SHORT-TERM MOBILE DATA TRAFFIC FORECASTING: A CASE STUDY OF KADUNA STATE, NIGERIA

PAUL, JOAN EZRA

(19PCK01988)

MAY, 2022

SHORT-TERM MOBILE DATA TRAFFIC FORECASTING: A CASE STUDY OF KADUNA STATE, NIGERIA

BY

PAUL, JOAN EZRA

(19PCK01988)

B.Eng Electrical And Electronics Engineering, Landmark University Omu-Aran

A DISSERTATION SUBMITTED TO THE SCHOOL OF POSTGRADUATE STUDIES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF MASTER OF ENGINEERING (M.Eng) DEGREE IN INFORMATION AND COMMUNICATION ENGINEERING, IN THE DEPARTMENT OF ELECTRICAL AND INFORMATION ENGINEERING, COLLEGE OF ENGINEERING, COVENANT UNIVERSITY

MAY, 2022

ACCEPTANCE

This is to attest that this dissertation was accepted in partial fulfillment of the requirements for the awards of the degree of Masters of Engineering in Information and Communication Engineering, Department of Electrical and Information Engineering, College of Engineering, Covenant University Ota, Nigeria.

Mr. Taiwo B. Erewunmi
(Secretary, School of Postgraduate Studies**)** Signature and Date

Prof. Akan B. Williams
(Dean, School of Postgraduate Studies**)** Signature and Date

DECLARATION

I, **Paul, Joan Ezra (19PCK01988)**, declare that this research was carried out by me under the supervision of Dr. Adeyinka A. Adewale of the Department of Electrical and Information Engineering, College of Engineering, Covenant University, Ota, Nigeria. I attest that the dissertation has not been presented either wholly or partially for the award of any degree elsewhere. All sources of data and scholarly information used in this dissertation are duly acknowledged.

PAUL, JOAN EZRA

Signature and Date

CERTIFICATION

We certify that this dissertation titled "**A SHORT-TERM MOBILE DATA TRAFFIC FORECASTING: A CASE STUDY OF KADUNA STATE, NIGERIA**" is an original research work carried out by **PAUL, JOAN EZRA (19PCK01988)** in the department of Electrical and Information Engineering, College of Engineering, Covenant University, Ota, Ogun State, Nigeria under the supervision of Dr. Adeyinka A. Adewale. We have examined and found this work acceptable as part of the requirements for the award of Master of Information and Communication Engineering.

Dr. Adeyinka A. Adewale
(Supervisor) Signature and Date

Prof. Emmanuel Adetiba
(Head of Department) Signature and Date

Prof. Adeseko Ayeni
(External Examiner) Signature and Date

Prof. Akan B. Williams
(Dean, School of Postgraduate Studies) Signature and Date

DEDICATION

I dedicate this work to God Almighty and the Ezra's family

ACKNOWLEDGEMENT

I want to thank God for the strength and grace to complete this dissertation; to Him alone be all the glory. I thank the Chancellor for yielding to God's calling and creating a platform for individuals like myself to learn, relearn, and unlearn. I appreciate my parent, Mr and Mrs Paul Ezra for their timeless support, financially and emotionally, and their prayers although my master's program. I would also like to appreciate my siblings for their words of encouragement anytime we communicate. I thank my supervisor, Dr Adewale Adeyinka A. for his advice, encouragement, and support and prayers throughout my project. I appreciate Prof. Francis Idachaba for his timeless support and mentoring during my project. I also appreciate Prof Emmanuel Adetiba and Prof Sanjay Misra for using his laboratory as my workspace during my project year. To Mr. Mattew Ibiwoye, thank you for helping me with the major resources I needed for this work to be a success. To Mr. Felix Ajibulu, thank you for your timeless support. To my esteem colleagues, I say a big thank you; you people made my master's years fun despite the work, thank you all for supporting me. You people are the best.

