Abstract—This paper presents a framework for short-term travel time prediction in a motorway with a three-stage architecture: traffic flow forecasting, traffic flow generation and travel time extraction. Traffic flow forecasting reads the historical traffic data and utilizes a forecasting model - Autoregressive Integrated Moving Average (ARIMA) to predict short-term traffic flow. The traffic flow generation utilizes the Cell Transmission Model (CTM) to generate outgoing flow of a road of interest based on the predicted incoming flow from ARIMA. Predicted short term travel times can then be obtained through N-Curve Analysis. Compared to most studies, this paper presents a historical data-driven framework for travel time prediction that can be trained based on specific profiles of routes and cities. The motorway M4 in Sydney, Australia was used to test this framework. It is shown that the predicted travel times can be used to anticipate congestion episodes at the network level.

I. INTRODUCTION

Since the early 1980s, traffic condition forecasting has been a part of an Intelligent Transportation System (ITS) [1]. As an encouragement of the deployment of ITS technology with public value, travel time prediction that can be displayed in Dynamic Messages Signs (DMSs) has been a focus. It can be used to optimize traffic efficiency and thus reduce traffic delays and emissions. The resulted algorithms can also benefit other ITS areas, such as automated highway, by providing real-time and accurate predictions on traffic conditions, which can be fed to semi- or full-autonomous vehicles to assist their decision making such as route choices.

In general, there are two types of methodologies for travel time prediction: one can be termed as mathematical models and another can be termed as computational models. Mathematical models concentrate on utilizing mathematical models to represent the evolution of traffic conditions (e.g., travel times) according to some contributing factors such as locations and time of the day. Computational models, in the other hand, put a focus on utilizing large volume of historical data to establish a heuristic model to predict travel time under different circumstances, such as under non-recurrent traffic jams.

In order to develop a methodology for real-world problems and robust enough for different application scenarios, a framework is developed based on the work of Juri et al. [2] to deal with the travel time prediction problem, especially short-term predictions. This framework, named as PreTravel, is presented in this paper with an example scenario involving a dataset from Sydney, Australia. Its architecture is elaborated along with its performance.

This paper is organised as follows: Section II will include a review of the-state-of-art literatures in the area of travel time prediction. Section III will include a description of the three-stage framework. Section IV will include a verification experiment along with some results and discussions. Section V will conclude this paper along with some discussions regarding future work.

II. LITERATURE REVIEW

When predicting travel times, researchers have found two types of models: mathematical models and computational models. These models can predict travel times directly or be used to predict traffic conditions such as traffic volume, which can be used by traffic simulations to prediction travel times.

The mathematical models are based on regression techniques or time series analysis to obtain predicted travel times. Those statistics-based methodologies have received attention due to their easy-to-implement and easy-to-understand features. Kalman Filters were used to predict travel times based on GPS (Global Positioning System) information and probe vehicles’ data [3]. The Bayesian dynamic linear model was used to predict real-time, short-term travel times by dividing the dataset into two: morning and afternoon [4]. As a time series analysis, ARIMA (Autoregressive Integrated Moving Average) utilised historical traffic conditions obtained in sequence or measured at uniform intervals to learn the trend and variation of them, and thus predict traffic conditions accordingly [2], [5]. Moreover, the process of selecting/fitting an ARIMA model can be automated [6], which means that the model can be tailored to a dataset. In general, the advantage of mathematical models is that the recursive processes do not require the analysis of all historical data, but its disadvantage is the lack of handling some specific non-recurrent traffic situations such as social events.

