SPI-LSTM APPROACH FOR ENHANCING TRAFFIC FLOW PREDICTION

NURAINI BINTI SHAMSAIMON

MASTER OF SCIENCE
(COMPUTER SCIENCE)

UNIVERSITI PERTAHANAN NASIONAL MALAYSIA

2022

SPI-LSTM APPROACH FOR ENHANCING TRAFFIC FLOW PREDICTION

NURAINI BINTI SHAMSAIMON

Thesis submitted to the Centre for Graduate Studies, Universiti Pertahanan Nasional Malaysia, in fulfilment of the requirements for the Degree of Master of Science (Computer Science)

2022

ABSTRACT

The idea of traffic flow prediction for congestion management have been proposed in order to improve traffic management. Machine learning and deep learning algorithms have enabled this idea to grow, as it involves an excessive amount of traffic data and variables. The usage and analysis of these data or variables are essential but vary among studies hence producing different outputs and results. The different types of traffic data may affect the accuracy of congestion calculations. This study proposes a conceptual model known as the Speed Performance Index and Long Short-Term Memory (SPI-LSTM) Approach for Enhancing Traffic Flow Prediction in Smart Cities model based on speed variable analysis, which is common in most traffic datasets. This study has explored the potential of adapting analysed traffic data, by calculating the Speed Performance Index, into Long Short-Term memory machine learning and deep learning algorithms in order to perform traffic prediction. To test the hypothesis of whether traffic data influence prediction outcomes, an experiment was conducted using the Python programme to generate the expected and predicted outcomes as well as performing results analysis. The results were validated using evaluation metrics and then, compared with other existing models in order to analyse the performance of the proposed model. The validation and comparison results illustrated a positive performance when compared with existing models, hence, showing the potential of the proposed model to improve traffic prediction.

ABSTRAK

Idea ramalan aliran trafik telah dicadangkan untuk pengurusan kesesakan demi menambah baik pengurusan lalu lintas. Pelbagai algoritma pembelajaran mesin (machine learning) dan pembelajaran mendalam (deep learning) telah berjaya mengembangkan idea ini kerana ia melibatkan jumlah data trafik dan pemboleh ubah yang besar. Penggunaan dan analisis data atau pemboleh ubah begini adalah penting, namun berbeza antara setiap kajian sekaligus menghasilkan keputusan yang berbeza. Kepelbagaian jenis data trafik mungkin akan mempengaruhi tahap ketepatan pengiraan kesesakan. Kajian ini mencadangkan model konseptual yang dikenali sebagai Model Adaptasi Indeks Prestasi Kelajuan dan Ingatan Jangka Pendek yang Panjang untuk Ramalan Aliran Trafik dalam Bandar Pintar (*Speed Performance Index and Long Short-Term Memory* (SPI-LSTM) *Approach for Enhancing Traffic Flow Prediction in Smart Cities)* berasaskan analisis pemboleh ubah kelajuan yang sering digunakan dalam kebanyakan set data trafik. Kajian ini telah meneroka potensi mengadaptasi data trafik yang dianalisis melalui pengiraan Indeks Prestasi Kelajuan ke dalam algoritma pembelajaran mesin dan pembelajaran mendalam Ingatan Jangka Pendek yang Panjang untuk membuat ramalan trafik. Untuk menguji hipotesis bahawa data trafik mempengaruhi keputusan ramalan, satu eksperimen telah dijalankan menggunakan program Python, untuk menjana hasil yang dijangka dan diramalkan. Keputusan eksperimen ini disahkan menggunakan metrik penilaian dan kemudian dibandingkan dengan model sedia ada lain untuk menganalisis prestasi model yang dicadangkan. Keputusan pengesahan dan perbandingan menunjukkan prestasi yang

positif jika dibandingkan dengan model-model sedia ada, justeru menunjukkan potensi model yang dicadangkan untuk mempertingkatkan ramalan trafik.

