Predicting Travel Times in Dense and Highly Varying Road Traffic Networks using STARIMA Models.

Islam-Khan, Rezwan; Landfeldt, Björn; Damdhere, Ashay

2012

SCHOOL OF INFORMATION TECHNOLOGIES

PREDICTING TRAVEL TIMES IN DENSE AND HIGHLY VARYING ROAD TRAFFIC NETWORKS USING STARIMA MODELS

TECHNICAL REPORT 685

REZWAN-AL-ISLAM KHAN, BJORN LANDFELDT, ASHAY DHAMDHERE

FEBRUARY 2012

Predicting Travel Times in Dense and Highly Varying Road Traffic Networks using STARIMA Models.

Rezwan-Al-Islam Khan[d,e], Bjorn Landfeldt[d], Ashay Dhamdhere[f]

[a] *School of Information Technologies, The University of Sydney and*
[b] *National ICT Australia(NICTA)*
[c] *School of Electrical Engineering and Telecommunications, University of New South Wales*

Keywords: Intelligent transport systems, Predictive modelling, Urban network, Motorway Network, Road traffic

Email addresses: `rkha8047@uni.sydney.edu.au` (Rezwan-Al-Islam Khan), `bjorn.landfeldt@sydney.edu.au` (Bjorn Landfeldt), `ashay@unsw.edu.au` (Ashay Dhamdhere)

Predicting Travel Times in Dense and Highly Varying Road Traffic Networks using STARIMA Models.

Rezwan-Al-Islam Khan[d,e], Bjorn Landfeldt[d], Ashay Dhamdhere[f]

[d] *School of Information Technologies, The University of Sydney and*
[e] *National ICT Australia(NICTA)*
[f] *School of Electrical Engineering and Telecommunications, University of New South Wales*

Abstract

As road networks are becoming increasingly utilised, it is increasingly important to be able to accurately predict travel times for disseminating information among road users and to support traffic management decisions and planning. Over the past few years, a new way of making such predictions utilising the space-time autoregressive integrated moving average (STARIMA) model has been introduced. The results have been very promising so far with very good accuracy reported for prediction times of several tens of minutes. However, so far, the literature only concerns steady state freeflow or Manhattan grid based scenarios. In this paper, we generalise on the previous work and investigate how this approach performs in urban traffic scenarios. We investigate the models prediction accuracy both on a data set of measured travel times in the metropolitan Sydney region as well as a faithful representation of a large section of Sydney's urban landscape. We analyse the performance of STARIMA under six level of service (LOS) in both settings and find that even though the model performs well in the steady state case, the basic approach is not suitable for modelling the urban traffic setting. We therefore propose to extend the STARIMA model with feedback control loops in order for the approach to be suitable for highly varying environments.

1. Introduction

Intelligent Transport Systems (ITS) combine a wide array of wireless and wire line communication-based technologies. Integrating these technologies into vehicles and transport infrastructure can significantly help in improving productivity and safety. One of the main issues in intelligent transport systems is to forecast various traffic conditions such as traffic flow, speed and travel time; and among these

Email addresses: `rkha8047@uni.sydney.edu.au` (Rezwan-Al-Islam Khan), `bjorn.landfeldt@sydney.edu.au` (Bjorn Landfeldt), `ashay@unsw.edu.au` (Ashay Dhamdhere)

travel time is considered to be the most important as it has direct implications for individual travellers and network authorities who manage the road networks. Providing travel time information to all users facilitates more informed decisions on route choices, thereby reducing the overall travel time on the road network. Similarly, a knowledge of travel times and traffic volumes can significantly aid in provisioning infrastructure (such as roadside access points) for providing communication services over vehicular ad hoc networks.