The computational models, on the other hand, are encouraged by both the spread of high performance computing devices and the advancing of computational intelligent technologies. They utilise computational intelligent algorithms to obtain predicted travel times or conditions such as traffic volumes based on large amount of historical data, which can include more real-world traffic situations, e.g., non-recurrent traffic jams due to social events. Neural networks were used to find hidden dependencies of variables in predicting future traffic states [7], [8]. Support Vector Regression (SVR) could be used to predict travel times and the results suggested that the application of SVR in travel time prediction has
a better empirical risk minimization principle [9]. K-Nearest Neighbour (KNN) was used to predict travel times from two sources of data (vehicle detectors and toll collection systems) with an improvement of the matching process [10]. KNN could be also used to predict travel times under different weathers [11]. In general, those methods can be tailored to specific sites of interest, but may require long training processes.

In order to take into account the traffic dynamics that can include traffic situations such as accidents or special events, a framework was developed to predict travel times by combining a statistical model with traffic simulation as demonstrated in [2]. However, the refinement/tuning of the statistical model, i.e., ARIMA, was not included and no real-world data has been used. Therefore, in order to not only evaluate this framework with real-world situations, but also enhance it with a robust architecture, this paper presents a novel three-stage framework PreTravel, which proposes a refinement procedure to tune the framework, especially the ARIMA model, based on a specific site. The framework will also be evaluated with a real-world data set.

### III. Framework Description

PreTravel is used to predict short-term travel time in a motorway in order to forecast congestion conditions. It has a three-stage architecture: 1) traffic flow forecasting based on a statistical model; 2) traffic flow generation based on traffic simulation and 3) travel time extraction based on N-Curve analysis. Traffic flow forecasting reads the historical traffic data and utilizes ARIMA to predict short-term traffic flow of a road or a road segment of interest. The traffic flow generation utilizes the Cell Transmission Model (CTM) to generate outgoing flow of a road or a road segment of interest based on the predicted incoming flow from ARIMA. Predicted short-term travel times can then be obtained through N-Curve Analysis based on the travel times reflected from the previous traffic simulation. The architecture of PreTravel has been illustrated in Fig. 1.

![Fig. 1. System Architecture of PreTravel](image)

#### A. Framework Workflow

The framework PreTravel follows a procedural process as illustrated in Fig. 2.

![Fig. 2. Flowchart of PreTravel](image)

Historical data with volume information are used to fit an ARIMA model first and then the resulted ARIMA model is used to predict inflows based on the volume data. These two procedures are included in the stage of traffic flow forecasting. Inflows are then used to run the traffic simulation to produce traffic flows in each road segment of interest. The resulted flows are used by the N-Curve analysis to produce travel times. In the rest of this section, each stage will be elaborated.

#### B. Traffic Flow Forecasting: ARIMA and its Fitting

The inflows, i.e., demand for traffic simulation, are predicted based on the data volumes from the database and ARIMA. Although this is a relatively simple time series model (readers might refer to any time-series analysis text, such as reference [12], for a detailed description of ARIMA models), its performance was found to be comparable to that of more complex methodologies. In addition, ARIMA models are relatively easy to implement, do not require extensive input data, and are computationally efficient, enabling their incorporation within a rolling horizon framework.

The traffic flow forecasting procedure is looped until the whole prediction horizon has been covered and the outputs, i.e. the inflows for the prediction horizon are transmitted to CTM in this framework. Before this looped prediction or the deployment of a specific ARIMA model, its order needs to be determined, i.e., $p$, representing the number of auto regression; $d$, representing the number of differentiation operations that can transfer the time series data into a
stationary one; \( q \), representing the number of moving average terms. The determination of the order can be performed automatically or manually. In this study, an automatic fitting process has been used by minimizing the RMSE (root mean square error) according to Equation 1.

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}}
\]  

where \( \hat{y}_i \) and \( y_i \) are the predicted and real traffic flows respectively. \( n \) represents the number of data points that are used to forecast flows, e.g., 10 data points can indicate a 5-minutes forecasting window if the data points representing traffic volume are collected every 30 seconds.

As a result, for every forecasting, ARIMA is tailored to fit the dataset provided so the model order \((p, d, q)\) can change with different RMSE errors.

