed and model selection is performed separately for time and space domains where later both were combined to develop a single mathematical model that can simultaneously capture the data traffic variation of any location in the city at any time of the day.

1.5 Scope and limitation

1.5.1 Scope of the thesis

In this research, a mathematical model based on Fourier Transform techniques is developed. As the research is a case study, Addis Ababa city's UMTS network traffic is considered in the modeling and from the perspective of the services delivered by the network, the modeling is performed only for the data traffic.

1.5.2 Limitation of the thesis

The following were the main limitations in conducting this research.

- ❖ The dataset from the network system has missing traffic information up to a week period for some sites; as a result the modeling is performed for selected weeks with full information. This also limits the forecasting process for future work.
- ❖ Previous works on the city's UMTS Traffic were more of towards forecasting; as a result lack of similar papers that are directly related to this research was the other limitation.
- ❖ Due to time limitation forecasting couldn't be incorporated in the modeling.

1.6 Contribution

The main contributions of this research are:

1. A mathematical model that can capture Addis Ababa city's traffic variation is developed and analysis on how the different parameters in the model affect or tune the model performance for the city's traffic is performed.
2. This research has presented a new parameter (Peak Hour Utilization) for evaluating the network performance and used it to assess the status of the network so that the company may take an appropriate action.

1.7 Thesis Layout

The rest of the paper is organized as follows. In Chapter 2 detailed description about UMTS network is presented. In Chapter 3 Fundamentals of Fourier Transform are discussed. In Chapter 4, explanation about the dataset used for the research and procedures followed in developing the model are presented in detail. In Chapter 5, how the model accuracy varies with truncation level on spectral components of spatial parameters and with different number of grouping is analyzed and discussed. Finally, conclusion points are discussed in Chapter 6.

2 Universal Mobile Telecommunications System

2.1 Introduction

In the last three decades, mobile cellular network technology has evolved through four generations with the fifth one under development. Except for the first, all generations coexist in the current cellular network scenario. The major factor for this, but not limited to, is the tendency of the operators to protect their investment in the old generations. This made the second generation (2G) to become the long serving technology from any of the others with a current share of 25% of the total connected devices in mobile network only overtaken by third generation (3G) and fourth generation (4G) since 2017 [1]. Even if 3G's current share doesn't exceed 33% in terms of the total connected devices in the network, its emergence has brought a dramatic growth in the mobile broadband data/subscription [1]. This was mainly due to the advancement in the air interface technology and the upgrade on the core part of 2G.

Universal Mobile Telecommunication System (UMTS), which is a 3G technology and a direct evolution of Global System for Mobile Communication (GSM), has two air interface technologies: Wideband Code Division Multiple Access (WCDMA) and Time Division Synchronous Code Division Multiple Access (TD-SCDMA). The difference between the two is that WCDMA utilizes much wider band and operates on both Time Division Duplex (TDD) and Frequency Division Duplex (FDD) whereas TD-SCDMA uses TDD only. Except in China, more specifically China Telecom, networks all over the world use WCDMA as the air interface technology of their UMTS system. The main reason China developed and deployed TD-SCDMA is to avoid paying highly expensive patent fees related with WCDMA. Even if they have a remarkable difference in their techniques, there is no difference between the architecture of the networks they belong to.

2.2 UMTS Network Architecture

The first UMTS network was developed as a universal infrastructure that can carry services in the past generations as well as services from future evolution. This was achieved through dividing the overall network into different technologies such as access, transport, service and user applications. Due to such subdivisions of the network technologies in the system, the architecture of UMTS network can be modeled in many ways. As shown in Figure 2.1, the more general architecture mainly has two parts: Core Network (CN) and UMTS Terrestrial Radio Access Network (UTRAN). Sometimes, User Equipment (UE) is also included in to the architecture as a compliment to the system.

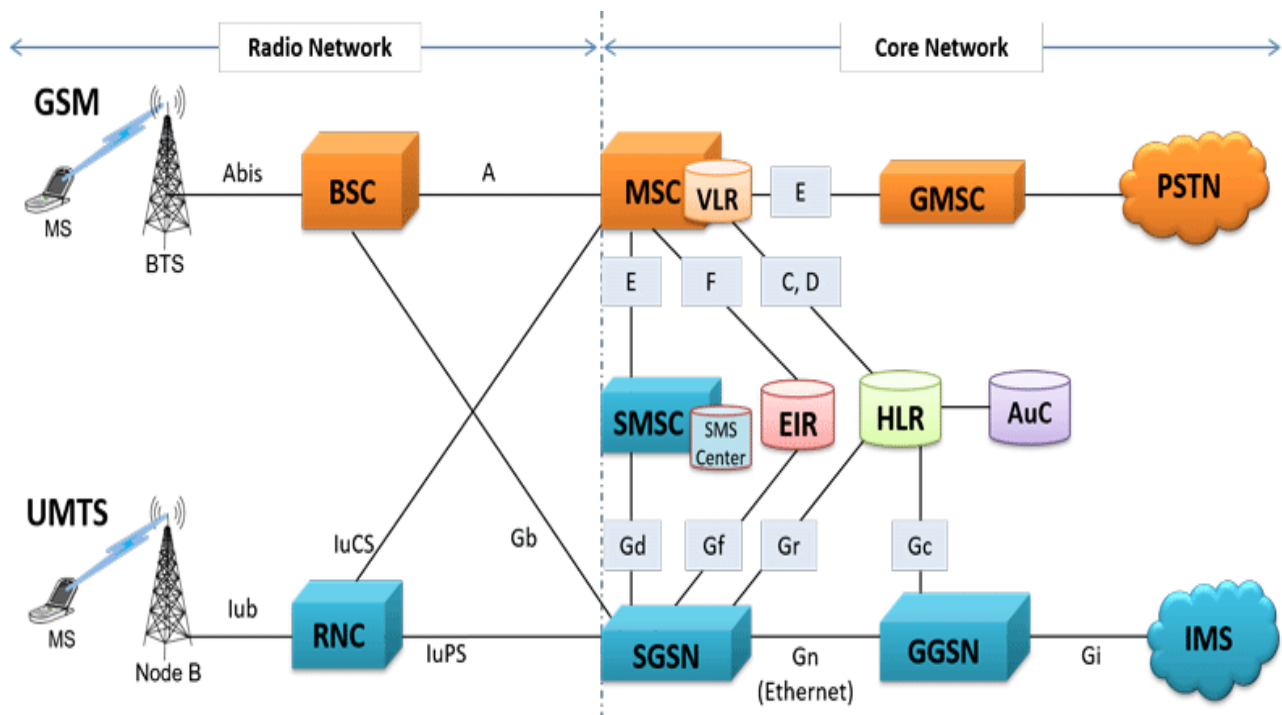


Figure 2.1 UMTS Network architecture [18].

2.2.1 UMTS Core network

The basic functions of the core network are switching and routing calls and data connections to and from external networks. It is divided into two functional modules which are circuit switched (CS) and packet switched (PS) domains. Network elements such as Mobile Switching Center (MSC), Gateway MSC (GMSC) and Visitor Location Register (VLR) are in the CS domain whereas Serving GPRS Support Node (SGSN) and Gateway GPRS Support Node (GGSN) belong to PS domain. Equipment Identity Register (EIR), Authentication Center (AuC) and Home Location Register (HLR) are common entities that belong to both domains. The core network uses GMSC as an attachment point of the CS part to connect Public Switched Telephone Network (PSTN) and GGSN to connect its PS part to IP Multimedia Subsystem (IMS).

2.2.2 UMTS Terrestrial Radio Access Network

The UMTS Terrestrial Radio Access Network (UTRAN) is the network part that facilitates the communication between UE and other network entities over the radio. It has two parts: Radio Network Controller (RNC) and Node B.

1. **RNC:** some of the responsibilities of this entity are radio resource management, a part of mobility management, data encryption/decryption and managing NodeBs which are connected to it. This network entity makes a communication with the core network, NodeB under it and with another RNC for the purpose of mobility management.
2. **Node B:** this is a base station in UMTS network where the air interface starts. It has transceivers for communicating user equipment. The air interface of UMTS has evolved through different releases without making any significant change on the CN part. Almost all upgrades are made on the data traffic serving section of the air interface and a summary is presented in Table 2.1.

Table 2.1 Releases of UTRAN air interface technologies.

Release	Technology	Maximum Up link Speed (Mbps)	Maximum Down link Speed (Mbps)
99	WCDMA	0.384	0.384
5	HSPA	0.384	14.4
6	HSPA	5.8	14.4
7	HSPA+	12	28
8	HSPA+	12	42
9	HSPA+	24	84
10	HSPA+	24	168
11	HSPA+	72	336

The two network entities, Node B and RNC, forms Radio Network Subsystem (NSS) and the general architecture for UTRAN looks as presented in Figure 2.2.

UTRAN interfaces

- Uu – This is an interface through which the UE accesses the fixed part of the system and it is usually called as ‘air interface’.
- Iub – an interface that connects NodeB with RNC
- Iur – this interface communicates one RNC with the other. Its main function is facilitating the soft handover between NodeB’s of different RNC.
- Iu-CS – an interface, that is in between RNC and the CS domain of the core network.
- Iu-PS – this is also for connecting RNC with core network but for PS domain.

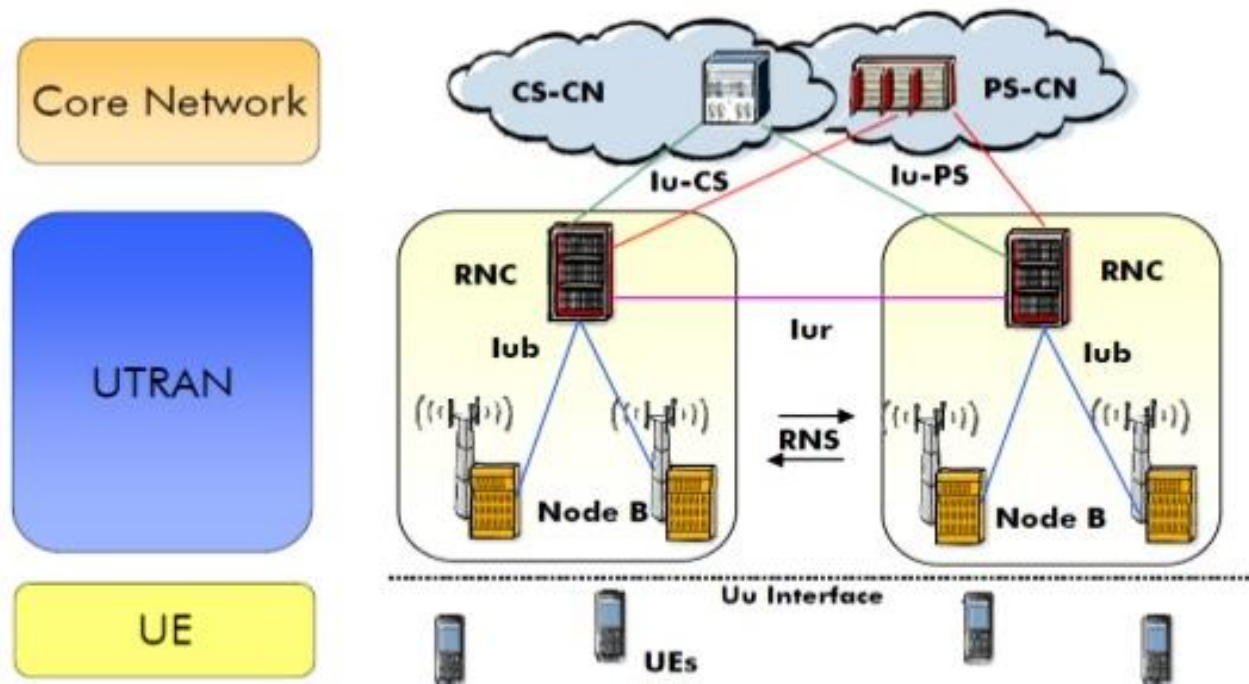


Figure 2.2 UTRAN Logical Architecture [19].

2.3 UMTS Network Functions and Protocols

2.3.1 UMTS Functions

Mobile networks are architecture-centric as architectures are to be designed first. Network architecture show what functions a network must provide, e.g., data forwarding or mobility support, and how these functions are grouped in the network [20]. It creates functional groups that should be connected through protocols which are to be designed later. The following are some of the functions in mobile network which can be grouped in UMTS architecture [20].

- *Transport* – is a network functionality which enables movement of information from one network element to the other.
- *Routing* – this function deals with selecting path for the transportation of traffic between different elements in the cellular network.

Universal Mobile Telecommunications System

- *Security* – solves safety issues from both the network and users perspective. From the network side it checks whether a user has the right to access or not and from the user side it filters trustworthy connections. Additionally security functions keep protected the users' privacy, such as identity and location.
- *Session control or call control* – are responsible for monitoring the communication between two users or devices after the call is set up or session is established. Most of the time word session is used in packet-switched communications whereas call is in circuit switched.
- *Quality of Service* – These functions ensure delivery of a service in the range of predefined agreement levels or network standards.
- *Radio resource control and provisioning* – mainly acts in the communication between the user equipment and NodeB. It is for setting up, modifying and releasing resources in different layer of the radio interface protocol stack.
- *Mobility* – this function enables the system to maintain a session or call while the user is moving across the network by making new connections.
- *Charging* – this is a billing function which controls the users' payment for the service they got from the network.

2.3.2 UMTS Protocols

Protocols are a set of rules and guidelines for communicating data whereas protocol stack is a group of protocols that perform concurrently for the purpose of understandable communication among the involved entities. As shown in Figure 2.3, UMTS protocol stack has two layers: Non Access Stratum (NAS) and Access Stratum (AS).

- **NAS** – is set of protocols that carry signals directly from UE to the core network and vice versa without being processed by the UTRAN. Functions such as

Universal Mobile Telecommunications System

Mobility Management (MM), Call Control (CC), and Session Management (SM) are in this category.

- **AS** – is for signaling between UE and UTRAN. Functions such as Radio Resource Connection and paging are among this type.

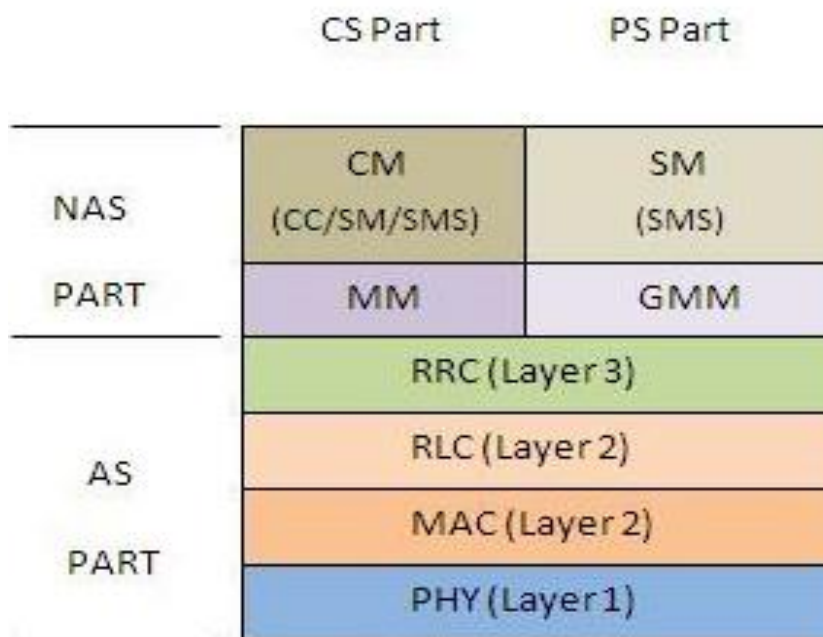


Figure 2.3 UMTS protocol stack [21].

From the overall UMTS protocol stack, only the first three layers belong to UTRAN. The general model for UTRAN protocol is as shown in Figure 2.4; it has two horizontal layers and two vertical planes. Planes and layers are independent to each other so that changes on one may not affect the other. The horizontal layer consists of radio link layer and transport link layer while the vertical plane consists of the control plane and user plane.

The protocols that constitute the UTRAN stack are:

Medium Access Control (MAC) protocol: This carries out low-level manipulation and control of the physical layer and has the following functionalities

- ❖ Mapping between logical channels and transport channels;
- ❖ Priority handling between UEs by means of dynamic scheduling;

Universal Mobile Telecommunications System

- ❖ Ciphering. This function prevents unauthorized acquisition of data;
- ❖ Traffic volume measurement. Measurement of traffic volume on logical channels and reporting to RRC.

