ADDIS ABABA UNIVERSITY

ADDIS ABABA INSTITUTE OF TECHNOLOGY

SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

TELECOMMUNICATION ENGINEERING GRADUATE PROGRAM

Predictive Analysis of Fixed Network Traffic in the City of Addis Ababa

By

Mesganaw Tefera

A Thesis Report Submitted to the School of Electrical and Computer Engineering in Partial Fulfillment of the Requirements for the Degree of Masters of Science in Telecommunications Engineering

Advisor: Yelemzewed Negash (PHD)

October 2021

Addis Ababa, Ethiopia

ADDIS ABABA UNIVERSITY

ADDIS ABABA INSTITUTE OF TECHNOLOGY

SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

TELECOMMUNICATION ENGINEERING GRADUATE PROGRAM

Predictive Analysis of Fixed Network Traffic in the City of Addis Ababa

By: Mesganaw Tefera

Approval by Board of Examiners

Dr. Yalemzewd Negash ______________

Advisor Signature

Dr. Ephrem Teshale ______________

Examiner Signature

Dr. Murad Ridwan ______________

Examiner Signature

Declaration

I, the undersigned, declare that this thesis is my original work, has not been presented for a degree in this or any other university, and all sources of materials used for the thesis have been properly acknowledged.

Mesganaw Tefera ______ ______________

Name Signature

____

Date of submission

Addis Ababa

This thesis has been submitted for examination with my approval as a university advisor.

Dr. Yalemzewd Negash ______________

Advisor's Name Signature

Abstract

The increase in the number of users and the high bandwidth demand of the Internet have increased network traffic at an alarming rate. Currently, social interaction among people has become highly dependent on Internet applications. In addition, government offices, schools, non-governmental organizations, and the private sector automate their services through digital applications. This increase in data demand, which is a source of revenue for Internet service providers (ISPs), must be managed properly in order for them to provide a higher level of service to their customers. One of the solutions to this issue is to conduct a predictive analysis of the network traffic to understand the traffic characteristics, explore the utilization levels of links and network elements' capacity to identify the bottleneck uplinks that need optimization to increase the network performance. Prediction and modeling research conducted on the Addis Ababa city network are based on mobile data, and the output may not entirely explain the unique characteristics of the fixed network behavior. Therefore, the aim of this research is to model and forecast the fixed access network data by exploring two well-known predictors: the recurrent neural network based Long-Short-Term Memory (LSTM) and the statistical Seasonal Auto-Regressive Integrated Moving Average (SARIMA) algorithm. In addition, a temporal traffic network characteristic pattern analysis has been conducted.

Six months of hourly data from a fixed access network was collected by adding sensors to the network management system. Seasonality, non-linearity, and trends were observed in the aggregated fixed network data tests. To evaluate forecast performance accuracy as well as to compare the two models, error performance measuring metrics such as R squared, Root Mean Squared Error (RMSE), Mean Average Percentage Error (MAPE), and Mean Absolute Error (MAE) were used. The thesis's results confirmed that both algorithms are effective and yield satisfactory performance in forecasting non-linear and seasonal fixed access network traffic data. However, LSTM was more accurate than the SARIMA model when evaluated using an error performance matrix. The result of this thesis could be used as an input for planning, designing, and optimization of fixed access networks.

Keywords: Fixed Access network, time series forecast, seasonality, LSTM, and SARIMA model

Acknowledgement

First, I would like to express my sincere gratitude to God for everything. Next, I want to give my endless gratitude to all my families, my Mother, my Wife, and the kids for their unprecedented help along the path of my life.

I would like to extend my sincere thankfulness to my advisor Dr. Yalemzewd Negash for his unreserved guidance, encouragement, and constructive comments. I deeply appreciate his extensive support and follow up in developing and finalizing this thesis.

I am also very grateful to my colleagues at Ethio Telecom for their unlimited support in providing the necessary information and data. Special gratitude goes to Yemeserach and Taye for sharing their expertise and helpful materials. I would like to thank Ethio telecom and Addis Ababa Institute of Technology for providing me the chance to pursue my education in MSC level.

Finally, I need to thank all my friends for their comments and suggestions while preparing this thesis.

Table of Contents

List of Figure

List of Table

List of Acronym

3G	3rd Generation
3GPP	The 3rd Generation Partnership Project
ADSL	Asynchronous Digital Subscriber Line
ANN	Artificial neural network
APON	ATM based Passive Optical Network
AR	Autoregressive
ARFIMA	Autoregressive Fractionally Integrated Moving Average
ARIMA	Autoregressive integrated moving average
BPON	Broadband Passive Optical Network
CNN	Convolution Neural Network
DSL	Digital Subscriber Line
DL	Deep learning
DPI	Deep packet inspection
DOCSIS	Data Over Cable System Interface Spectrum
DSLAM	Digital Subscriber Line Access Multiplexer
eNodeB	Evolved Node B
EPON	Ethernet Passive Optical Network
ETSI	European Telecommunications Standards Institute
FNN	Feed-forward networks
FFNN	Feed-forward neural networks
FL NGN	Fixed Line Next Generation Network
FTTB	Fiber-To-The-Building
FTTC	Fiber-To-The-curb
FTTH	Fiber-To-The-Home
FTTx	Fiber-To-The-x
GPON	Gigabit Passive Optical Network
HFC	Hybrid fiber coaxial
ICMP	Internet Control Message Protocol
IP	Internet Protocol
IPv4	IP version 4
IPv6	IP version 6
ISDN	Integrated Services Digital Network
ISUP	Integrated Services Digital Network User Part
ITU	International Telecommunication Union
KPSS	Kwiatkowski-Phillips – Schmidt -Shin
LAN	Local Area Network
LSTM	Long Short-Term Memory
LTE	Long Term Evolution
MA	Moving Average
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error

MSE	Mean Squared Error
MGW	Media gateway
MGCP	Media Gateway Control Protocol
ML	Machine learning
MSAG	Multi-Service Access Gateway
MSAN	Multi-service Access Node
MSC	Master of scince
MRA	Multi Resolution Analysis
NGN	Next generation network
NGN	Next Generation Network
NMS	Network Monitoring System
ODN	Optical Distribution Network
OLT	Optical Line Terminal
PCM	Pulse-Code Modulation
PDCCH	Physical Downlink Control CHannel
PoP	Points of presence
PtMP	Point-to-multipoint
PtP	Point-to-point
PON	Passive Optical Network
PRTG	Paessler Router Traffic Grapher
PSTN	Public Switched Telephone Network
QoE	Quality of Experience
QoS:	Quality of Service
RNC	Radio Network Controllers
RTP	Real-time Transport Protocol
RMSE	Root Mean Squared Error
RNC	Radio Network Controller
RNN	Recurrent Neural Network
SARIMA	Seasonal Auto-Regressive Integrated Moving Average
SIP	Session Initiation Protocol
SNMP	Simple Network Management Protocol
SPSS	Statistical Package for the Social Sciences
TCP	Transmission Control Protocol
TDM	Time Division Multiplexed/Multiplexing
TTC	Telecommunication Technology Committee
UDP	User Datagram Protocol
UMTS	Universal Mobile Telecommunications System
VDSL	Very High Speed Digital Subscriber Line
VPN	Virtual Private Network

Chapter One
Introduction

1.1 Background

The use of the Internet has increased the network traffic volume tremendously as the Internet has marked its path in every aspect of technology [1]. The growth and popularity of the Internet has been sparked by newly emerging socially acceptable applications such as Torrent, Facebook, Tiktok, Netflix, etc., which enable users to make video calls, share any data, voice or video, attend online classes, or socialize. Furthermore, governments and business owners are automating their services to be available online. As a result, dependable telecom infrastructure is required to meet the needs.

In Ethiopia's case, as of ethio telecom 2013 E.C. (2020/21) annual business performance summary report [2], the number of fixed access customers reached 1.3 million. This number is incomparable with the number of connected mobile customers, which is 54 million, but still, the fixed access network contributes a considerable amount of traffic volume to the backbone network. Furthermore, because of their relative high capacity and low cost, mobile customers prefer fixed access networks, such as Wi-Fi, when transferring large amounts of data or using bandwidth-intensive applications.

Because of the booming data traffic, network operators are trying to look for solutions in which they can optimize the utilization of their network infrastructure. One of the approaches to tackling such challenges is to conduct predictive analysis of the operational network to understand the traffic characteristics, explore the utilization levels of links, the network elements' capacity, and identify the bottleneck links and routers that need a capacity upgrade or optimization to increase the network performance. Traffic analysis on an operational network can also aid the operator in knowing how the network can cope up with high data traffic imposed on the backbone. Furthermore, developing traffic models can help an operator to characterize the traffic trends, assess the performance of the network links and devices to understand and forecast the future capacity demand.

Ethio telecom, a government owned service provider in Ethiopia, is deploying different expansion projects to address telecom services throughout the country. Having the full figure of the current network traffic volume will help the company to deploy new telecom services for its future expansion works. Currently, the company provides triple play services through various fixed access last mile technologies such as Digital Subscriber Line (xDSL) and Passive Optical Network (PON). In addition, Virtual Private Network (VPN) and dedicated point-to-point services are provided for enterprise customers. The majority of the fixed access infrastructure deployment is copper-based Asynchronous Digital Subscriber Line (ADSL)/ Very High Speed Digital Subscriber Line (VDSL) technology, replacing the old Public Switched Telephone Network (PSTN) voice network using multiple access high capacity devices such as Multi-service Access Node (MSAN), Multi-Service Access Gateway (MSAG), and Digital Subscriber Line Access Multiplexer (DSLAM). In major cities like Addis Ababa, there are optical fiber Fiber-to–the-Building (FTTB) and Fiber–to-the-Curb (FTTC) deployments for residential and enterprise customers mainly using Ethernet PON and gigabit PON technology connected to Optical Line Terminal (OLT), MSAG, or MSAN. These IP-based fixed access devices are connected to the fixed aggregation layer and core layer using the Fixed Line Next Generation Network (FL NGN) infrastructure.

There are different factors for traffic volume change and it is mainly dependent on the user's data usage behavior. Thus, exactly knowing where to reach the traffic volume level in the next period could be difficult. Many academics argue that if correlations are observed in historical data, they can be used to make forecasts, so fixed network traffic forecasting based on past data is preferable. There are several forecasting model algorithms available in different domains for different types of data. For time-series traffic data, statistical modeling, machine learning, and deep learning forecasting techniques are the most common and relevant one's. The International Telecommunication Union Recommendation (ITU-R) E.507 explains many techniques for evaluating forecasting models and choosing the model. However, it does not provide any guidance on selecting which models are more appropriate for forecasting the telecommunications series [3]. Thus, the selection of suitable modeling and forecasting methods or algorithms is vital, depending on the deep knowledge of the behavior of given data at hand. Fixed access network traffic exhibits nonlinear, seasonal, and trend characteristics. For this research, to incorporate these behaviors into the forecast model, the Seasonal Auto-Regressive Integrated Moving Average (SARIMA) based model algorithm from the statistical forecasting methods and Long-Short Term Memory (LSTM) from the deep learning recurrent neural network method were selected. Both algorithms are capable of capturing nonlinear characteristics of the data, and in addition, LSTM captures short and long term trends of the traffic into the proposed model

1.2 Statement of the Problem

Over the last two decades or so, the telecom industry has experienced rapid changes and advancements over wired and wireless broadband particularly on Internet data services. Even though wired technology paves a way to the telecom industry the relative technological maturity, mobility, and ease in the deployment of the cellular networks contributes the majority of the connected customers. On the other hand, for high bandwidth and reliable system requirements, fixed access network via fiber or copper as a last-mile is still the best solution and contributed a significant amount of data volume on the backbone network. In the ever-growing telecommunications industry, the ability to determine future trends is an important endeavor [3]. Forecasting of network traffic is very important for tasks such as resource allocation, network design and planning, and optimization. In bandwidth allocation schemes, better traffic forecasts can help to reduce congestion or waste of resources [4]. Moreover, precisely predicted future traffic loads will help us save equipment power and design greener traffic-aware networks as well.

Several studies have been conducted in order to forecast Ethiopian Telecom's network traffic data.[8] Capable of forecasting video streaming traffic on the 3G network. Similarly, [5] modeled and forecasted aggregate UMTS traffic using a hybrid SARIMA-ELM algorithm. [24] manages to model the flow length and flow size of different applications. The researchers used traditional statistical modeling as well as machine learning modeling algorithms to predict the specific parts of the cellular network. However, to the best of my knowledge, this thesis is the first to conduct prediction and modeling research on ethio telecom's fixed access network.

Hence, it is evident that every network or sub-network has a distinctive characteristic and trend [26]. Likewise, fixed-access network traffic data behaves differently from that of cellular network data because the network architecture, technology, and mainly the types of users served by the two networks are quite different from each other. As a result, previously conducted research (prediction and modeling) may not fully describe the fixed access network, and it surely necessities a separate study. Therefore, this thesis aims to conduct a characteristics study and modeling of fixed network data by zooming in into the Addis Ababa city network area.

1.3 Objective

1.3.1 General Objective

The main goal of this research is to make predictive analysis and study the characteristics of ethio telecom fixed access network traffic in Addis Ababa city. Prediction and modeling will be conducted using machine learning algorithm LSTM and SARIMA model.

1.3.2 Specific Objectives

Particularly, the study has the following specific objectives:

- To collect, process, and visualize fixed network data traffic in Addis Ababa city.
- To study the behavior of the fixed network traffic, such as:
    - Peak and off peak hour traffic characteristics
    - Daily pattern
    - Weekly pattern
    - Monthly and other special occasion behaviors
- To review appropriate prediction models for fixed network data traffic.
- Conduct traffic characteristics test on fixed access network data.
- Select appropriate modeling and forecasting algorithms depending on the tested network data behavior.

- Apply modeling and forecasting process on the fixed network data using selected algorithm.
- Evaluate the performances of the implemented models with forecast error metrics (RMSE, MAE, MAPE, R-squared, and MSE).
- Compare the performance results of the proposed models.

1.4 Methodology

The methodology used to achieve the general and specific objectives of this thesis are, as described in figure 1.1.

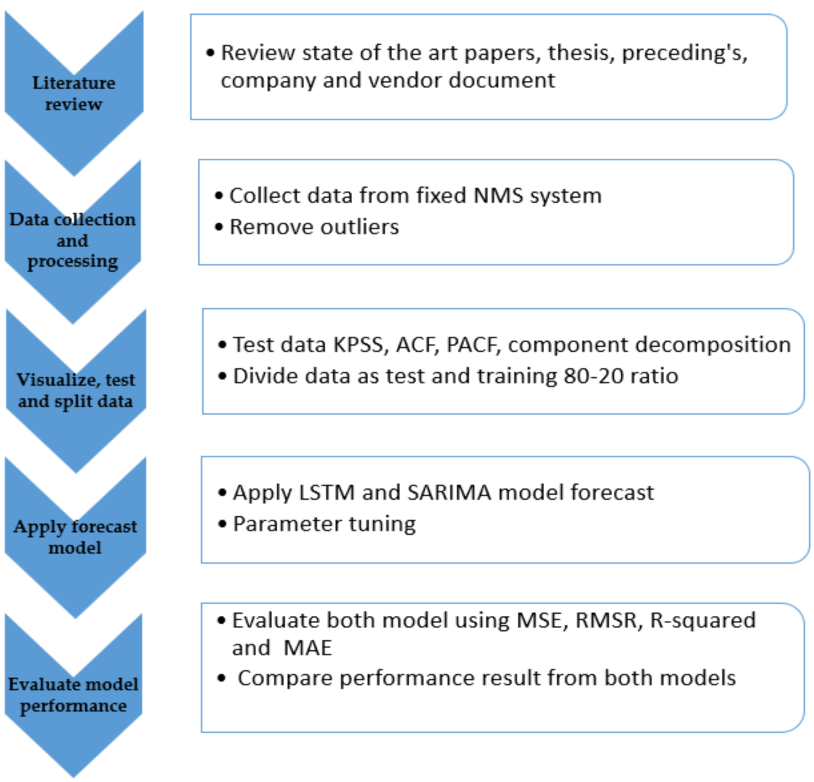

Figure 1.1 Methodology flow chart summary

- ***Literature Review***

Related research studies on traffic prediction and analysis, books, different IEEE articles, conference proceedings, and industry white papers reviewed to use for the thesis.

- ***Data Collection***

To perform the prediction and analysis, fixed access network raw data was collected from 11 network devices in Addis Ababa starting from February 25th 2021, where 4465 hourly-based traffic volume data was collected from each devices. Data collection was performed by adding a sensor on each switches, where the access devices are connected on, using Simple Network Management Protocol (SNMP) based Paessler Router Traffic Grapher (PRTG) software via Network Monitoring System (NMS).

- ***Data Analysis, modeling and forecasting***

The collected time-series raw data was first cleaned, processed, and interpreted. Following data preparations, the behavior of data such as linearity, seasonality, stationarity, and trend was examined using different test procedures. Depending on the behavior of the data, appropriate traffic modeling and prediction using machine learning algorithms, such as artificial neural networks and their variants, LSTM were selected. Furthermore, from statistical network traffic prediction and modeling algorithm, SARIMA's was used and its performance results are used to compare the accuracy of the model with that found using the LSTM algorithm. Temporal network traffic behavior, daily pattern (peak hours), weekly pattern (weekend trend) and monthly pattern analysis are also conducted using hourly collected aggregate data. To evaluate the model's forecasting performance, error metrics such as MAE, MSE, RMSE, MAPE, and R-squared were used. The Python programming language and the Jupyter Notebook editor tool were used for processing, modeling, and predicting the fixed network traffic data. Additional tools, such as Microsoft Excel and SPSS software were used for data processing, analysis, and to visualize the output.

1.5 Literature Review

Network traffic forecasting is a crucial task for any medium-to-large network provider that has received little attention from the computer network community. However, there are a significant number of tasks that have to be done by network providers that would gain from using forecasting methods [28]. This section presents some relevant results and methodology from previous work concerning traffic modeling, prediction, and network traffic characterization.

The author in [8] collected ten months' of data on video streaming data traffic information from five Radio Network Controllers (RNCs) of the Universal Mobile Telecommunication System (UMTS) network in the city of Addis Ababa. The aim was to predict the video streaming data traffic by using the Deep Learning, Long Short Term Memory (LSTM), model that incorporates self-similarity and long term dependence. The author demonstrated that the LSTM model is an effective method for predicting video streaming traffic to reflect temporal patterns with 57.8% of MAE (mean average error) improvement of forecasting error compared to the hybrid model. The author also noted that accuracy is vital to provide a better dynamic resource allocation for video streaming traffic.

In [9] Hoang Duy Trinh et al designed a traffic prediction system using Recurrent Neural Networks (RNN). The design of the prediction system includes LSTM units. They collected mobile traffic information from the Physical Downlink Control CHannel (PDCCH) of the LTE eNodeB using the passive tool at 1ms resolution. For their data, they stated the problem as a supervised multivariate prediction of the mobile traffic, where their objective is to minimize the prediction error given the information extracted from the PDCCH. They are able to run the model and evaluate the one-step prediction and the long-term prediction errors of the proposed methodology. In the end, they compared their proposed architecture with the ARIMA and the deep Feed-Forward Neural Network (FFNN) model. As expected, the LSTM algorithm captures the temporal characteristics of the mobile traffic, and it provides superior accuracy with respect to the Feed-Forward Neural Network or to the classic ARIMA model.

Apart from prediction, categorizing and studying the network data is crucial for network administrators. The authors in [29] collected two months' of data from different parts of a campus network. They generate statistics for 5-minute intervals chosen to generate aggregated statistics in five record category of the volumetric properties of the traffic: basic volume, advanced volume, Layer 3, Layer 4, and Layer 7 statistics. They analyzed the acquired data and divided it into seven areas of network behavior to study the day and night pattern, weekday pattern, flow characteristics, IPv6 utilization, protocol share, and most frequent ports used. As a conclusion, flow-based statistics are more stable than byte or packet-based statistics. Still, IPv4 is dominant, TCP and UDP are the dominating protocols, while ICMP and other protocols are so negligible.