C. Traffic Flow Generation: Cell Transmission Model

A traditional CTM [13] has been developed to produce the output flow from a set of input flows, provided by the traffic flow forecasting module. The decomposition of a transportation network into a cell network is performed by selecting a discretization time-step which represents the smallest amount of time that is observable. Each link of the network can then be decomposed into multiple cells, according to its length and the free-flow speed on this section of road. Fig. 3 shows each type of cell that can be used to represent a network.

![Fig. 3. Types of Cells](image)

Each cell is defined by several design parameters such as the jam density that represents the local capacity of the road (e.g., 1600 vehicles per hour per lane) and the maximum inflow and maximum outflow that represent the maximum amount of vehicles that can enter or leave the cell during a time-step. Furthermore, the cell decomposition must keep track of the topology of the network and therefore correctly introduce merging and diverging cells to represent the successors and predecessors of each node in the network.

D. Travel time Extraction: N-Curve Analysis

Travel time on a freeway under work zone conditions can be predicted with N-Curve analysis [14]. This approach can be described by modifying input-output diagrams to measure the time and distance spent by vehicles in a queue in a simpler manner than using a time-space diagram. This process requires the construction of a curve depicting the cumulative number of vehicles reaching the back of queue as a function of time.

The advantage of utilizing the N-Curve model is that it requires only traffic flow counts for the upstream, downstream, and ramp points. In the case of ramps existing between the upstream and downstream detectors, the highway segment loses its conservation of flow and the cumulative curve at the downstream detector cannot be directly used for travel time extraction without further modifications. In the event that both on- and off-ramps exist in the segment of interest, the situation could become more complicated and virtual detectors would need to be adopted in the N-Curve analysis.

IV. Case Study

PreTravel has been implemented in C++ based on the QT 4.8.3 64bit library\(^1\) under Windows 7 Enterprise 64bit together with R forecast package [6]. A section of the motorway M4 in Sydney, Australia was used to test this framework.

A. Data Description

The data used in the case study were obtained from 246 monitor sites along the M4 motorway, which is a motorway in Sydney, New South Wales, Australia. M4 connects the inner-west of Sydney with the outer western suburbs and is 46 kilometres long. This dataset was provided by the Transport for New South Wales (TfNSW) in Sydney, Australia.

Each monitor site records traffic conditions, e.g., average speed, traffic flow rate, occupancy, in 24 hours each day. The dataset in February, 2013 was initially used in this case study to test the framework and includes traffic conditions of M4 in every 30 seconds from 246 monitor sites. The raw dataset consists of 6914 data files in binary format throughout 28 days and 246 monitor sites. Those raw data was stored into a structured SQL database by transferring the Hexadecimal-based raw data into meaningful human-readable M4 flow data, which contain fields such as monitor site and time stamp. The locations of each monitor site were also identified with a location record file. Each record includes one lane’s states in 30 seconds at one monitor site, including monitor site ID, time stamp, etc.

B. Experiment Procedure

One section, with seven monitor sites (from east to west in sequence: No. 20a - No. 26a) that recorded traffic from Church Street, Granville to Pearson Street, South Wentworthville on 1/2/2013 between 14:30pm to 15:34pm, was used. They have formed a road segment with a total length of 3.1km. Some pilot studies have been used in this experiment to determine the parameters of the CTM model so that the predicted travel time distribution can reflect the original traffic conditions within the road segments without ramps. The CTM model in this study used a 4-second interval with a

\(^1\)http://qt-project.org/
free flow speed of 90 kmh, which resulted in 31 cells in total. The capacity is 1600 vehicles per hour per lane and the jam density is 180 vehicles per kilometre. The observed inflow and speed on 1/2/2013 between 14:00pm to 20:00pm for monitor site No. 22a included in this case study is illustrated in Fig. 4.