TABLE OF CONTENTS

Content	Pages
COVER PAGE	**i**
TITLE PAGE	**ii**
ACCEPTANCE	**iii**
DECLARATION	**iv**
CERTIFICATION	**v**
DEDICATION	**vi**
ACKNOWLEDGEMENT	**vii**
LIST OF FIGURES	**xi**
LIST OF TABLES	**xiii**
ABBREVIATIONS	**xiv**
ABSTRACT	**xviii**
CHAPTER ONE: INTRODUCTION	**1**
1.1 Background of Study	1
1.2 Statement of the Problem	2
1.3 Aim and Objectives	3
1.4 Justification for the Research	3
1.5 Scope of Study	4
1.6 Limitation of the Research	4
1.7 Dissertation Organization	4
CHAPTER TWO: LITERATURE REVIEW	**6**
2.1 Introduction	6
2.2 Global Estimation	6
2.3 Cellular Network Architecture	7
2.3.1 Mobile Station (MS)	7
2.3.2 Base Station Subsystem (BSS)	7
2.3.3 Network Switching Subsystem (NSS)	8
2.3.4 Operation and Support System (OSS)	9
2.4 Evolution of Cellular Network	9
2.4.1 1^{ST} Generation (1G)	10
2.4.2 2^{nd} Generation (2G)	10
2.4.3 3^{rd} Generation (3G)	11
2.4.4 4^{th} Generation (4G)	11
2.4.5 5^{th} Generation (5G)	12
2.5 Telecommunication Links	13
2.6 Multiple Access Technique	16
2.6.1 Frequency Division Multiple Access (FDMA)	17
2.6.2 Time Division Multiple Access (TDMA)	17

2.6.3 Code Division Multiple Access (CDMA) 19
2.6.4 Space Division Multiple Access (SDMA) 19
2.6.5 Orthogonal Division Multiple Access (OFDMA) 19
2.6.6 Beam Division Multiple Access (BDMA) 20
2.7 Mobile Traffic 20
2.7.1 Mobile Voice Traffic 21
2.7.2 Mobile Video Traffic 21
2.7.3 Mobile Web Traffic 22
2.8 Types of Traffic Forecasting 23
2.8.1 Short-Term Forecasting 23
2.8.2 Medium-Term Forecasting 23
2.8.3 Long-Term Forecasting 23
2.9 Time Series 23
2.10 Models for Data Forecasting 23
2.10.1 Traditional Method 24
2.10.2 Neural Networks 28
2.10.3 Support Vector Machine (SVM) 35
2.11 Boosting Ensemble Learning 37
2.11.1 Adaboost 37
2.11.2 Gradient Boosting (GB) 37
2.11.3 Extreme Gradient Boosting (XGBoost) 40
2.11.4 LightGBM 41
2.11.5 Category Boosting (CATBoost) 41
2.11.6 Accelerated Gradient Boost (AGB) 42
2.12 Review of Related Works 42
2.12.1 Traditional Method 42
2.12.1.1ARIMA Model 42
2.12.1.2 Holt-Winter’s Exponential Smoothing 43
2.12.2 Neural Network 44
2.12.2.1 MLP, MLPWD, AND SVM 45
2.12.2.2 Long-Short Term Memory Unit (LSTM) 45
2.12.2.3 3D Convolutional Network 47
2.13 LSTM With Ensemble Learning 47
2.14 Chapter Summary 52
CHAPTER THREE: METHODOLOGY 53
3.1 Introduction 53

3.2 Dataset 53
3.3 Data Gathering 53
3.4 Data Cleaning and Preprocessing. 54
3.5 Algorithm used for the Training 55
3.5.1 The Auto-Regressive Integrated Moving Average (ARIMA) Algorithm 57
3.5.2 Long-Short Term Memory (LSTM) 60
3.5.3 Accelerated Gradient Boosting (AGB) 62
3.6 Performance Evaluation Criteria 64
3.6.1 The Mean Absolute Error (MAE) 64
3.6.2 Root Mean Squared Error (RMSE) 64
3.6.3 Mean Absolute Percentage Error (MAPE) 65
3.7 System Requirement 66
3.7.1 Hardware Requirement 66
3.7.2 Software Requirement 66
3.8 Chapter Summary 66

CHAPTER FOUR: RESULTS AND DISCUSSION 67

4.1 Introduction 67
4.2 Data Preprocessing 67
4.3 Result 67
4.3.1 Forecast for Kaduna South 68
4.3.2 Forecast for Kaduna North 77
4.4 Model Evaluation Result 85
4.5 Discussion 90
4.5.1 Forecasting of Mobile Traffic using ARIMA 91
4.5.2 Forecasting of Mobile Traffic using LSTM and LSTM-AGB 92
4.6 Chapter Summary 92