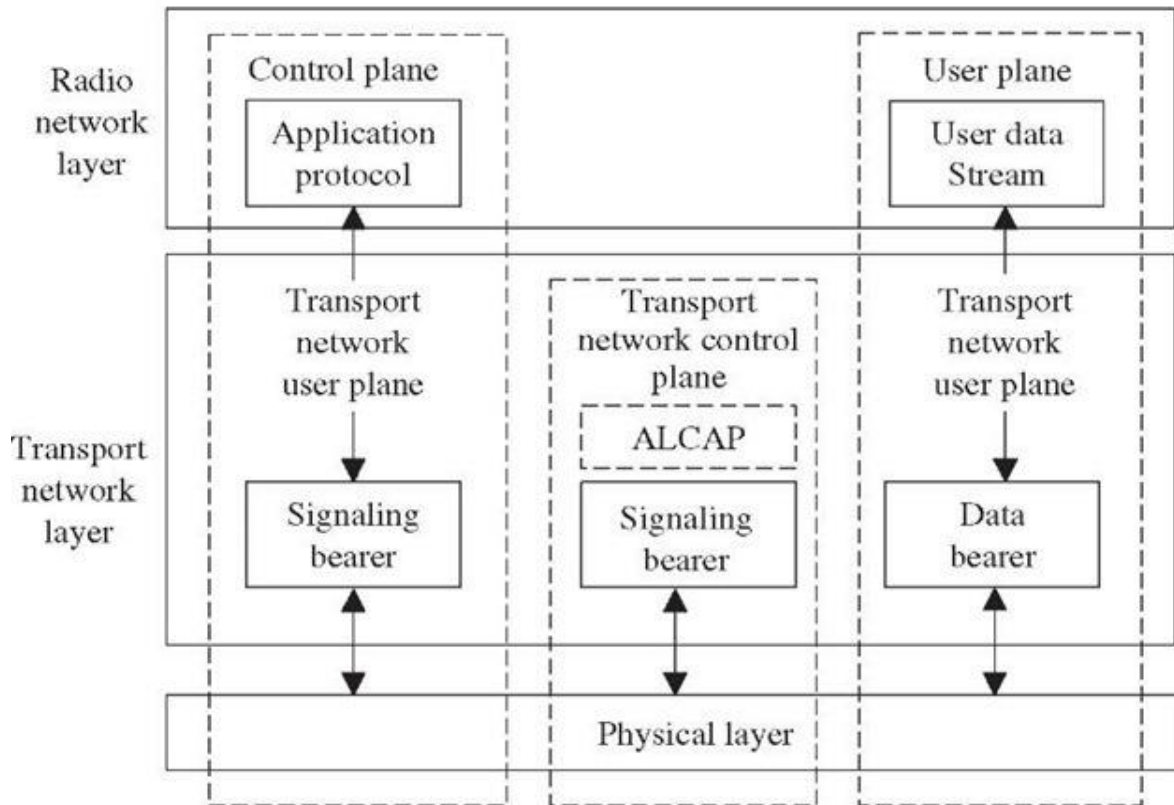


Figure 2.4 UTRAN general protocol model [22].

Radio Link Control (RLC) protocol: manages the radio connection between user equipment and the network. Some of its functions are:

- ❖ Making data suitable for transmission over air interface by segmenting and reassembling in the reverse.
- ❖ Error correction
- ❖ Padding
- ❖ Sequence number check
- ❖ Flow control

Packet Data Convergence (PDC) protocol: this is used as an interface for packet data and its main functions are:

Universal Mobile Telecommunications System

- ❖ Transfer of user and control plane data
- ❖ Header compression and decompression of IP data streams
- ❖ Ciphering
- ❖ Maintenance of PDCP sequence numbers

Broadcast/Multicast Control (BMC) protocol: this protocol is used in broadcasting messages from the network to UEs and collecting them at the UE. Its functionalities include:

- ❖ Storage of Cell Broadcast Messages
- ❖ Scheduling of BMC messages
- ❖ Transmission of BMC messages to UE
- ❖ Delivery of Cell Broadcast messages to upper layer

Radio Resource Control (RRC) protocol: It sets up, reconfigure and reestablish the signalling messages that are exchanged between the mobile and the radio access network. Its main functions are

- ❖ Cell Broadcast Service (CBS) control
- ❖ Initial cell selection and cell re-selection
- ❖ Sets up and release RRC connections and paging
- ❖ Controls Timing advance
- ❖ Outer loop power control
- ❖ Slow Dynamic Channel Allocation (DCA)

2.4 UMTS Quality of Services

There are two kinds of services in Telecommunication: Teleservice and Bearer service. Teleservice is a telecommunication service provided by the network for the communication between end users whereas bearer service is a telecommunication service that is used for the transfer of user and control data between network equipment's. The layered architecture for UMTS bearer service is given in Figure 2.5.

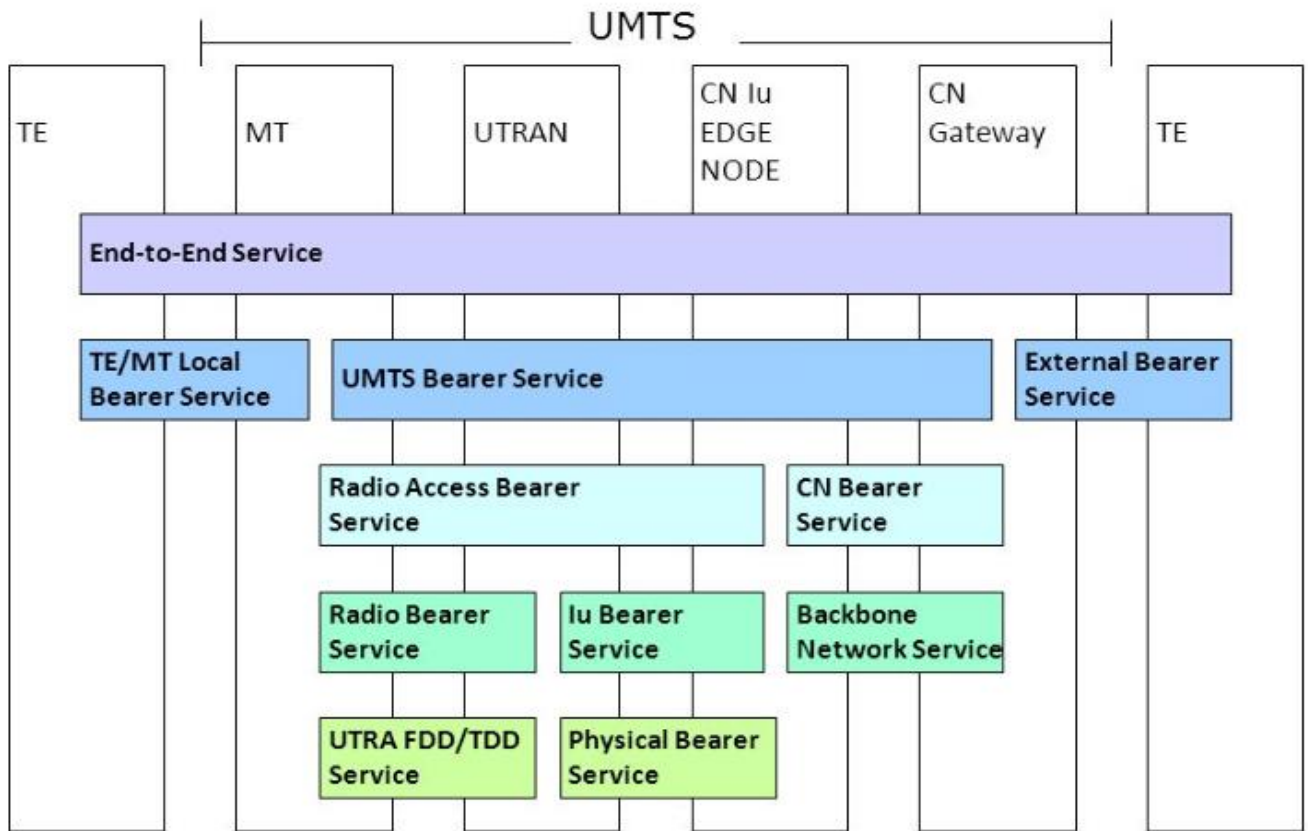


Figure 2.5 UMTS bearer service architecture [23].

A bearer service includes all aspects to enable the provision of a contracted QoS. As defined in [23], Quality of Service (QoS) is the mechanism of ensuring that a service can be delivered to the end user in an acceptable time-frame and that the service properties are stable overtime within predefined boundaries. QoS in UMTS has parameters and major ones are explained as given below in Table 2.2.

Table 2.2 UMTS QoS Parameters

Parameters	Explanation
Maximum bit rate	Defines the maximum bit rate when delivering information between end points of UMTS bearer
Guaranteed bit rate	Defines the bit rate that the UMTS bearer must carry between its end points
Allowed transfer delay	Set the limits of the delay
QoS negotiable	Describes the negotiability level of the service

Universal Mobile Telecommunications System

Based on the above mentioned basic and other parameters, for UMTS bearer services there are four different QoS classes (or traffic classes). They are summarized in Table 2.3 with their properties and application services.

Table 2.3 UMTS Traffic classes' properties and application.

QoS Classes	Properties	Example application
Conversational	Minimum fixed delay, no buffering, symmetric traffic, guaranteed bit rate	Speech and video call
Streaming	Minimum variable delay, buffering allowed, asymmetric traffic, guaranteed bit rate	Real time streaming video
Interactive	Moderate variable delay, buffering allowed, asymmetric traffic, no guaranteed bit rate	Web surfing
Background	Big variable delay, buffering allowed, asymmetric traffic, no guaranteed bit rate	File downloading, e-mail

3 Fundamentals of Fourier Analysis

3.1 General Introduction

Fourier analysis is a method of defining signals as a sum of trigonometric functions where signals are transformed from time to frequency domain and vice versa. Mathematical approaches such as Fourier Series and Fourier Transform are among the many to convert signals from time domain to frequency. Both are under Fourier Analysis in which signals are defined as a sum of trigonometric functions. The difference between the two is that the former is for periodic signals whereas the latter is for non-periodic signals.

3.1.1 Fourier Series

The basis for Fourier to show that there are infinite series expansions over other functions such as sines and cosines, is the study of heat and wave equations [24]. In addition to periodic function, Fourier Series is also applicable for a function that is finite in length. Values in the series that are out of the finite interval are less significant. Its basics can be summarized as follows:

A periodic function $f(t)$ with a period of T is defined as

$$f(t) = f(t + T) \quad (3.1)$$

As $f(t)$ is a periodic function, analyzing its behavior over the interval of one period (t_0 to t_0+T) is enough. So the Fourier Series of $f(t)$ is expressed as the infinite sum of sines and cosines as given in (3.2)

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos \frac{2\pi nt}{T} + b_n \sin \frac{2\pi nt}{T} \right) \quad (3.2)$$

Where a_n and b_n are coefficients of the function and can be evaluated using (3.3) and (3.4)

$$a_n = \frac{1}{T} \int_{t_0}^{t_0+T} f(t) \cos \frac{2\pi nt}{T} dt \quad (3.3)$$

$$b_n = \frac{1}{T} \int_{t_0}^{t_0+T} f(t) \sin \frac{2\pi nt}{T} dt \quad (3.4)$$

And the more compact exponential form of the Fourier Series and the coefficients is given below in (3.5) and (3.6) respectively

$$f(t) = \sum_{n=-\infty}^{\infty} F_n e^{i\omega n t} \quad (3.5)$$

$$F_n = \frac{1}{T} \int_{t_0}^{t_0+T} f(t) e^{-i\omega n t} dt, \text{ where } \omega = \frac{2\pi}{T} \quad (3.6)$$

The main advantage of expressing functions in terms of sum of cosines and sines, even in exponentials is their simplicity for mathematical operations such as derivation, integration and differentials.

For a discrete function $x[n]$ with N number of elements, the discrete Fourier series and its coefficients are as given in (3.7) and (3.8)

$$x[n] = \sum_{t=0}^{N-1} X_t e^{i\omega n t} \quad (3.7)$$

$$X_t = \frac{1}{N} \sum_{n=0}^{N-1} x[n] e^{-i\omega n t} \quad (3.8)$$

3.1.2 Fourier Transform

Fourier Transform can be expressed as a generalization of a Fourier Series by extending its period of the integral to infinity, which enables it to represent non periodic signals with a series of periodic functions.

For a continuous function $f(t)$, the Fourier Transform and its inverse are as given in (3.9) and (3.10)

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \quad (3.9)$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{-i\omega t} d\omega \quad (3.10)$$

For a discrete function $x[n]$ with N number of elements, the Discrete Fourier Transform and its inverse are as given in (3.11) and (3.12) respectively

$$x[n] = \sum_{k=0}^{N-1} X_k e^{i2\pi k n / N}, k = 0, 1, 2 \dots N - 1 \quad (3.11)$$

$$X_k = \frac{1}{N} \sum_{n=0}^{N-1} x[n] e^{-i2\pi kn/N}, k = 0, 1, 2 \dots N - 1 \quad (3.12)$$

Overall, based on the function's periodicity and continuity in time domain, Fourier transforms can be classified in to four types.

1. Aperiodic continuous time signal and continuous, aperiodic spectrum
2. Periodic continuous time signal and discrete, aperiodic spectrum
3. Aperiodic discrete time signal and continuous, periodic spectrum
4. Periodic discrete time signal and discrete periodic spectrum

As computer calculations usually require digital inputs, most signals are presented in discrete form. As a result Discrete Fourier Transform (DFT) is relatively more applicable than the others. There are different version of DFT with their own upgrade on the original approach, such as Fast Fourier Transform (FFT) and Discrete Cosine Transform (DCT).

Fast Fourier Transform: this is an upgrade of DFT which reduces the required computational step from N^2 to $N \log_2 N$.

Discrete Cosine Transform: Basically, DFT is a combination of even and odd functions. When the odd part is zero DFT becomes DCT. There are eight variants of DCT; however the most common one is DCT-II. As FFT and DCT-II are the core of this research, the rest of this chapter is dedicated for discussing them.

3.2 Fast Fourier Transform

Technologically, Fourier transform is one of the core sciences in cellular systems as it makes filtering, modulation, sampling and digitizing processes to be realizable. The latest application to be mentioned is the deployment of FFT and its inverse algorithms for Long Term Evolution's OFDM scheme. The following mathematical explanation shows how FFT has got its computational speed for enhancing DFT.

DFT's equation in (3.13) can be rewritten as:

$$X[k] = \begin{bmatrix} 1 & e^{-i2\pi k/N} & e^{-i4\pi k/N} & \dots & e^{-i2\pi k(N-2)/N} & e^{-i2\pi k(N-1)/N} \end{bmatrix} \begin{bmatrix} x[0] \\ x[1] \\ \vdots \\ \vdots \\ x[N-1] \end{bmatrix} \quad (3.13)$$

By substituting $e^{-i2\pi/N} = W_N$, we will have W_N matrix as given below in (3.14)

$$W_N = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & W_N & W_N^2 & \dots & W_N^{(N-1)} \\ 1 & W_N^2 & W_N^3 & \dots & W_N^{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & W_N^{(N-1)} & W_N^{(N-1)*2} & \dots & W_N^{(N-1)(N-1)} \end{pmatrix} \quad (3.14)$$

Since W_N is an N by N matrix and $x[n]$ is a 1 by N matrix, their product requires N^2 multiplication and N^2 additions. This makes DFT to be impractical as it has a slow computation and got complicated when length of the sequence increases. As a result FFT, which is a numerically efficient method to calculate DFT, is developed. Even if there are many algorithms in FFT, here Decimation in Time (DIT) is used to prove. In this algorithm a big DFT is built from smaller ones and initially assumes N to be a power of 2.

First it separates the time sequence data $x[n]$ into even and odd-indexed subsequence as given in (3.15).

$$X[k] = \sum_{r=0}^{N/2-1} x[2r]W_N^{k2r} + \sum_{r=0}^{N/2-1} x[2r+1]W_N^{k(2r+1)} \quad (3.15)$$

This can also be expressed as

$$X[k] = \sum_{r=0}^{N/2-1} x[2r](W_N^2)^{kr} + W_N^k \sum_{r=0}^{N/2-1} x[2r+1](W_N^2)^{kr} \quad (3.16)$$

But $W_N^2 = e^{-\frac{i2\pi}{N}2} = e^{-i2\pi/(\frac{N}{2})} = W_{N/2}$. As a result (3.16) will be a sum of even sequence DFT and odd sequence DFT as shown in (3.17)

$$X[k] = X_{ek} + W_N^k X_{ok} \quad (3.17)$$

(3.17) is the sum of two $N/2$ point DFTs, this means two $(N/2)^2$ multipliers corresponding to each sequence and N multipliers corresponding to their addition in total $(2*(N/2)^2 + N)$ multipliers are required. By applying the same approach of splitting up to n levels where $2^n=N$ (as this is initially assumed), the required number of multipliers will be $2^n *(N/2^n)^2 + n * N$. Substituting $2^n= N$ and $n = \log_2 N$, the final number of multipliers will be $N+N\log_2 N$. For example as discussed in [23] for $N=8$ the first and the complete flow graph of the DIT decomposition is presented in Figure 3.1 and Figure 3.2.