The authors in [11] collected data from the fixed access network of a municipality in Sweden having 2600 households with FTTH broadband access to analyze Internet usage in terms of traffic volume, pattern, applications, and user behavior based on traffic measurement performed using Packet Logic. Their work shows that traffic patterns for weekdays and weekends are almost similar, having peak time from 7 PM-10 PM. In addition, the traffic is asymmetrical with more outbound traffic than inbound traffic for high bandwidth households but the reverse behavior is observed for lower bandwidth users. The households with higher bandwidth (100 Mbps) were found to be high users of the Internet than those with less bandwidth (with respect to traffic volumes and time).

The authors in [13] attempted to conduct a temporal analysis on a set of one-year network traffic data collected from a real IP backbone node in China. In their experiment, they utilized the traffic decomposition method to divide the IP backbone traffic into long-term trends and temporal fluctuations. Their results reveal that the traffic volume grows a little higher than normal days for short holidays (as long as three days). For the long holidays (as long as seven days), the traffic volume grows greatly. They adopted an exponential-fitting model to describe the long-term network traffic nature to characterize the data traffic growth.

The authors in [12] tried to analyze three years of SNMP data collected throughout the IP backbone network and presented research methods that relied on wavelet Multi Resolution Analysis (MRA) and linear time series models for predicting when and where link upgrades or additions have to take place. The authors also computed aggregate demand between any two adjacent Points of Presence (PoPs). They showed that IP backbone traffic exhibits long-term trends and strong periodicities at multiple time scales. They tried to model weekly approximations of the two methods described above using autoregressive integrated moving average (ARIMA) models and developed a prediction scheme. They showed that forecasting network backbone traffic based on their model could result in accurate estimation for at least 6 months.

Literature Summary

From the literature reviewed so far, I have grasped the concept that statistical modeling and prediction methods AR, MA, ARIMA, and SARIMA are the best for modeling time-series network traffic data [8], [9]. These models are computationally simple, but they require to follow a sequence of steps and assumptions, such as model building and manual parameter tuning, unlike machine learning algorithms. To incorporate nonlinear parts of the data into a model, such as self-similarity and dependency, deep learning and machine learning algorithms (CNN, RNN, LSTM,...) are used by researchers. A recent research result in [8] shows the ML neural network algorithm (LSTM) performs better efficiency than traditional statistical methods when predicting video stream traffic in ethio telecom's UMTS network. The authors in [30] claim that statistical forecasting methods are as accurate as sophisticated ML forecasting algorithms. They argue that understanding the reasons for their underperformance is the only way to improve them. Furthermore, the literature indicates that both statistical and ML algorithms provide satisfactory accuracy in modeling and forecasting time series. To get the best result out of it, knowing the data set is critical.

Depending on the reviews, I selected the SARIMA and LSTM algorithms to model and forecast the network traffic data. Analyzing network behavior by classifying the network data in [5], [29], and [31] gives a better understanding of the network, but it requires special data collection tools in addition to the tools used (SNMP-based PRTG software), such as deep packet inspection, or DPI.

1.6 Scope and Limitation

1.6.1 Scope

The research focuses on the analysis and forecasting of fixed access network data on the metropolitan city of Addis Ababa considering the relative network utilization exists in the country. In order to analyze, model, and predict the fixed access network traffic volume, this thesis analyzes the daily, weekly, and monthly characteristics of fixed access network traffic volume pattern flowing on fixed network routers of Ethio Telecom, based on the data collected from NMS.

1.6.2 Limitation

There are different types of network traffic characteristics predictive analysis but due to data unavailability and time constraints, traffic volume predictive analysis and modeling was conducted using the two famous time series algorithm for this research. In addition, this thesis is limited to analyzing and predicting of the fixed access network data in metropolitan city of Addis Ababa, and the result can be mostly applied to the city network.

1.7 Contribution

Predictive analysis on affixed access network helps the ISP to model, predict, and understand the characteristics trend of the network traffic for the optimal use of the available resource. Hence, this thesis forecast output using SARIMA model with ((0,1,1)(1,1,1)24) and LSTM algorithm Hyperparameters to find best performance and aids ethio telecom to:

- Depict peak and off peak hour traffic.
- Know where and when uplink device capacity upgrades are required.
- Understand the traffic volume growth pattern on fixed access network and use the input for optimization.
- Provide insight for temporal properties of traffics that can help in traffic planning, design and operation of the fixed access networks.
- Identify problems associated traffic congestion to improve QoS in fixed broadband services provisioning.

1.8 Thesis Organization

This thesis contains four chapters. Chapter 1 deals with the introduction of the whole thesis. It clearly cites the statement of the problem, the literature review, the objective of the study, methodologies used, its contribution, scope, and limitations of the thesis.

Chapter 2 covers the overview of the fixed access network. The latest access technology exists in the network to provide triple play broadband services, such as xDSL copper-based technology and fiber access FTTx introduction, presented. The last part of this chapter is about the Next Generation Network (NGN) concept.

Chapter 3 presents a brief description of time series modeling and forecasting methods and algorithms. Specifically, statistical and neural network modeling and forecasting methods. Scientific and mathematical approaches to traffic modeling are described, specifically the SARIMA and LSTM algorithms.

Chapter 4 describes the temporal traffic characteristics of fixed access networks, followed by data tests for different characteristics such as seasonality, model selection validation and hyper parameter tuning on both LSTM and SARIMA criteria, and finally, experimental modeling and forecasting results are presented with discussion.

Chapter 5 of the thesis is all about conclusion, recommendation, and possible future work that can be potential research area on fixed access network.

Chapter Two

Fixed Access Network Overview

Fixed access networks are generally referred to as all of the wired and fixed wireless networks that are used for voice, video, and data communications. A fixed line can be seen as a connection to end users by means of a fixed media, through which a user can make phone calls, transmit any data or connect to the Internet. In this chapter, the overview of the general fixed access network is presented. The developmental, architecture, technology, and service overview of the fixed access network will be discussed in the next section, and finally, the concept of the next generation network will be explained.

2.1 The Developments of the Fixed-line Network

The fixed network was initially a purely circuit-switched telephone network. With the development of digital technology and the demand for high speed transmission, the fixed-line network has advanced into a universal, integrated services network. Fixed-line broadband Internet access with a bandwidth of many gigabits per second is now possible with DSL or GPON. Standard protocols are used to deliver end-to-end services and integrate all network elements, maintaining interoperability. Standards, recommendations, and specifications on all matters prepared by organizations, mainly ITUT and others, such as DOCSIS and ETSI, are available.

The first generation of broadband fixed access networks used asynchronous digital subscriber line (ADSL) technology for twisted pair last-mile networks and DOCSIS 1.0/2.0 for coaxial cable networks. Newer technologies, network upgrades, and competition are pushing these speeds up. By upgrading access devices and selectively deploying them in the distribution network by shortening loop lengths, carriers in many countries are able to offer DSL-based broadband services at rates of up to 100 Mbps, depending on the distance to the central office [14] – [17]. In addition to better broadband performance, these speeds have allowed many operators to deploy bundled services including voice and multi-channel video alongside broadband data. Vectoring could double those rates, but it requires relatively substantial central office upgrades. Besides, with current technology, it makes it difficult to offer unbundled access to the local loop, which could reduce the level of competition in multiple ISP environments [21]. However, the ease of upgrading relative to full fiber deployments makes the technology attractive in many countries. Vectoring may make FTTN architectures more interesting in the short- to medium-term strategy, potentially as an initial step toward later deployment of FTTH solutions. On the other hand, it may enable operators to expand coverage plans by using a mixture of FTTH and FTTN. However, it raises important considerations for regulators (status of contacting state and finance). In developing countries like Ethiopia, the choice of developing fiber access infrastructure, especially in new cities where no copper access is available, is the best strategy. Whereas in main cities, a mixture of FTTH and FTTN strategies is considered as a profitable option by retaining already deployed copper infrastructure.

Developments in fixed broadband networks remain important despite the growth of mobile data. Wireless broadband networks still carry far less traffic than fixed networks, and they generally offer lower speeds and reliability. Moreover, the growth of cellular data actually increases demand for fixed access networks [21]. Mobile connections travel over the air medium for a shorter distance, after which they are carried on high capacity wired back-hauling connections. The growth of Wi-Fi connections and other mechanisms for offloading cellular traffic will also place greater demands on wired networks [21].

2.2 Fixed Access Network Architecture

Network architecture selection mainly depends on the type of devices, access networks, core networks, services, and application types. Other than operator requirements, vendor capability to provide specific operational scenarios determines the type of architecture to be implemented. A typical fixed access network architecture including core and access parts is depicted in figure (2.1) below.

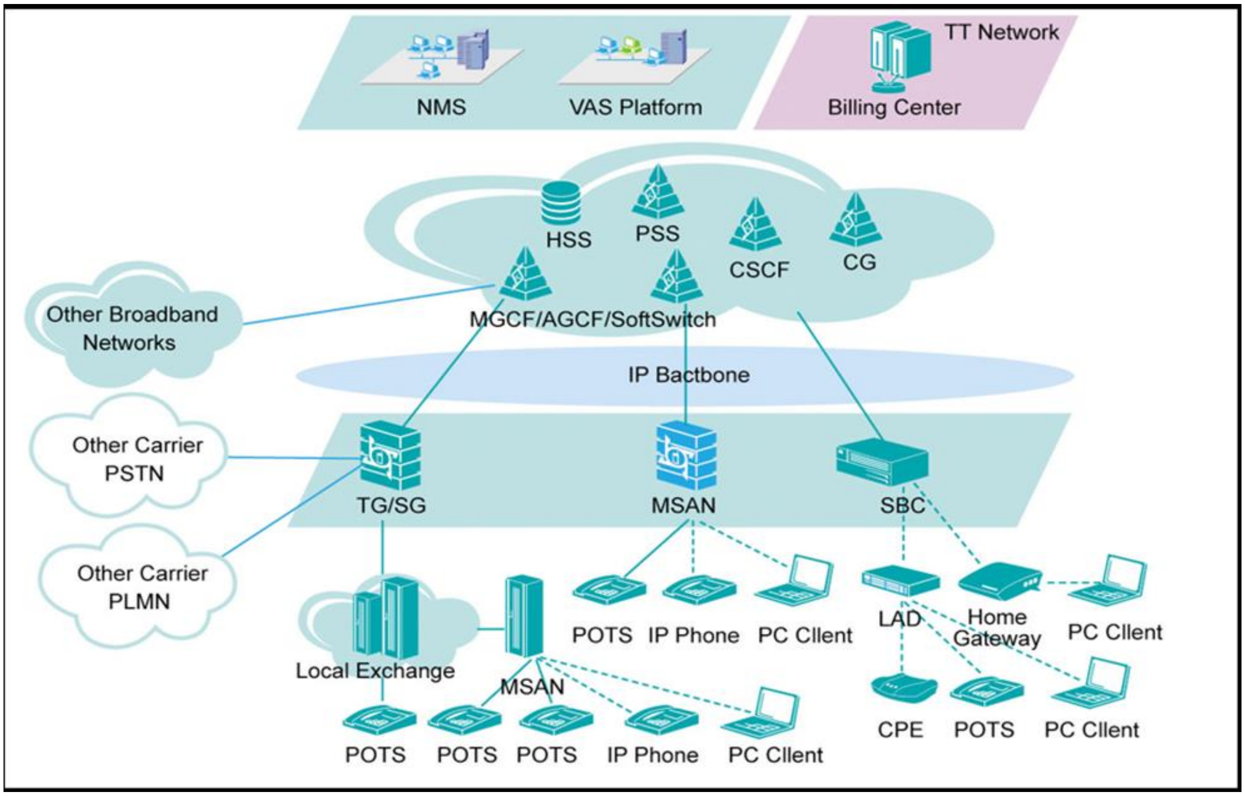

Figure 2.1 Example of fixed access network architecture [15]

From an operator perspective, excluding the user terminal, the fixed access network can be divided into three major network layers, the core network layer, the aggregate layer, and the access network layer [15].

The access network layer has blanket coverage of mostly twisted-pair wires, coaxial cable, and fiber optic lines that ensures many subscriber terminals are connected to the nearby operator's access equipment such as DSLAM, MSAN.

The convergence layer constitutes equipment for the convergence layer, mostly servers and routers, such as Radius servers and user database servers. It realizes the connecting function between access equipment such as MSAN and core network devices using the IP NGN network, providing user accounting and administration.

The core network layer contains lines that mainly have high bandwidth for connecting different network elements such as application servers, switching devices, billing servers, operation and management servers, etc., using layered IP routers.

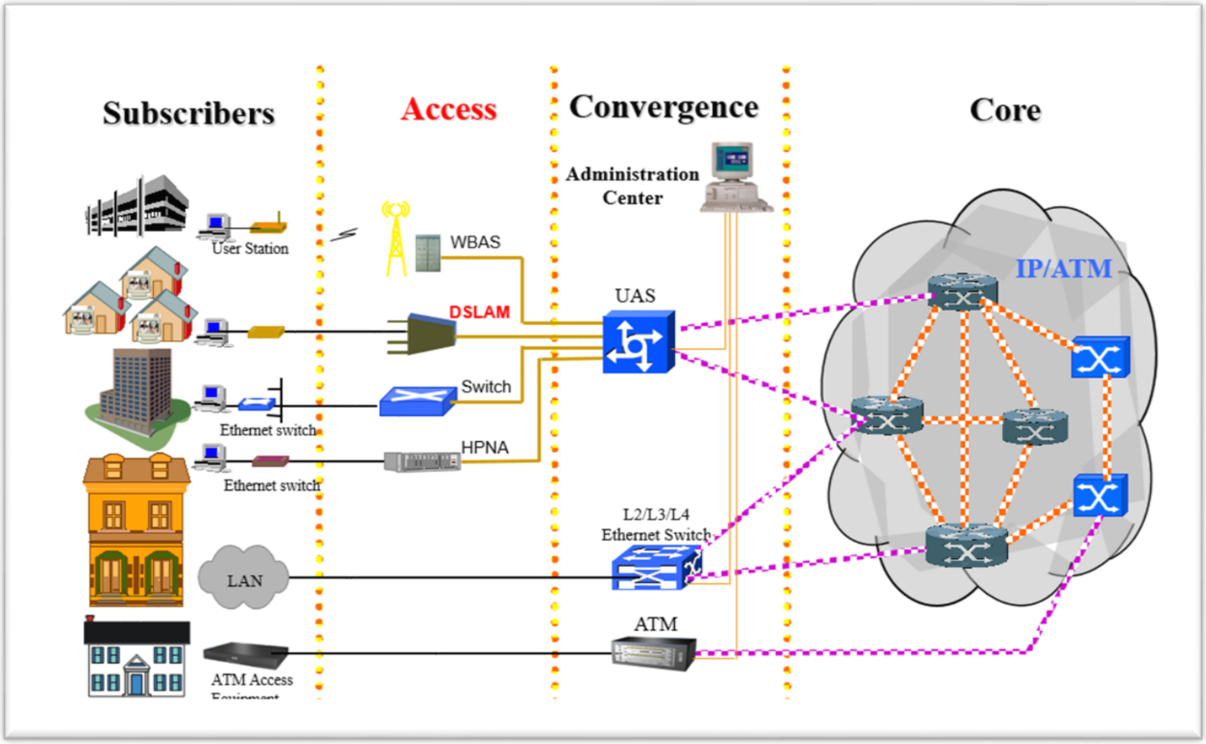

Figure 2.2 Three layered architecture of a fixed access network

2.3 Fixed Access Service Type and Technology

This section discusses the overview of fixed access technologies, mainly broadband access, to deliver triple-play services such as voice, video, and data to subscribers as seen from a last-mile deployment scenario. The section considers the main wired technologies, such as copper and optical technologies used in the access networks. Other technologies such as Broadband Power Line (BPL) and public WiFi have not been included in this study as they are not deployed in the current networks.

2.3.1 Copper Access

2.3.1.1 Digital Subscriber Line (DSL)

Digital Subscriber Line (DSL) refers to copper-based standards that reuse legacy infrastructure to connect customers to the telephony grid, such as ADSL, VDSL, HDSL, SHDSL, and SDSL [14] - [16].The general idea is to utilize and leverage the old and already existing telephony infrastructure to provide customers with broadband services, such as Internet access.

In general terms, all DSL access techniques can be described as in figure 2.3 below. In the telephony grid, each customer is connected in a star-like network to an access device such as DSLAM or MSAN in the central node. The access node is located in the Central Office (CO), where the local telephone switch is also located, or in any continent location according to the deployment scenario. The voice and data signals are physically separated, both in the access device and in the Customer Premises Equipment (CPE), but both signals are sent in a Frequency-Division Multiplexing (FDM) transmission over the twisted-pair copper line in the telephony loop to deliver Internet access.

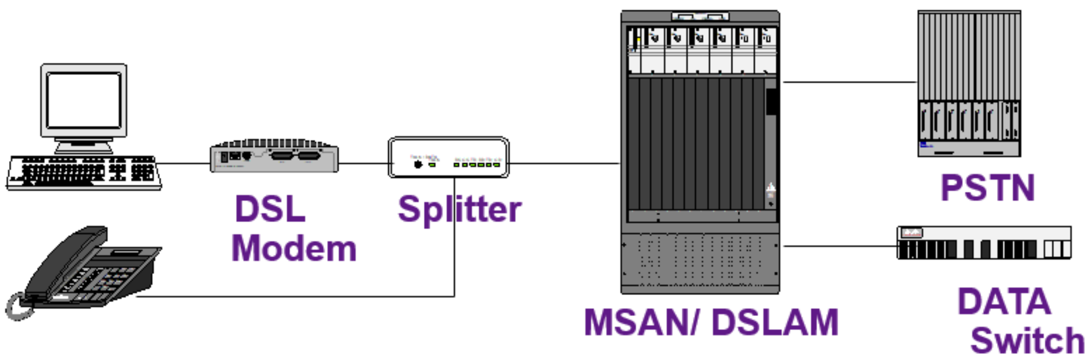

Figure 2.3 A general DSL connection to the customer from a nearby access device

Several DSL technologies have been standardized up to now, to mention a few currently deployed ones. The Asymmetric Digital Subscriber Line (ADSL) was standardized in 1999 ITU-T G.992.1 [32] can provide up to 8Mbps downlink and a maximum range of 5km. ADSL2plus in 2002 G.993.5 [34] can provide up to 24Mbps downlink twice to that of ADSL2. Very High Bit Rate DSL (VDSL2) in 2006 G.993.2 [33] provide up to 100Mbps within 300m reach. For this technique, the DSL signals can use a much higher frequency band, up to 30 MHz, to reach bit rates in the order of a hundredth of a Mb/s. To make use of these frequencies, the length of the cable loop must be shorter than ADSL. For loops that are shorter than 1.5 km VDSL provides more capacity than ADSL. That means, as the loop distance and operating frequency rate increase, cable attenuation, interference, and cross talk increase exponentially that reduces the performance of VDSL to reaches to that of the ADSL. Figure (2.4) indicates the throughput performance of various DSL technologies against distance.

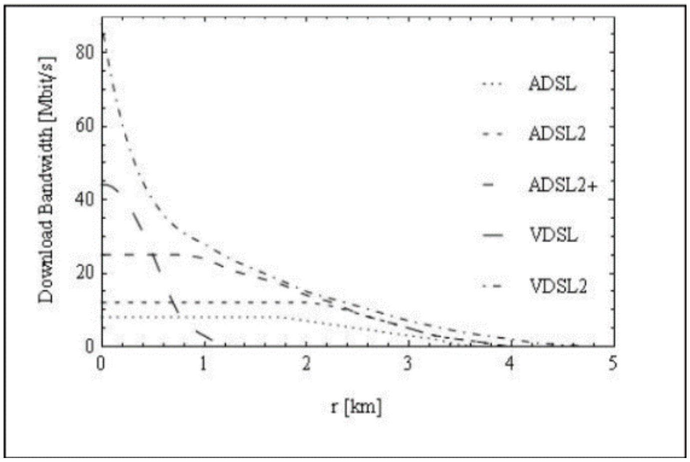

Figure 2.4 Theoretical DSL throughput versus distance curve

To overcome the limitations incurred, techniques like bonding and vectoring were introduced based on the existing DSL technology, which reduces the interference as distance increases and is standardized by ITUT [15], [16].