Fig. 4. Observed Inflow and Speed Used in Case Study

Fig. 4 shows that the detected speed in monitor site No. 22a is firstly decreasing from around 80 kmh to around 30 kmh, and then increasing back to around 80 kmh. At the same time, the detected inflow is fluctuating around 1500 vehicles per hour. The relationship between inflow and speed indicate that congestion is forming during this time period, especially around 15:00 in the afternoon. In the next part of this paper, we will test the PreTravel in tracking a congestion condition. Therefore, travel times predicted by PreTravel will be analysed to investigate any predicted congestion conditions.

In addition, the number of data points that should be used to fit the ARIMA model was determined by a pilot study, which compared the RMSE associated with each fitted ARIMA model that is based on different number of data points used. A 30-minutes window was selected and includes 60 data points. A 5-minutes prediction window was also selected. Therefore, the data points recorded between 15:00pm to 15:30pm will be replaced with predicted values and the analysis will be concentrated on this time period.

C. Results

The observed data from 15:00pm to 15:30pm is shown in Fig. 5, which is generated by N-Curve analysis. From Fig. 5, we can see that the travel times between monitor sites are increasing as the result of different vehicles arriving rate.

As low speeds and high volumes can be found in Fig. 4, the forming procedure of the congestion is the key reason for the increasing of the travel times. To compare with the real congestion situation, the following part of this paper will use PreTravel to predict travel times for the same road section.

Firstly, the output of the ARIMA model is presented in Fig. 6, which shows the evolution of the observed volumes together with predicted ones. Table I summaries the resulted ARIMA models, whose RMSE ranges from 409.2729 to 560.2724.

Fig. 6. Observed and Predicted Inflow

<table>
<thead>
<tr>
<th>Fitted Period</th>
<th>Predicted Period</th>
<th>ARIMA Model</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:00:00</td>
<td>15:00:00</td>
<td>6,1,9</td>
<td>485.4158</td>
</tr>
<tr>
<td>14:05:00</td>
<td>15:05:00</td>
<td>8,1,9</td>
<td>487.968</td>
</tr>
<tr>
<td>15:04:30</td>
<td>15:09:30</td>
<td>2,1,9</td>
<td>560.2724</td>
</tr>
<tr>
<td>15:40:00</td>
<td>15:10:00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15:09:30</td>
<td>15:14:30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14:45:00</td>
<td>15:15:00</td>
<td>9,1,9</td>
<td>519.3689</td>
</tr>
<tr>
<td>15:14:30</td>
<td>15:19:30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14:50:00</td>
<td>15:20:00</td>
<td>7,1,9</td>
<td>515.6163</td>
</tr>
<tr>
<td>15:18:30</td>
<td>15:24:30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15:45:00</td>
<td>15:25:00</td>
<td>5,1,9</td>
<td>495.4232</td>
</tr>
<tr>
<td>15:24:30</td>
<td>15:29:30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15:00:00</td>
<td>15:30:00</td>
<td>8,2,7</td>
<td>409.2729</td>
</tr>
<tr>
<td>15:20:30</td>
<td>15:34:30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It is shown that the inflows predicted by automatic fitting ARIMA model reflect the evolution of the observed data during the whole prediction horizon of 30 minutes (six 5-minutes prediction windows), although in some shorter period (1-3 minutes), the predicted inflows does not follow the pattern of observed ones.

As the free travel time should be nearly 124 seconds with 90 km/h free flow speed, the predicted travel times from 15:00:00 to 15:00:30, ranging from 124 seconds to 272 seconds, show that a congestion is forming in the road section. The evolution of travel time distribution in Fig. 7 shows the increasing of travel times from a near-unified free flow travel times to larger travel times, which is a firm demonstration of the existence of congestion as shown.
Caused by the congestion, the predicted travel times are larger than those in the 124 seconds group at 15:00:00, which is consistent with the inflow observed around that time slot (Fig. 4)

![Travel Time Prediction 15:00:00 (1/2/2013)](image)

![Travel Time Prediction 15:00:30 (1/2/2013)](image)

Fig. 7. The Distribution of Predicted Travel Times

V. CONCLUSIONS AND FUTURE WORK

In this paper, a framework for short-term travel time prediction was presented. It is aimed to deal with real-world problems with a modular, extensible architecture for different application scenarios. This framework includes a traffic flow forecasting procedure with automatic fitting ARIMA, which feeds the traffic simulation model CTM with inflows. CTM then produces outflows, which are used to produce predicted travel times based on N-curve analysis.