CHAPTER FIVE: CONCLUSION AND RECOMMENDATION 93

5.1 Summary 93
5.2 Contributions of the Research to Knowledge 93
5.3 Recommendation for Future Works 93

REFERENCES 95
APPENDIX A 105
APPENDIX B 106
APPENDIX C 107
APPENDIX D 108
APPENDIX E 113

LIST OF FIGURES

Figures	Title of Figures	Pages
2.1	Global Mobile Data (Cisco VNI, 2019)	6
2.2	Cellular Architecture (Vohra, Dubey, & Vachhhani, 2016)	9
2.3	Uplink and Downlink (Liu, Li, Luo, & Jiang, 2018)	16
2.4	TDMA Frame Structure (I. Wang, 2014)	18
2.5	ARIMA Model forecasting Process. (Xie et al., 2020)	26
2.6	Neural Network Structure. (Khashei & Bijari, 2010)	29
2.7	A repeating module in a standard RNN (Fu, Chu, Guo, & Chen, 2019)	32
2.8	A repeating module in an LSTM contains four interacting layers (Fu et al., 2019)	34
2.9	A structure of the LSTM unit (Yuan, Li, & Wang, 2020)	34
2.10	A Support Vector Machine (Sathyanarayana et al., 2014)	35
3.1	A block diagram of the method employed in this study	53
3.2	Raw dataset	54
3.3	Processed Data	55
3.4	A block diagram of the method employed in this study	57
3.5	Data Preprocessing and Model Pipeline	57
3.6	Flowchart for ARIMA Model	59
3.7	Schematic diagram of LSTM Architecture (Siami-namini & Tavakoli, 2018)	62
4.1	Processed Data	67
4.2	Total downlink data volume for HLBE001	69
4.3	Total Uplink data volume for HLBE001	69
4.4	Total downlink data volume for HLBE002	71
4.5	Total uplink data volume for HLBE002	71
4.6	Total Downlink data volume for HLBE016	72
4.7	Total Downlink data volume for HLBE016	73
4.8	Total downlink data volume for HLBE052	74
4.9	Total uplink data volume for HLBE052	74
4.10	Total downlink data volume for HLBE120	76
4. 11	Total uplink data volume for HLBE120	76
4.12	Total downlink data volume for HLBE019	77
4.13	Total uplink data volume for HLBE019	78
4.14	Total downlink data volume for HLBE029	79

4.15 Total uplink data volume for HLBE029 79
4.16 Total downlink data volume for HLBE 098 81
4.17 Total uplink data volume for HLBE 098 81
4.18 Total downlink data volume for HLBE117 82
4.19 Total uplink data volume for HLBE117 82
4.20 Total downlink data for HLBE146 84
4.21 Total uplink data for HLBE146 84
4.22 Performance evaluation of LSTM-AGM, LSTM, ARIMA in Total DL data volume for HLBE 019 85
4.23 Performance evaluation of LSTM-AGM, LSTM, ARIMA in Total UL data volume for HLBE 019 85
4.24 Performance evaluation of LSTM-AGM, LSTM, ARIMA in Total DL data volume for HLBE 146 86
4.25 Performance evaluation of LSTM-AGM, LSTM, ARIMA Total UL data volume for HLBE 146 86

LIST OF TABLES

Tables	Title of Tables	Pages
2.1	Evolution of Cellular network	13
2.2	Parameter that affects voice quality	21
2.3	Parameter that affects video quality	22
4.1	HLBE 001 Prediction result	87
4.2	HLBE 002 Prediction result	87
4.3	HLBE 016 Prediction result	87
4.4	HLBE 052 Prediction result	87
4. 5	HLBE 120 Prediction result	88
4.6	HLBE 019 Prediction result	88
4.7	HLBE 029 Prediction result	88
4.8	HLBE 098 Prediction result	88
4.9	HLBE 117 Prediction result	89
4.10	HLBE 146 Prediction result	89
4.11	Forecasted Valued for Kaduna South	89
4.12	Forecasted Value for Kaduna North	90