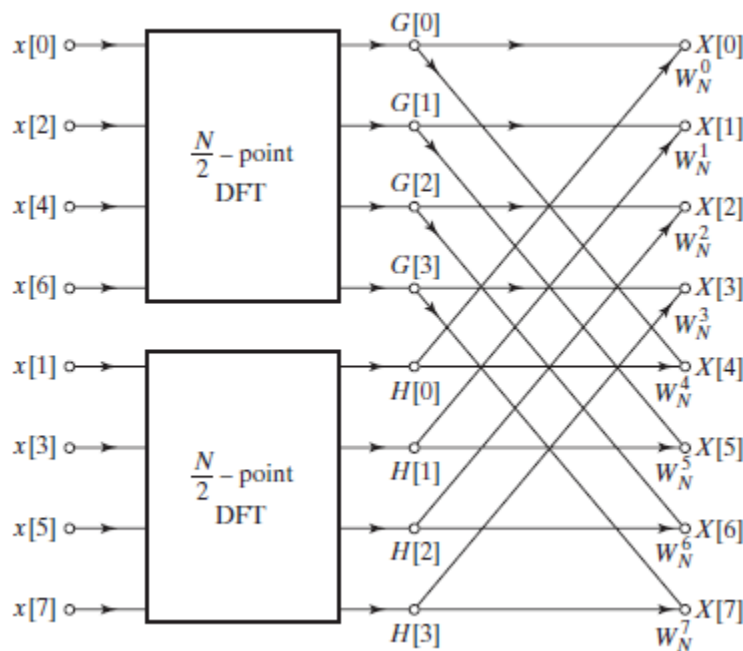


Figure 3.1 Flow graph of the decimation-in-time decomposition of an N -point DFT computation into two $(N/2)$ -point DFT computations ($N = 8$) [25].

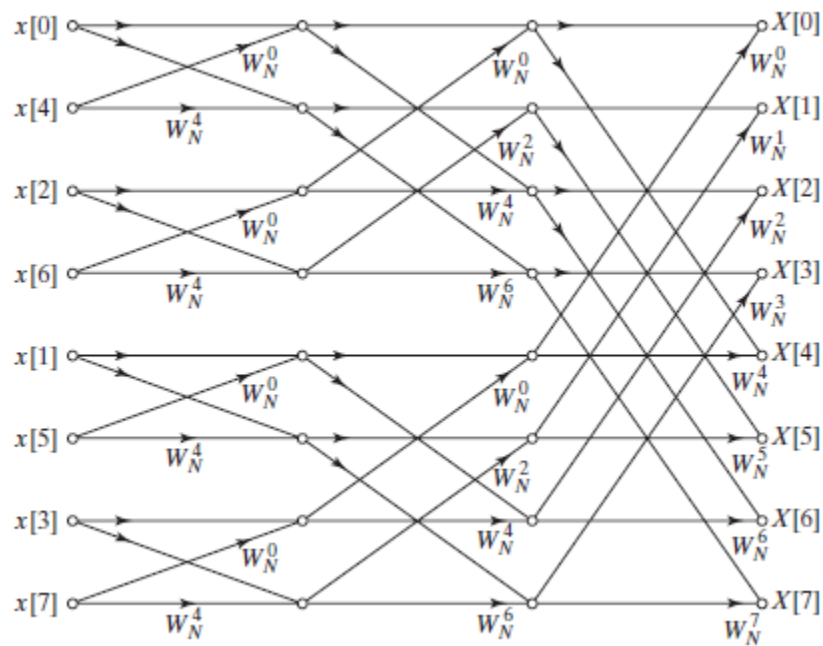


Figure 3.2 Flow graph of complete decimation-in-time decomposition of an 8-point DFT computation [25].

For a larger N it approximately equals to $N \log_2 N$, which is much less than N^2 . This computational efficiency and capability to process multiple sequences simultaneously enabled to be adopted in LTE's air interface multiple access technology.

3.3 Two dimensional Discrete Cosine Transform

Two dimensional discrete cosine transform is a modification of two dimensional DFT where only the cosine (or real) part of the equation is used. It is a highly applicable mathematical approach in multimedia (i.e. image and video) compression algorithms. As image is the basis for video, explanation related with image is enough for the discussion at this level. The image compression process which is used by standards such as JPEG is presented as follows.

1. The image is broken in to multiple blocks where every block is again divided in to an 8×8 pixels.
2. Each pixel of the block will be represented with their respective color value and this creates an 8 by 8 matrix of discrete numbers corresponding to each block.

3. Applying 2D – DCT, each block of the image will be transformed to frequency domain and coefficient of the frequency components will be determined.
4. Each block's coefficient in the frequency domain will be compressed by applying quantization and coding.
5. The array of compressed blocks that constitute the image is stored in a drastically reduced amount of space.
6. And finally if a need is there to get the image back, the inverse 2D-DCT will be applied on the stored data.

The forward and reverse equation of 2D-DCT is as given in (3.18) and (3.19) respectively

$$F(n, m) = \frac{2C(n)C(m)}{\sqrt{NM}} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f(x, y) \cos \left[\frac{(2x+1)\pi n}{2N} \right] \cos \left[\frac{(2y+1)\pi m}{2M} \right] \quad (3.18)$$

$$f(x, y) = \frac{2}{\sqrt{NM}} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} C(n)C(m)F(n, m) \cos \left[\frac{(2x+1)\pi n}{2N} \right] \cos \left[\frac{(2y+1)\pi m}{2M} \right] \quad (3.19)$$

$$\text{Where } C(n), C(m) = \begin{cases} \frac{1}{\sqrt{2}}, & \text{for } n, m = 0 \\ 1, & \text{otherwise} \end{cases}$$

In image compression, as the blocks are divided in to 8 by 8 pixels, both N and M are equal to 8. The corresponding frequency components are a combination of f_n and f_m

$$f_x = \frac{x}{2N} \quad \text{Where } x=0, 1, 2, \dots, 7 \quad \text{for both } n \text{ and } m \quad (3.20)$$

Table 3.1 The horizontal and vertical frequency components of a block

n=x	Frequency component(f_x)
0	0
1	0.0625
2	0.1875
3	0.25
4	0.3125
5	0.375
6	0.4375
7	0.5

Fundamentals of Fourier Analysis

The frequency components in Table 3.1 are numerical representation in one dimension but practically as images are two dimensional, there should be spatial frequency components and it is a combination of horizontal and vertical components which looks like the one displayed in Figure 3.3 these are commonly called as basis functions of the image in frequency domain and output of 2D-DCT provides coefficients for these basis functions.

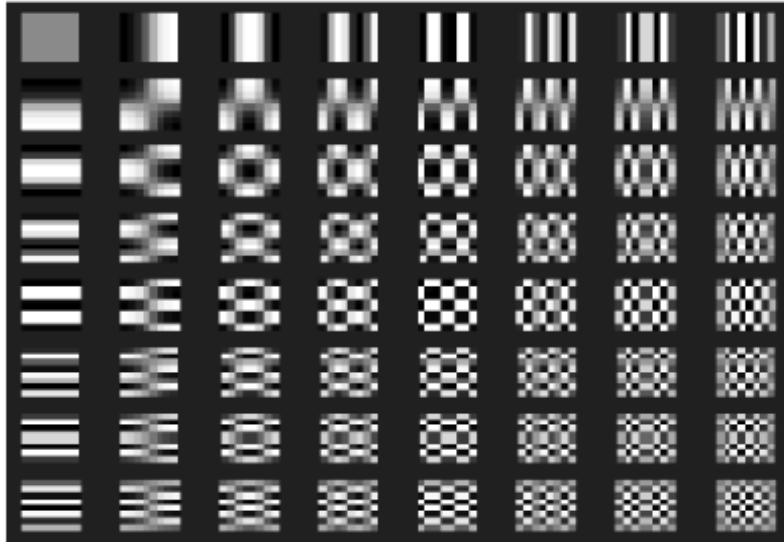


Figure 3.3 Basis functions of 2D-DCT with n and m equals to 8 [26]

After extracting coefficients of the basis functions using the transform equation, the two processes namely quantization and coding will be performed on the output values consequently. For example JPEG uses a standard quantization matrix which can vary according to the different combination need of quality and compression. Subjective experiments involving the human visual system has resulted in the JPEG standard quantization table that has 100 quality levels. A sample quantization matrix with quality level 50 is presented below in Table 3.2.

Table 3.2 Quantization matrix for quality level of 50

$$Q_{50} = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}$$

Finally by applying one of the source coding techniques, a compression process will be completed.

4 UMTS Data Traffic Modeling Using Fourier Transform

4.1 Basics of Modeling

Modeling is a simplified representation of a system at some particular point in time or space intended to promote understanding of the real system [27]. The representation could be physical, mathematical or otherwise logical. In addition to a system, the representation could be for an entity, phenomenon or a process. In this research, since the dataset used is a time series data and as the modeling is performed in both space and time, their basics are discussed in the below sections.

4.1.1 Time Series Analysis and Modeling

Time series is a sequential set of data points related with a variable and measured at equally spaced time intervals. Time series analysis accounts for the fact that data points taken over time may have an internal structure (such as autocorrelation, trend or seasonal variation) that should be accounted for [28]. Time series data can either be univariate when recorded sequence is for a single observation or multivariate when it is for multiple set of observations. Time Series Analysis is used for many applications such as: Economic forecasting, sales forecasting, stock market analysis and so on.

Time series modeling is a way of developing an equation that fits a time series data and according to [28] its usage is twofold:

1. Obtain an understanding of the underlying forces and structure that produced the observed data.
2. Fit a model and proceed to forecasting, monitoring or even feedback and feed forward control.

There are a number of approaches to modeling time series. Here a few of the most common approaches are outlined.

1. One approach is to decompose the time series into a trend, seasonal and residual component.
2. Frequency-based methods: This approach is commonly used in scientific and engineering applications by analyzing the series in the frequency domain.
3. Autoregressive (AR) model: This is used for predicting a time series data based on past behaviors of the dataset.
4. Moving average (MA) model: in simple term it is a past error multiplied by a coefficient.
5. Box–Jenkins Approach: this is a combination of moving average and autoregressive approaches.

From the above list of approaches, the second one is used for time domain modeling in this research.

4.1.2 Spatial Modeling

Spatial analysis is a study that provides a unique set of techniques and methods for analyzing events that are located in geographical space. It is widely involved into different number of research areas; as a result it is difficult to get a generalized classification of the approaches in the field. So, it is good to narrow the discussion towards the interest of this research. In mobile network technology, most of the spatial modeling studies concentrate on the statistical distribution of users, base stations and traffic. The four most common ones as summarized in [9] are presented here.

1. Spatial Poisson Point Process distribution model: Spatial point process is a random pattern of points in d -dimensional space and it will be SPPP when the pattern takes poisson distribution. Using the approach in different manner enables the modeling to capture both uniform and non-uniform distribution of a system.

2. Lognormal distribution model: In probability theory a log-normal distribution is a continuous probability distribution of a random variable whose logarithm is normally distributed. Models that use this distribution in cellular technology can be applied to generate traffic for the purpose of simulation, to estimate user distribution in a network and also to determine traffic distribution in a coverage area.
3. Exponential distribution model: The exponential distribution is used to model the time between the occurrence of events in an interval of time, or the distance between events in space.
4. Gaussian distribution model: in this model the system under study has a normal distribution. The system can be analyzed in d-dimensions and the model will be called d-dimensional Gaussian distribution model.

4.2 Model Development procedure

For this research, the dataset is collected from Ethio telecom's UMTS network in Addis Ababa city. It contains data traffic volume (MB) served by each NodeB for every hour of a week period and the coordinate information (in degrees) of 734 base stations located in the city.

The modeling process has gone through the following steps.

1. Firstly, as shown in Figure 4.1, the minimum rectangular area that can fit all the boundaries of the city is selected. The dimension of the city is 30km by 30km as a result the inscribing area is a square of the same dimension.
2. Each side of the square is equally divided into N parts creating a matrix of N by N, which reduced the space domain by a factor of $1/N^2$ and also the average number of sites in each cell by the same factor.

- NodeB located in the same cell of the matrix are grouped together and their data traffic of every hour is summed up so that groups may be represented with a single traffic profile.

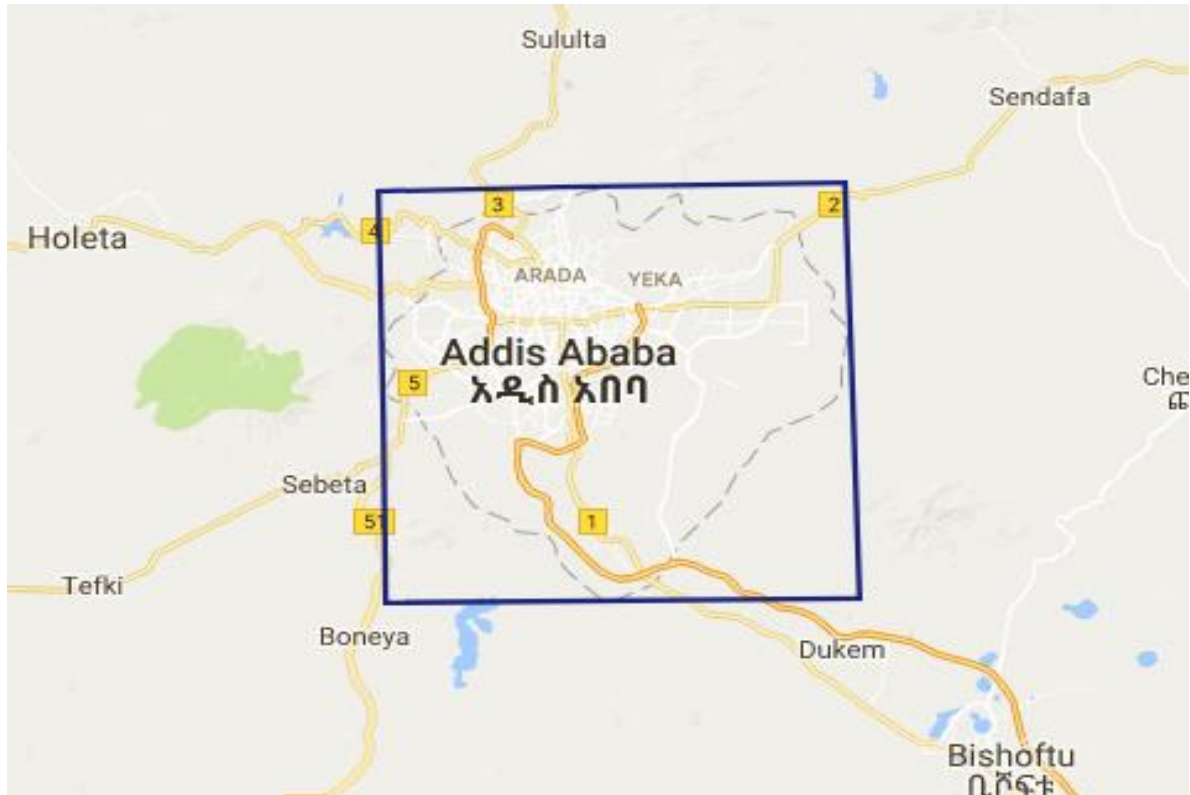


Figure 4.1 Map of the city inscribed in square.

4.2.1 Traffic Model in Time

By applying FFT, each group's data traffic is transformed from time to spectral domain where major frequency components are extracted and used in the inverse transform to develop a Fourier series which is the temporal model as shown in (4.1).

$$Tg(t) = a_0 + \sum_{i=1}^4 (a_i \sin(2\pi f_i t + \alpha_i)) \quad (4.1)$$

Where $Tg(t)$ is the total traffic of a specific group, a_0 is the average traffic of the group, f_i is the frequency components of traffic variation, a_i and α_i are the amplitudes and phases corresponding to the major frequency components, i is the number of major frequency components.

UMTS Data Traffic Modeling Using Fourier Transform

This made the groups to be identified with three parameters: frequency, coefficient and phase.

Frequencies: determine the hidden pattern of customers' data usage behavior in a day.

Coefficients: determine the maximum data traffic volume of the corresponding patterns (or Amplitude of the frequency components)

Phases: give us information about the position of the starting point of the patterns in a day. These are the parameters that help in knowing peak hours of the different areas in the city.

As shown in Figure 4.2, four major frequency components which are similar for all groups are identified from the spectral domain analysis. These are: $f_1=1/24$ (or $T=24$ hour), $f_2=1/12$ ($T=12$), $f_3=1/8$ ($T=8$) and $f_4=1/6$ ($T=6$).