ACKNOWLEDGEMENTS

I would like to express my most profound appreciation to my main supervisor, Assoc. Prof Ts. Dr. Noor Afiza Binti Mat Razali, for her genuine passion, guidance, wisdom, and most importantly, her utmost patience in guiding me throughout my study journey. Her dedication and kindness always keeping me on track and motivated me in finishing this study. Thank you to my co-supervisors Assoc. Prof Ts. Dr. Suzaimah Binti Ramli and Assoc. Prof Dr. Mohd Fahmi bin Mohamad Amran for sharing their best knowledge and valuable advice to improve my study. Special thanks to Mister Khairul Khalil, and Assoc. Prof Dr Muslihah Binti Wook for the precious insight on smart transportation, machine and deep learning. Not forgetting to the examiners for their comments and suggestion which means a lot to me to improve this thesis.

A special thanks I would give to both my parent, Shamsaimon Bin Ghazali and Kamariah Binti Sharif, with their sincerest support, guidance, prayers, inspiration and education for me have contributed to where I am right now. Not to forget all my siblings (Along, Ika, Nabila, Asiah and Afiq), especially Norhidayah which have always been supportive throughout my life and inspired me to keep moving forward. The sacrifices all of my family members have given is something I appreciate with my whole heart and will never be forgotten.

In addition, I would want to express my gratitude to my loving husband Ammar Ashraf Bin Narul Akhla, who has been an incredible resource and teacher to me throughout my journey at this university. It is hoped that Allah SWT will bless him with the highest possible recompense for all of the services he has rendered to Him.

Not to forget to my girls Nur Atiqah Binti Malizan and Normaizeerah binti Mohd Noor. These people have been of great partners throughout the duration of this journey by offering assistance, support, and suggestions. Also, to everyone who has been involved, either directly or indirectly, in contributing insights and being willing to share knowledge throughout the course of my research: Thank You.

"So, when you have made a decision, then put your trust in Allah."

[Surah Al Imran Verse 159]

APPROVAL

The Examination Committee has met on **4 October 2022** to conduct the final examination of **Nuraini Binti Shamsaimon** on his degree thesis entitled **"SPI-LSTM APPROACH FOR ENHANCING TRAFFIC FLOW PREDICTION".**

The committee recommends that the student be awarded the of Master of Science (Computer Science).

Members of the Examination Committee were as follows.

Assoc. Prof Dr. Zuraini Binti Zainol
Research Fellow, Department of Knowledge Technology & Security
Universiti Pertahanan Nasional Malaysia
(Chairman)

Ts. Dr. Mohd Sidek Fadhil Bin Mohd Yunus
Senior Lecturer of Faculty of Defence Science and Technology
Universiti Pertahanan Nasional Malaysia
(Internal Examiner)

Prof. Dr. Chiew Kang Leng
Senior Lecturer
Universiti Malaysia Sarawak
(External Examiner)

APPROVAL

This thesis was submitted to the Senate of Universiti Pertahanan Nasional Malaysia and has been accepted as fulfilment of the requirements for the degree of **Master of Science (Computer Science)**. The members of the Supervisory Committee were as follows.

Assoc. Prof Ts. Dr. Noor Afiza Binti Mat Razali
Faculty of Defence Science and Technology
Universiti Pertahanan Nasional Malaysia
(Main Supervisor)

Assoc. Prof. Ts. Dr. Suzaimah Binti Ramli
Faculty of Defence Science and Technology
Universiti Pertahanan Nasional Malaysia
(Co-Supervisor)

Assoc. Prof Dr. Mohd Fahmi Bin Mohamad Amran
Faculty of Defence Science and Technology
Universiti Pertahanan Nasional Malaysia
(Co-Supervisor)

UNIVERSITI PERTAHANAN NASIONAL MALAYSIA

DECLARATION OF THESIS

Student's full name : NURAINI BINTI SHAMSAIMON

Date of birth : 27TH JUNE 1996

Title : SPI-LSTM APPROACH FOR ENHANCING TRAFFIC FLOW PREDICTION

Academic session : JAN 2021/2022

I hereby declare that the work in this thesis is my own except for quotations and summaries which have been duly acknowledged.

I further declare that this thesis is classified as:

☐ **CONFIDENTIAL** (Contains confidential information under the official Secret Act 1972)*

☐ **RESTRICTED** (Contains restricted information as specified by the organisation where research was done)*

☐ **OPEN ACCESS** I agree that my thesis to be published as online open access (full text)

I acknowledge that Universiti Pertahanan Nasional Malaysia reserves the right as follows.