Predicting travel time is difficult since traffic conditions can vary widely in both spatial and time domains (Hall and Persaud (1989)). Different travel-time prediction models have been proposed depending on the complexity and size of road network and the size of the prediction window. On one end of the scale, complex networks and larger prediction horizons need to be considered for planning purposes, while simpler networks and immediate travel-time prediction is suitable for real-time guidance purposes (Liu (2008)). A variety of techniques can be used for measuring travel times in road networks, including loop detectors, image processing cameras placed at intersections, automatic vehicle identification, floating cars, cellular phone tracking and GPS-based locating (Rad (2007)). Conventional loop detectors are the most commonly used form of measurement; flow, occupancy and travel time data from these detectors for freeway networks is widely availableRegiolab-Delft (2010); kaggle.com (2010). This data has been used as a source in many travel time prediction models. Spectral analysis (Nicholson and Swann (1974)) and Kalman filtering (Okutani and Stephenades (1984)) have been used to predict traffic flow on very short time horizons, but the flow predictions do not necessarily lead to accurate travel time prediction. Therefore, recent research efforts using neural networks or statistical models mostly focus on direct prediction of travel time using historic and current levels of flow and occupancy data (Nookala (2006)). Popular models use a variety of statistical methods including parametric linear/non-linear and non-parametric regressions, Auto Regressive Integrated Moving Average (ARIMA), Space-Time Auto Regressive Integrated Moving Average (STARIMA) and ATHENA (Danech-Pajouh and Aron (1991)). Among these models STARIMA is especially notable because of its applicability in spatio-temporal domains and relatively small number of parameters needed to estimate the model.

In the context of traffic forecasting, STARIMA models can be used when there are spatial time series collected at a specific location at constant time intervals. They can be used to model the effects of shocks that are introduced at a given spatial point in the system (e.g. an accident on a link that affects the traffic conditions on neighbouring links). STARIMA models can model an entire network with only one model while ARIMA needs a separate model for each link. For example, Kamarianakis and Prastacos

(2002) presents a STARIMA model that requires only 7 parameters while the corresponding ARIMA models require 75 parameters. Although STARIMA models have been used for forecasting a plethora of multivariate data including crime rates, potato prices, Server load and flour prices, they have not been used to predict travel-time and traffic flow until relatively recently (Lin et al. (2009); Min et al. (2009); Min and Wynter (2011); Min et al. (2010); Cheng et al. (2011); Ding et al. (2010)). These papers reported impressive accuracy in relatively long-term prediction. In all cases the models were estimated either for strictly small scale Manhattan grid urban networks with traffic signals, or networks of highways or freeways. According to Min et al. (2010) the multivariate nature and the ability to capture spatio-temporal behaviour makes STARIMA models the preferred choice for urban traffic networks and very promising for planning and resource allocation. We return to these works in detail in section 3 after we have presented an overview of the STARIMA model itself in section 2 .

Even though the above-mentioned results are encouraging, the investigations were limited to scenarios with low complexity and resulting stable traffic conditions. It has been shown in Torday and Dumont (2004) that estimation of travel time depends strongly on the type of traffic network being analysed. Congestion and traffic control methods (signals/giveway/stop signs/roundabouts) are the two main components that influence travel time in urban networks. Congestion introduces long-term variability while traffic control methods introduce short-term variability. Because of the short term variability, it is more difficult to estimate travel time in urban networks than on freeways or motorways.

Therefore, in order to evaluate the suitability of using the STARIMA technique in general traffic scenarios, in this paper we investigate how well STARIMA models can perform in both urban/mixed urban and freeway networks. We also investigate the tolerance of these models to changes in traffic volumes and evaluate their accuracy in predicting travel times under different congestion levels. In order to provide results that are easy to interpret and to be used for comparison in future studies, we define congestion levels in accordance with the Level of Service (LOS) categories in the Highway Capacity Manual 2000 standard (HCM 2000)(TransportationResearchBoard (2000)). We use data from a traffic simulator (for urban/mixed urban case) as well as from New South Wales traffic authorities to build and test our STARIMA models and provide insight into the models' performance in the general traffic case.

This paper is organised as follows. Section 2 introduces a brief description of the STARIMA model. In section 3, we present related work that use this model and variations of this model. Methodologies for data collection, model estimation and prediction are described in Section 4 . Section 5 details our experiments and results from which we draw conclusions about the performance of the STARIMA model. We conclude and discuss future research directions in Section 6.