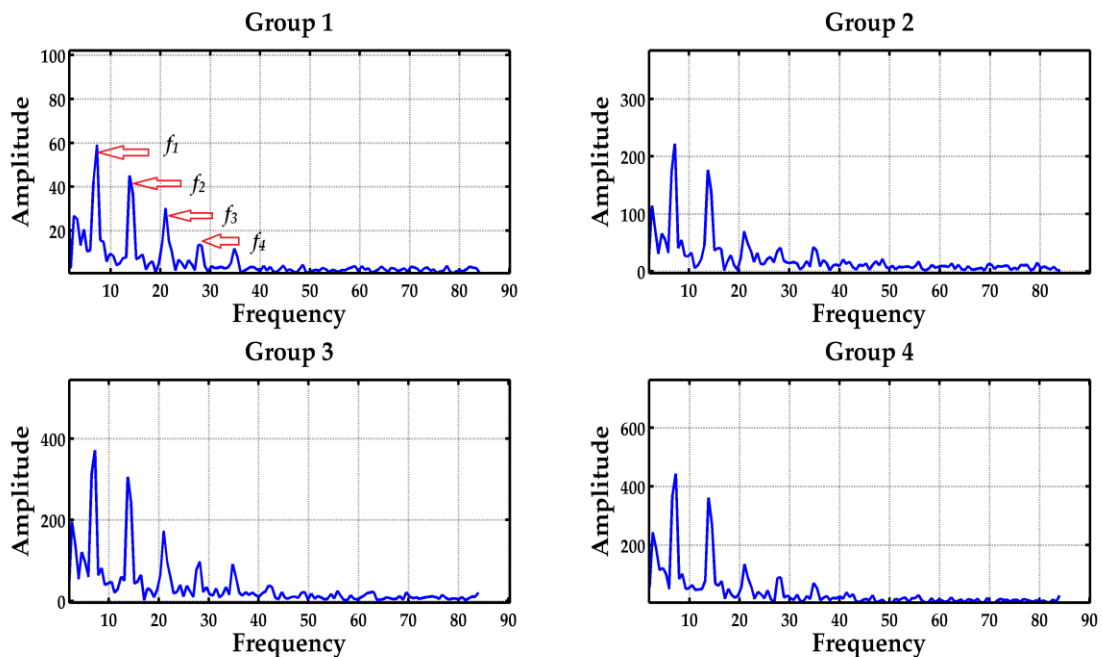


Figure 4.2 Traffic of sample groups in spectral domain

This implies that there are four major customers' traffic usage patterns in the city that contribute to the daily traffic variation. The dominant pattern which corresponds to f_1 repeats itself every 24 hour whereas the others have an impact in a decreasing order of

$f_2, f_3,$ and f_4 Using these components in (4.1), the temporal model is plotted against the real data and four sample groups are presented in Figure 4.3.

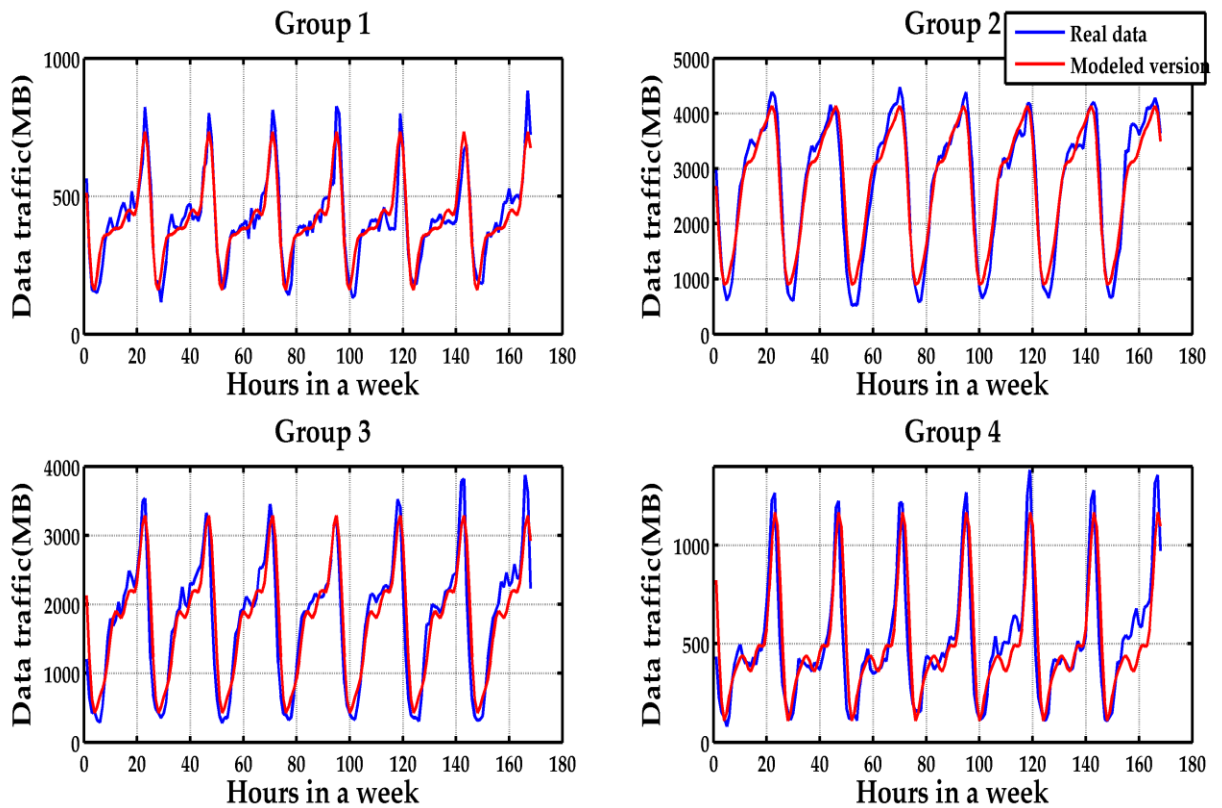


Figure 4.3 Plot of real data vs. temporal model

4.2.2 Traffic Model in Space

In order to incorporate space domain in the temporal model, parameters in the equation that vary across groups are analyzed and identified. As summarized in Table 4.1, since frequencies are similar for all groups they are left to be time domain parameters while coefficients and phases are taken as space parameters as they vary across the groups.

Table 4.1 Parameters variation in space domain

Parameters	Frequency	Coefficient	Phase
In space (across groups)	Constant	Vary	Vary

UMTS Data Traffic Modeling Using Fourier Transform

As a result coefficient and phase values of each group are prepared as a matrix as shown in Table 4.2 and Table 4.3. A total of 9 matrices corresponding to the two spatial parameters are prepared. For the ease of discussion only for a_0 and α_1 is presented here and for the others it can be referred in the Appendix section.

Table 4.2 Matrix for average traffic (coefficient a_0) in MB of each group

13	24	1	22	16	28	27	4	5	0	3	0	0	97	0	0
348	560	718	320	612	730	645	208	85	12	14	0	444	358	76	0
20	101	921	685	1415	1488	1028	447	82	179	300	502	811	170	23	0
0	93	287	1509	2554	2474	1561	1128	310	275	1574	288	349	0	3	20
192	273	467	883	1643	2165	1329	1759	1532	1062	1204	878	892	387	25	6
0	1369	1277	913	550	1366	2077	3317	2154	1518	866	736	107	52	0	2
283	252	906	710	1206	1458	2070	2711	658	607	127	601	51	261	0	10
280	414	1172	2171	1312	1428	1414	453	160	45	20	78	9	78	0	0
79	322	940	621	865	1951	1392	411	36	11	17	37	0	20	0	0
553	388	171	43	774	592	373	63	54	0	14	0	1	0	0	0
293	0	21	0	146	173	461	37	210	46	29	0	2	0	0	0
19	11	0	0	0	0	505	317	0	69	26	29	0	0	0	0
0	0	17	0	0	50	731	311	379	0	0	0	0	0	0	0
0	0	0	0	0	0	0	337	159	510	0	0	0	0	0	0
0	0	0	0	0	0	0	110	0	0	0	0	0	0	0	0
0	0	0	0	0	0	37	22	0	0	0	0	0	0	0	0

Table 4.3 Matrix for α_1 (in rad) of each group

0.7	1.4	1.1	1.1	2.1	1.0	1.1	1.1	1.1	0.0	0.6	0.0	0.0	1.3	0.0	0.0
1.4	0.7	0.9	0.7	1.0	0.8	0.8	1.0	0.9	0.9	0.7	0.0	-5.8	1.2	0.7	0.0
0.9	0.8	0.6	0.7	0.9	0.9	0.9	0.7	0.8	1.1	0.7	0.7	0.8	1.0	0.9	0.0
0.0	1.2	0.5	0.8	1.0	1.8	1.6	1.1	0.8	0.7	1.0	1.0	0.8	0.0	1.0	1.3
-5.8	0.7	0.6	0.7	1.0	1.6	1.7	1.2	1.3	1.4	1.3	0.9	1.0	0.6	1.1	6.8
0.0	0.7	0.9	0.8	1.1	1.3	1.7	1.4	1.1	0.9	0.9	0.9	0.8	1.0	0.0	6.7
0.8	0.9	0.9	0.8	1.1	1.4	1.0	1.5	1.3	0.8	1.2	0.8	1.0	0.5	0.0	1.5
0.4	-5.3	-0.2	0.8	0.9	0.9	1.3	0.9	0.9	1.0	2.5	2.1	2.0	1.2	0.0	-0.3
0.8	1.0	0.9	0.8	1.0	0.8	0.8	0.8	1.0	1.5	1.8	1.0	0.0	1.2	0.0	0.0
1.2	1.0	0.9	1.3	0.6	1.2	1.5	1.1	0.7	0.0	1.8	0.0	-0.1	0.0	0.0	0.0
1.0	0.0	1.0	0.0	1.2	1.0	1.4	1.7	1.6	2.0	1.4	0.0	-1.2	0.0	0.0	0.0
0.9	0.7	-0.2	0.0	0.0	0.0	1.4	1.0	0.0	0.8	1.1	1.0	0.0	0.0	0.0	0.0
0.0	0.0	1.9	0.0	0.0	1.0	0.8	0.9	1.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.3	0.9	0.8	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	-1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

UMTS Data Traffic Modeling Using Fourier Transform

The matrices match the geographical area in such a way that the bottom left corner of the table corresponds to south-west of the city whereas the upper right corner refers north-east.

By applying 2D-DCT on each of the matrices, all are transformed from spatial to spectral domain. In spectral domain, as shown in Figure 4.4, energy of the low frequency components is much bigger than that of the higher components. As a result frequency components with very low energy content are truncated. By using the remaining components in the inverse 2D-DCT, a simplified spatial equation is obtained for both parameters as given in (4.2).

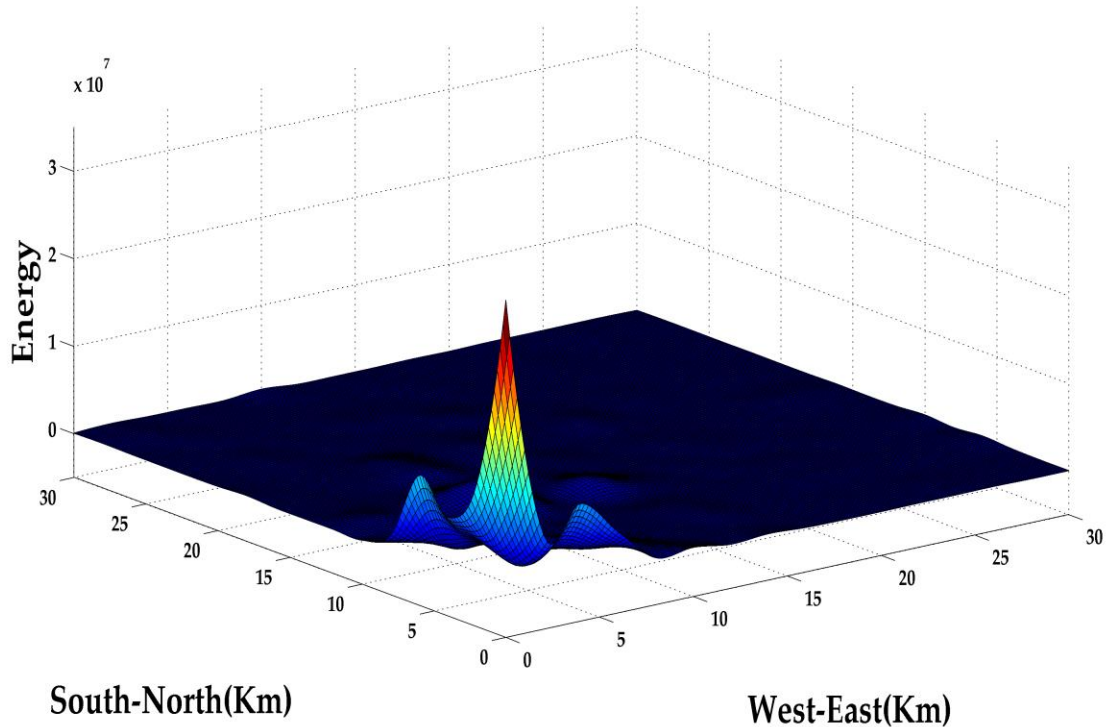


Figure 4.4 Energy content of spatial parameters' spectrum

The function a_n represents the spatial parameters whereas A is the spectral domain content after truncating frequencies with less energy value. X and Y are respectively the city's width and length while x and y are the local distances.

UMTS Data Traffic Modeling Using Fourier Transform

$$a_n(x, y) = \frac{2}{N} \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} C(n)C(m)A(n, m)\cos(w_n x + \theta_n) \cos(w_m y + \theta_m) \quad (4.2)$$

$$\text{Where } C(n), C(m) = \begin{cases} \frac{1}{\sqrt{2}}, & \text{for } n, m = 0 \\ 1, & \text{otherwise} \end{cases}, \quad w_n = \frac{(N-1)n\pi}{N.X}, \quad w_m = \frac{(N-1)m\pi}{N.Y}, \quad \theta_n = \frac{\pi n}{2N} \quad \text{and} \quad \theta_m = \frac{\pi m}{2N}$$

Table 4.4 Matrix of coefficient values for a0 in spectral domain

5794	2414	2906	2050	577	779	144	536	249	41	24	179	191	260	28	78
3384	1133	1478	1038	963	265	200	143	59	135	215	239	347	244	213	30
2501	1497	1345	1084	174	272	115	508	126	215	29	154	33	49	93	286
3142	697	2274	545	89	138	374	436	209	216	353	288	209	219	231	544
1334	440	1487	196	95	29	250	109	247	33	297	388	67	112	231	185
4	158	456	427	1303	82	1232	668	85	336	108	23	280	75	288	518
140	724	543	1128	1425	6	613	201	211	39	393	198	70	76	76	209
58	334	345	507	315	230	271	439	43	358	340	47	176	67	278	83
91	98	21	596	843	189	612	277	56	182	189	136	135	366	409	241
452	11	209	421	754	253	599	442	339	563	163	148	32	43	129	30
194	40	268	225	112	138	333	38	202	343	195	308	36	163	190	36
0	2	440	132	313	89	62	532	243	25	200	112	17	180	54	164
91	14	166	230	122	437	365	239	156	181	150	274	110	97	108	127
161	65	379	284	198	196	244	73	277	111	108	260	136	228	3	190
266	455	176	224	189	196	47	237	229	213	23	66	23	358	276	206
119	435	251	188	1	599	127	58	17	222	31	250	228	250	271	24

Table 4.5 Matrix of coefficient values for α_1 (rad) in spectral domain

9.9	0.1	3.3	2.1	1.3	1.6	1.1	1.3	0.5	0.4	0.2	0.7	1.4	0.9	1.2	0
3.8	0.6	0.4	2.3	0.1	1.6	0.8	0.3	0.3	0.4	0.7	0.8	0.3	1.3	0.6	0.5
3.1	2.1	1.5	0.2	0.9	0.3	1.2	0.9	1.1	0.2	1.7	0.1	1	0.5	0.4	0.6
0.7	4.5	0.8	2.5	0.6	1.1	0.5	2.3	1.6	2.6	1.3	0.8	0.3	0.1	0.8	0.1
1	1.6	0.4	1.7	1.2	1.4	0.1	2.4	0.3	2.5	0.2	0.7	1	0	0.6	0.4
0.1	0.5	0.2	0.3	0.1	0	0.4	0.1	1.2	0.4	0.5	0.2	1.1	0.1	0.7	0.1
1.3	0.2	0.7	2	1.6	1.2	0.6	1.8	0.2	1.3	0	1.9	0.2	0.8	0.2	0
0.7	0.9	0.7	1.8	0.1	1.7	0.4	1.3	0.2	2.1	0.3	1.3	0.2	0.9	0.8	0.1
0.1	1.9	1.1	0.5	1.9	0.6	0.6	1	0.3	0.4	0.3	0.1	0.3	0.5	0.4	0.5
0.4	0.7	1.3	0.9	1.7	1.1	0.1	1.1	0.7	0.2	0.5	0.5	0.5	1.1	0.7	0.8
1	2.2	0.4	0.6	0.2	1.8	1	0.1	0.6	0.6	0.1	0	0	0.9	1.5	0.6
0.8	1.1	2.3	1.7	0.4	1.5	0.2	0	0.4	1	0.3	0.7	1	0.3	0.8	0.3
1	1.2	0.6	0.4	0	0.4	0.3	0.5	1.2	0.4	1	0.4	0.1	1	0.4	0.1
0.2	1.9	0.3	0.8	1.8	0.3	0.3	0.3	0.7	0.5	1.3	0.6	0.6	0.6	0.4	1.3
0	1.3	0.2	0	0.8	0.1	0.6	0.7	0.1	0.6	0.1	1.2	1.1	0.1	1.1	0.1
0.5	0.4	0	0.3	0.1	1.1	1.1	0.1	1.1	1.7	1	0.6	0.3	0.1	0	0.1

4.2.3 Traffic Model in Space and Time

As explained in the above sections the two models are combined in such a way that coefficients and phases from the time modeling are further modeled spatially by using 2D-DCT which made the model to be a complete. So by inserting (4.2) in to (4.1) the formula gets completed by giving a spatiotemporal model as given in (4.3).

$$T(x, y, t) = a_0(x, y) + \sum_{n=1}^4 (a_n(x, y) \sin(2\pi f_n t + \alpha_n(x, y))) \quad (4.3)$$

5 Results and Discussion

Because of the following two reasons, this research has focused its analysis mainly on the relation between the level of frequency truncation in the spectral domain of the spatial parameters and the model accuracy.