ABBREVIATIONS

M2M- Machine-to-Machine

VNI-Visual Networking Index

ANN-Artificial Neutral Network

LTE-Long Term Evolution

BS-Base Station

MS-Mobile Station

MLP-Multilayer Preceptor

MLPWD-Multilayer perceptron with weight delay

SVM-Support Vector Machine

SRM-Structure Risk Minimization

ARIMA-Autoregression Integrated Moving Average

FARIMA-Fractional Autoregression Integrated Moving Average

LSTM-Long-Short Term Memory

AGB-Accelerated Gradient Boost

GB- Gradient Boost

ConvoLSTM- Convolution Long-Short Term Memory

RNN-Recurrent Neural Network

EOM-Equipment out of Maintenance

EOS- Equipment out of Service

DL-Downlink

UL-Uplink

PS-Packet Switch

CS-Circuit Switch

AIC-Akaike Information Criteria

BIC-Bayesian Information Criteria

SGD-Stochastic Gradient Descent

HSUPA-High-Speed Uplink Packet Access

HSDPA- High-Speed Downlink Packet Access

AMPS-Advance Mobile Phone Service

GSM-Global System for Mobile Communication

FDM-Frequency Division multiplexing

FDMA-Frequency Division Multiple Access

TDMA-Time Division Multiple Access

CDMA-Code Division Multiple Access

SDMA-Space Division Multiple Access

OFDMA-Orthogonal Division Multiple Access

BDMA-Beam Division Multiple Access

MMS-Multimedia Message

GPRS-Global Packet Radio Access

EDGE-Enhance Data for Global Evolution

ITU-International Communication Union

IMT-International Mobile Communication

PSK-Phase-Shift Key

WCDMA-Wideband Code Division Multiple Access

MIMO-Multiple Input-Multiple output

FBMC-Filter Bank Multi-Carrier

CP-OFDM- Cyclic Prefix- Orthogonal Frequency Division Multiplexing

DFT-OFDM-Direct Fourier Transform Spread OFDM

NR-New Radio

GMSK-Gaussian Modulation Shift Keying

SINR-Signal Interference for Noise Ratio

E-UTRAN-Evolved-Universal Terrestrial Radio Access Network

MME-Mobility Management Entity

SGW-Serving Gateway

PGW (PDGW)-Packet Data Network Gateway

EPC-Evolved Packet Core

RAN- Radio Access Network

HSS-Home Subscriber Server

PCEF-Policy Control Enforcement Function

PCRF- Policy Control Regulation Function

EnB- Evolved NodeB

UE-User Equipment

KPI-Key Performance Indicator

QoS-Quality of Service

MAPE-Mean Absolute Percentage Error

RMSE-Root Mean Square Error

MAE-Mean Absolute Error

FDD-Frequency Division Duplex

TDD-Time Division Duplex

LoS-Line of Sight

DL-Downlink

UL-Uplink

TUL-Total Downlink

TDL-Total Uplink

NMT- Nordic Mobile Telephone

NOMA-Non- Orthogonal Division Multiplexing

ABSTRACT

Mobile networks are essential for today's seamless communication. As more individuals subscribe to mobile networks, the need for mobile networks has increased significantly. The network operators must devise strategies to handle the enormous demand of mobile network resources, such as spectrum, which are costly. There is a need for effective network resource management as well as a mechanism to predict future networks that can be used for network management and planning. This study uses real-life data to forecast mobile traffic using Kaduna State as a case study and compared prediction algorithms with the hybrid. The data set was gotten from a network provider. The hybrid combination of LSTM and AGB has been proposed in this study, and its performance has been compared with LSTM and ARIMA using MAE, RMSE and MAPE as evaluation metrics. The prediction performance of the algorithms was carried out on ten base stations with both the highest and lowest traffic from two local government areas, which are Kaduna South and Kaduna North. The LSTM-AGB outperformed LSTM and ARIMA. From the performance evaluation, the RMSE, MAPE and MAE of all the selected base stations in LSTM-AGB have a lower value than LSTM and ARIMA, which indicates a good fit of the model. It was observed that the hybrid algorithm performed better in base stations with high traffic.

Keywords: Mobile traffic, Long-Short Term Memory, Traffic forecasting, Cellular Network.