This framework was tested in a case study with a road segment from M4 motorway in Sydney, Australia. As the predicted travel times can be used to forecast congestion condition, this research is promising. However, PreTravel should be further developed to improve its performance in order to deal with other road conditions, such as non-recurrent accidents or traffic lights. Basically, a module can be developed to calibrate ARIMA and CTM models, e.g., if the capacity of each lane should be reduced if a non-recurrent accident has been found by either recognising the flow pattern or importing real-time event alerts. Moreover, other methodologies can be adopted by the traffic flow forecasting procedure to deal with complicated road infrastructures, e.g., with several on/off-ramps. This travel time prediction framework, termed as PreTravel, can be used by relevant agencies to support the deployment of the ITS with a need of real-time information.

ACKNOWLEDGMENT

This research received support from the Transport for New South Wales (TfNSW), Australia.

REFERENCES


Zhitao Xiong is a post-doctoral researcher at the Research Center for Integrated Transport Innovation (rCITI) in the University of New South Wales (UNSW). He started his post-doc appointment at rCITI in his fourth year of PhD at the Institute for Transport Studies in the University of Leeds. His PhD research focused on creating repeatable and realistic driving simulation for human factor researches. His research interests include driver models, driving behaviours, knowledge representation and automated action planning.

David Rey is a post-doctoral researcher in transport engineering at the Research Center for Integrated Transport Innovation (rCITI) at the University of New South Wales (UNSW). He joined UNSW after obtaining his PhD in operations research from the Université de Grenoble during which he studied conflict resolution strategies arising in air traffic management. His key research interests lie in resource allocation schemes in transportation networks, mathematical programming as well as fair optimization.

Tuo Mao is a PhD candidate in University of New South Wales (UNSW) as well as a researcher at Research Center for Integrated Transport Innovation (rCITI) in UNSW. His PhD research focuses on Highway Traffic Signal Control and Management including Transit Signal Priority (TSP), Coordinated Ramp Metering (CRM) and Variable Speed Limit (VSL). He is interested in developing efficient Intelligent Transport Systems (ITS) and improving them.

Haiyang Liu is a visiting PhD student at the Research Center for Integrated Transport Innovation (rCITI) in the University of New South Wales (UNSW). He received the B.E. Degree in traffic engineering from Northeast Forestry University, Harbin, China, in 2009 and M.E. in traffic engineering information and control from Harbin Institute of Technology, where he is a candidate for the Ph.D. degree in transportation planning and management. His research interests include traffic flow theory, evacuation management and intelligent transportation system.

Vinayak V. Dixit is a Senior Lecturer at the Department of Civil and Environmental Engineering at the University of New South Wales (UNSW). His key research interest lies in studying risk in the transportation infrastructure system as it relates to highway safety, travel time uncertainty, as well as natural and man-made disasters.

S. Travis Waller is the Evans & Peck Professor of Transport Innovation and Director of the Research Centre for Integrated Transportation Innovation (rCITI) at the University of New South Wales (UNSW) in Civil and Environmental Engineering. Professor Waller was named one of the top 100 young innovators in science and engineering in the world in 2003 by MIT’s Technology Review magazine, was awarded the National Science Foundation’s CAREER award in 2004, received TRB’s Fred Burggraf research award in 2009 and the 2011 Hojjat Adeli Award for Innovation in Computing.