1. The model is composed of multiple parameters with the possibility of having different truncation level which affects the accuracy in different manner.
2. A range of application areas for the model in the network are identified but their sensitivity for the accuracy of the model is not similar. For example application areas such as planning, dynamic resource allocation and energy efficient design are more sensitive to its accuracy whereas applications like dynamic pricing and knowing customer's usage behavior are less sensitive to accuracy.

The model development has required transformation of the traffic from time and space to spectral domain followed by frequency truncation. In time domain, as it is observed that all groups have the same major frequency components, there is only one truncation level but in space domain there are many parameters (five related with coefficients, four from phases and the number of grouping) whose truncation level affects the overall model performance. The result has shown that determining the level of truncation is a trade-off between accuracy and simplicity. In image compression, which also applies a 2D-DCT approach for transforming matrix of colour values to spectral domain, a standard quantization table developed by industrial researches is used for truncating less valued frequencies. As there is no a similar study in this area that presented quantization table by analysing the value of frequencies for traffic variation, this research applied the truncation directly on the magnitude of frequencies amplitude. Based on this, first analysis was performed on how level of truncation on frequencies of transformed coefficients affects the model performance while fixing grouping number (N) to sixteen and keeping phases untruncated. Coefficient of determination (R^2) is

Results and Discussion

taken as a measure of accuracy and the first result is as presented in Figure 5.1. Coefficient of determination is a statistical measure that is used to evaluate the goodness of fit for a model. It is the ratio of the sum of two squared errors as given in (5.1)

$$R^2 = \frac{\sum_{i=1}^n (x_i - y_m)^2}{\sum_{i=1}^n (y_i - y_m)^2} \quad (5.1)$$

Where x_i and y_i are the i^{th} elements of the model and the real data respectively and y_m is the average of the real data.

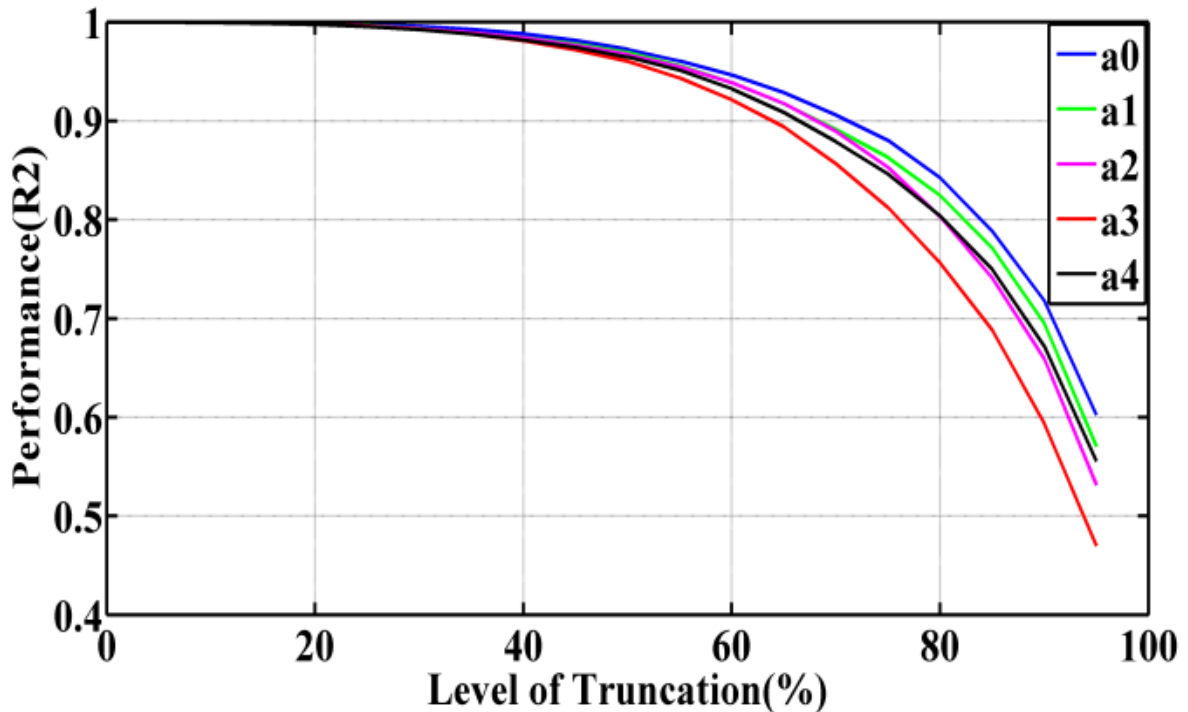


Figure 5.1 Model accuracy vs. Transformed coefficients' frequency truncation level

Secondly, as shown in Figure 5.2, how level of truncation on frequencies of transformed phases affects the model performance is analysed by taking grouping number (N) to be 16 and coefficients un truncated.

Results and Discussion

The two results show that as the magnitude of frequency in time domain increase, model performance of the corresponding coefficients and phases gets more sensitive to the level of truncation.

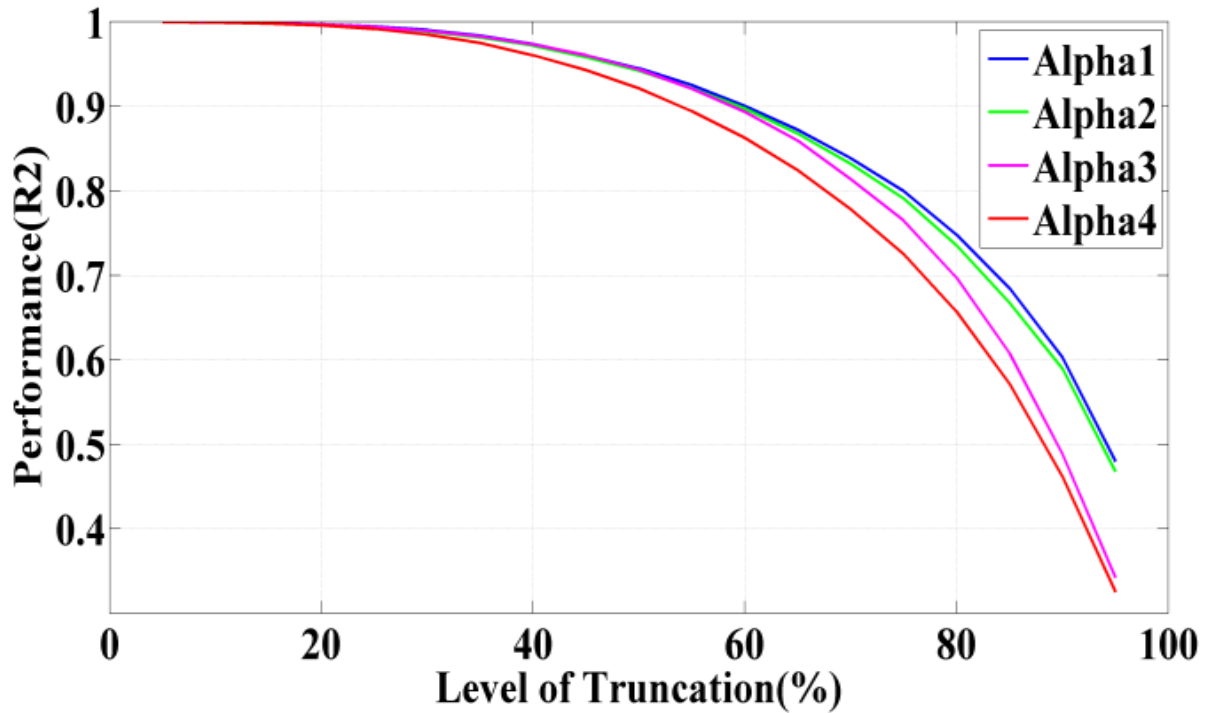


Figure 5.2 Model accuracy vs. Transformed phases' frequency truncation level.

Thirdly, how its accuracy changes with the size of grouping while applying a uniform truncation over all parameters is checked and the result is as shown in Figure 5.3. Increasing the grouping number makes the traffic distribution to be uniform spatially which decreases the required number of frequencies in spectral domain to be used for the reverse transformation. As a result for higher grouping the model's simplicity increases while keeping accuracy level constant.

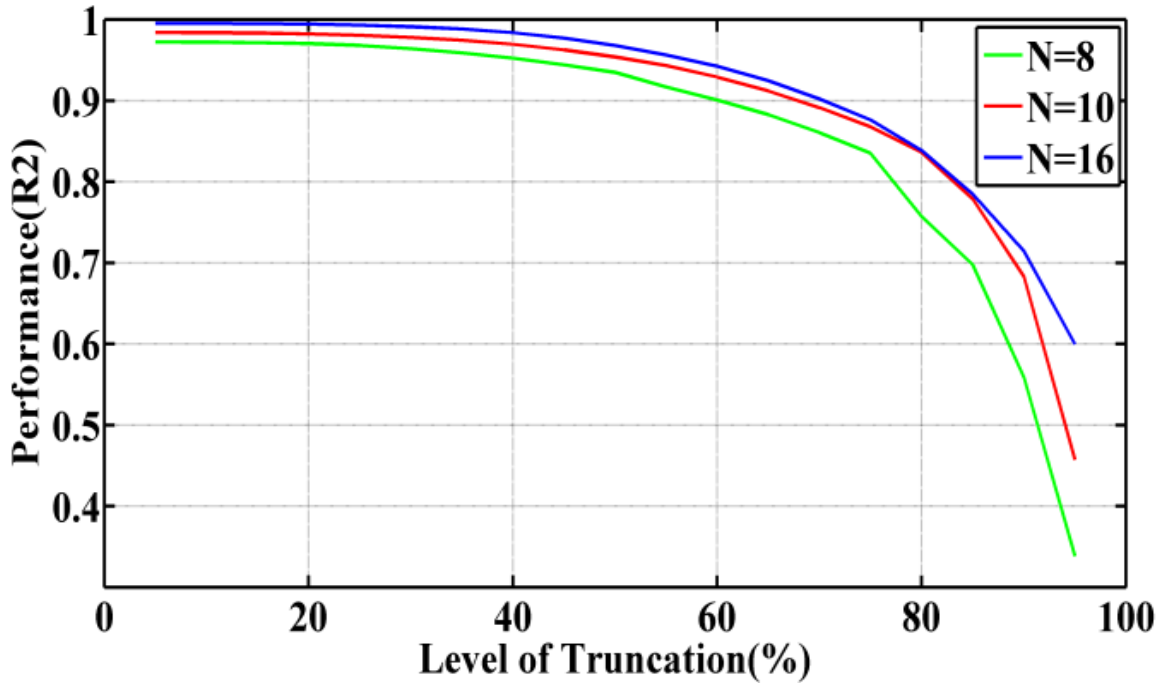


Figure 5.3 Performance of the model for different grouping

Finally, with a uniform truncation on all parameters except keeping number of grouping to be sixteen, the model performance variation is as plotted in Figure 5.4.

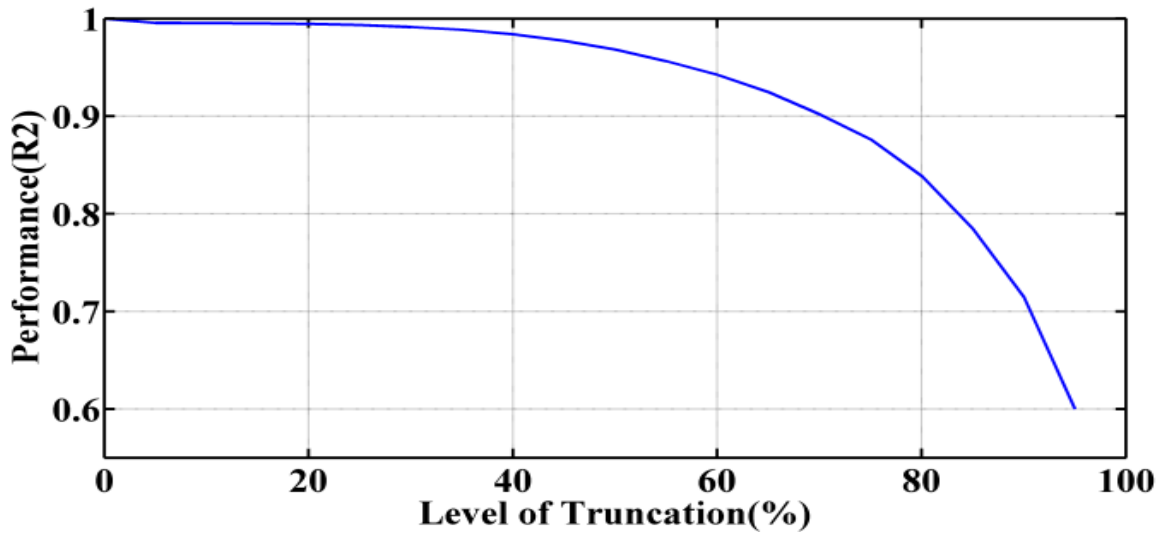


Figure 5.4 Model accuracy vs. spatial frequency truncation level

Results and Discussion

In order to compare the model with the real data by simulation, setting target performance to be 0.9 and taking the grouping size as 16, the combination of truncation level for different parameters of the model is presented in Table 5.1.

Table 5.1 Truncation level of spatial parameters

Coefficients	a0	a1	a2	a3	a4	α_1	α_2	α_3	α_4
Truncation level (%)	71	68	64	66	64	60	59	58	54

And using the above values in the model, the traffic of the city is simulated for a given day and three sample hours are presented in Figure 5.5, Figure 5.6 & Figure 5.7.