2. Brief intro to STARIMA

The STARMA model attempts to describe and forecast spatial interrelationships between different regions with a set of N observable time series that represent patterns in N regions. It is characterised by linear dependence lagged in both space and time (Pfeifer and Deutsch (1980)). This model is used to fit and describe $X_i(t)$random variables observed in N linear combination of past observation $\mathbf{x}\left(t-k\right)$ and errors $\boldsymbol{\epsilon}\left(t-k\right)$ which may be lagged in both space and time. The weighted linear combination mechanism is achieved by hierarchical ordering of the neighbours of each region and sequence of a $N \times N$ weight matrix. This weight matrix may represent any spatial relationship between observations in different regions depending on the nature of the observation. The STARIMA model is expressed as follows:

$$\nabla^d \mathbf{x}\left(t\right) = \sum_{k=1}^{p} \sum_{l=0}^{\lambda_k} \phi_{kl} \mathbf{W}^{(l)} \nabla^d \mathbf{x}\left(t-k\right) - \sum_{k=1}^{q} \sum_{l=0}^{m_k} \theta_{kl} \mathbf{W}^{(l)} \boldsymbol{\epsilon}\left(t-k\right) + \boldsymbol{\epsilon}\left(t\right) \tag{1}$$

where p is the autoregressive order, q is the moving average order,λ_k is the spatial order of the kth autoregressive term, m_k is the spatial order of the kth moving average term, ϕ_{kl} is the autoregressive parameter(A.R.) (to be estimated) at temporal lag k and spatial lag l, θ_{kl}, is the moving average(M.A.) parameter (to be estimated) at temporal lag k and spatial lag l, $\mathbf{W}^{(l)}$is the $N \times N$ matrix of weights for spatial order$l$. $w_{i,j}^{l}$ is non-zero if there is a spatial relationship between location i and j,$\boldsymbol{\epsilon}\left(t\right)$ is random error vector at time t with an expected value:

$$E\left[\boldsymbol{\epsilon}\left(t\right)\boldsymbol{\epsilon}\left(t+s\right)'\right] = \begin{cases} \sigma^2 \mathbf{I}_N & \text{if } s = 0; \\ -x & \text{otherwise.} \end{cases} \tag{2}$$

The space-time auto covariance function is expressed as:

$$\gamma_{lk}\left(s\right) = E\left\{ \frac{\left[\mathbf{W}^{(l)} \nabla^d \mathbf{x}\left(t\right)\right]' \left[\mathbf{W}^{(l)} \nabla^d \mathbf{x}\left(t+s\right)\right]}{N} \right\} \tag{3}$$

and the Space Time Autocorrelation Function (STACF) $\rho_l\left(s\right)$ and $\gamma_{h0}\left(s\right)$ Space Time Partial Autocorrelation Function (STPACF) respectively are expressed as:

$$\rho_l\left(s\right) = \frac{\gamma_{lk}\left(s\right)}{\left[\gamma_{ll}\left(0\right)\gamma_{kk}\left(0\right)\right]^{\frac{1}{2}}} \tag{4}$$

$$\gamma_{h0}(s) = \sum_{j=1}^{k}\sum_{i=0}^{\lambda} \phi_{jl}\gamma_{hl}(s-j) \tag{5}$$

The conditional likelihood estimation function as proposed in Pfeifer and Deutsch (1980) for Φ, Θ and σ^2 is expressed as:

$$L\left(\Phi, \Theta, \sigma^2 \mid \mathbf{x}\right) = (2\Pi)^{\frac{-TN}{2}} \left(\sigma^2\right)^{\frac{-TN}{2}} \exp\left(-\frac{S_*(\Phi, \Theta)}{2\sigma^2}\right) \tag{6}$$

where $S_*(\Phi, \Theta)$ is the conditional sum of squares function:

$$S_*(\Phi, \Theta) = \hat{\epsilon}'\hat{\epsilon} \tag{7}$$

The vector $\hat{\epsilon}$ is calculated using the expression:

$$\boldsymbol{\epsilon}(t) = \nabla^d \mathbf{x}(t) - \sum_{k=1}^{n}\sum_{l=0}^{\lambda_k} \phi_{kl}\mathbf{W}^{(l)}\nabla^d \mathbf{x}(t-k) + \sum_{k=1}^{q}\sum_{l=0}^{m_k} \theta_{kl}\mathbf{W}^{(l)}\boldsymbol{\epsilon}(t-k) \tag{8}$$