In VDSL2, crosstalk is a primary limitation and its effect dominates loops shorter than one kilometer. The most relevant type of crosstalk in VDSL2 and ADSL2+ is Far End Cross Talk (FEXT), consisting of the interference of one line's signal over the signal that travels in the neighbor pairs along the cable propagating in the same direction. By means of interference cancellation, vectoring minimizes the effect of FEXT in both transmission directions, downstream and upstream. Vectoring in VDSL2 systems is defined in the ITU-T Recommendation G.993.5 [34]. Vectoring is optimal for cables shorter than 1 km and without FEXT coming from non-vectored systems (without unbundling). Literature suggests that for poorly isolated cables and little twisting, some improvement can still be achieved for loops between 1 and 1.4 km. If noise other than FEXT interference is very high, the positive effects of vectoring are extremely reduced. Besides, if a cable has multiple binders, the vectoring gain will be higher when applied to the whole cable since coupling from different binders, although slightly, affects the victim pair [15] and [16].

Bonding is the simultaneous use of multiple DSL pairs to enhance total throughput. In bonding, the data channel of multiple DSL pairs is actually bonded and the throughput of a set of n bonded DSL pairs is approximately the sum of the rates of the individual DSL pairs. The ITU has standardized bonding in the G.998 (or G.bond) series. G.998.1 [38], G.998.2 [39], and G.998.3 [40] are standards for ATM-based multi-pair bonding, Ethernet-based multi-pair bonding (ADSL2+, VDSL2), and time-division inverse multiplexing, respectively.

Fast access to subscriber terminals (G.fast) is the short name given by the ITU and the standard defined in G.9701 [42] to define a new copper transmission technology focused on providing much higher bandwidth up to 1Gbps over copper loops out to 200 meters. G.fast achieves this performance by increasing the signal spectrum to 200 MHz.

2.3.1.2 Power Line and DOCSIS

Cable TV, communication over coaxial cable, and power line communication are unreached technologies in our country. The overview of these two technologies is discussed as follows. Originally, cable TV networks were only configured for one-way transmission of TV and radio programs from the content provider to subscribers using coaxial cable. In order to offer telephony and broadband Internet access, the operators upgraded their networks to allow two-way transmission, creating next generation cable TV networks, which are typically hybrid fiber coaxial (HFC) networks in recent times. This technology is popular in developed countries such as Europe and America [16]. ITUT has approved a global physical transmission standard as DOCSIS (Data over Cable Service Interface Specification). The latest DOCSIS standard was released in 2017, as DOCSIS4.0 and can deliver up to 6 Gbits/s upstream and 10 Gbits/s downstream. The European also has their own annexed standard, published under the name EuroDOCSIS, with a little modification from the global standard.

Power Line communication is an emerging broadband fixed access technology and is currently under various development and field trials. Its major advantage is its ability to reach almost every household in the country since the power lines are installed virtually everywhere. The service is estimated to be gradually available to consumers at an affordable price in the coming years [15], [16].

2.3.2 Optical Fiber Access

An optical fiber is an optical transmission media that works on the principle of total internal reflection. Because of the low cost, light weightness, small loss, long transmission distance, and easy transportation and deployment, optical fibers have been rapidly developing and replacing copper since the 1980s. Optical fibers are progressively extended from cells to buildings, offices, and homes to desktops, making FTTH develop in full swing and all-fiber home coverage a reality. In a broadband access network, the coverage distance of optical fibers can reach dozens of kilometers. Single-mode optical fibers are generally used in the downlink to meet the coverage requirements of different distances [14]. Single mode or multi-mode optical fibers are selected in accordance with upper layer equipment connectors and are adapted in accordance with different connectors. When selecting optical fiber, the operating wavelength, system capacity, and transmission distance should be considered comprehensively. In a real network, from the operator's perspective, the initial fiber deployment construction work cost is a huge factor in the return on investment. As a result, a careful FTTx network planning design is required.

2.3.2.1 FTTx Introduction

FTTx is an acronym referring to all possible optical fiber topologies that access network service providers can use to implement an access network from the central office (CO) to their customers. FTTx means Fiber-to-the X, and depending on the destination or deployment scenario (far or close to the customers' premises) it is possible to have different network topologies. Currently, there are many different FTTx alternatives; some of the most common are Fiber to the Home (FTTH), Fiber to the Building (FTTB), Fiber to the Curb (FTTC), Fiber to the Node (FTTN), etc. FTTx networks typically use equipment in the CO, such as OLT or MSAN, which is shared among the subscribers connected to it through an Optical Distribution Network (ODN).

Depending on the powering requirements of the equipment installed in the optical distribution network, it is possible to differentiate between Passive Optical Networks (PON, without elements inside the ODN requiring electrical power) and Active Optical Networks (AON, with some elements that require it). Additionally, FTTx networks can be deployed with a dedicated fiber for each subscriber, using a point-to-point (PtP) connection, or with a fiber that is shared by multiple subscribers, using a point-to-multipoint (PtMP) connection. AON can cover a maximum range of 100km and encompasses an active component like amplifiers, repeaters, or shaping circuits. Whereas PON has a maximum range of 20km and is significantly less expensive due to simple passive components such as splitters and combiners [14], [16], [17]. A typical FTTx type and implementation diagram is shown in figure 2.5.

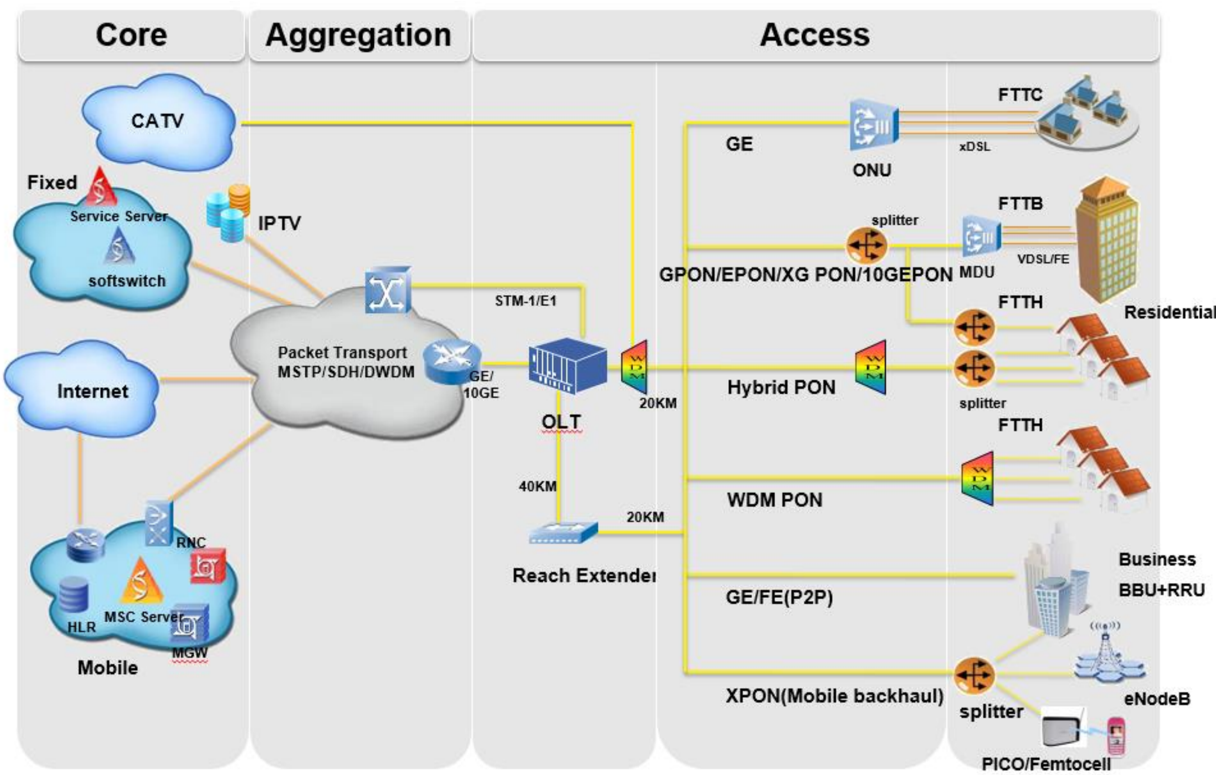

Figure 2.5 FTTx implementation scenario on optical access network [15]

Ethernet-based PtP is popular in the core, metro networks, LANs, datacenters, etc. Whereas, PON in PtP/PtMP is used in geographically distributed access part implementation like FTTH, PtP and PtMP technologies standard carried out by IEEE as a part of Ethernet technology as Ethernet-based optical access system or IEEE 802.3ah. The technical standardization section of the International Telecommunication Union (ITU-T) has a standard such as ITU-T Recommendation G.986, G.988, G.997.1, and other new standards, which include network monitoring and management.

2.3.2 PON Technology

A passive optical network (PON) based access network is an efficient and vibrant technology, which can meet the ever-increasing bandwidth demand. PON has been extensively investigated due to its high bandwidth, cost-sharing of infrastructure and absence of active components [17]. PON components, the optical fiber and splitters are truly 'passive', with no electrical power required. That makes a PON inherently efficient from an operating cost standpoint. In a PON network, an optical line terminal (OLT) is placed at the head end of the network. A fiber optic cable runs from the OLT to an optical beam splitter, which multiplies the signal and relays it to many optical network terminals (ONTs). The passive splitter did not re-shape the signal, boost optical power, or select a specific wavelength. The passive network can be deployed in three connection types: point-to-point (PtP), point-to-multipoint (PtMP), and ring [19] as shown in figure 2.6.

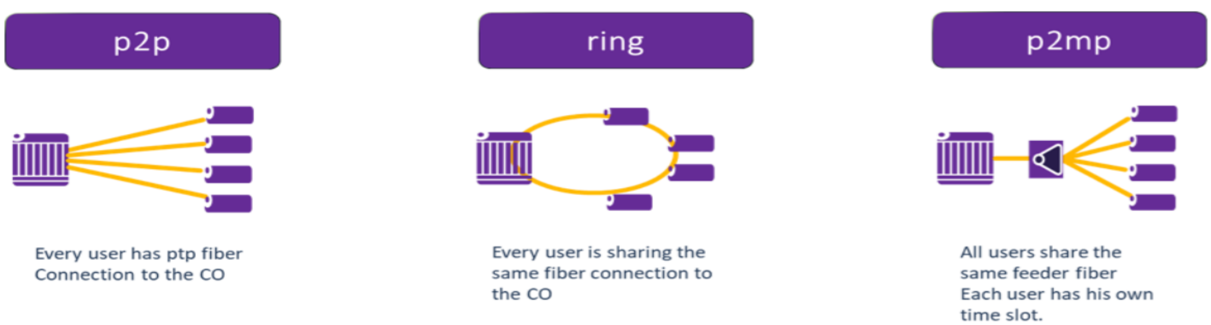

Figure 2.6 Three types PON connection architecture

So far, there are some different PONs that have been deployed or are being deployed, including asynchronous transfer mode PON/broadband PON (APON/BPON), Gigabit PON (GPON) and Ethernet PON (EPON). Standards for PON technologies come from ITU-T, FSAN, and parallel efforts by IEEE [18]. A typical APON/BPON provides 622Mbit/s of downstream bandwidth and 155Mbit/s of upstream traffic using G.983.1 to G.983.5 standards [18], [17].The GPON standard (G.984.1 to G,984.4) represents a boost in both the total bandwidth and bandwidth efficiency through the use of larger, variable-length packets. A GPON network has a downstream data rate of up to 2.5Gbit/s and an upstream data rate of 1.25Gbit/s [17], [18], and [20]. EPON employs standard 802.3 Ethernet frames with symmetrical up- and down-stream data rates of 1.25Gbit/s [18], [20].All these PONs use the time division multiplexing (TDM) access technology, where the bandwidth of a single wavelength is shared among all users (typically 32 users in a PON), and hence are referred to as TDM-PON. As a result, although GPON and EPON can offer an aggregated bandwidth of over 1Gbit/s, the bandwidth for each user may be a maximum of 100Mbit/s [17]. The evolution and data rate of the different PON technologies are summarized in figure 5.7.

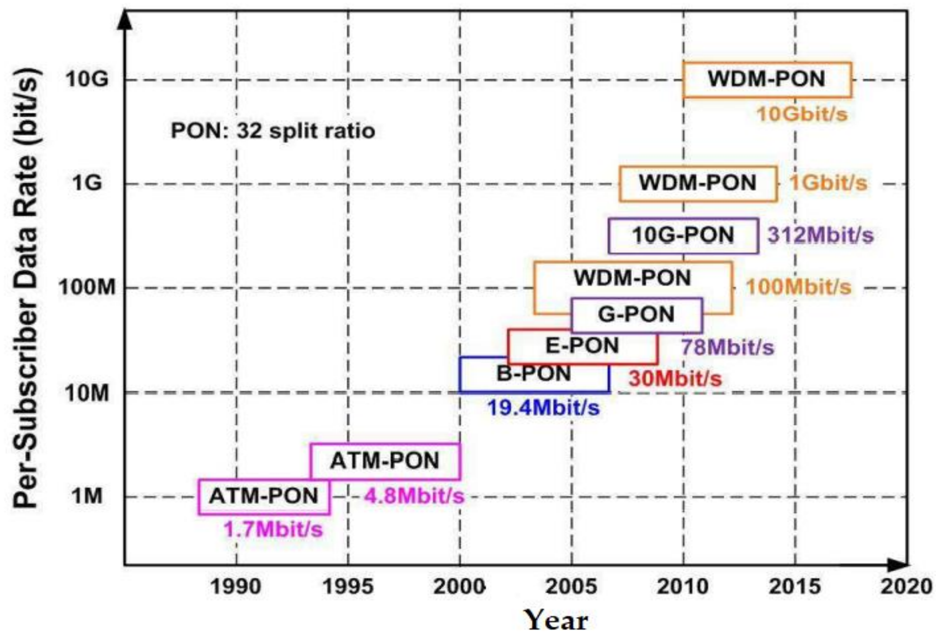

Figure 2.7 Evolution of optical access networks [17]

2.4 Concept of Next Generation Network

According to the ITU-T definition, a Next Generation Network (NGN) is a packet-based (IP) network that utilizes various broadband, QoS-enabled transport technologies, separates service and signaling functions from transport functions, offering users unrestricted access to services and various service-providers, supporting mobility that allows for consistent and ubiquitous service provisioning [14]. The NGN definition integrates various functionality into IP-based networks, including legacy fixed access services, and the term describes major architectural evolutions in telecommunication networks. NGN is a technology that will enable a single network to transport multiple traffic formats: voice, data, video, by containing these in packets. It is a packet-based network, which has its functionality that is service-related and it is independent of any primary transport-related technology.

NGN is a fusion of various service provider networks, such as the PSTN, the World Wide Web(WWW), the wireless network, etc. A principal objective of NGN is the support for high bandwidth, high mobility, and most importantly, the provision of a composite communication model [16]. A major motivation for the NGN is to provide new and diverse services, value-addition, IP-centered services, and applications, as well as reduce capital expenses and operational expenses, by utilizing the network resources more efficiently [14].

Initially, the telecom trend was to develop separate applications for both transport protocols (primarily Real-Time Protocol-(RTP) as the primary voice transport protocol) and signaling protocol stacks (primarily H.323, but also Session Initiation Protocol-SIP and Media Gateway Control Protocol – (MGCP) in recent times). Somewhere between these two concepts, a new model has appeared, with the design and creation of the first IP based phone exchanges SS (soft switches) capable of processing large numbers of VoIP calls. Deployed on standard server hardware (either in discrete servers or recently on blade-based systems), these systems came at the fraction of the cost of classical phone exchanges, and provided scalability beyond the capacities of classical exchange.

With IP-based NGN networks, it was possible to separate the voice traffic (RTP) from signaling (H.323/SIP/MGCP) in the voice core network, so Softswitch only processes signaling, which is not extremely traffic-intensive, while RTP voice flows cross the voice core network only between their entry points (either VoIP terminals or media gateways).

On the other hand, a classical switched phone network (PSTN) still exists, and these entry points were the development and deployment bottlenecks, as both media and signaling traffic had to be bridged between two networks based on a disparate set of standards. Here, two systems can be identified, a Media Gateway (MGW) which converts PCM encoded voice to RTP-based IP voice;

and a Signaling Gateway (SG) which converts messages from the classical Signaling System 7 stack (SS7) into the H.323 (or SIP) signaling stack used by Softswitch [14], [15]. Signaling gateways comprise of hardware systems that insert and extract SS7/ISUP (ISDN User Part) messages from El/T1 trunks, and systems talking ISUP on one side and H.323 or other IP-based signaling protocol stack on the other, recognizable by central Softswitch system.

The NGN functional model is organized into a three-layered architecture: the Transport layer, the Service Control layer, and the Application layer. As mentioned above, the complex multitasking future is made possible with typical elements such as Softswitch, media gateway, and signaling gateway [16]. The main advantage of NGN is enabling operators to provide any service on alternative access platforms. It is known that access technology has its own separate evolution apart from NGN development but to provide alternative access solutions for end customers such as DSL, Ethernet PtP, and PON, especially after the introduction of the famous IP based access device MSAN operators migrated their legacy network to NGN. The MSAN is an ideal device for the FTTx scenario, which is capable of providing various services, ranging from legacy narrowband (POTS and ISDN) to the most recent xDSL broadband solutions such as VDSL2+ over existing copper local loop and fiber access solutions in PON technology such as the EPON and GPON series.

In June 2004, Recommendations Y.2001 "General overview of NGN" and Y.2011 "General principles and the general reference model for Next Generation Networks" were accepted, and 13 papers describing the outline part of NGN release 1 began the approval procedure for becoming recommendations at the SG13 meeting held in July 2006. The use of the IP Multimedia Subsystem (IMS) designed by the 3rd Generation Partnership Project (3GPP) [16] is a characteristic feature of the NGN stated by Supplement 1 to Y.2000-series "NGN release 1 scope" and Y.2201 "NGN release 1 requirements."

The Japanese and the Europeans also have their own annexation of NGN standards and prepared several documents through Telecommunication Technology Committee (TTC) and (European Telecommunications Standards Institute (ETSI).

Though it has huge advantages and applications, NGN has some major disadvantages, such as migration complexities, not all existing infrastructure can be shut down, and regulatory restrictions for critical services.

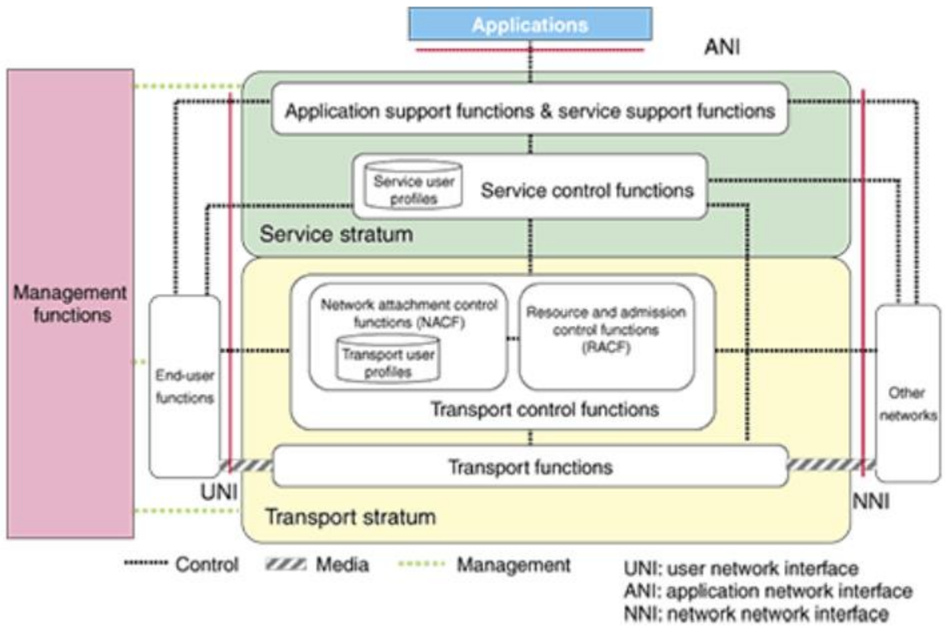

Figure 2.8 ITUT NGN architecture overview

Chapter Three

Concepts of Time Series Modeling and Forecasting

3.1 Definition of a Time Series

Time series arise as recordings of processes that vary over time. The records are a sequential set of data points, measured over consecutive times. It is mathematically defined as a set of vectors X(t), t=0,1,2,3,...(where t represents the time elapsed). The variable X(t)is treated as a random variable [44]. The measurements during an event in a time series arranged in chronological order. Many kinds of data can be gathered into time series, as long as there is a time dependence in the data. A time series that contains records of a single variable is termed a univariate time series. However, if a record contains more than one variable, it is referred to as a multivariate time series. Regular time series are defined as time series that are generated at regularly spaced intervals of time (e.g., daily temperature) [44].The data in a time series, on the other hand, does not have to be collected at regular intervals. It's referred to as an irregular time series in this circumstance.