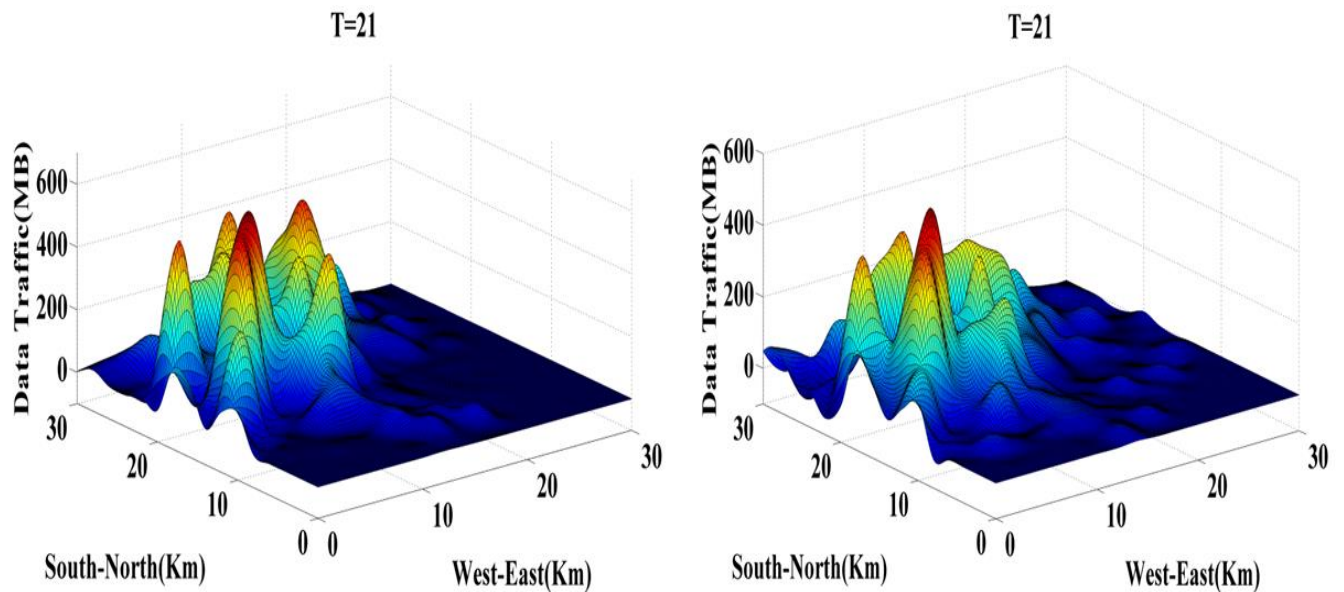


Figure 5.5 Real data traffic vs. model output for 9 pm

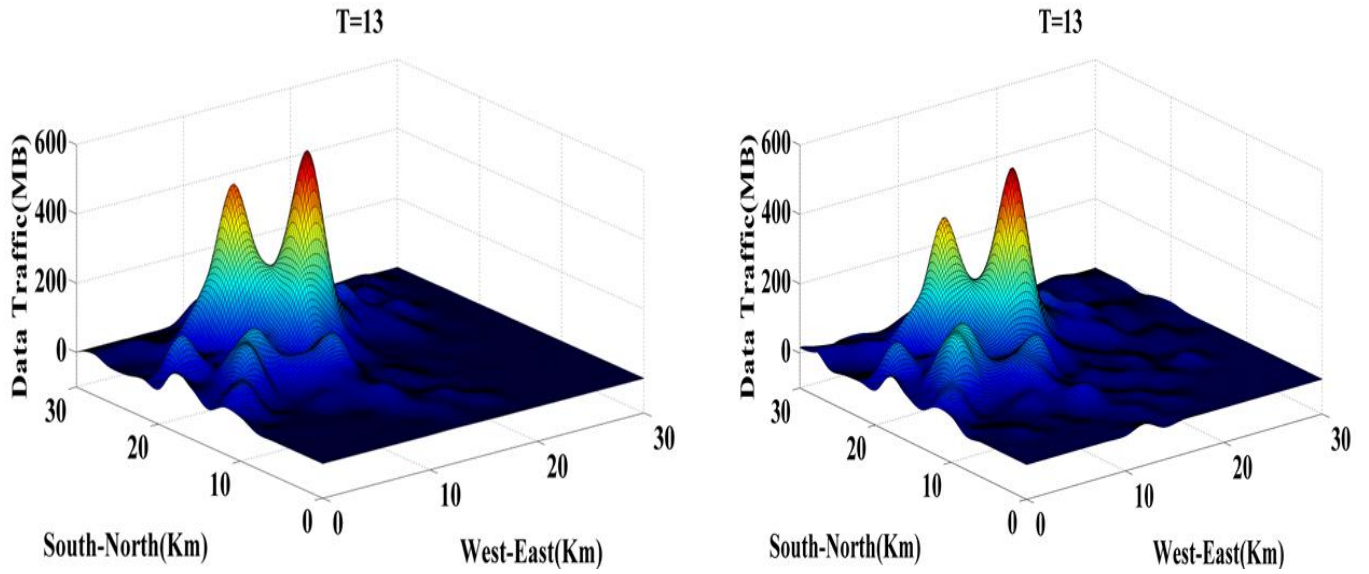


Figure 5.6 Real data traffic vs. model output for 1 pm

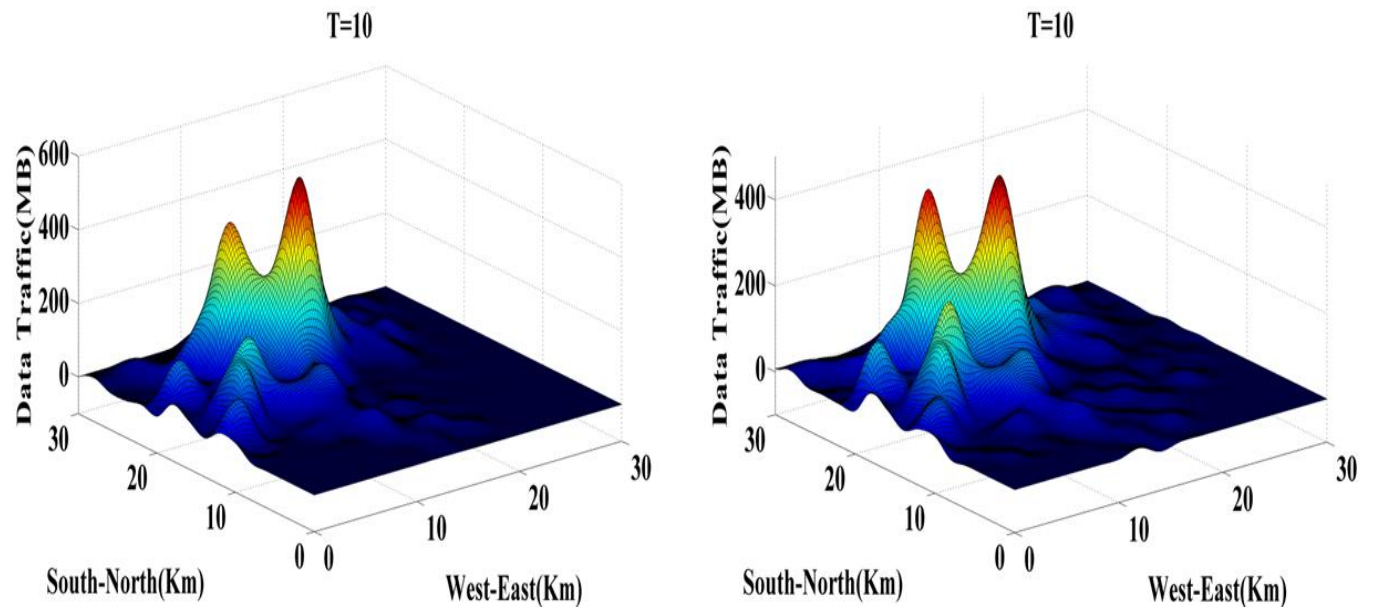


Figure 5.7 Real data traffic vs. model output for 4AM

Using the same values, the overall traffic distribution of the city is calculated and displayed in Figure 5.8.

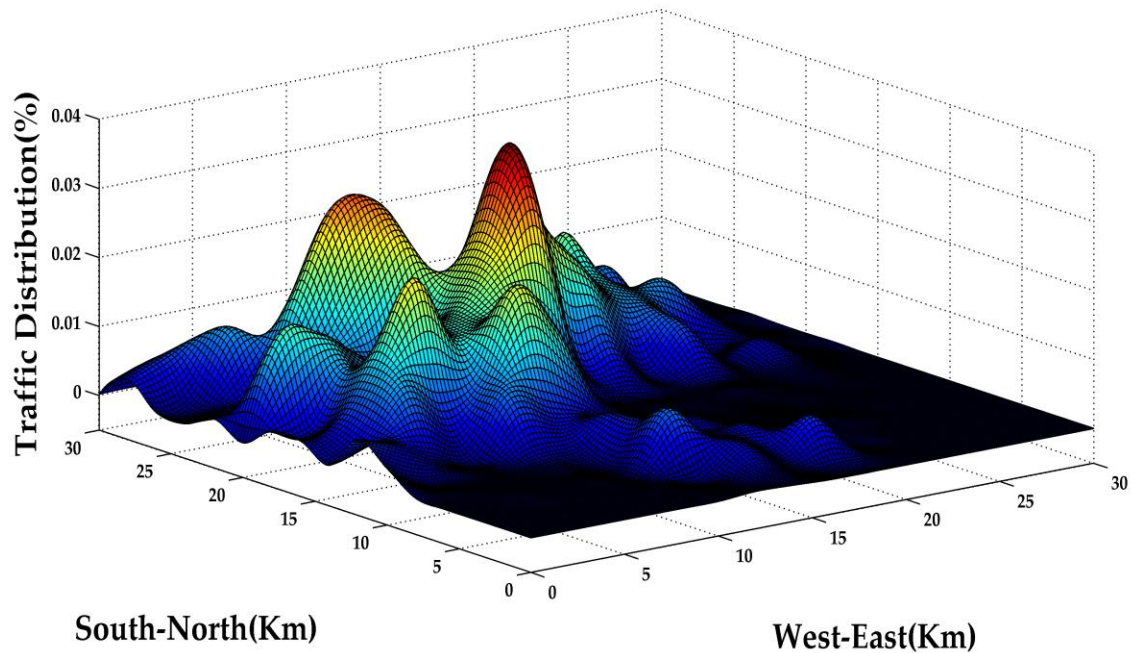


Figure 5.8 Traffic distribution of the city

In addition to using the dataset for the modeling, it is also used to find out the spatial distribution of the city's NodeBs busy hour traffic. For that both probability density function and cumulative density function are calculated as follows:

1. A new dataset is generated by taking the logarithmic scale of each NodeB's busy hour traffic
2. Elements of the new dataset are binned into 15 equally spaced containers where each one is represented with a bar as shown in Figure 5.9
3. The normal distribution of the logarithmic scale data is calculated by using its mean and standard deviation and plotted against the histogram developed in step2. X-axis represents the logarithmic scale of base stations' busy hour traffic whereas the Y-axis represents its probability density function.

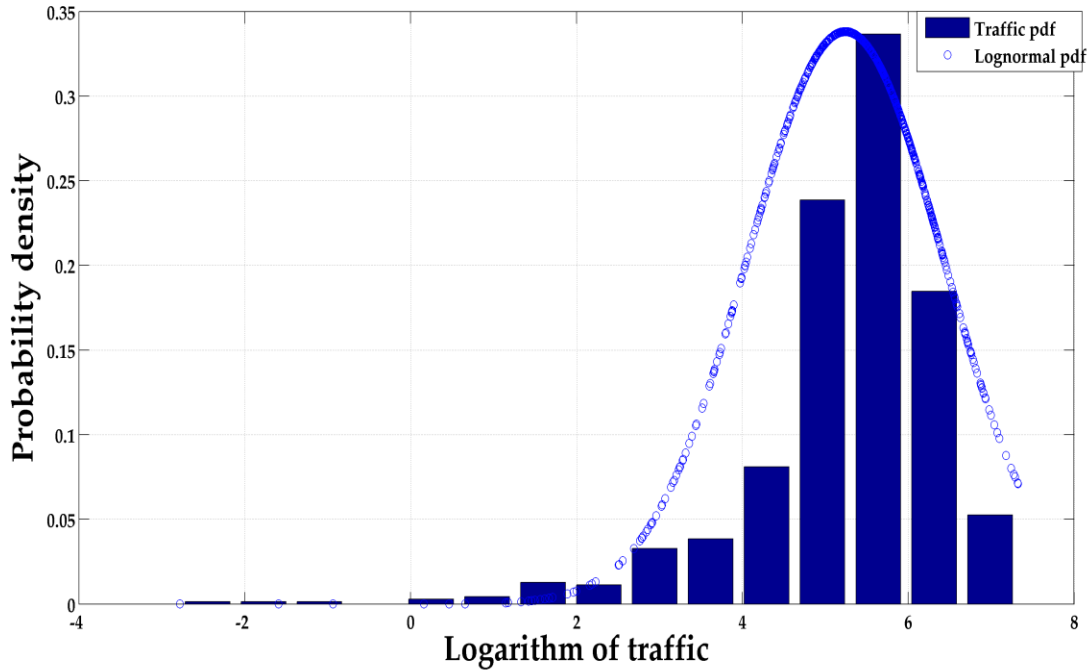


Figure 5.9 Probability density function of base stations peak hour traffic

And the cumulative distribution function of the UMTS base stations' peak hour traffic is as displayed in Figure 5.10.

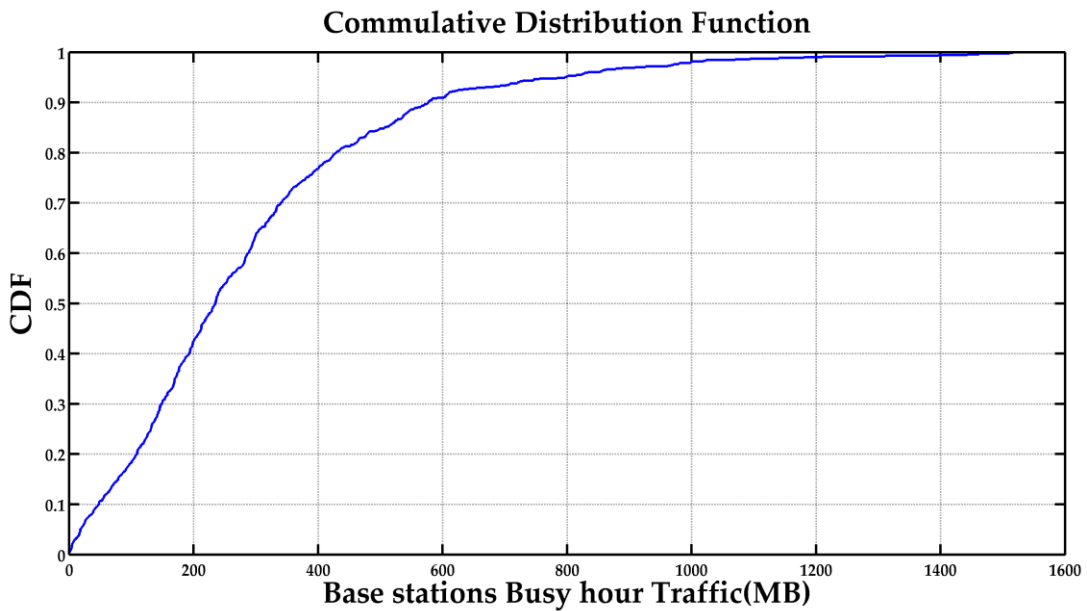


Figure 5.10 Cumulative Distribution Function of NodeB's busy hour Traffic

5.1 Network Utilization

As discussed earlier the model has got different application areas in the network but there is a need to define a parameter that can determine the network utilization and be used as an input for its applicability. Some of the application areas to remind are Energy efficient network design, dynamic resource allocation, developing marketing strategies and network planning. In this research, two metrics related with utilization are defined as given in (5.2) and (5.3).

$$\text{Installed Capacity Utilization} = \text{Traffic served} / \text{Installed Capacity} \quad (5.2)$$

$$\text{Peak Hour Utilization} = \text{Traffic served} / \text{Peak Hour Traffic} \quad (5.3)$$

Installed Capacity Utilization (ICU) is the ratio of current network traffic to the maximum traffic that the network can serve whereas Peak Hour Utilization (PHU) refers to the percentage of the average traffic served with respect to the peak hour traffic. As installed capacity is a network parameter, for finding ICU it is enough to know the amount of traffic served which can be provided by the model. But for calculating PHU, additionally it is needed to calculate the busy hour of the different areas so that peak hour traffic may be determined. By applying local maxima formula on the mathematical model as given in (5.4), the busy hour distribution for different areas is calculated and presented in Figure 5.11 the result shows most areas have a peak hour between 21:00 and 22:00 hour (9 pm and 10 pm).

$$\frac{dT(x,y,t)}{dt} = 0 \quad (5.4)$$

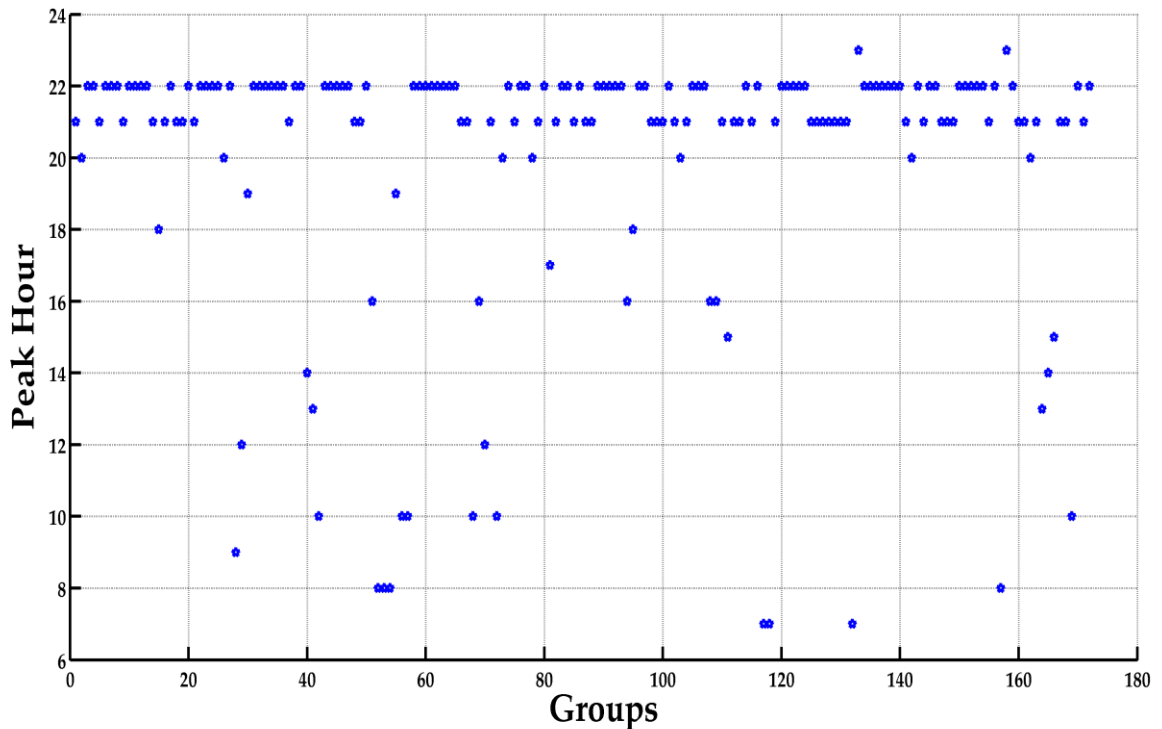


Figure 5.11 Busy Hour Distribution of the groups

Assuming that traffic served is a constant value, ICU doesn't change while PHU varies with the amount of peak hour traffic. In other word any mechanism by the operator to distribute the peak hour traffic to other hours of the day improves network efficiency by maximizing PHU while keeping ICU constant. This implies that PHU explains network efficiency much better than the way ICU does. As a result using the model, PHU of the network is analysed in different combination of space and time as follow:

1. In space and time – in this approach network utilization for every hour of the day and every group of the geography is obtained.
2. Average in time and distributed in space – in this way average utilization of the day for each group is provided
3. Average in space and distributed in time – this provides the average utilization of the whole city for every hour of the day.