1. The thesis is the property of Universiti Pertahanan Nasional Malaysia.
2. The library of Universiti Pertahanan Nasional Malaysia has the right to make copies for the purpose of research only.
3. The library has the right to make copies of the thesis for academic exchange.

_____ _____

Signature **Signature of Supervisor/Dean of CGS/ Chief Librarian

_____ _____

IC/Passport No. **Name of Supervisor/Dean of CGS/ Chief Librarian

Date: Date:

*If the thesis is CONFIDENTAL OR RESTRICTED, please attach the letter from the organisation with period and reasons for confidentiality and restriction.

** Witness

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3G	Third Generation
4G	Fourth Generation
5G	Fifth Generation
ANN	Artificial Neural Network
ARIMA	Autoregressive Integrated Moving Average
C-LSTM	Convolutional – Long Short-Term Memory
CBR	Case-Based Reasoning
CNN	Convolutional Neural Network
DE-LSTM	Deep Ensemble-stacked Long Short-Term Memory
DL	Deep Learning
DNN	Deep Neural Network
ETA	Estimated Time of Arrival
FDCN	Fuzzy Deep Convolutional Network
GCN	Graph Convolution Network
GRU	Gated Recurrent Unit
HA	Historical Average
HCM	Highway Capacity Manual
IoT	Internet of Things
IT	Information Technology
ITS	Intelligent Transportation Systems
KNN	K-Nearest Neighbour
LGSTN	Local Global Spatial-Temporal Network
LSTM	Long Short-Term Memory
LTE	Long Term Evolution
MF-CNN	Multi Features – Convolutional Neural Network
ML	Machine Learning
Mres-RGNN	Multiple Residual Recurrent Graph Neural Network

MTGCN	Multi Task Graph Convolution Network
RBF	Radial Basis Function
RCI	Relative Congestion Index
RF	Random Forest
Ri	Road Segment Congestion Index
RNN	Recurrent Neural Network
RO	Research Objectives
RQ	Research Questions
SAE	Stacked Auto Encoder
SDLTFP	Supervised Deep Learning Traffic Flow Prediction
SG-CNN	Segment Grouping Convolution Neural Network
SPI	Speed Performance Index
SRI	Speed Reduction Index
STRCN	Spatio Temporal Recurrent Convolutional Network
SVM	Support Vector Machine
SVR	Support Vector Regression
TKGNN	Transfer Knowledge Graph Neural Network
VANET	Vehicular Ad-Hoc Network
VSTGC	Varying Spatiotemporal Graph-Based Convolution
WiMAX	Worldwide Interoperability for Microwave Access
XGBoost	Extreme Gradient Boosting

LIST OF SYMBOLS

υ	-	Average
F_t	-	Actual Value
X_t	-	Prediction Value
n	-	Total number

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Chapter Overview

This chapter explains the background of this study by focusing on the concept of smart city, ITS, machine and deep learning technology for traffic prediction, the problem statement, research objectives and questions, scope and limitations, the definition of terms, and the thesis structure.

1.2 Background of Study

Technological advancements in the current era of digitalisation have produced beneficial outcomes for society in various aspects of life. An example of such advancements includes the introduction of the 'Smart City' concept where various aspects, such as health, education, environment, safety, and transportation have seen significant improvements due to the implementation of technological innovations.

The concept of 'Smart Transportation' was introduced as a branch of the Smart City, which thrives on the application of the Internet of Things (IoT) in the transportation sector. The idea behind Smart Transportation was to solve urbanisation

issues that involve human mobility, with the aims of improving user experience, and providing a safe and organised commute (An & Wu, 2020).

The Machine Learning (ML) and Deep Learning (DL) algorithms have led to numerous opportunities for the implementation of an Intelligent Transportation System (ITS). For example, fleet management, vehicle-to-everything (V2X) communications, intelligent traffic management and monitoring, and various other mobility solutions for roads, rails, air and sea transportation. This system can provide better traffic management, solve traffic problems, enhance traffic efficiency, reduce travel time, and provide comfort and safety to travellers (Karami & Kashef, 2020).

The introduction of smart or connected vehicles requires a medium for wireless communication among these vehicles in order to share valuable information regarding mobility and safety during commute (D. G. Yang et al., 2018). This concept is beneficial for managing and solving traffic problems, taking congestion as one of the main problems faced by travellers all around the world (Makino et al., 2018).