where $\mathbf{x}(t) = 0$ and $\boldsymbol{\epsilon}(t) = 0$ for $t < 1$. The Maximum Likelihood Estimation (MLE) function for sample values $\hat{\Phi}$, $\hat{\Theta}$ and $\hat{\sigma}^2$ is finally formulated as:

$$\hat{\sigma}^2 = \frac{S_*\left(\hat{\Phi}, \hat{\Theta}\right)}{TN} \tag{9}$$

Two special sub-class models of STARIMA are STAR ($q = 0$) and STMA ($p = 0$). As STARIMA is non-linear in nature ,in order to estimate the ϕ_{kl} and θ_{kl} coefficients, nonlinear optimisation techniques must be used (Pfeifer and Deutsch (1980)). The following steps are performed in order to build the model proposed in Pfeifer and Deutsch (1980); Zhou and Mitchell (2005):

1. Identification; determine the number of auto regression coefficients should be used to fit the data by using STACF (4) and STPACF (5).
2. Estimation of Auto Regressive and Moving Average parameters
3. Model diagnostics; investigate the error obtained by comparing the original time series with the fitted time series. If the random error test is not passed, return to step 1 for re-identification of the model.
4. Prediction; applying the derived model to the traffic flow data for travel time estimation.

We refer the interested reader to Pfeifer and Deutsch (1980) for further details on the STARIMA model.

3. Related STARIMA based prediction work

The STARIMA model has been applied to traffic flow and travel time prediction relatively recently. In Lin et al. (2009) the STARIMA model was built to predict traffic flow on freeways around Beijing and it was reported that this model suitably fit traffic data. A slight variation of this model named Dynamic STARIMA was introduced in Min et al. (2009), in which the authors used a dynamic turn ratio prediction (DTRP) model to improve the performance of traffic flow prediction. Another proposition by the same authors in Min et al. (2010) was the Generalised STARIMA (GSTARIMA), where an increased number of parameters were estimated using DTRP depending on the flow pattern of various observation points to improve traffic flow prediction and yielded a 0.2% improvement with increased complexity in parameter estimation as an added cost. It was also reported in Cheng et al. (2011) that the GSTARIMA model may not be functioning properly because of temporal shifts in prediction. Both DSTARIMA and GSTARIMA works used the same 12×2 arterial Manhattan grid network in Beijing with 50 sensors. An extensive study on spatio-temporal correlations was presented in Cheng et al. (2011) where travel-time data from automatic number plate recognition (ANPR) cameras was provided by the London Congestion Analysis Project (LCAP). The authors used data from a small section of 22 links comprising freeway and urban roads for analysis but no prediction performance was presented. In Min and Wynter (2011) the STARIMA model was used to predict travel times using speed and volume data from 502 links and the authors reported a maximum accuracy of 88.6% for prediction 60 minutes into the future, though the type and congestion level of the road network was not clear. In the paper, $w^l_{i,j}$ was set to 1 if a vehicle originating from location i may reach link j within l time interval(s), rather than describing the weighted spatial relationship between different sites. This change was crucial in order to capture dynamic change in traffic conditions at the cost of added complexity in the model identification. In Ding et al. (2010), the authors used real expressway data from 13 free flow links in Beijing from to predict traffic volume and compared the results with the ARIMA model. The authors also reported that volume prediction was sufficiently accurate for 30 minute intervals. In this paper, the $\mathbf{W}^l$ matrix was constructed similar to the basic principal used in Min and Wynter (2011) with an added subcategory based on the average speed of all six LOS levels. In our approach, we also followed the same design principle for representing $\mathbf{W}^l$ as explained in sections 4.2 and 4.3.

4. Methodology

In this paper we follow the same methodology as in Min and Wynter (2011) in order to verify that our model conforms and yields similar results as previously presented. However, since we are investigating an urban scenario with significantly increased complexity we need to extend the procedures and model as described below.