A time series can be categorized as continuous or discrete. In a continuous time series, observations are undertaken at every instance of time, but in a discrete time series, observations are measured at discrete points of time. Commonly, in a discrete time series, the successive observations are recorded at equally spaced time intervals such as hourly, daily, weekly, monthly, or yearly time separations.

A time series can be stationary or non-stationary. A statistical property of a stationary time series like mean, variance, autocorrelation, etc., does not depend upon time. Whereas a statistical property of a non-stationary time series changes over time. Predicting a stationary time series is relatively easy because the statistical property of stationary time series will be similar in the future, as it has been in the past. The existence of properties such as trend, changes in variance, seasonality, etc., is an indication of non-stationarity in a given time series data. Furthermore, depending on randomness or probabilistic aspects, a time series can also be divided into deterministic time series and non-deterministic/stochastic time series.

3.2 Time Series Analysis

Time-series analysis process contains methods that attempt to understand the nature of the series and are often useful for modeling, future forecasting, and simulation. To investigate the nature of a given time series data, which affects the values of an observation, the analysis process mainly depends on the analysis of the major four components of the time series. These main components are secular trends, seasonal variations, cyclic variations, and irregular or random variations.

3.2.1 Secular Trend

Secular trend shows the overall tendency of the records to increase or decrease during an extended period of time. A trend may be a smooth, general, long-term, or average tendency. It is not always necessary that the rise or fall within the same direction throughout a given period of time. It is observable that the tendencies may increase, decrease, or be stable in several sections of time. However, the general trend must be upward, downward, or stable. The population, agricultural production, the number of births, and the number of deaths are examples of some kinds of tendencies of movement.

The trend may behave linearly or non-linearly. When plotting the time-series values on a graph in accordance with time, the pattern of the data shows the type of trend. If the set of data clustering is a straight line or roughly a straight line, then the trend is linear, otherwise it is non-linear.

3.2.2 Seasonal Variations

Seasonal variations are the rhythmic forces that operate in a regular and periodic manner. The pattern repeats itself in a similar manner over a period mostly less than one year. This variation will be noticeable in a time series when the data is recorded hourly, daily, weekly, quarterly, or monthly.

Seasonal variations come into play, either because of natural causes or man-made conventions. Climatic conditions play an important role in seasonal variations. Such as the production of crops depends on seasons, the sale of umbrellas and raincoats in the rainy season, and the sales of electricity and charcoal increase exponentially in summer seasons.

The effect of artificial conventions like festivals, customs, habits, and a few occasions like marriage month is definitely noticeable. They recur themselves year after year. An upswing during a season should not be considered as an indicator of higher business conditions.

3.2.3 Cyclic Variations

Cyclic Variations are variations in a time series that maneuver themselves over a span of more than one year. This cyclic movement has a period of oscillation of more than a year and is usually called the 'Business Cycle'.

Prosperity, recession, depression, and recovery are the four phases of the cycle constituents. The cyclic variation may be regular but not periodic. The upswings and the downswings in business depend upon the joint nature of the economic forces and the interaction between them.

3.2.4 Irregular or Random Variation

There is another factor which causes the variation in the variable under study. They are not regular variations and are purely random or irregular. These fluctuations are unpredictable, unforeseen, uncontrollable, and erratic. Some examples of random or irregular variation are earthquakes, wars, floods, famines, etc.

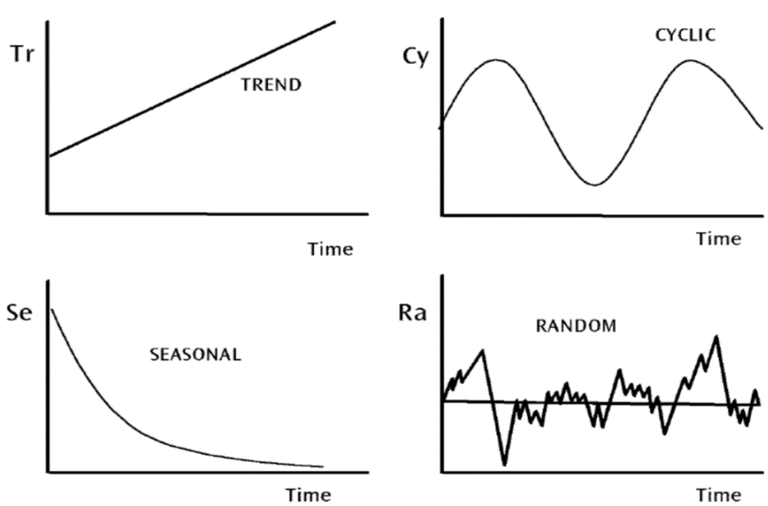

Figure 3.1 Four component of time series graphical example

3.3 Model for Time Series Analysis

In time-series, quantitative data is arranged in accordance with its occurrence and resulting statistical series. The quantitative values are usually recorded over equal time intervals like daily, weekly, monthly, quarterly, half-yearly, yearly, or another type of time measurement. Major forces that influence time series analysis are four components of the time series. Some are continuously effective, while others make themselves felt at recurring time intervals. Components of time series provide a basis for the explanation of behavior in the past time and help us predict the future.

Modeling a time series is to estimate and separate the four types of variations and to bring out the relative effect of each on the overall behavior of the time series. Additive, multiplicative, and sometimes mixed models are used for the decomposition of time series into their four components.

1. **Additive model**

   The additive model assumes that a given data is the summation of all four components of a time series.

   Mathematically denoted as $F(t) = T(t) + S(t) + C(t) + I(t)$ (3.1)

   Where: T denotes the trend component, S denotes seasonality, C is for cyclic variation, and I is for irregular variation.

2. **Multiplicative model**

   The multiplicative model assumes that any given data is a product of all four components of time series.

   Mathematically denoted as $F(t) = T(t) \times S(t) \times C(t) \times I(t)$ (3.2)

3. **Mixed model**

   Depending on the different assumptions inferred, the combination of additive and multiplicative models will result in different mixed model. The mathematical notation also differs based on the assumption.

$$F(t) = T(t) \times S(t) + C(t) \times I(t) \text{ or } F(t) = T(t) + S(t) \times C(t) \times I(t) \text{ or}$$
$$F(t) = T(t) \times S(t) \times C(t) + I(t) \quad (3.3)$$

3.4. Time Series Modeling and Forecasting Methods

Forecasting, in general, is a process of predicting some future event or events. Danish physicist Neil Bohr said in the book "Bad Predictions" that making predictions is not an easy task. Therefore, a proper selection of forecasting methods or algorithms is vital, depending on the deep knowledge of the behavior of the given data at hand. The employment of a model to predict future values based on previously observed values is known as time series forecasting. There are several forecasting model algorithms available in different domains for different types of data. For time-series traffic data, statistical modeling, machine learning, and deep learning forecasting techniques are the most relevant ones for this thesis. The following sections will elaborate them in detail.

3.4.1 Statistical Methods for Forecasting Time Series

In simple terms, statistical forecasting is the use of statistics based on historical data to anticipate what could happen in the future. For time series forecasting, there are a variety of statistical techniques available. However, the few effective and relevant for traffic modelling and forecasting that were selected are: Auto Regression (AR), Moving Average (MA), Auto Regression Moving Average (ARMA), Autoregressive Integration Moving Average (ARIMA), Seasonal Autoregressive Integration Moving Average (SARIMA). The first three, AR, MA, and ARMA, are stochastic processes and are used for stationary time series. Stochastic means that the values come from a random probability distribution, which can be analyzed statistically, whereas the last two models incorporate non-stationary components of a time series. The concept of stationarity and autocorrelation functions will allow us to explain the models' predictive properties mathematically.

3.4.1.1 Auto Regressive (AR)

The AR model establishes that a realization at time t is a linear combination of the previous realization plus some noise term. The autoregressive process of order p, AR (p), defined by equation 3.4

$$X_t = \sum_{j=0}^{p} \emptyset_j X_{t-j} + E_t \tag{3.4}$$

Where the noise term denoted by $E_t \sim n(0, \sigma^2)$ and $\emptyset = (\emptyset_1, \emptyset_2, \emptyset_3, \ldots, \emptyset_n)$ is a vector model coefficients.

P is a positive integer and indicates the order of AR. AR(1),AR(2) …is termed as first order, second order auto regression, and so on.

The 1st order Autoregressive model, denoted by AR(1) is:

$$X_t = \emptyset_1 X_{t-1} + E_t \tag{3.5}$$

The 2nd order Autoregressive model, denoted by AR(2) is:

$$X_t = E_t + \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} \tag{3.6}$$

The qth order Autoregressive model, denoted by AR(p) is:

$$X_t = E_t + \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \ldots \emptyset_j X_{t-j} \tag{3.7}$$

The lag values of the process are expressed by the lag operator, which is denoted by B

$$BX_t = X_{t-1}, B^2 X_t = X_{t-2}, B^3 X_t = X_{t-3}, \ldots, B^d X_t = X_{t-d}, \tag{3.8}$$

The characteristic polynomial $\Phi(B)$ of the process AR (p) defined by equation 3.9

$$\Phi(B) = 1 - \sum_{j=1}^{p} \emptyset_j B^j = 1 - \emptyset_1 B^1 - \emptyset_2 B^2 - \emptyset_3 B^3 - \cdots \emptyset_p B^p \tag{3.9}$$

AR(p) process is rewritten in the following equation.

$$X_t \Phi(B) = E_t \quad where \quad t = 1, 2, 3\ldots \text{ n} \tag{3.10}$$

The characteristic polynomial root determines whether it is stationary or not.

3.4.1.2 Moving Average (MA)

The moving average model accounts for very short-run autocorrelation and states that the next observation is the mean of every past observation. A Simple moving average term in a time series model is a past error (multiplied by a coefficient).

The moving average process of order q, MA (q), defined by equation 3.11

$$X_t = E_t + \sum_{j=1}^{q} \theta_t E_{t-j} \tag{3.11}$$

Where the noise term is denoted by $E_t \sim n(0, \sigma^2)$ and

The lag operator notation, the MA(q) process is given by the equation

$$X_t = \Theta(B) E_t \tag{3.12}$$

$$\text{Where,} \quad \Theta(B) = 1 + \sum_{j=1}^{q} B^j \theta_j \tag{3.13}$$

Similarly, the first order of moving average is denoted as MA(1), the second is MA(2) and so on.

The 1st order moving average model, denoted by MA(1) is:

$$X_t = E_t + \theta_1 E_{t-1} \tag{3.14}$$

The 2nd order moving average model, denoted by MA(2) is:

$$X_t = E_t + \theta_2 E_{t-2} \tag{3.15}$$

The qth order moving average model, denoted by MA(q) is:

$$X_t = E_t + \theta_1 E_{t-1} + \theta_2 E_{t-2} + \ldots \theta_j E_{t-j} \tag{3.16}$$

3.4.1.3 Autoregressive Moving Average (ARMA)

An ARMA(p, q) model is a combination of AR(p) and MA(q) models and is suitable for univariate time series modeling.

Mathematically, an ARMA(p, q) model is represented as

$$X_t = E_t + c + \sum_{j=1}^{q} \theta_t E_{t-j} + \sum_{j=0}^{p} \emptyset_j X_{t-j} \tag{3.17}$$

Where E_t is the noise term with $n(0, \sigma^2)$ and c is a constant term and mostly omitted for model simplicity. The model order ARMA(p, q) p and q refer to p autoregressive and q moving average terms.

3.4.1.4 Autoregressive Integration Moving Average (ARIMA)

The AR, MA, and ARMA described above are stochastic processes and are used for stationary time series. Stochastic means that the values come from a random probability distribution, which can be analyzed statistically. Time series, which contain trends and seasonal patterns, are non-stationary in nature. To incorporate a non-stationary component of time series data into the model, integrated ARMA (ARIMA) was introduced. The model has a differencing process that effectively transforms the non-stationary data into a stationary one by applying finite differencing of the data points [3].

Mathematically, an ARIMA (p, d, q) model is represented by Lag polynomial as

$$X_t \Phi(B)(1-B)^d = \Theta(B)E_t \tag{3.18}$$

$$X_t\left(1-\sum_{j=1}^{p} \emptyset_j\, B^j\right)(1-B)^d = \left(1+\sum_{j=1}^{q} B^j\, \theta_j\right)E_t \tag{3.19}$$

Where the terms p, d, and q refer to the order of the autoregressive, integrated, and moving average parts of the model respectively, and are integers greater than or equal to zero. The integer d controls the level of differencing, usually d=1 first order difference. When differencing parameter "d" allows non-integer values, the ARIMA model is represented by a generalized model as Autoregressive Fractionally Integrated Moving Average (ARFIMA).

3.4.1.5 Seasonal Autoregressive Integrated Moving Average (SARIMA)

As discussed above, one of the main components of time series is seasonal variation. To incorporate seasonal components into a time series data forecast and modeling, additional seasonal differences are included in the ARIMA model. Therefore, SARIMA Seasonal Autoregressive Integrated Moving Average, or Seasonal ARIMA, is an extension of the ARIMA model that supports time series data with seasonal components.

This model is generally termed as the SARIMA (p, d, q) × (P, D, Q) m model.

Where the term (p, d, q) is similar to that of ARIMA model parameters. Whereas (P, D, Q) m hyperparameters are representations of the four seasonal components. Seasonal differencing parameter orders P,D,Q and m are integers greater than or equal to zero. The P,D,Q and m referred to as the order of the seasonal-autoregressive, seasonal-difference, seasonal-moving average, and the number of time steps for a single seasonal period part of the model, respectively.

Mathematically, the SARIMA (p,d,q)×(P,D,Q)m model is represented by the Lag polynomial as:

$$\Phi_P (B^s)\, \phi_p (B)\, (1 - B)^d\, (1 - B^s)^D X_t = \Theta_Q (B^s)\theta_q (B)E_t \tag{3.20}$$

$$\text{i.e.} \quad \Phi P\, (Bs)\phi\, p\, (B)\, Z_t = \Theta Q\, (Bs)\theta q\, (B)\, E_t \tag{3.21}$$

Here Zt is the seasonally differenced series.

3.4.2 Machine Learning Methods for Forecasting Time Series

3.4.2.1 Artificial Neural Network

A neural network is a typical machine learning system designed to operate like a human brain, with the objective of mimicking the intelligence of the human brain in a machine. Human information processing takes place through the collaboration of many billions of neurons, interconnected and sending signals to each other. Similarly, a Neural Network is a network of artificial neurons, as found in human brains, for solving AI problems such as image identification, classifications, forecasting, etc. ANNs try to recognize regularities and patterns in the input data, learn from experience and then provide generalized results based on their known previous knowledge. Information that flows through the network determines the structure of the ANN, because the neural network learns according to the input data. Artificial neural networks are the best modeling tools used for nonlinear statistical data, where a complex relationship between input and output patterns is found or modeled. Due to this flexible capability, ANN is considered an innate candidate for forecasting time series data.

One of the major advantages of neural networks over traditional expert systems is the ability to learn from experience, which is data, and alter its behavior accordingly without expert intervention or previous sets of rules. Here, learning process for ANNs can be viewed as a problem of updating network architecture and connection weights so that a network can efficiently perform a specific task.

The three types of learning methods are:

- ***Supervised Learning:*** we are given a data set and already know what our correct output should look like, having the idea that there is a relationship between the input and output.
- ***Unsupervised Learning:*** to approach problems with no idea about what the results look like. It derives structure from data where we do not necessarily know the effect of the variables.
- ***Reinforcement Learning:*** aims at using observations gathered from the interaction with the environment to take actions that would maximize the reward or minimize the risk.

The choice of selecting the type of learning method depends on the type of task intended to accomplish and the selection of model algorithm required to solve the task. For this thesis, supervised learning methods are used for modeling, forecasting, and evaluation of the model by dividing the data into training and test sets.

Neural Network architectures

ANNs are weighted directed graphs in which artificial neurons are nodes, and directed edges (with weights) are connections between neuron outputs and neuron inputs. Based on the architecture pattern, ANNs can be grouped into two major categories [45]:

- Feed-forward networks (FNN), graphs without loops, and
- Recurrent networks (RNN), loops occur because of feedback connections.

Feed-forward networks are the most frequent type of neural network and neurons are grouped into layers with unidirectional connections between them. Different connectivity yields different network behaviors. Feed-forward networks are static, in the sense that they produce a single set of output values rather than a sequence of values based on the input. Feed-forward networks are memory-less in the sense that their response to an input is independent of the previous network state. Recurrent networks, in contrast, are dynamic systems. When the first input pattern is presented, the neuron outputs are computed. Due to the feedback paths for every other later input, neurons are then modified, which leads the network to enter a new state for every input pattern in a sense that memories will store in neurons.

An example of the two category of ANN shown in figure 3.2-A and 3.3-B

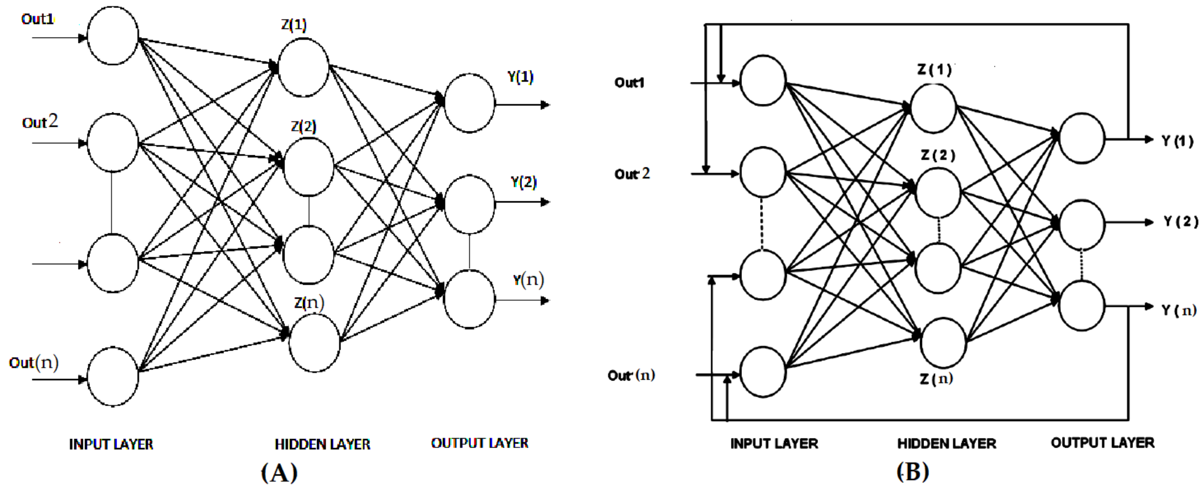

Figure 3.2 (A) ANN Feed-forward graph; (B) ANN recurrent (with loop) graph

When comparing the forecasting and modeling efficiency of the above two ANN categories. According to [8], [46], both FNN and RNN networks produce satisfactory results when forecasting time series. However, Danko Brezak and et al. [46] confirmed that RNN has great effectiveness and potential of dynamism in modeling and predicting nonlinear processes rather than feed-forward neural networks. Furthermore, the feed forward neural networks constrain the inputs and targets to be vectors of fixed length [50] and are not recommended for forecasting problems, as it fails to show long-term dependencies in data. Therefore, considering the highest efficiency of the RNN algorithm, I selected RNN and specifically LSTM (RNN variant) for modeling and forecasting the data. The following section will discuss Deep Learning (DL) RNN and its continuation LSTM model in detail.

3.4.2.2 Recurrent Neural Networks

A recurrent neural network, as defined above, is a deep learning (DL) special type of feedback artificial neural network suitable for modeling and forecasting nonlinear processes [7], [4]. In theory, RNNs employ their feedback connections to store activations or memory representations of recent input events. These deep learning algorithms are commonly used for ordinal or temporal problems and have potential significance for many applications, including speech recognition, forecasting, non-Markovian control, and music composition [8], [47], [48]. RNN is considered deep learning due to the fact that it has more complex ways of connecting layers, automatic extraction of features, computing power, and has more neurons count to express complex models [48].