Results and Discussion

- Average in both time and space – the overall utilization of the network is presented in this way.

The first way exhibits a similar result as the model but for the other three is as shown below.

Table 5.2 PHU (%) using average of traffic in time and distributed in space

35	38	30	40	55	48	46	43	27	0	41	0	0	64	0	0
46	38	41	37	35	37	34	35	31	33	34	0	48	49	42	0
36	40	33	35	38	41	36	33	39	37	37	39	40	45	46	0
0	35	32	38	38	58	48	47	33	38	40	46	44	0	34	54
43	50	37	35	41	56	63	50	65	54	47	47	46	29	47	58
0	47	41	39	43	47	60	63	55	45	46	47	38	41	0	41
45	46	40	39	48	47	49	64	58	37	40	42	45	31	0	54
42	52	53	43	45	42	49	46	49	43	58	46	48	39	0	24
41	45	53	42	47	43	38	46	47	51	49	58	0	48	0	0
53	44	53	44	39	43	49	33	47	0	55	0	33	0	0	0
42	0	40	0	47	35	41	45	56	50	58	0	33	0	0	0
34	44	31	0	0	0	50	41	0	45	40	44	0	0	0	0
0	0	50	0	0	52	41	42	50	0	0	0	0	0	0	0
0	0	0	0	0	0	0	49	40	43	0	0	0	0	0	0
0	0	0	0	0	0	0	69	0	0	0	0	0	0	0	0
0	0	0	0	0	0	51	46	0	0	0	0	0	0	0	0

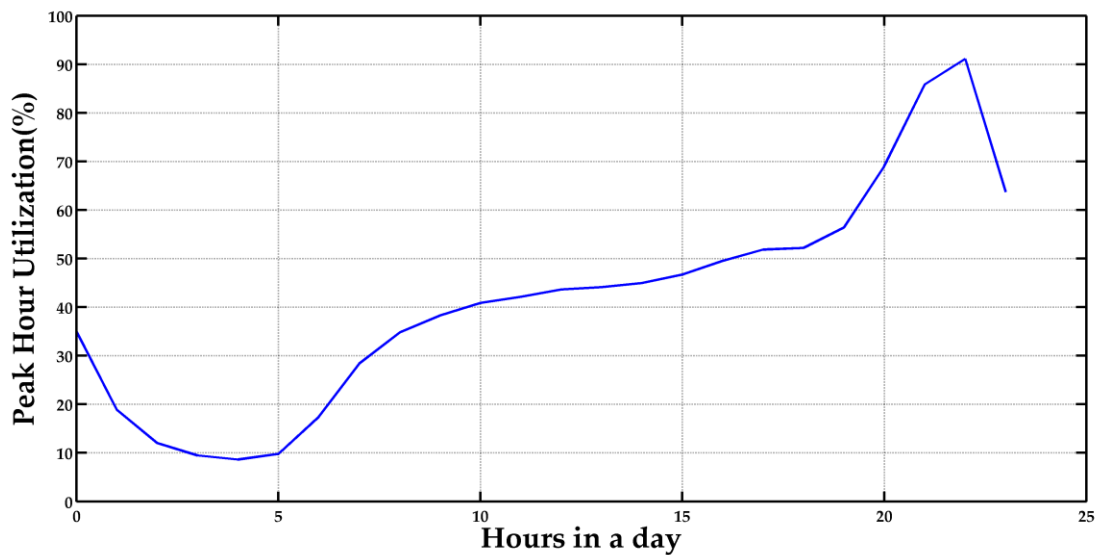


Figure 5.12 PHU (%) using average of traffic in space and distributed in time

Like the busy hour distribution, the PHU over 24 hour of the day also shows that the highest value to be in between 9 pm and 10 pm whereas the least value of the city is at 4 am

The result in Figure 5.12 shows that for most of the hours in a day, the network is underutilized which enables the company to implement Energy efficient strategies and dynamic resource allocation by using the result of this analysis as an input.

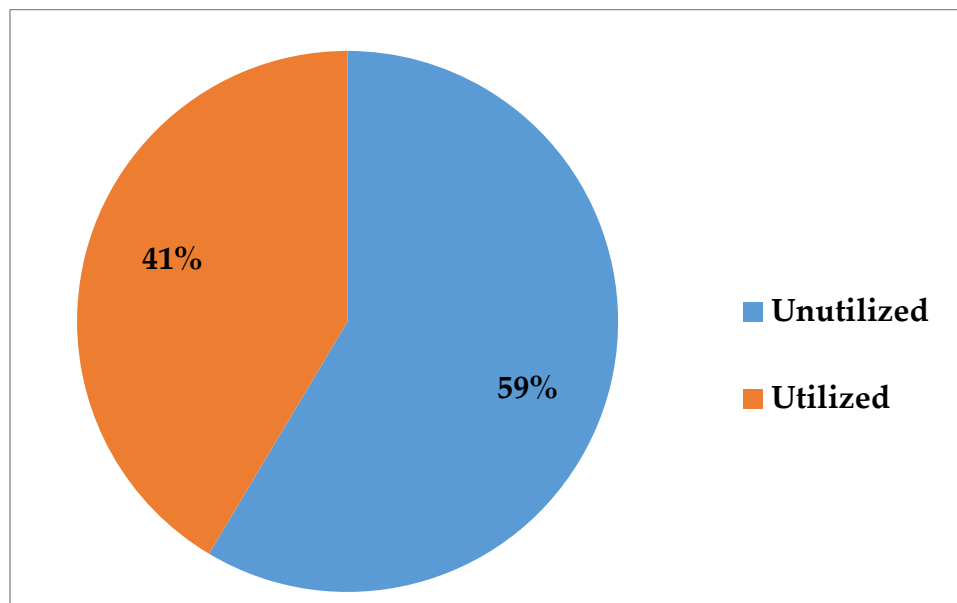


Figure 5.13 PHU of the whole city (using average of traffic in both space and time)

The overall PHU of the city indicates that much of the network (59%) is unutilized which can get even higher if it were evaluated with ICU. Even if network utilization is few, the key performance indicators show that some coverage areas in the network experience congestion at some point of the day especially at their busy hour. This merely implies inefficiency of the network. Based on this, how an ideal network with an optimum utilization looks like is analysed as follows:

Results and Discussion

The identified four major patterns of each group, that determine the overall variation of daily traffic, oscillate about the average traffic a_0 . Any strategy to decrease size of these oscillations around the offset (which also decreases peak hour traffic) creates an optimal network with the same ICU while improving PHU. That means bringing the data usage to a_0 for every hour of the day makes the traffic variation to be uniformly distributed as indicated in Figure 5.14. Taking amplitude as a metrics for measuring the effect of the frequency components on the traffic variation, for all groups the effect decreases from f_1 to f_4 and a typical group's value is presented in Table 5.3.

Table 5.3 Amplitude value of major frequency components

Frequency components	f_1	f_2	f_3	f_4
Amplitude (MB)	263	243	140	69

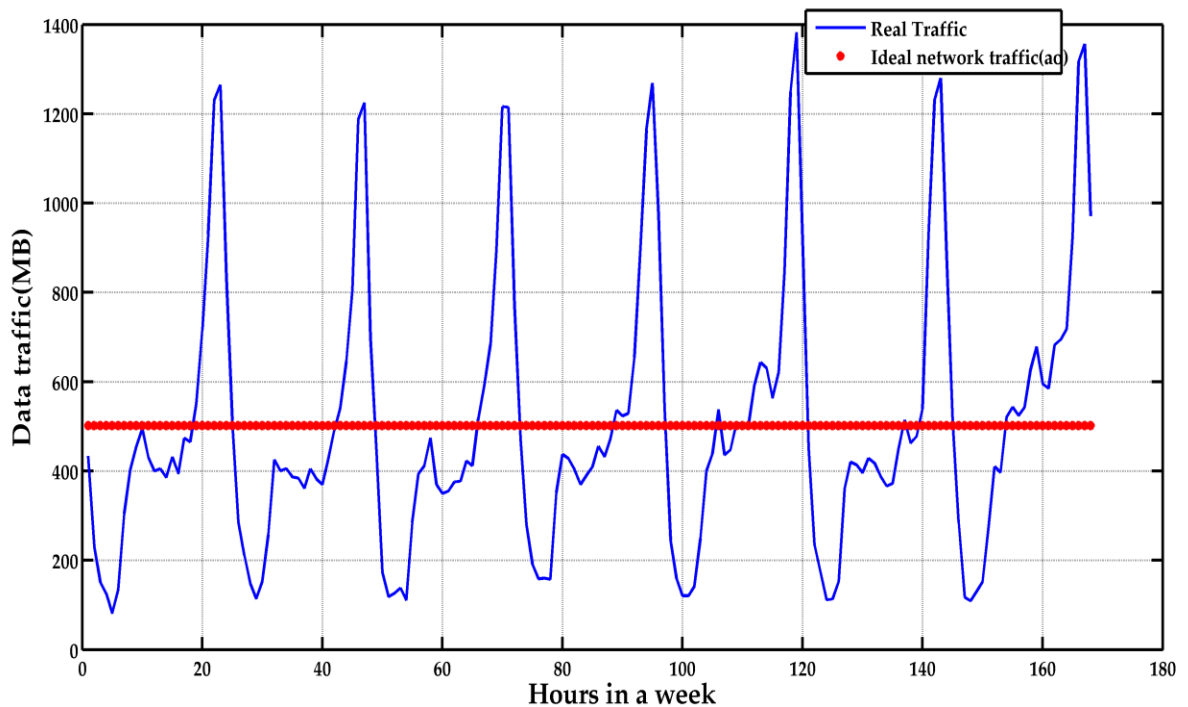


Figure 5.14 Plot of Ideal (optimal) network vs. Real network traffic

Similarly how PHU looks like for real and ideal (optimal) network is presented in Figure 5.15

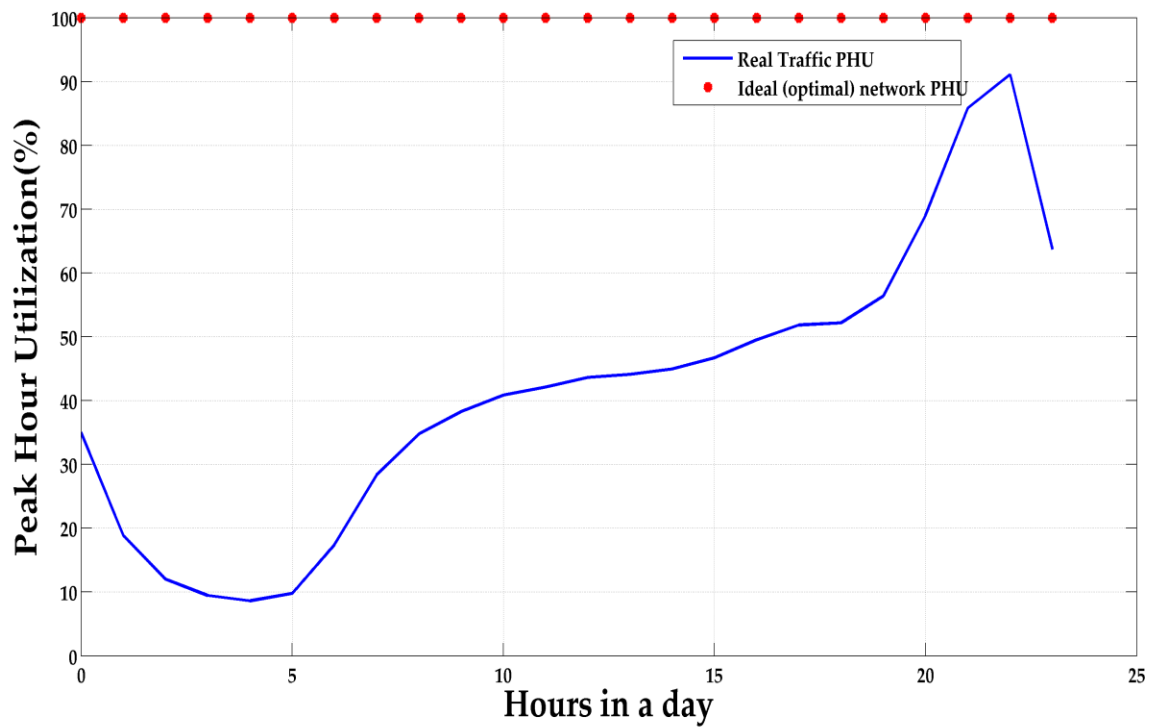


Figure 5.15 PHU of real and ideal network

These results and discussions show that optimizing networks' performance should be performed through a measure of maximizing PHU.

6 Conclusions and Recommendations

6.1 Conclusions

Even though the dataset collected for this research is from Addis Ababa city's UMTS network, the approach followed in developing the model has neither technological nor geographical limitation in its applicability. A similar approach can be followed for different technologies (i.e. GSM and LTE) as well as for different coverage areas of the operator even as big as the whole country. But as a case study, the output provided by the model in this research is specific to the city and more of the following conclusions are generalized for the city's UMTS data traffic.

- ❖ Related with the users' behavior, the model has identified four customer data usage patterns that determine the daily traffic variation in the network and also impact its performance. This gives an insight for the operator where and how to apply its strategies of optimizing the network.
- ❖ This research has shown that devising strategies for improving PHU can significantly minimize network congestion levels. As congestion is one factor that drives operators towards either expansion or execution of new projects, the operator can relief the network from it for long by combining the usual optimization trend with PHU analysis and measures. This avoids unnecessary operational and expenditure cost and also enhances quality of service which in turn brings customer satisfaction
- ❖ It is also shown that an ideal network has a uniform traffic distribution in time. Mathematically, improving the network towards this level requires actions to reduce the amplitude level of the identified harmonics whereas in practical term the users' behavior should be diverted from peak hour to other hours of the day. This makes the effect of some patterns on the traffic variation to vanish and reduces their number. Similarly, achieving traffic uniformity in space domain

also reduces the required number of spatial frequencies in the inverse transform. Increasing the number of grouping is one way of keeping the traffic to be uniform in space but technically it can be achieved through optimization measures by forcing edge users from high to less traffic areas for neighboring groups. Over all, bringing uniformity in both time and space enhances network performance by avoiding the effect of multiple patterns and also reduces the model complexity.

6.2 Recommendations

One of the spaces in this research for a further study is incorporating prediction in to the model so that future traffic variation of the city may be captured in advance. Based on the observation while conducting this research, the direction for forecasting the traffic by using the developed model is presented in the following way:

Among the three parameters in the model only coefficients and phases show change with time. As they are spatial parameters in the model representing different areas of the city, they are multiple in their number. So the forecasting task should consider applying multivariate prediction algorithms. But due to the role they have in the model, the way they change in time differs one from the other. Coefficients represent magnitude of the identified traffic patterns in the city whereas phases represent the starting position of the patterns. As a result, forecasting on the coefficients should be performed on their growth and on phases it should be positional prediction.

The other study areas to be pointed in this research can be covered with two assumptions. The first one, assuming congestion is not the case in the network; the model has shown a low utilization (more specifically PHU as described in the analysis part) which can be considered as wastage in terms of some resources such as energy. So developing an approach for the network to use energy in its utilization range can be a

Conclusions and Recommendations

promising study. Additionally, making a research on how to devise a technique for circulating scarce resources according to the utilization profile of the different coverage areas (as provided in this research) is also another study area that can base this research.

Secondly, assuming congestion as a critical issue in the network, future studies can be conducted on developing strategies that distribute peak hour traffic to other hours of the day. For example dynamic pricing is one aspect which increases traffic uniformity as well as service performance by improving PHU.