Thus, a technique that can make highly accurate predictions of congestion or traffic flow can be beneficial for traffic management in smart cities. However, currently in Malaysia, only a few studies have focused on traffic flow prediction techniques, which if implemented, could enhance travel experiences and management.

1.2.1 Congestion Problems in Smart Transportation Environment

According to Koźlak, Weisbrod and others, traffic congestion is associated and defined as a situation where the volume of vehicles surpasses the capacity of the road, causing decrease in speed, mutual obstruction, and prohibition of free movement (Koźlak & Wach, 2018; Weisbrod et al., 2003). Traffic congestion has always been related to the travel demands, and the supply of the transport system. The growth of the supply of vehicles on the road, particularly in urban cities have contributed to a long recurring traffic congestion problem.

Various governments (Afrin & Yodo, 2020; Mohan Rao & Ramachandra Rao, 2012; Shamsher & Abdullah, 2013; Ukpata & Etika, 2012) in different countries have come up with various measures but to no avail. Congestion problems have caused an inconvenience for most road users such as the unreliable travel time, higher cost spent during transportation through high fuel consumption, as well as environmental noise and air pollution have all degraded the transportation experience of all road users (Ukpata & Etika, 2012).

1.2.2 Causes of Traffic Congestion

Traffic congestion originates from various reasons and factors. These factors can both be classed as avoidable and unavoidable, but the scenarios of these factors occurring will always result in traffic congestion. Listed down below are some of the main causes and factors of congestion based on the findings of other researchers (Koźlak & Wach, 2018; Litman, 2013; Sarda et al., 2018; Suresh et al., 2018; Ukpata & Etika, 2012).

One of the main causes of traffic congestion is accidents. Accidents can be caused by various factors such as excess speed, vehicular malfunction, driver's condition and tiredness and many more reasons. When an accident occurs, the situation requires attention and cannot be moved at an instance due to the condition of the victims involved and these will cause congestion.

Besides, the condition of the road also contributes to the cause of congestion. Potholes, uneven road surfaces, slippery surfaces, all are possible to cause the driver to lose control of their vehicle which causes congestion.

Furthermore, congestion also mostly occurs due to the driving behaviour and habit of drivers, such as excessive speed and brakes, sudden lane change, no direction change indication, and road rage. This causes other drivers to be alerted and aware, which causes them to slow down and causes congestion.

Next, the weather conditions, such as rain or seasons such as winter, affects the state of the road being slippery and icy, which causes drivers to easily lose control of their vehicle. As drivers are usually more careful during these conditions, congestion occurs due to vehicles slowing down due to carefulness and visibility.

Moreover, a sudden car stopping by the roadside to park inappropriately and illegally can cause congestion, as it interferes with the fluidity of other cars' movement.

Likewise, the increase of the number of vehicles on the road also contributes to congestion. With a lot more vehicles on the road due to an increasing population of people, congestion is almost certain to occur. This is due to the volume of vehicles overwhelming the road capacity, hence causing an abundance of vehicles during a time which the road could not occupy.

Next, during rush hour times in the morning and evening where people commute to and from their workplace, as well as during festive seasons, where people commute to their hometowns causes congestion as the volume of vehicles on the road increases and causes congestion.

Lastly, heavy vehicles such as lorries, trucks, trailers, and buses are able to cause congestion due to the driving capability of their vehicles. Take an example of a lorry driving through a narrow road, and causing other vehicles behind unable to overtake due to the narrow road.

1.2.3 Measures Taken to Prevent Congestion

Multiple initiatives and measures have been taken by governments and organisations to tackle and help solve the ongoing traffic congestion problems. Although congestion is a temporary situation, it has been a concurrent event which causes inconvenience every time it occurs. Addition to the growth of the number of vehicles on the road, initiatives and efforts on traffic congestion management are highly appreciated.

Several studies have listed down solutions and remedies which are suggested to assist in handling traffic congestion (AKYÜZ, 2015; Koźlak & Wach, 2018; Lindley, 1987; Popoola et al., 2013; Shamsher & Abdullah, 2013; Ukpata & Etika, 2012). For example, proper and strict lane management, encourage use of public transportation, public education and awareness campaign, road widening, surveillance and control system, increase of provisional parking space and higher penalty for traffic law infringement.