RNN architecture

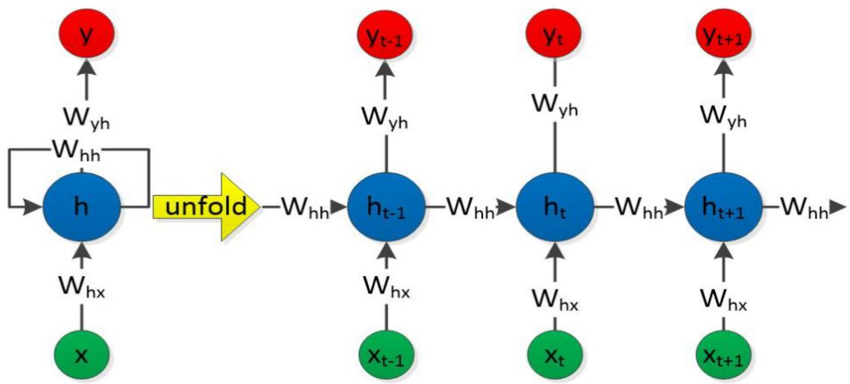

Figure 3.3 Unfolded RNN architecture

Given an input time series X = $\{x_1, x_2, \cdots, x_T\}$, the RNN computes the hidden state sequence H = $\{h_1, h_2, \cdots, h_T\}$ as well as the output sequence Y = $\{y_1, y_2, \cdots, y_T\}$ iteratively using the following set of equations [49].

$$H_t = f\,(W_{hx} t_x + W_{hh} h_{t-1} + b_h) \tag{3.22}$$

$$Y_t = g\,(W_{yh} h_t + b_y) \tag{3.23}$$

In the above two equations (3.22) and (3.23), W_{hx}, W_{hh} and W_{yh} denote the input-hidden weight matrix, the hidden-hidden weight matrix, and the hidden output weight matrix, respectively. Whereas the vectors and represent the bias of the hidden layer and the output layer, respectively. In addition, f(·) and g(·) are the activation functions for the hidden layer and the output layer, respectively. The RNN uses the hidden state at time step t to memorize the network. The hidden state captures all the information included in the previous time steps. In simple RNN, when the interval of data dependencies in time series increases, the model fails to predict multi-step ahead due to the vanishing gradients problem. As a result, since I am using long network traffic data, I skipped using RNN and chose LSTM, which is an extension of RNN and capable of learning long-term dependencies for forecasting fixed network data. The following section explains the working principle of LSTM.

3.4.2.3 LSTM -Long Short Term Memory

LSTM network is an extension of recurrent neural network that designed explicitly to overcome failure in long term dependency problem (vanishing/exploding gradient) of RNN [47]-[49]. Hochreiter (1991) and Bengio, et al. (1994) explored the problem in standard RNN in depth, later, Hochreiter & Schmidhuber (1997) introduced LSTM for its excellent ability to memorize long-term dependencies, and it has been refined by many people since then.

The LSTM network has a series of repeating modules similar to that of a standard RNN network, but the hidden layers of LSTM have a more complicated structure. In each hidden layer, LSTM specifically offers the concepts of gates and memory cells.

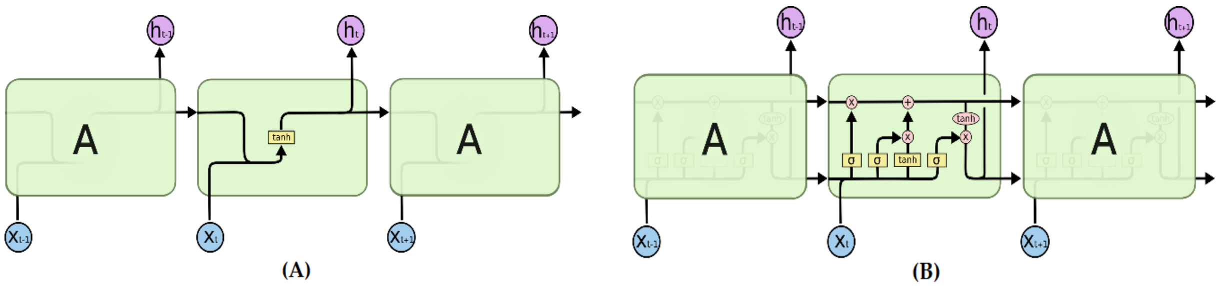

Figure 3.4 (A) RNN single layer; (B) LSTM four interacting layers

As shown in figure 3.4-A and 3.4-B above, the repeating module in a standard RNN contains a single layer, there are four, interacting in a very special way.

An input gate **I**, a forget gate **f**, an output gate **o**, and self-connected memory cells **C,** represent the majority of a memory block. The input gate controls the entry of the activations to the memory cell. The output gate learns what cell activations to filter and output to the successive network. The forget gate helps the network to forget the past input data and reset the memory cells. In addition, multiplicative gates are applied carefully to make it possible for the memory cells to access and store the information over a long time interval. This structure can effectively mitigate the vanishing gradient problem and makes LSTM an architecture suitable for problems with long-term dependencies [49]. Depending on the applications, selecting a specific architecture, output layer, activation functions, and so on varies. For example, in traffic flow prediction modeling, the linear regression layer is best applied to the output layer of the LSTM cell [50]. Although to get the advantage of time series forecasting, LSTM architecture with peephole connections scheme is applied [49].

Given an input time series as X = {x_1, x_2, · · · , x_T}, the hidden state sequence as H = {h_1, h_2, · · · , h_T } the output time sequence as Y = {y_1, y_2, · · · , y_T}. LSTM computation defined in the following set of equations.

$$h_t = \mathrm{H}\,(W_{hx}x_t + W_{hh}h_{t-1} + b_h) \tag{3.24}$$

$$P_t = W_{yh}h_{t-1} + b_y \tag{3.25}$$

Where W is a weight matrix and b is a bias vector

The hidden state of memory cell is denoted in the following equations

$$i_t = \sigma\,(W_{ix}x_t + W_{hh}h_{t-1} + W_{ic}c_{t-1} + b_i) \tag{3.26}$$

$$f_t = \sigma\,(W_{fx}x_t + W_{hh}h_{t-1} + W_{fc}c_{t-1} + b_f) \tag{3.27}$$

$$c_t = f_t * c_{t-1} + i_t * g\,(W_{cx}x_t + W_{hh}h_{t-1} + W_{cc}c_{t-1} + b_c) \tag{3.28}$$

$$O_t = \sigma\,(W_{ox}x_t + W_{hh}h_{t-1} + W_{oc}c_{t-1} + b_o) \tag{3.29}$$

$$h_t = O_t * h(c_t) \tag{3.30}$$

$$\sigma = (\mathrm{x}) = \frac{1}{1+e^x} \tag{3.31}$$

Where σ is a standard sigmoid function , ∗ denotes the scalar product of two vectors or matrixes, and g and h are extends of stand sigmoid function with the range changing [-2,2] and [-1,1]. The squared loss function or root mean squared loss function is used for the objective function e, by the following formula:

$$e = \sum_{t=1}^{n}(y_t - p_t)^2 \tag{3.32}$$

Where y_t is real output value and p_t is predicted traffic volume.

Chapter Four
System Modeling and Traffic Analysis

This section studies and identifies the network traffic characteristics, articulates the test results on the given dataset, and discusses the system modeling for both models. Understanding the behavior of traffic is important, so traffic behavior will be presented first.

4.1 Daily, Weekly, and Monthly Network Traffic Trends

Day and night traffic pattern

The day-night behavior of the network discloses a lot of information about the purpose of the network. As shown in Figure 4.1, the network traffic does not behave at night as it does during the daytime, resembling the natural human activity trend. The aggregate data traffic in figure 4.1 starts on Monday, March 1, 2021, labeled as Day-1 and continues up to the 7th of March, Day-7. The data indicates that traffic usage on weekdays starts to increase in the early morning and reaches its first peak around 09:00. The second and the highest peak is around 13:00 in the afternoon. The last and smallest peak reached around 19:00 in the late evening, followed by the minimum traffic volume usage during the nighttime. The daily weekend traffic, as shown in figure 4.3, is different in volume as well as in peak pattern. Figure 4.2 shows that on Saturday, the highest peak is around 10:00 in the morning, whereas on Sunday the highest peak is around 18:00 in the late evening. This trend, the weekend and weekday pattern, is the same for every day.

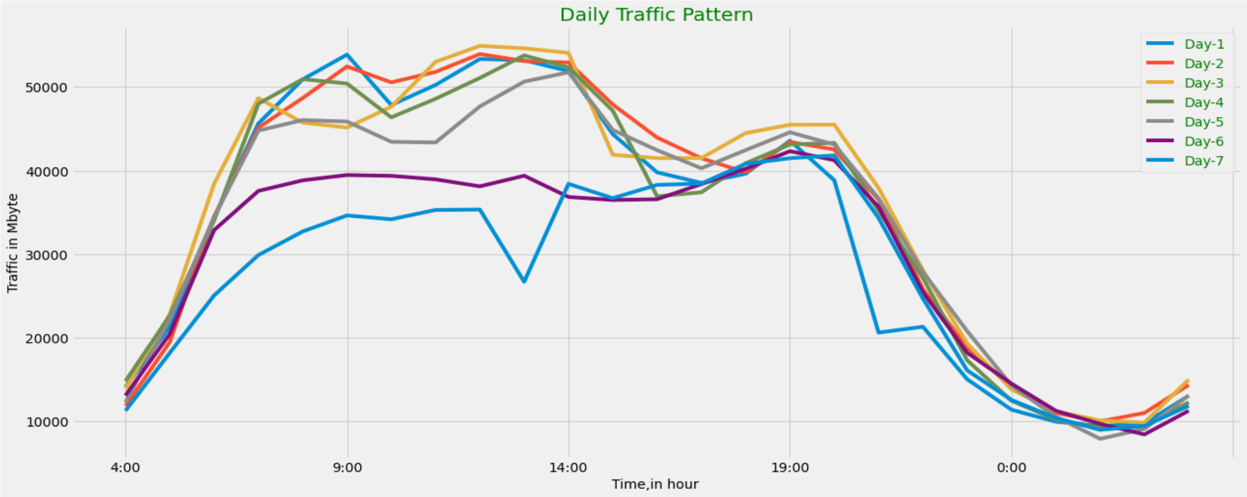

Figure 4.1 Aggregate daily traffic pattern consecutive 7 days in month of March

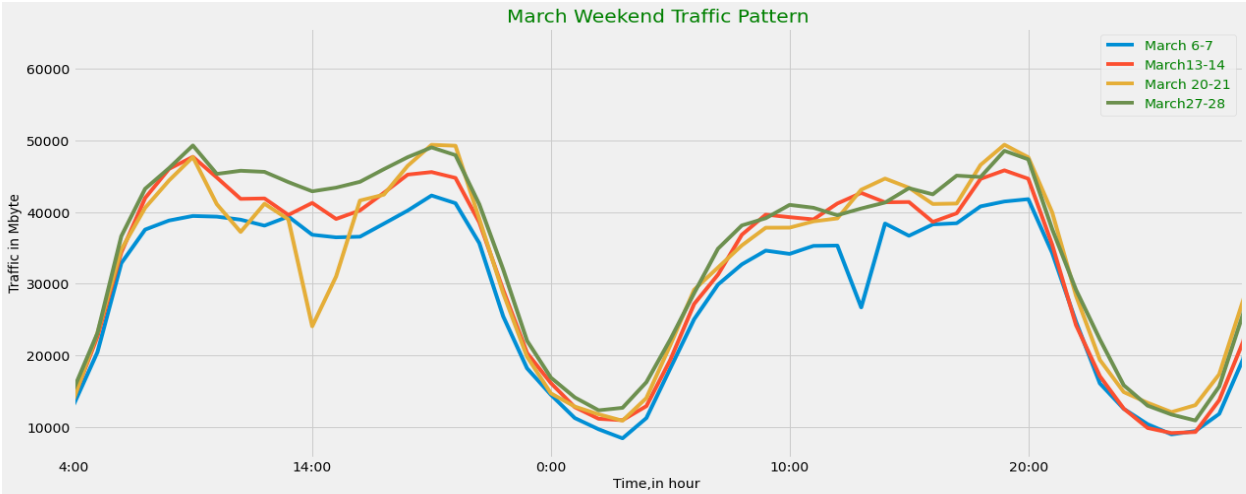

Figure 4.2 March weekend traffic pattern

Weekly and monthly traffic pattern

Another significant pattern that can be observed in fixed network traffic is the weekend pattern. Figure 4.3 shows the aggregate traffic pattern for the consecutive four weeks of the month of March 2021, which shows an almost similar pattern repeated every week throughout the month. Furthermore, as depicted in figure 4.4 circled in red, the traffic pattern on weekdays has a similar pattern, but during the weekend, the traffic volume reduces considerably. A similar pattern was observed for the rest of the weeks in all six months.

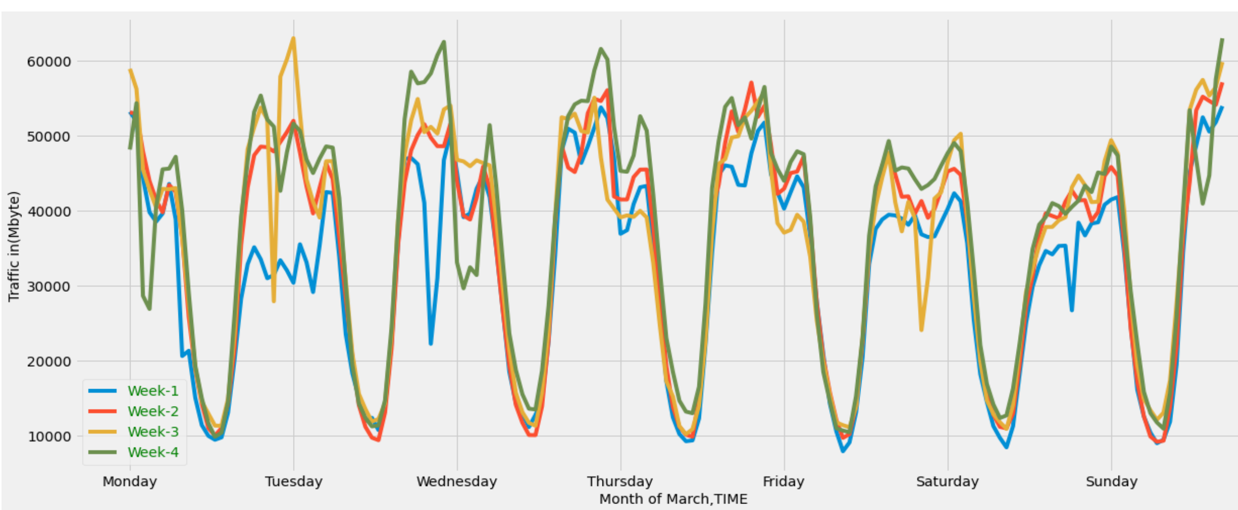

Figure 4.3 Weekly aggregate traffic pattern

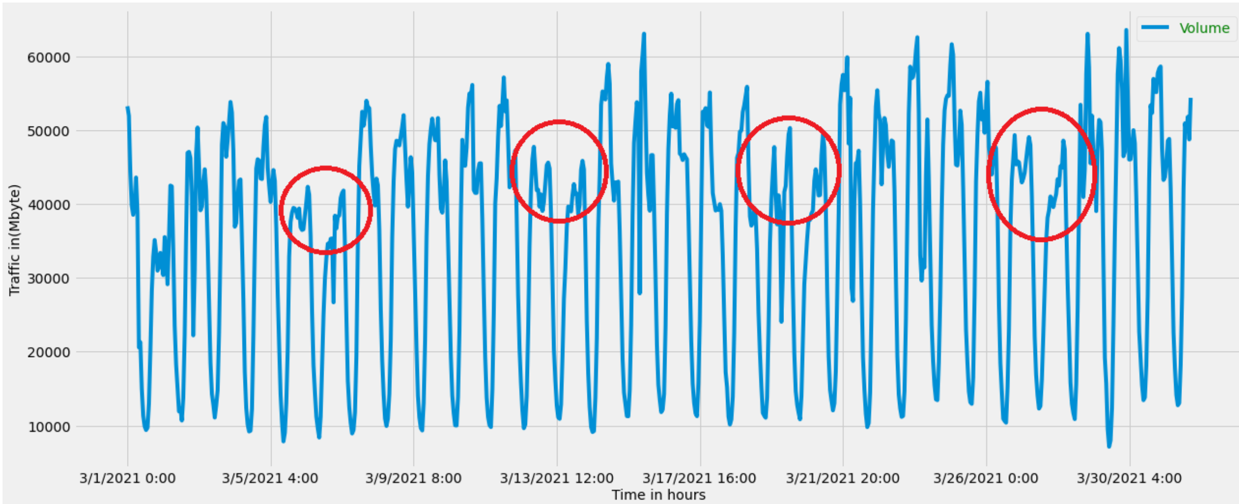

Figure 4.4 Aggregate traffic pattern for the month of March 2021

4.2 Dataset

When comparing the average utilization of fixed access network devices in ethiotelecom nationwide, data card level as well as network equipment level; Addis Ababa city/region network has the highest level of utilization and also generates the highest data traffic. Therefore, the raw data was collected from 11 access network devices located in Addis Ababa starting from February 25th, 2021, up to August 30th, 2021, and was able to collect 4465 hourly-based traffic volume data (in Mbytes) observations from each access device. The data collection is performed by adding a sensor to each switch where the access device is connected to, using SNMP-based PRTG software via Network Monitoring System (NMS).

4.3 Data Preprocessing

In the raw data on the two monitored access devices (code 841 and 851), 2, 32, and 30 rows of missed observations values were found in the month of March As presented in figure 4.5, the device 841, the missing value is circled in red. Thus, the exponential-smoothing algorithm is used to estimate and impute missed values of the time series data in order to maintain consistency and completeness of the data for analysis. Exponential smoothing is a time-series forecasting method for univariate data that can be extended to support data with trend or seasonal components. It is a powerful forecasting method that may be used as an alternative to the popular Box-Jenkins ARIMA family of methods [53]. After estimation and filling in the missing observations, the dataset is presented in figure 4.6. Furthermore, the significance of these estimated and replaced values is normalized during the aggregation process with the rest of the other raw data. Finally, the nature of the aggregate six month traffic data, as presented in figure 4.7, has no significant outliers or missing values.