Finally, the traffic modelling in space and time can be done by any combination of the spatial and temporal models currently available in the scientific world. Any future study on a similar modelling can consider using a dataset that include GSM and LTE traffic but it will limit its applicability (as some applications such as dynamic resource allocation require technology specific output).

References

- [1] Cisco Whitepaper, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016-2021," 28 March 2017.
- [2] Jihwan Kim, Hyang-Won Lee and Song Chong, "Traffic-Aware Energy-Saving Base Station Sleeping and Clustering in Cooperative Networks," IEEE Transactions on Wireless Communications, 2017.
- [3] Aleš Švigelj, Radovan Serbec, and Kemal Alic, "Network Traffic Modeling for Load Prediction: A User-Centric Approach," IEEE Network, 2015.
- [4] Ethio telecom, "Performance Reporting System", accessed on July 15, 2018.
- [5] S. Wang, X. Zhang, J. Zhang, J. Feng and W. Wang, "An Approach for Spatial-temporal Traffic Modeling in Mobile Cellular Networks," Wireless Signal Processing and Network Laboratory, Beijing, China, May 2017
- [6] E. Oh and B. Krishnamachari, "Energy savings through dynamic base station switching in cellular wireless access networks," IEEE GLOBECOM, 2010.
- [7] M.F.Hossain, K. S. Munasinghe and A. Jamalipour, "A protooperation-based sleep-wake architecture for next generation green cellular access networks," Proc. of IEEE Int. Conf. ICSPCS, pp.1-8, Dec. 2010.
- [8] M. A. Marsan and M. Meo, "Energy efficient management of two cellular access networks," ACM SIGMETRICS Performance Evaluation Review, vol. 37, no. 4, Mar. 2010.
- [9] L.Guan et al., "Spatial Modeling and Analysis of Traffic Distribution Based on Real Data from Current Mobile Cellular Networks", Ministry of Education Beijing University of Posts and Telecommunications, 2013.
- [10] D. Lee et al., "Spatial modeling of the traffic density in cellular networks," IEEE Wireless Communications, 2014.

References

- [11] J. G. Andrews, F. Baccelli and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks", Nov. 2011.
- [12] H. S. Dhillon, R. K. Ganti, F. Baccelli, and J. G. Andrews, "Modeling and Analysis of K-Tier Downlink Heterogeneous Cellular Networks," Selected Areas in Communications, IEEE Journal on, vol. 30, no. 3, pp, 550-560. Apr. 2012.
- [13] B. Ahn, H. Yoon and J. W. Cho, "A design of macro-micro CDMA cellular overlays in the existing big urban areas," Selected Areas in Communications, IEEE Journal on, vol. 19, no. 10, pp, 2094-2104. Oct. 2001.
- [14] S. Almeida, J. Queijo and L. M. Correia, "Spatial and temporal traffic distribution models for GSM," Proc. IEEE VTS 50th Vehicular Technology Conference, vol. 1, pp. 131-135. Sep. 1999.
- [15] N. Son and A. Robert, "Approximating user distributions in WCDMA networks using 2-D Gaussian," International Conference on Computing, Communications and Control Technologies, Jul. 2005.
- [16] A. Salem, "Seasonal Auto-regression integrated moving average-based data traffic forecasting: The case of UMTS network in Addis Ababa, Ethiopia", MSc. Thesis, Addis Ababa University, 2016.
- [17] D. Gemechu, "UMTS Traffic Model using artificial neural network: the case of Addis Ababa, Ethiopia", MSc. Thesis, Addis Ababa University, 2017
- [18] "UMTS Protocol Analysis", Internet: <https://www.gl.com/umts-network-protocol-analyzer.html>, [September, 2018].
- [19] Mushtaq Yousafzai, "UMTS System Architecture, protocols and processes" 2014 [Online]. Available: <https://www.slideshare.net/mushtaqyousafzai1/umts-system-architecture-protocols-processes>. [Accessed: 16-September-2018].
- [20] Cornelia kappler, "UMTS Networks and Beyond," John Wiley & Sons, 2009.

References

- [21] Naveen Jakhar, "UMTS protocol stack," 2016 [Online]. Available: <https://www.slideshare.net/naveenjakhar12/umts-protocols-66308012>. [Accessed: 16-September-2018]
- [22] Harri Holma and Antti Toskala, "WCDMA FOR UMTS HSPA Evolution and LTE," John Wiley & Sons, 2010.
- [23] Claudiu Chioariu, "QoS In UMTS," Helsinki University of Technology, 2004.
- [24] Russel L.Herman, "An Introduction to Fourier Analysis ," Taylor & Francis Group, 2017
- [25] Alan V. Oppenheim and Ronald W. Schaffer, "Discrete-Time Signal Processing," Pearson higher education, 2010.
- [26] G.K. Wallace et al., "The JPEG Still Picture Compression Standard," IEEE Transactions on Consumer Electronics, 1991.
- [27] Bellinger, G. 2004. "Modeling & Simulation: An Introduction," in *Mental Model Musings*. Available at <http://www.systems-thinking.org/modsim/modsim.htm>.
- [28] C. E. Chuka et al., "Regression and Time Series Analysis of Diesel Product Sales in Masters Energy oil and Gas," International Journal of Science, Engineering and Technology Research, 2013.
- [29] M. Zubair Shafiq, L. Ji, A. X. Liu and J. Wang, "Characterizing and Modeling Internet Traffic Dynamics of Cellular Devices," Michigan state university, 2011
- [30] Saeed Aghabozorgi, AliSeyedShirkhorshidi and TehYingWah, "Time-series clustering - A decade review," Elsevier Information Systems, 2015.
- [31] Syed Ali khayam, "The Discrete Cosine Transform (DCT): Theory and Application," Michigan State University, 2003.

Appendices

Matrix for average traffic (coefficient a1) in MB of each group

8	15	1	11	9	12	15	2	4	0	2	0	0	33	0	0
202	328	371	205	399	431	393	128	65	9	8	0	147	168	44	0
14	59	613	415	823	873	602	301	51	119	192	263	426	72	10	0
0	57	190	852	1308	1383	878	510	184	169	908	144	168	0	2	9
64	107	281	513	847	1210	767	926	862	545	624	443	460	340	12	2
0	532	723	518	297	736	1064	1892	1002	808	426	372	60	25	0	1
129	109	504	388	635	803	1132	1367	226	357	77	311	27	180	0	5
144	131	349	1000	666	820	764	222	72	25	13	52	4	44	0	0
44	145	394	356	418	1136	838	203	16	5	6	19	0	9	0	0
288	227	65	20	409	307	162	43	24	0	7	0	1	0	0	0
169	0	14	0	77	103	237	22	84	23	9	0	2	0	0	0
9	6	0	0	0	0	260	169	0	37	16	11	0	0	0	0
0	0	9	0	0	25	395	167	170	0	0	0	0	0	0	0
0	0	0	0	0	0	0	194	89	206	0	0	0	0	0	0
0	0	0	0	0	0	0	41	0	0	0	0	0	0	0	0
0	0	0	0	0	0	16	11	0	0	0	0	0	0	0	0

Matrix for average traffic (coefficient a2) in MB of each group

6	11	1	10	4	10	10	2	3	0	2	0	0	22	0	0
125	256	306	156	335	365	338	114	53	7	8	0	134	131	35	0
12	45	499	370	655	663	505	243	41	87	151	243	339	66	10	0
0	45	157	685	1123	623	542	399	169	136	706	118	150	0	2	8
65	90	221	439	698	563	311	599	362	343	487	362	358	256	11	2
0	499	543	415	227	502	522	634	653	624	364	296	51	24	0	1
107	90	385	339	454	529	666	599	219	306	60	291	22	155	0	4
116	115	366	941	541	615	502	177	57	22	5	15	2	40	0	0
33	135	292	267	359	844	673	185	13	5	7	10	0	10	0	0
173	160	57	20	367	281	160	36	21	0	6	0	0	0	0	0
134	0	10	0	66	100	217	16	72	13	10	0	0	0	0	0
10	5	0	0	0	0	190	155	0	32	13	13	0	0	0	0
0	0	4	0	0	18	328	148	152	0	0	0	0	0	0	0
0	0	0	0	0	0	0	117	75	207	0	0	0	0	0	0
0	0	0	0	0	0	0	35	0	0	0	0	0	0	0	0
0	0	0	0	0	0	15	8	0	0	0	0	0	0	0	0

Appendices

Matrix for average traffic (coefficient a3) in MB of each group

3	6	0	5	2	5	6	1	2	0	1	0	0	7	0	0
70	156	173	89	187	218	222	66	34	4	5	0	134	64	15	0
7	30	324	219	370	315	312	150	20	48	92	140	185	35	4	0
0	27	115	394	669	247	263	251	105	74	378	49	67	0	1	2
58	48	136	276	361	242	88	236	67	130	189	136	156	153	5	1
0	292	283	235	106	229	162	155	229	292	156	115	28	15	0	0
70	49	215	196	194	272	276	136	124	166	32	125	9	96	0	2
73	87	317	517	227	304	188	70	32	11	2	4	2	20	0	0
20	68	154	114	116	358	377	70	8	1	3	3	0	1	0	0
73	73	32	9	206	129	54	20	11	0	1	0	0	0	0	0
71	0	4	0	28	60	102	9	27	1	3	0	1	0	0	0
8	2	0	0	0	0	65	75	0	12	6	6	0	0	0	0
0	0	2	0	0	5	149	65	60	0	0	0	0	0	0	0
0	0	0	0	0	0	0	40	37	127	0	0	0	0	0	0
0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0
0	0	0	0	0	0	4	2	0	0	0	0	0	0	0	0

Matrix for average traffic (coefficient a4) in MB of each group

2	3	0	6	1	5	5	1	1	0	0	0	0	6	0	0
50	69	98	59	118	140	148	43	20	2	3	0	54	38	11	0
4	13	178	138	253	216	221	103	10	30	50	69	101	18	3	0
0	15	57	248	479	289	224	129	75	53	251	35	39	0	1	2
27	22	61	159	241	298	152	205	131	103	125	91	99	93	4	0
0	119	178	150	76	211	182	162	143	169	81	71	18	8	0	0
30	22	134	110	125	203	183	128	55	100	20	77	7	60	0	1
36	38	61	248	163	172	135	41	10	4	2	3	2	15	0	0
9	37	92	87	89	238	231	43	1	2	2	4	0	2	0	0
59	38	12	6	85	81	51	14	3	0	1	0	1	0	0	0
42	0	2	0	14	36	74	4	18	2	3	0	1	0	0	0
4	1	0	0	0	0	53	41	0	7	4	3	0	0	0	0
0	0	2	0	0	3	94	51	42	0	0	0	0	0	0	0
0	0	0	0	0	0	0	51	21	68	0	0	0	0	0	0
0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0
0	0	0	0	0	0	2	1	0	0	0	0	0	0	0	0

Appendices

Matrix for α_2 (in rad) of each group

0.6	1.7	0.7	0.6	1.0	0.3	0.7	0.4	1.6	0.0	0.3	0.0	0.0	0.7	0.0	0.0
1.2	0.7	0.7	0.6	1.1	0.6	0.6	1.1	1.2	1.3	0.2	0.0	-12.2	0.8	0.2	0.0
1.1	1.2	0.5	0.6	0.5	0.6	0.5	0.6	0.7	1.3	0.6	0.6	0.6	0.6	0.2	0.0
0.0	1.7	0.3	0.5	0.3	0.8	1.0	0.5	0.6	0.7	0.7	0.6	0.6	0.0	1.1	0.6
-12.1	0.5	0.5	0.5	0.5	0.4	0.4	0.6	-0.2	0.7	1.2	0.7	0.7	0.7	0.5	12.0
0.0	0.8	0.5	0.6	0.6	0.5	1.0	0.5	0.4	0.6	0.6	0.6	0.7	0.5	0.0	11.8
0.3	0.5	0.6	0.6	0.6	1.0	0.5	0.7	0.6	0.6	1.2	0.7	0.7	0.7	0.0	1.7
-0.1	-11.8	1.2	0.5	0.5	0.4	1.1	0.5	0.7	0.8	1.3	0.0	0.1	1.1	0.0	-4.3
0.5	0.5	0.5	0.6	0.6	0.5	0.6	0.5	0.6	1.1	0.8	-1.5	0.0	0.2	0.0	0.0
0.5	0.6	0.5	1.2	0.6	1.2	0.5	1.2	0.5	0.0	1.4	0.0	-3.3	0.0	0.0	0.0
0.6	0.0	1.1	0.0	1.2	1.2	1.2	1.7	1.1	-0.2	0.7	0.0	-5.5	0.0	0.0	0.0
1.2	0.0	-1.8	0.0	0.0	0.0	0.6	1.1	0.0	0.3	0.8	0.6	0.0	0.0	0.0	0.0
0.0	0.0	1.7	0.0	0.0	0.7	0.6	0.6	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6	0.7	0.5	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	-5.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Matrix for α_3 (in rad) of each group

0.7	2.1	0.5	0.1	-1.3	0.0	0.5	1.0	2.2	0.0	0.7	0.0	0.0	-0.2	0.0	0.0
1.1	0.4	0.2	0.4	1.3	0.4	0.5	1.3	1.5	1.5	-0.1	0.0	-19.2	0.2	0.0	0.0
1.2	1.1	0.3	0.4	0.4	0.4	0.5	0.5	0.8	1.3	0.5	0.3	0.2	0.0	0.3	0.0
0.0	1.8	0.1	0.2	0.2	1.0	1.3	0.4	0.5	0.6	0.5	0.5	0.2	0.0	1.1	0.7
-19.4	-0.3	0.2	0.2	0.2	0.0	0.2	0.5	-0.5	0.3	1.2	0.3	0.4	0.7	0.7	17.3
0.0	-0.1	0.4	0.4	0.5	0.3	1.0	0.6	0.0	0.3	0.3	0.2	0.3	0.4	0.0	17.8
-0.2	0.0	0.3	0.4	0.4	1.2	0.4	0.6	-0.2	0.3	1.4	0.2	0.1	0.5	0.0	1.4
-1.0	-18.4	-0.2	-0.1	0.2	0.2	1.1	0.0	-0.1	0.8	-0.5	-1.2	0.3	1.6	0.0	-10.4
0.3	0.1	-0.3	0.3	0.3	0.2	0.5	0.2	-0.1	-0.1	1.5	-3.7	0.0	-0.5	0.0	0.0
0.2	0.3	-0.1	1.2	0.1	1.2	0.7	1.5	-0.3	0.0	1.4	0.0	-5.1	0.0	0.0	0.0
0.6	0.0	1.4	0.0	1.2	1.4	1.4	2.3	0.7	-2.0	0.8	0.0	-12.8	0.0	0.0	0.0
1.2	-0.4	-4.5	0.0	0.0	0.0	0.6	1.2	0.0	-0.1	1.1	0.1	0.0	0.0	0.0	0.0
0.0	0.0	1.8	0.0	0.0	0.1	0.4	0.5	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.6	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	-9.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	-0.1	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Appendices

Matrix for α_4 (in rad) of each group

1.8	3.0	1.3	0.9	1.0	1.0	1.0	-0.8	2.9	0.0	-3.9	0.0	0.0	0.1	0.0	0.0
1.8	0.6	0.3	0.7	1.9	0.7	0.7	1.8	2.0	1.7	0.3	0.0	-26.1	1.0	-0.2	0.0
1.7	1.9	0.5	0.7	0.7	0.6	0.6	0.7	1.2	1.9	0.7	0.6	0.2	0.2	0.2	0.0
0.0	2.7	0.2	0.4	0.4	1.5	1.7	0.8	0.6	1.0	0.7	0.7	0.4	0.0	1.5	1.2
-26.3	-0.5	0.3	0.5	0.5	0.6	0.7	0.7	-0.7	0.6	1.7	0.6	0.7	0.7	0.9	23.3
0.0	0.2	0.5	0.6	0.5	0.6	1.6	0.3	0.4	0.7	0.7	0.5	0.7	0.9	0.0	18.2
-0.2	0.2	0.6	0.7	0.7	1.7	0.5	0.4	1.1	0.5	2.1	0.5	0.8	0.7	0.0	2.9
-1.3	-24.7	-1.0	0.3	0.3	0.0	1.6	0.2	0.2	1.4	2.7	-0.9	0.3	1.8	0.0	-9.6
0.2	0.7	-0.2	0.2	0.2	0.1	0.8	0.4	-0.4	-4.0	-2.5	-3.3	0.0	-0.3	0.0	0.0
0.5	0.4	-0.2	2.0	0.0	1.7	0.7	2.0	0.0	0.0	-1.9	0.0	-5.7	0.0	0.0	0.0
0.9	0.0	1.4	0.0	1.6	2.0	1.9	3.0	1.3	0.0	-3.8	0.0	-13.2	0.0	0.0	0.0
2.4	-0.2	-4.6	0.0	0.0	0.0	0.9	1.4	0.0	0.0	1.2	0.5	0.0	0.0	0.0	0.0
0.0	0.0	3.6	0.0	0.0	0.6	0.4	0.6	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.6	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	-10.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0