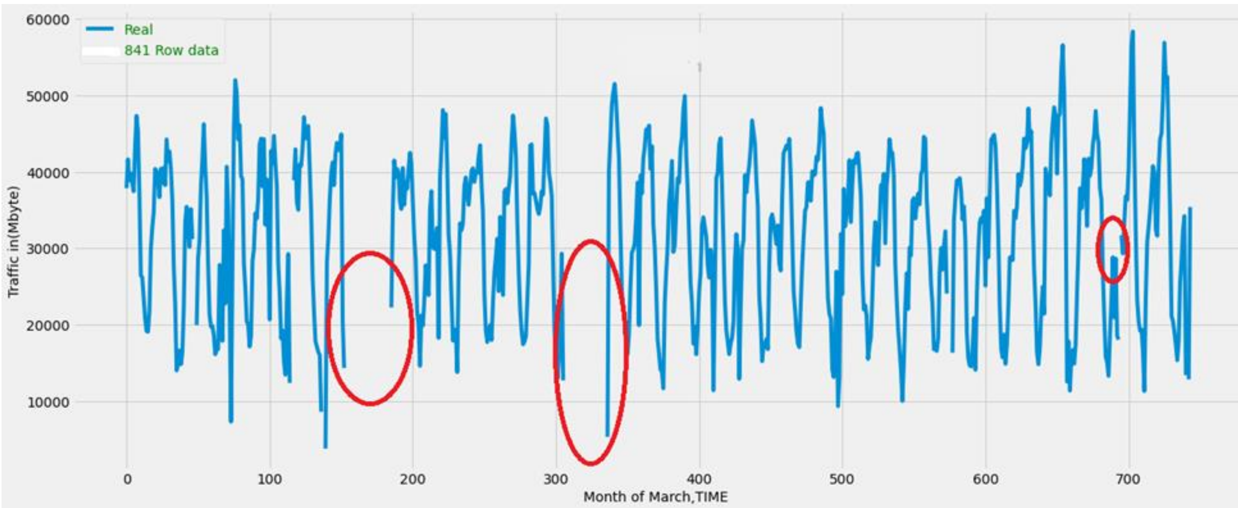

Figure 4.5 Missing values in device 841 circled in red

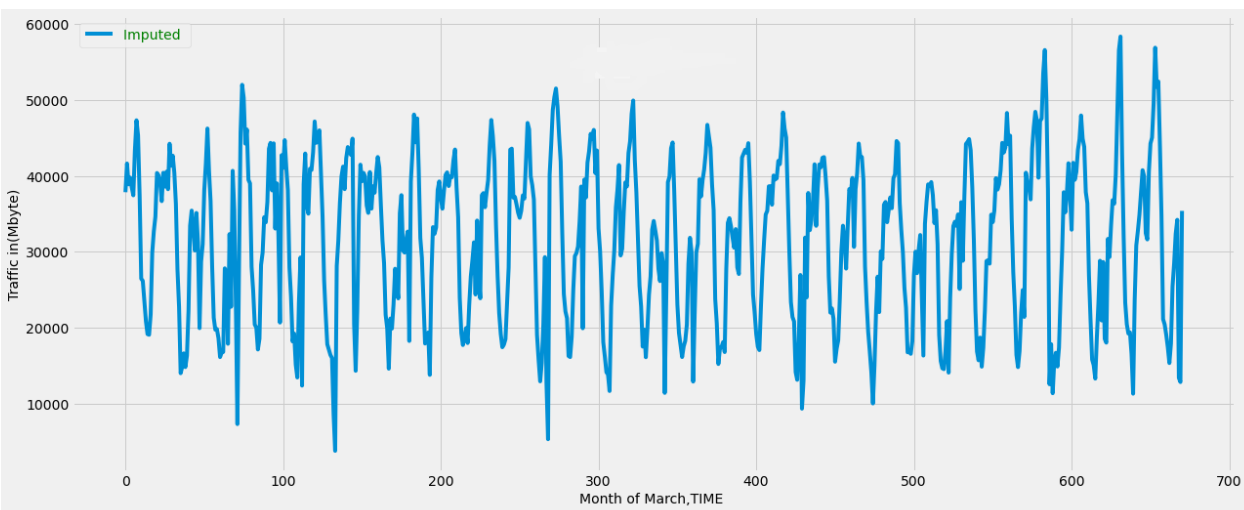

Figure 4.6 Dataset plot after imputed

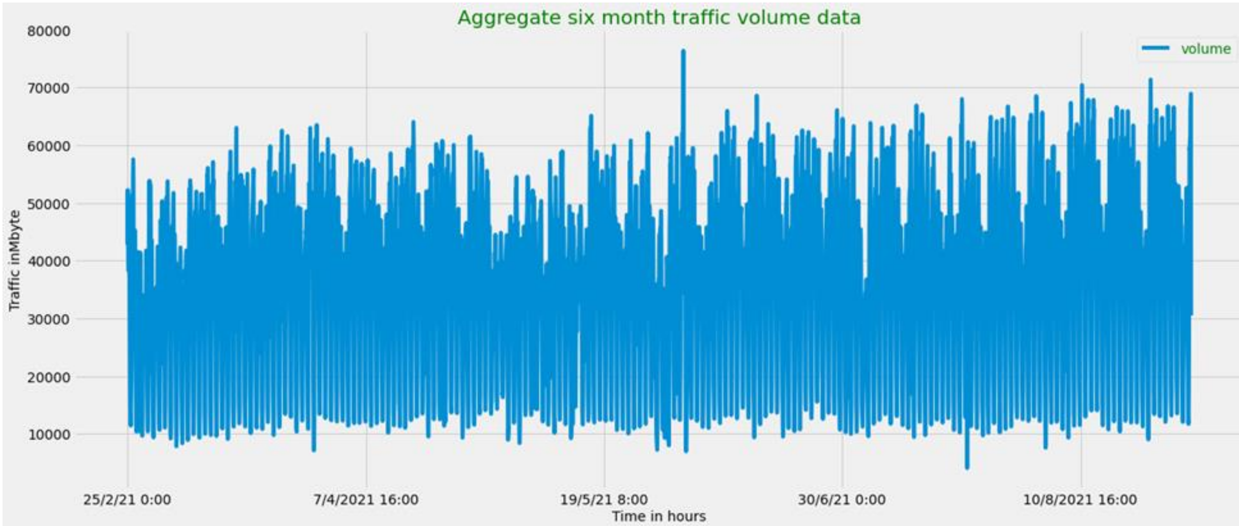

Figure 4.7 Monthly traffic pattern from March to August

4.4 Dataset Test - Stationary, trend, and seasonality

Figure 4.8 below presents the seasonal decomposition of fixed access data for the six months, divided as a trend, seasonal, and residual. The pattern in the seasonal category clearly shows that the data exhibits seasonal characteristics with a period of 24 or one day.

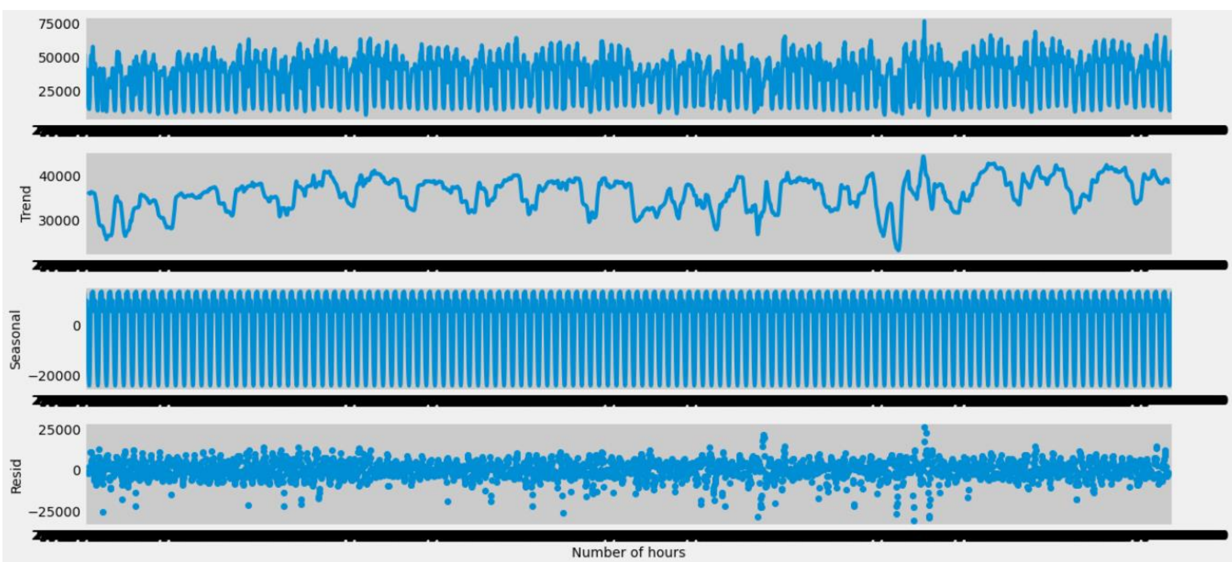

Figure 4.8 Seasonal decomposition of fixed access traffic

Stationarity

As discussed in chapter 3, time series data exhibiting trends and seasonality are not stationary. As shown in the seasonal-decomposition diagram trend part, the traffic trend has an indication of an increasing tendency. The seasonal part of the graph also has a pure seasonal pattern repeating every 24 hours. Thus, the traffic characteristics show a non-stationary property. In addition, the Kwiatkowski-Phillips–Schmidt-Shin (KPSS) test (trend stationary) was used to examine the behavior of the data against stationarity and linearity. A special feature of the KPSS test is checking for stationarity in the presence of a deterministic trend.

Table 4.1 KPSS stationary test result output

KPSS simulation result		
KPSS Statistics	0.618	-
P-value	0.021	0.05 Reference
Critical values: 10%	0.347	-
Critical values: 5%	0.463	-
Critical values: 1%	0.739	-

The test result demonstrates that the p-value is significant, with p-value (0.02) < 0.05, and hence, rejects the null hypothesis and leads to the conclusion that the series is NOT stationary.

4.5 Forecast Performance Evaluation Matrix

There are several prediction error-measuring matrices used to evaluate the performance of a model. To rigorously evaluate the performance of time series forecasting algorithms (LSTM and SARIMA), Mean Squared Error (MSE), Mean Average Error (MAE), Root Mean Squared Error (RMSE), Normalized Root Mean Squared Error (NRMSE), R-Squared, and Mean Absolute Percentage Error (MAPE) are proposed. These performance evaluation matrices compare the model's predicted values with the real test data values and the mathematical formula are presented in equation 4.1 up to 4.6 [27].

$$\text{Root mean square error RMSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(y_j - y_j'\right)^2} \tag{4.1}$$

$$\text{Normalized root mean square error } NRMSE = \frac{RMSE}{y^*} \tag{4.2}$$

$$\text{Mean absolute error MAE} = \frac{1}{n}\sum_{j=i}^{n}\left|y_j - y_j'\right| \tag{4.3}$$

$$\text{Mean square error } MSE = \frac{1}{n}\sum_{j=i}^{n}\left(y_j - y_j^*\right)^2 \tag{4.4}$$

$$\text{R-squared } R^2 = 1 - \frac{\sum_{j=1}^{n}\left(y_j - y_j'\right)^2}{\sum_{j=1}^{n}\left(y_j - y_j^*\right)^2} \tag{4.5}$$

$$\text{Mean absolute percentage error MAPE} = \frac{100}{n}\sum_{j=1}^{n}\frac{\left|y_j - y_j'\right|}{y_j'} \tag{4.6}$$

Where, y_j' is predicted value, y_j is real value and y_j^* is mean observed value

4.6 System Model

In this thesis, two forecasting algorithms, LSTM and SARIMA, are used. The collected 6-month aggregate dataset was divided into two parts, the training and test dataset, with an 80:20% ratio respectively. In the training set, 80% of the data (3572 observations) were fed to both predictor models. The remaining 20% of the data (893 observations) is the test data set used to evaluate the performance of the model, actual value against the predicted value. Finally, the models are evaluated based on different performance metrics such as MSE, RMSE, MAE, and R-squared. The overall objective of the forecasting principle is to minimize the loss function, which is the error metric. Therefore, parameter tuning is based on minimum matrices. The following flow chart explains the proposed system model.

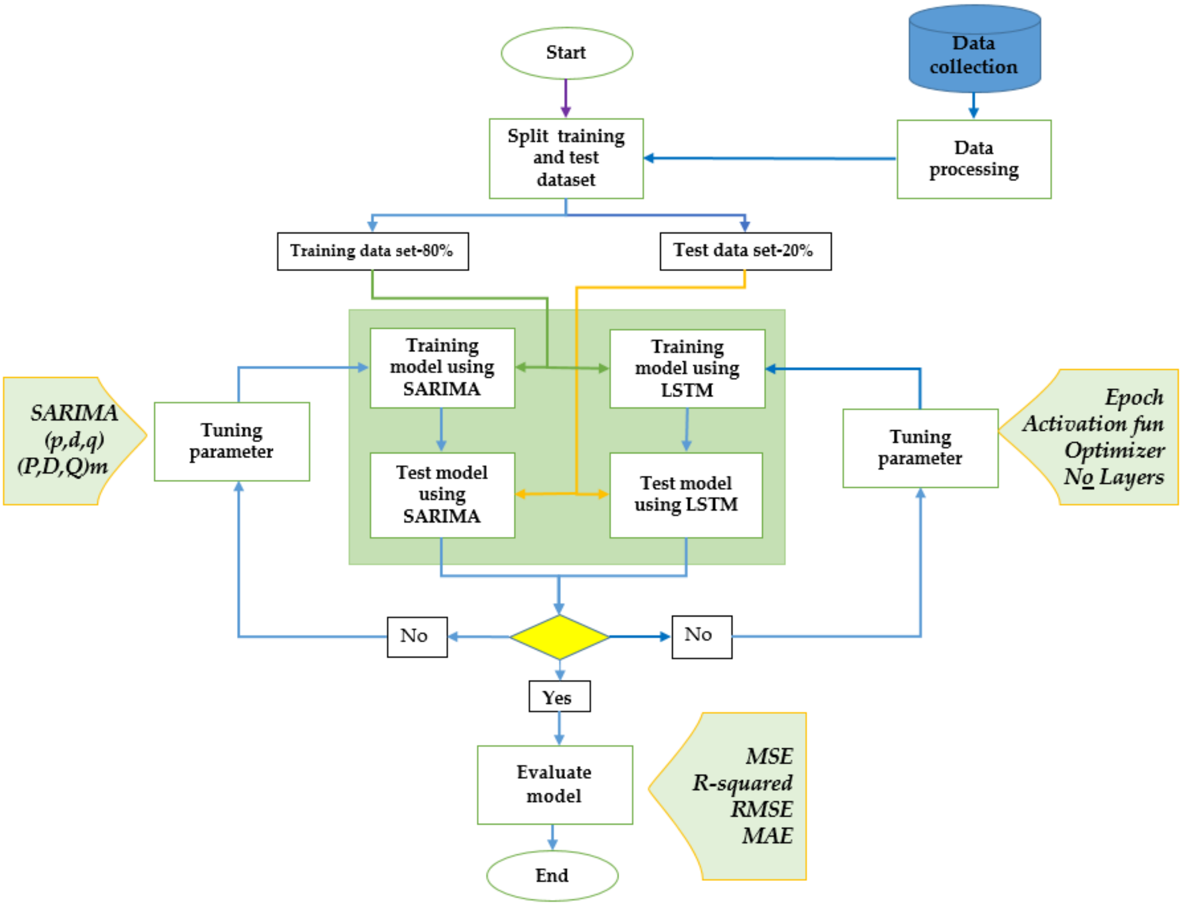

Figure 4.9 System-model flaw chart

4.7 Model Identification and Parameter Tuning

4.7.1 SARIMA Model Fit and Parameter Selection

As illustrated in section 4.4, using component identification and the KPSS test, the data exhibits non-stationarity and seasonality. In order to include these non-linear components of the traffic into the proposed architecture, the seasonal ARIMA model or SARIMA (p,d,q)(P,D,Q)m selected for modeling and forecasting the fixed access network data. The parameter initial estimation can be performed with an Autocorrelation Function (ACF) and a Partial Autocorrelation Function (PACF) analysis. From the plots presented in figure 4.10, the ACF plot shows a cluster of positive spikes at lags 1, 24, and 48 indicating a seasonality of m=24. The PACF tapers in multiples of 24; that is, the PACF has significant and dying out lags at 24, 48, and so on.

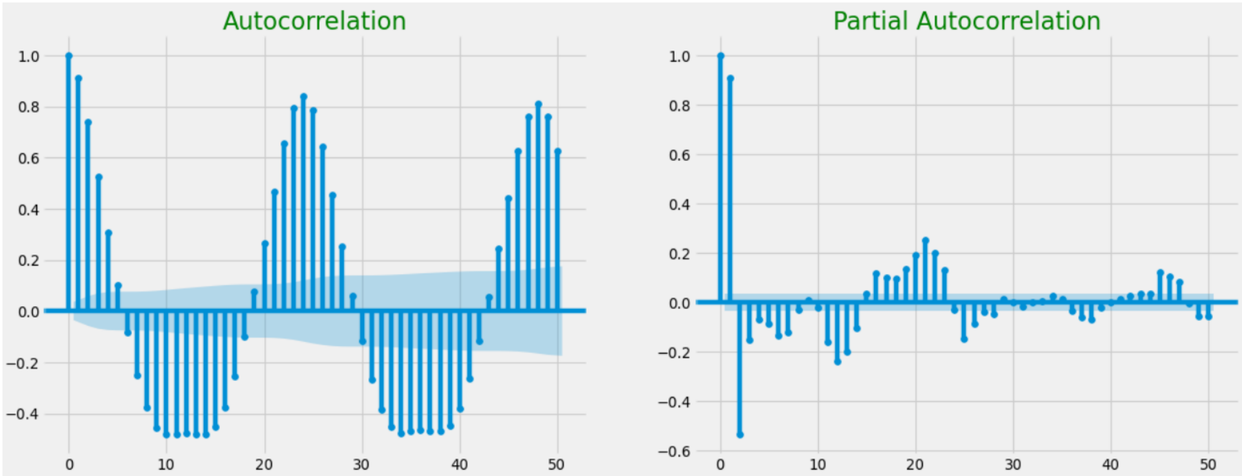

Figure 4.10 Autocorrelation ACF and partial autocorrelation PACF function plots

In order to find the best model parameter, the Akaike Information Criterion (AIC) value is used as a comparison matrix. The AIC measures how well a model fits the data while taking into account the overall complexity of the model. The grid search method is used to iteratively explore different combinations of SARIMA's (p,d,q)(P,D,Q)m parameters against AIC value, simulated using python code SARIMAX() library function from the statsmodels module. Based on the minimum values of AIC, SARIMA (0,1,1) (1,1,1)-24 was selected as the best fit model for the fixed access data traffic (see Annex-1).

Diagnosis of model fit

To investigate the model fit further, a model diagnostics test is used to ensure that the residual resembles white noise and none of the assumptions made by the model have been violated. As depicted in the model diagnostic plot figure 4.11 (up left side), the residual is uncorrelated and normally distributed. In addition, the Q-Q plot on the bottom left of figure 4.11, shows that the ordered distribution of residuals (the blue line) follows the linear trend of the samples taken from a standard normal distribution with N (0, 1). Uncorrelated and normal distributed residual is an indication of a model perfectly fitted with the training data.

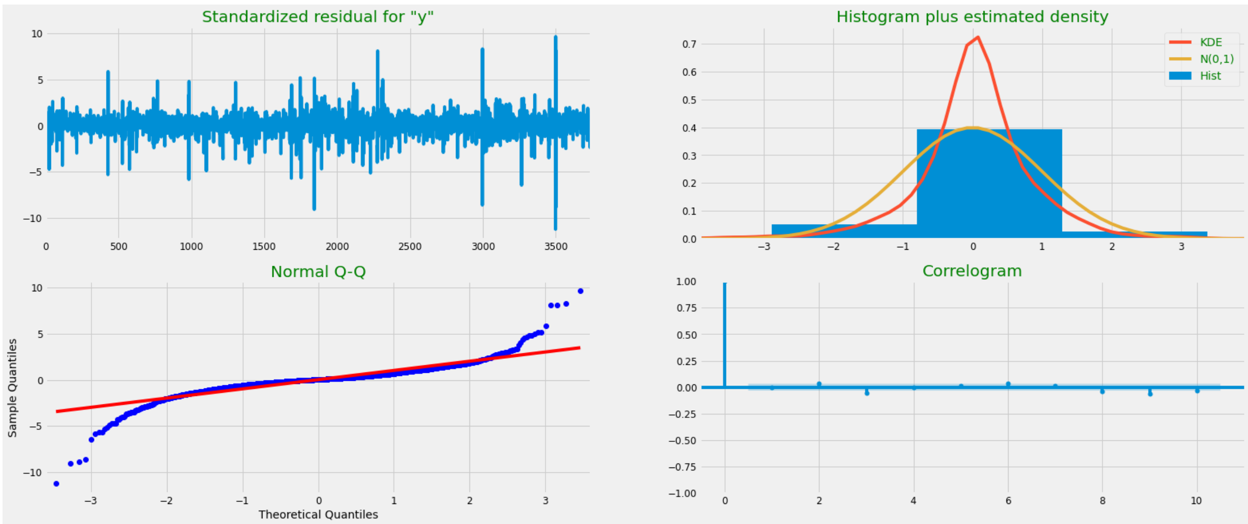

Figure 4.11 Model diagnostic plot

4.7.2 LSTM Model Fit and Parameter Selection

The aggregated and normalized data set was divided into two parts; the training and test data set, with an 80:20% ratio, and passed to the sequential mode of the LSTM model algorithm. Python-based Anaconda library functions such as 'Keras' and 'Tensorflow' were used to compile the model. Selecting the Hyperparameters of LSTM is a tradeoff between model complexity and accuracy [8]. Increasing the number of epochs and number of layers increases the complexity of the model by upturning the number of trainable parameters and computational time.

As shown in figure 4.12, initially, the LSTM model loss (Mean squared error) reduced exponentially with the number of epochs, but the loss function graph flattened with the higher number of epochs. The other LSTM Hyperparameters are data type dependent and the overall parameter tuning process is performed manually until optimal values are reached with a simple to complex approach. The 'Adam' optimization algorithm and 'Relu' as an initial activation function were used for the LSTM model. As shown in Table 4.2, the number of layers is set to 2, the batch sizes to 1 and 32, and the learning rate is set to 0.001.

Table 4.3 shows the model summary of LSTM sequential mode. For a hidden state of 300 and a number of epochs of 1000, the number of trainable parameters reached 390,301. The number of parameters shows the complexity and capacity of the model, and it has information such as input size, output size, and number of neurons in the hidden layer.

Table 4.2 LSTM Training Hyperparameters

Hyper parameter	Remark
Initial learning rate	0.001
Number of Epochs	100,300,1000
LSTM hidden layers	2
LSTM hidden states	300
Optimization algorithm	Adam
Activation function	Relu
Loss function	MSE
Batch size	1,32

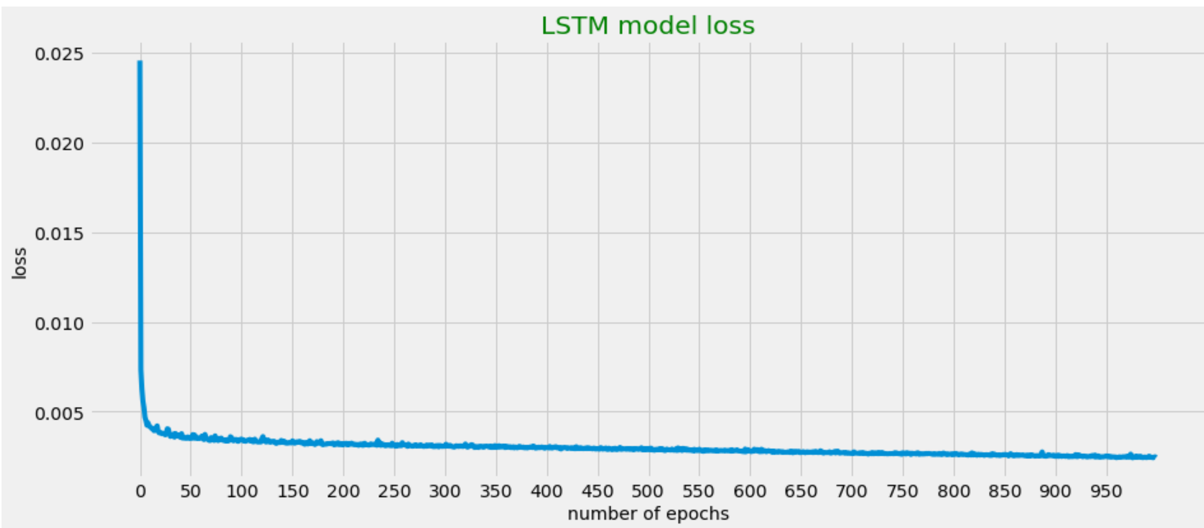

Figure 4.12 LSTM model loss against 1000 epochs

Table 4.3 LSTM Sequential model summary

LSTM Model summary: "Sequential"		
Layer (type)	Output shape	Parameter Number
lstm_6 (LSTM)	(None, 300)	390,000
dense_6 (Dense)	(None, 1)	301
activation_6 (Activation)	(None, 1)	0
Total parameters : 390,301 Trainable parameters : 390,301 Non-trainable parameters : 0		

Chapter Five

Results and Discussion

This section discusses the numerical results from both prediction models and compares the performance results based on the error matrix defined in section 4.6.

5.1 SARIMA Result Analysis

To evaluate the model's performance with different data sizes, the aggregated six-month fixed network data is divided into three parts. The first dataset includes the month of March and April traffic data. The second dataset contains four months of data from March up to June, and the third dataset contains all six months of the dataset. The fixed access data used has non-stationary, seasonality, increasing trend, and nonlinear properties as per the test in section 4.4. Therefore, the proposed modeling algorithm requires the incorporation of such an effect into the forecasting model. As discussed in section 4.7.1, the best model (with the lowest AIC values) is selected and then the forecasts are produced in the Python environment using the SARIMAX() library function from the statsmodels module. The result shows that the data fitted well in SARIMA (0,1,1)(1,1,1)24. Table 4.4 shows the results of the forecasting error matrix for three datasets that were defined earlier and for SARIMA parameters with AIC values nearer to the selected model.

Figure 5.1 shows the comparison of forecasting results for the two datasets, four and six months of data. When using six months of data, the SARIMA model produces relatively higher errors in terms of MAE and RMSE. Whereas in the percentage error matrix (R-squared, MAPE, and MDAPE), the model produces nearly the same results. The performance reduction when using a higher dataset is a data-dependent property and may not apply in a general sense. Figure 5.2 shows a plot of SARIMA (0,1,1)(1,1,1) 24 forecast plotted against the real data volume with an 90% confidence interval. By visually investigating the graph, the predicted value precisely follows the oscillating behavior of the fixed traffic pattern.

Table 4.4 SARIMA model forecast error performance matrix results

Dataset	SARIMA (p,d,q)(P,D,Q) m	R-Squared -%	MAPE -%	MDAPE -%	MAE	RMSE	N RMSE
Two Month	SARIMA(1, 0, 1)(0, 1, 1)24 AIC:30772	74.48	11.17	6.82	3220	4707	0.129
	SARIMA(1, 0, 1)(1, 1, 1)24 AIC:30774	74.55	11.14	6.58	3213	4699	0.128
	SARIMA(1, 1, 1)(0, 1, 1)24 AIC:30783	76	10.81	7.08	3115	4576	0.125
	SARIMA(1, 1, 1)(1, 1, 1)24 AIC:30785	76.02	10.81	7.11	3113	4574	0.125
Four Month	SARIMA(0, 1, 0)(0, 1, 1)24 AIC:58248	95.22	6.31	4.22	2247	3508	0.097
	SARIMA(0, 1, 0)(1, 1, 1)24 AIC:58247	95.24	6.3	4.22	2242	3500	0.097
	SARIMA(0, 1, 1)(1, 1, 1)24 AIC:58184	95.2	6.32	4.25	2249	3510	0.098
	SARIMA(1, 0, 1)(0, 1, 1)24 AIC:58287	95.17	6.31	4.46	2251	3429	0.095
Six Month	SARIMA(0, 1, 0)(0, 1, 1)24 AIC:86755	95.42	6.13	3.39	2274	3678	0.092
	SARIMA(0, 1, 0)(1, 1, 1)24 AIC:86753	95.42	6.12	3.95	2271	3677	0.092
	SARIMA(0, 1, 1)(0, 1, 1)24 AIC:86651	95.3	6.22	3.99	2308	3720	0.094
	SARIMA(0, 1, 1)(1, 1, 1)24 AIC:86651	95.3	6.22	4.03	2305	3719	0.093

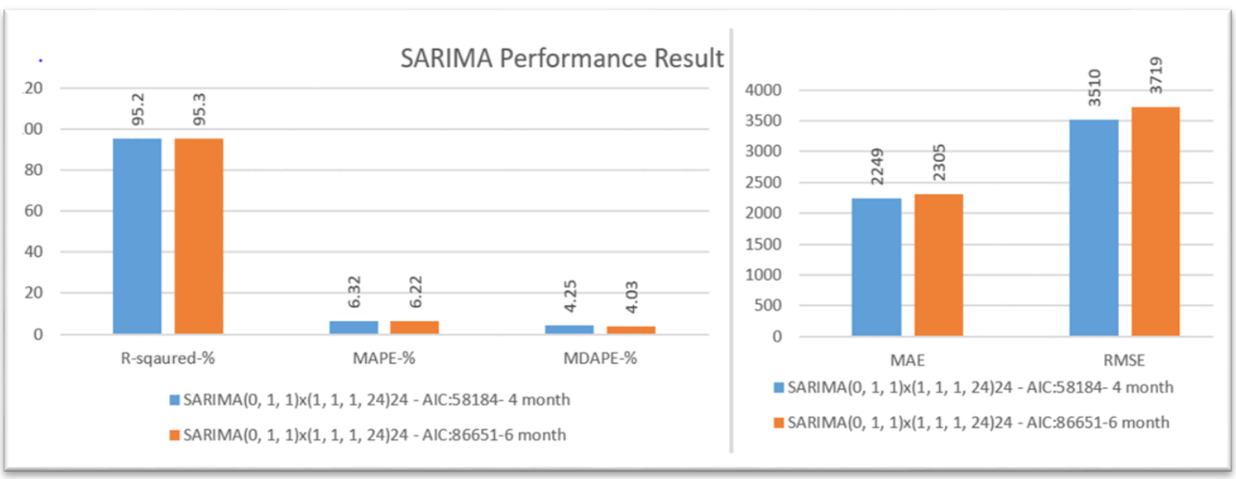

Figure 5.1 SARIMA (0,1,1)(1,1,1)24 forecast error results for six and four month dataset

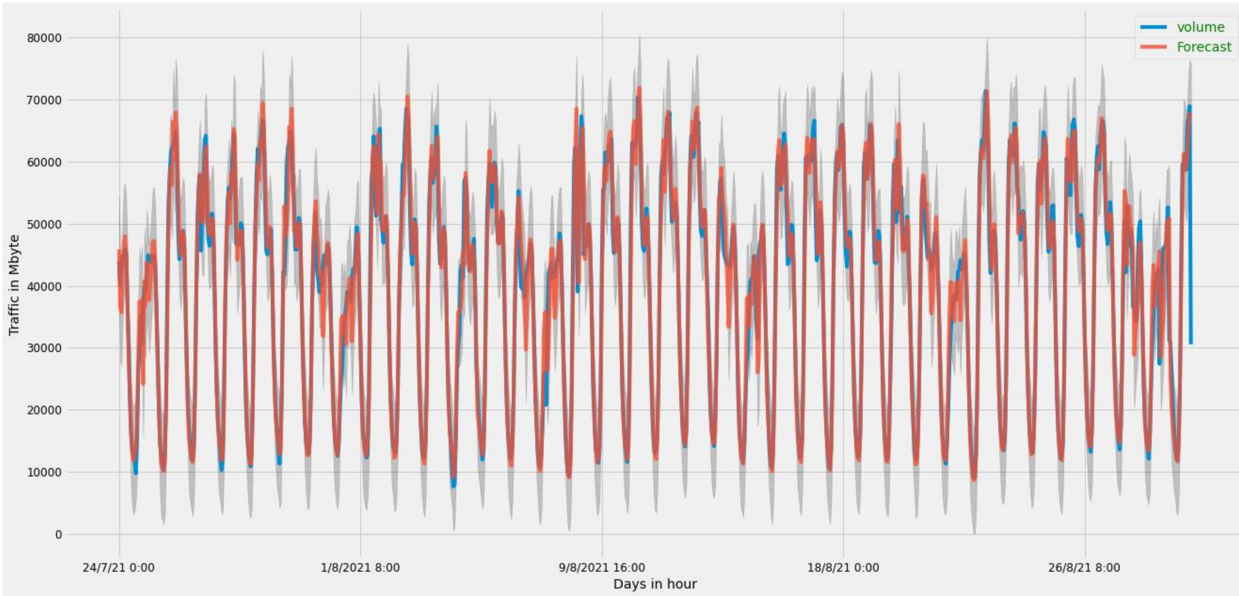

Figure 5.2 SARIMA (0,1,1)(1,1,1)24 forecast versus Observed traffic volume plot

5.2 LSTM Result Analysis

To evaluate the performance of the proposed architecture, the aggregated fixed access traffic data was used, which was collected for six months. A similar dataset that was used for the SARIMA model, with 80% of the data for training and the rest of the data for testing the LSTM model, was used. Three datasets from the main aggregate data were prepared in the evaluation setup to identify the effects of the number of datasets on the performance of the forecasting model, which determines the quantity of information needed to be memorized and utilized by the network. The three datasets are two, four, and six months of the main aggregate data, as was done for the SARIMA modeling setup. The implementation of the fixed traffic prediction algorithm is done in Python, using Keras, Tensorflow (as backend), and other supportive library functions.

The chosen Hyperparameters are reported in Table 4.2. The number of hidden layers is fixed at 2 and the number of hidden states at 300. For my dataset, choosing a higher number of layers did not increase the performance of the prediction, and I focused on the relationship between the number of past observed values and the number of epochs against the precision of the prediction. Furthermore, the Adam optimizer and Relu activation function are used in the Sequential LSTM model compilation process. The experimental setup uses different batch sizes, number of Epochs, and dataset. The result of the forecast error matrix using the three datasets is shown in table 4.5 with an optimal 32 batch size and in table 4.6 with a single batch size.

Table 4.5 LSTM forecast error performance matrix results for batch size of 32

LSTM-batch size 32	Number of Epochs	R-squared-%	MAE	MAPE-%	MDAPE-%	RMSE	N RMSE
Two Month	100	75.99	3304	11.46	7.25	4780	1.307
	300	76.12	3252	11.68	7.21	4767	1.303
	1000	75.21	3348	11.69	6.93	4858	1.328
Four Month	100	96.45	2213	6.51	4.85	3007	0.084
	300	96.37	2306	7.05	5.4	3042	0.085
	1000	96.3	2261	6.21	5.23	3071	0.085
Six Month	100	96.16	2350	6.2	4.59	3350	0.084
	300	96.44	2200	5.7	4.3	3224	0.081
	1000	96.91	2053	5.92	4.17	3005	0.076

Table 4.6 LSTM forecast error performance matrix results for batch size of 1

LSTM-batch size 1	Number of Epochs	R-squared-%	MAE	MAPE-%	MDAPE-%	RMSE	N RMSE
Two Month	100	76.04	3309	11.85	7.29	4775	0.131
	300	74.2	3337	11.7	7.11	4955	0.135
	1000	72.1	3541	12.49	7.47	5162	0.141
Four Month	100	96.81	2068	5.72	4.38	2852	0.079
	300	96.34	2204	6.04	4.61	3054	0.085
	1000	94.25	2739	7.18	5.37	3831	0.106
Six Month	100	96.81	2139	6.61	4.5	3055	0.077
	300	97.06	1995	5.6	3.97	2931	0.074
	1000	95.72	2369	6.12	4.54	3539	0.089

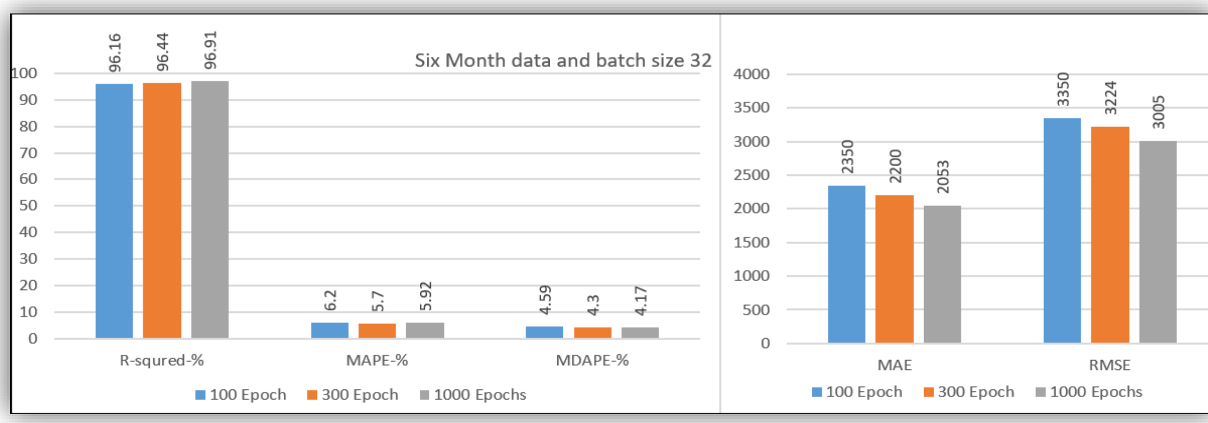

Figure 5.3 LSTM forecast error results for six-month data and Batch size 32

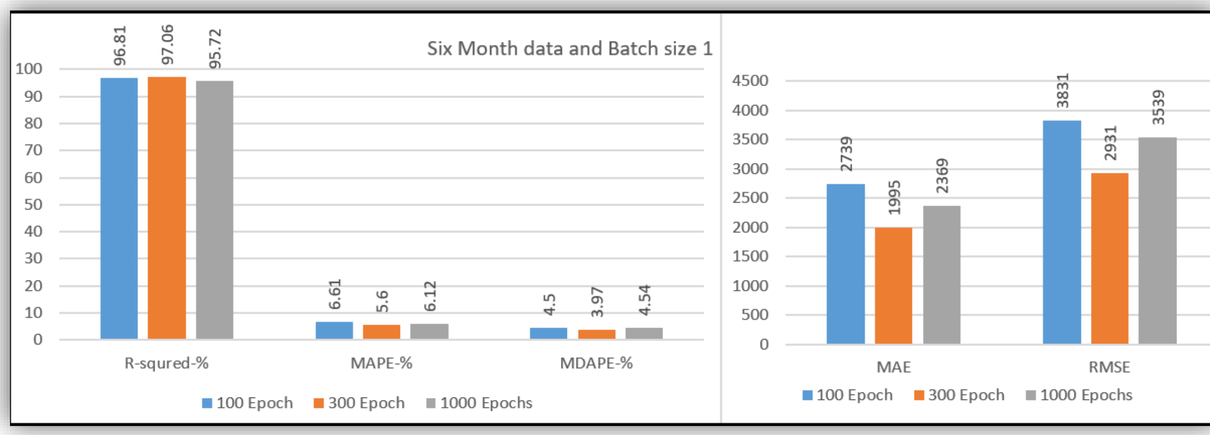

Figure 5.4 LSTM forecast error results for six-month data and Batch size 1

In general, as shown in tables 4.5 and 4.6, and as expected, the forecast error decreases with a larger number of observations for both batch sizes of the experimental setup. The model was trained using 100, 300, and 1000 epochs. It is possible that by augmenting the epochs and reducing the losses considerably, the time series forecast is becoming more accurate. However, regardless of what one intuitively thinks, the results show that this is not entirely certain. I ought to keep increasing the epochs until I reach the convergence point where increasing the epochs further results in increasing errors (i.e RMSE). Figure 5.3 and Figure 5.4 show the results of the LSTM forecast error for the six month dataset. For a batch size of one, increasing the number of epochs (i.e. from 100 to 300), as shown in figure 5.4, decreases the forecast error, but increasing the number of epochs further to 1000 reduces the performance. Therefore, 300 is an optimal number of epochs for the dataset with a single batch size. In contrast, the training error for the 32-batch size setup keeps decreasing up to 1000 epochs. It seems that, for a 32-batch size, the model converges around 1000 epochs. Therefore, the best results were found in the 32-batch size with 1000 epochs and the single batch size with 300 epochs combination. The latter combination results in a slightly smaller error. As shown in figure 5.5, the LSTM forecast for higher and lower peaks precisely follows the real fixed access data despite the oscillating behavior of the traffic pattern.

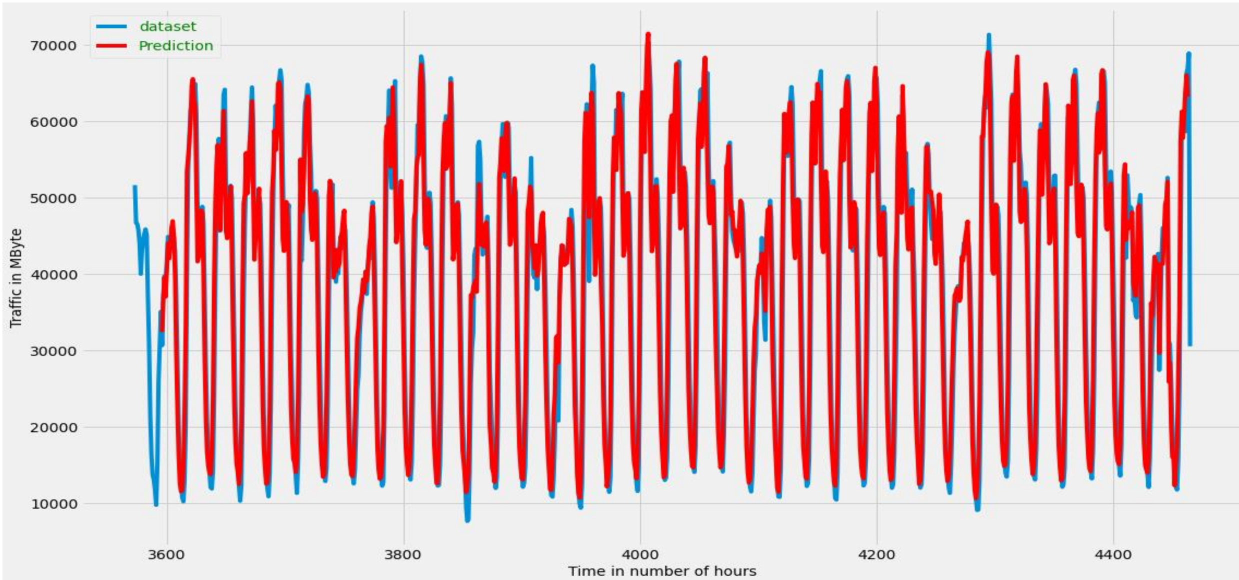

Figure 5.5 LSTM forecast versus observed plot

5.3 Model Comparison and Analysis

The objective of machine learning (LSTM) forecasting methods is the same as that of SARIMA. They both aim at improving forecasting accuracy by minimizing some loss function, typically the RMSE. Their difference lies in how such minimization is done with their respective forecasting algorithm utilizing non-linear time series data.

To make a fair comparison of both models' performance, identical datasets were used and the highest performing forecasting algorithm result was selected. Figure 5.6 shows that the LSTM-based forecasting method performed a maximum accuracy result of 97% in R-Squared value, 5.7% in MAPE, and 2931 in RMSE error matrix. Similarly, the forecast results of SARIMA reached 95.3% in R-squared, 6.3% in MAPE, and 3510 in RMSE. These results were found when using six months' data for both models; LSTM with 300 epochs and batch size 1; and SARIMA (0,1,1)(1,1,1)24. When considering RMSE as a comparator, the LSTM forecasting model was enhanced by 19.7%. The R-squared value of LSTM increased by 1.68%. In general, in all forecast error matrixes, LSTM performs better than SARIMA for this research dataset. Figure 5.7 shows the prediction from the LSTM and SARIMA models against 891 real traffic observation data, and as depicted in the graph, both models' predictions follow the oscillating real network traffic accurately.

Scholars indicate that ML methods such as LSTM are computationally more demanding than statistical ones and require greater dependence on computer science to be implemented. However, experimental results indicate that the simple SARIMA model performs at a satisfactory level when compared to the complex LSTM method.

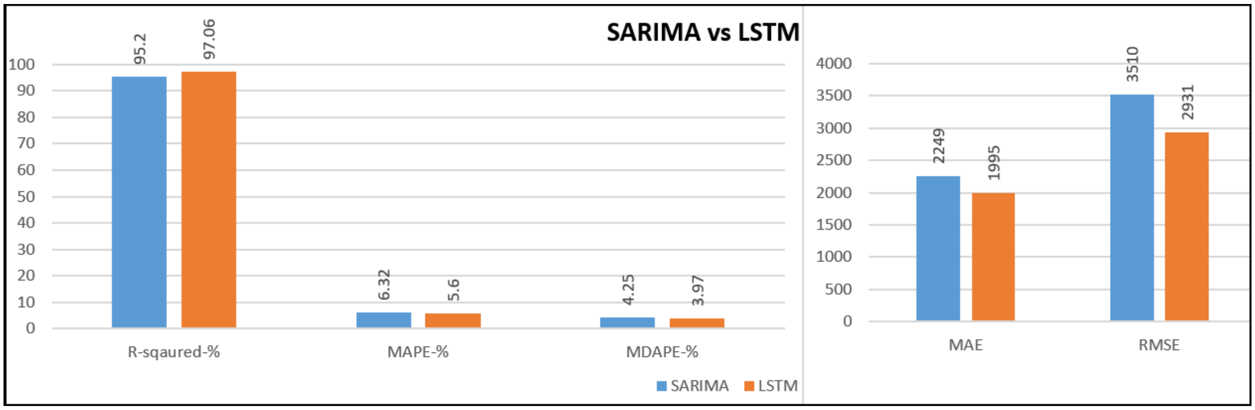

Figure 5.6 LSTM versus SARIMA forecast error result

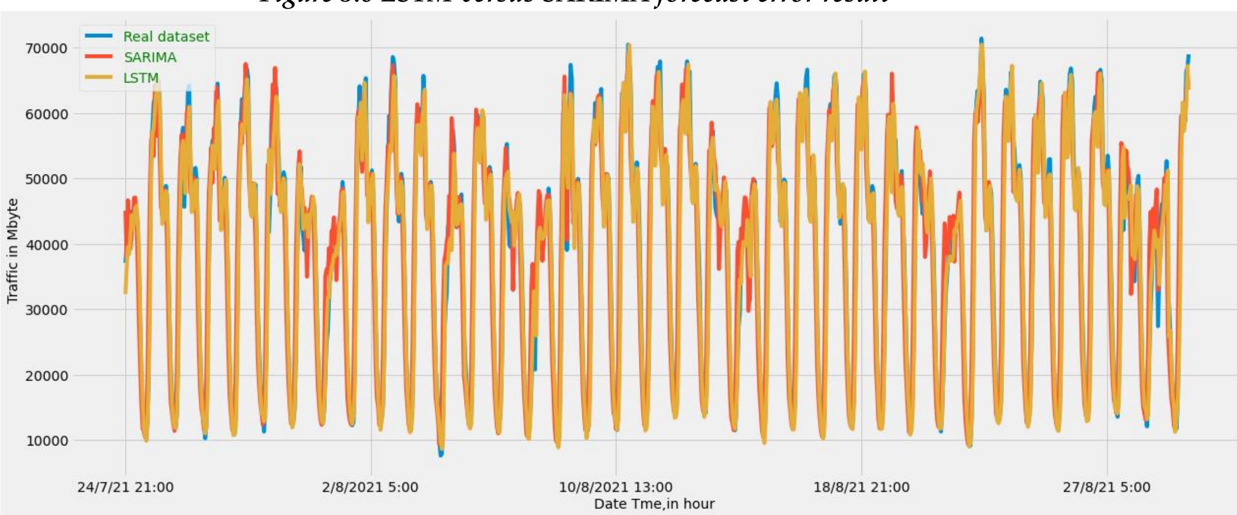

Figure 5.7 SARIMA, LSTM, and real data forecast plot

Chapter Six

Conclusion and Future Work

6.1 Conclusion

In this thesis, the recurrent neural network based Long-Short-Term Memory (LSTM) and the Seasonal-ARIMA model were used to conduct predictive analysis on fixed access network traffic data collected from the city of Addis Ababa. By exploiting the long term dependencies in the time series, the LSTM model forecasts complex univariate fixed network traffic time series with non-stationarity and seasonality. Seasonal ARIMA is also a famous statistical modeling algorithm suitable for nonlinear and seasonal time series data.

The experimental setup was conducted with six months of fixed access network data and by dividing it into three datasets (2, 4, and 6 month's data). Similar datasets were used for the proposed algorithms to train and test the models with an 80-20% ratio. Data processing and model parameter tuning were also the core parts of the research to find accurate forecasting results.

Based on the research results, both algorithms performed satisfactory forecasting results with MAPE ranging from 3–6% and RMSE ranging from 2931–3719 when using all six months of the time series dataset. Furthermore, the result showed that in a smaller number of datasets setup (such as two months), the forecasting performance reduced significantly in both forecasting algorithms, but increasing the number of datasets further (i.e. six months) led to better performance for the LSTM than SARIMA model.

In general, comparison among the forecasting methods shows that the LSTM produces the lowest errors when compared to the SARIMA model. Considering RMSE as a comparator matrix, an average of 19.75% improvement was observed in the LSTM forecasting method. Nevertheless, both forecasting algorithms performed satisfactory results with an R-squared value greater than 95% and MAPE less than 7%. Therefore, the use of LSTM and SARIMA algorithm applications in the design of fixed network traffic forecasting models is recommended.

6.2 Future work

The fixed access traffic modeling and forecasting were conducted using two famous and distinct algorithms (SARIMA and LSTM). Satisfactory results on both algorithms were found with the dataset and parameters used. However, as a recommendation, research needs to be conducted on fixed network traffic using other standalone or hybrid models for possibly better results. In addition, the temporal traffic analysis was conducted considering normal days of the year. However, during special days such as holidays, the traffic behaves differently, so detailed analysis on special days needs also to be conducted in the future.

References

[1] A. Jenefa and M. B. Moses, "Multi level statistical classification of network traffic," 2017 International Conference on Inventive Computing and Informatics (ICICI), 2017, pp. 564-569.

[2] "Ethio telecom 2020/2021 Annual business performance" www.ethiotelecom.et July 2021.

[3] N. Nalawadea and M. Pawar "Forecasting telecommunications data with autoregressive integrated moving average models" International Journal of Engineering Research and General Science, 2015.

[4] C. Katris and S. Daskalaki "Comparing forecasting approaches for Internet traffic" Department of Electrical & Computer Engineering, University of Patras, Greece, 2015.

[5] G. Tesfaye "Hybrid SARIMA-ELM-based Data Traffic Forecasting: The Case of UMTS Network in Addis Ababa, Ethiopia" Thesis, Masters of Science in Telecommunications Engineering, AAU, 2018.

[6] L. Vidmar, B. Peternel, M. Štular and A. Kos, "Broadband access network investment optimization in rural areas," Melecon 2010 - 2010 15th IEEE Mediterranean Electrotechnical Conference, 2010, pp. 482-486.

[7] S. Gul and J. Gutierrez "Evolution of Broadband Communication Networks: Architecture and Applications" Auckland University of Technology, Auckland, New Zealand, 2017.

[8] B. Tamene "Video Streaming Data Traffic Prediction by Using Long Short Term Memory (LSTM) Model: In the case of UMTS Network in Addis Ababa" Masters of Science in Telecommunications Engineering, AAU, 2020.

[9] H. Trinh, L. Giupponi and P. Dini "Mobile Traffic Prediction from Raw Data Using LSTM Networks" IEEE 29th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), Barcelona, Spain, 2018.

[10] G. Tesfaye "Hybrid SARIMA-ELM-based Data Traffic Forecasting: The Case of UMTS Network in Addis Ababa, Ethiopia" Thesis, Masters of Science in Telecommunications Engineering, AAU, 2018.

[11] M. Kihl, C. Lagerstedt, A. Aurelius and P. Ödling ''Traffic Analysis and Characterization of Internet User Behavior'', Dept. of Electrical and Information Technology, Lund University, Sweden, 2010.

[12] K. Papagiannaki, N. Taft, Zhi-Li Zhang and C. Diot, "Long-term forecasting of Internet backbone traffic," in IEEE Transactions on Neural Networks, vol. 16, no. 5, pp. 1110-1124, Sept. 2005.

[13] H. Tang and L. Ou, "A temporal analysis of holiday effect on IP backbone traffic," 2017 IEEE 9th International Conference on Communication Software and Networks (ICCSN), 2017, pp. 1042-1046.

[14] Z. Zelenika and K. Magdic, "Network Building Blocks of Fixed NGN Telecom Operator," 2007 9th International Conference on Telecommunications, 2007, pp. 317-324.

[15] "White Paper on Next-Generation Fixed Network Access Technologies" https://www.zte.com.cn, ZTE, 2020.

[16] D. Hindocha, A. Bagga, N. Atmande and D. Rojatkar "Next Generation Network: An Overview A Future Telecom Network" International Journal of Engineering Research & Technology, Electronics and Telecommunication dept. Government College of Engineering Chandrapur, India, 2015.

[17] M. Zhu." Research on key techniques in passive optical networks and optical grid applications" Doctorial thesis, ENS Cachan; Shanghai Jiao Tong University, 2012.

[18] M. Elaydi "Next Generation Passive Optical Network Stage Two NG-PON2" Thesis, Master in Electrical Engineering, The Islamic University – Gaza, 2014.

[19] "The Future of Passive Optical Networking is Here" NG-PON2 Marketing report, https://www.broadband-forum.org.

[20] D. Nowak and J. Murphy "FTTH: The over view of existing Technologies " Dublin City University and University College Dublin, Dublin, Proceedings Volume 5825, Opto-Ireland, 2005.

[21] K. Werbach "The Development of Fixed Broadband Networks" Directorate for Science, Technology And Industry, Committee on Digital Economy Policy, 2015.

[22] DSL Forum, [online] Available: www.dslforum.org

[23] A. Alamirew "An Assessment of Fixed Broadband Internet Customer Satisfaction in Ethio Telecom: The Case of Addis Ababa" Master's thesis, St. Mary's University, Addis Ababa, Ethiopia, 2013.

[24] S. Tadesse "Statistical Modeling of Internet Traffic Flow Length and Flow Size" Master Thesis, Addis Ababa University, May 2019.

[25] M. Daka "Traffic Analysis of IP Core Network: The Case of Ethiotelecom" Master Thesis, Addis Ababa University, May 2017.

[26] M. Heikkinen and W. Berger "Comparison of User Traffic Characteristics on Mobile Access versus Fixed Access network" Computer Science and Artificial Intelligence Laboratory Technical Report, MIT-CSAIL-TR-2011-028, May, 2011.

[27] P. Brockwell and R. Davis, "Introduction to Time Series and Forecasting," springer texts in statistics.

[28] P. Cortez, M. Rio, M. Rocha and P. Sousa "Internet Traffic Forecasting using Neural Networks" 2006 International Joint Conference on Neural Networks, Vancouver, BC, Canada, 2006.

[29] P. Velan, J. Medková, T. Jirsík and P. Čeleda, "Network traffic characterisation using flow-based statistics," NOMS 2016 - 2016 IEEE/IFIP Network Operations and Management Symposium, 2016, pp. 907-912.

[30] S. Makridakis, E. Spiliotis and V. Assimakopoulos "Statistical and Machine Learning forecasting methods: Concerns and ways forward" PLoS ONE 13(3): e0194889. https://doi.org/10.1371/journal.pone.0194889, 2018.

[31] J. Yang, Y. He, P. Lin and G. Cheng, "Characterizing Internet backbone traffic from macro to micro," 2009 IEEE International Conference on Network Infrastructure and Digital Content, 2009, pp. 139-143.

[32] "ITU-T Recommendation G.992.1: Asymmetric digital subscriber line (ADSL) transceivers", http://handle.itu.int/11.1002/1000/4718.

[33] "ITU-T Recommendations G.993.2 (02/2019) Very high speed digital subscriber line transceivers 2 (VDSL2)" http://handle.itu.int/11.1002/1000/13747.

[34] "ITU-T Recommendation G.993.5 (02/2019) Self-FEXT cancellation (vectoring) for use with VDSL2 transceivers" http://handle.itu.int/11.1002/1000/13748.

[35] "ITU-T Recommendation G.986 (01/2010), 1 Gbit/s point-to-point Ethernet-based optical access system" http://handle.itu.int/11.1002/1000/10408.

[36] "ITU-T Recommendation G.988 (11/2017), ONU management and control interface (OMCI) specification" http://handle.itu.int/11.1002/1000/13291.

[37] "ITU-T Recommendation G.997.1 (02/2019), Physical layer management for digital subscriber line transceivers" http://handle.itu.int/11.1002/1000/13751.

[38] "ITU-T Recommendation G.998.1 (01/2005), ATM-based multi-pair bonding" http://handle.itu.int/11.1002/1000/3395.

[39] "ITU-T Recommendation G.998.2 (11/2018), Ethernet-based multi-pair bonding" http://handle.itu.int/11.1002/1000/13753.

[40] "ITU-T Recommendation G.998.3 (01/2005), Multi-pair bonding using time-division inverse multiplexing, http://handle.itu.int/11.1002/1000/7493.

[41] "ITU-T Recommendation G.9701 (03/2019), Fast access to subscriber terminals (G.fast) – Physical layer specification" http://handle.itu.int/11.1002/1000/13772.

[42] "802.3ah-2004 - IEEE Standard for Information technology-- Local and metropolitan area networks-- CSMA/CD Access Method and Physical Layer Specifications Amendment: Media Access Control Parameters, Physical Layers, and Management Parameters for Subscriber Access Networks" https://ieeexplore.ieee.org/servlet/opacpunumber=9283.

[43] The High Level Design for Fixed Line Network, ethio telecom internal document, 2014.

[44] R. Adhikari and R. Agrawal "An Introductory Study on Time Series Modeling and Forecasting" arXivLabs, LAP Lambert Academic Publishing, Germany, 2013.

[45] A. Jain, J. Mao and K. Mohiuddin, "Artificial neural networks: a tutorial," in Computer, vol. 29, no. 3, pp. 31-44, March 1996.

[46] D. Brezak, T. Bacek, D. Majetic, J. Kasac and B. Novakovic, "A comparison of feed-forward and recurrent neural networks in time series forecasting," 2012 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFEr), 2012, pp. 1-6.

[47] S. Hochreiter and J. Schmidhuber "Long Short-Term Memory" Neural Computation 8(9), 1997, pp. 1735-1780.

[48] O. Abiodun, A. Jantan, A. Esther, K. Victoria, N. Mohamed and H. Arshad "State-of-the-art in artificial neural network applications: A survey" Heliyon 4 (2018) e00938, Published by Elsevier Ltd., 2018.

[49] J. Zheng, C. Xu, Z. Zhang and X. Li "Electric Load Forecasting in Smart Grids Using Long-Short-Term-Memory based Recurrent Neural Network" The 51st Annual Conference on Information Systems and Sciences Proceedings, USA, 2017.

[50] R. Fu, Z. Zhang and L. Li, "Using LSTM and GRU neural network methods for traffic flow prediction," 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC), 2016, pp. 324-328.

[51] M. Iqbal, M. Zahid, D. Habib and L. John "Efficient Prediction of Network Traffic for Real-Time Applications" Journal of Computer Networks and Communications, Article ID 4067135, 2019.

[52] P. Cortez, M. Rio, M. Rocha and P. Sousa "Internet Traffic Forecasting using Neural Networks" 2006 International Joint Conference on Neural Networks, Vancouver, BC, Canada, 2006.

[53] R. Hyndman and R. Athanasopoulos "Forecasting: principles and practice", 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2, 2018.

Annex

SARIMA candidate model parameter with last 22 smallest AIC values

SARIMA(p, d ,q)(P, D, Q)m	AIC value
SARIMA(0, 1, 0)x(0, 1, 1, 24)24 - AIC:	86755.55424
SARIMA(0, 1, 0)x(1, 0, 1, 24)24 - AIC:	87170.16373
SARIMA(0, 1, 0)x(1, 1, 0, 24)24 - AIC:	87934.78601
SARIMA(0, 1, 0)x(1, 1, 1, 24)24 - AIC:	86753.36583
SARIMA(0, 1, 1)x(0, 1, 1, 24)24 - AIC:	86651.68007
SARIMA(0, 1, 1)x(1, 0, 1, 24)24 - AIC:	87869.08517
SARIMA(0, 1, 1)x(1, 1, 1, 24)24 - AIC:	86651.77044
SARIMA(1, 0, 0)x(0, 1, 1, 24)24 - AIC:	86802.4838
SARIMA(1, 0, 0)x(1, 1, 0, 24)24 - AIC:	87542.65294
SARIMA(1, 0, 0)x(1, 1, 1, 24)24 - AIC:	86804.18291
SARIMA(1, 0, 1)x(0, 1, 1, 24)24 - AIC:	86778.48156
SARIMA(1, 0, 1)x(1, 0, 1, 24)24 - AIC:	87746.01866
SARIMA(1, 0, 1)x(1, 1, 0, 24)24 - AIC:	87542.16929
SARIMA(1, 0, 1)x(1, 1, 1, 24)24 - AIC:	86780.19581
SARIMA(1, 1, 0)x(0, 1, 1, 24)24 - AIC:	87184.08932
SARIMA(1, 1, 0)x(1, 0, 1, 24)24 - AIC:	87765.71074
SARIMA(1, 1, 0)x(1, 1, 0, 24)24 - AIC:	88002.47358
SARIMA(1, 1, 0)x(1, 1, 1, 24)24 - AIC:	87185.35974
SARIMA(1, 1, 1)x(0, 1, 1, 24)24 - AIC:	86865.57309
SARIMA(1, 1, 1)x(1, 0, 1, 24)24 - AIC:	87512.86213
SARIMA(1, 1, 1)x(1, 1, 0, 24)24 - AIC:	87627.92415
SARIMA(1, 1, 1)x(1, 1, 1, 24)24 - AIC:	86866.43892