4.1. Data collection

For our extended investigations we sourced data using two different methods. Due to the scarcity of suitable publicly available urban network data, extensive simulations were performed using the Quadstone Paramics simulator. The urban traffic network model designed using Paramics is a faithful recreation of an open (i.e. vehicles travel in and out of the road network) segment of the greater Sydney metropolitan region including all traffic lights, turning rules, roundabouts and speed limits including variable speed limits for school zones. Loop detectors were used at different points to collect data about speed and volume at fixed intervals. In order to generate vehicular traffic in Paramics, the road network was divided into multiple zones. An origin-destination matrix (demand matrix in paramics) was defined to generate traffic volume from one zone to the next. The demand matrix for different time periods reflected changes in traffic volumes according to changes in traffic situation (e.g. peak, off-peak, event, holiday, etc.). We introduced a load shift parameter in the predefined demand matrix to increase or decrease the overall traffic volume in the network. After establishing the baseline demand, the load shift parameter was incrementally changed in the demand matrix to achieve the desired load. The generated LOS levels and the corresponding load levels generated by the demand matrix in Paramics are listed in Table 1. All the simulation data for a particular load/scenario was based on 10-30 runs with random seeds and 4 hours duration, where the first hour was used for warming up the simulation and letting the system reach steady state. The route selection for vehicles was dynamic except for public transport units (e.g. buses). The first set of data generated from the simulator was used for model identification, and the second set of data was used as real-time input to fit the model and make travel time predictions. For free flow scenarios, we used a data set captured on the M4-Western Motorway provided by Roads and Traffic Authority (RTA), in New South Wales (NSW), Australia kaggle.com (2010). This data set contained measurements of the time required to travel different sections of the motorway network, captured every 3 minutes.

Table 1: LOS Levels of HCM2000

LOS Level	Traffic Condition (Rodrigue et al. (2009))	Corresponding Load in Paramics
A	Free Flow Traffic	0-24%
B	Steady Traffic	25-36%
C	Steady Traffic but Limited Mobility	37-54%
D	Steady Traffic at High Density	55-62%
E	Traffic at Saturation	63-78%
F	Congestion	>79%

4.2. Model building

Two factors need to be considered to use this model for a vehicular traffic network. First, there is a need to determine what type of data $\mathbf{x}(t)$ should be used to fit, and secondly how to define the spatial relationship through the $\mathbf{W}$ matrix. $\mathbf{x}(t)$ is a $N \times T$ matrix of time series of data (speed,volume and travel time) observed in N different locations (from here on referred to as links) at T points in time. A differencing operator ∇^d was used on data d times to yield an approximate stationary stochastic process. The design and assignment of weights in the $N \times N$ $\mathbf{W}^l$ matrix governs the time interval and spatial relation within N different links. In Min et al. (2009, 2010) defines $w^l_{i,j}$ as the turn ratio from link i to link j or the flow rate from link i to j link (Lin et al. (2009)). As mentioned above, in this paper we followed the weight assignment proposed in Min and Wynter (2011). The value of $w^l_{i,j}$ was set to 1 if a vehicle originating from link i may reach link j within l time interval(s) denoted by $w^l_{i,j} = 1$ or $w^l_{i,j} = 0$. By doing so, it captures the dynamic effective range of the spatial neighbourhood Cheng et al. (2011). Thus, if there will be two separate matrices $\mathbf{W}^1$, $\mathbf{W}^2$ describing links reachable from origin to destination within 1 and 2 time-steps. For example, if each time step is 5 minutes long, $\mathbf{W}^1$ represents links reachable within 5 minutes, $\mathbf{W}^2$ links reachable within 10 minutes etc. In the classic STARIMA model definition (Pfeifer and Deutsch (1980)) where $\sum_{j=1}^{N} w_{ij} = 1$, model identification is easily done using the STACF and STPACF. As $\sum_{j=1}^{N} w_{ij} \geq 1$ in this particular way of weight assignment, these functions were used as an indication of the number of AR and MA parameters that will be needed to fit the observed data. Finding the coefficients for a candidate model is solved using the MATLAB non-linear optimization tool set. To avoid over-fitting the data, at least 20 data points per parameter were used for coefficient estimation. A candidate model was then chosen by performing an F test to check the relationship between the response and predictor variables (with large F and $p < 0.05$) and t-testing for significance in the order of AR and MA coefficients (when $t > +2.0$ or $t < -2.0$ and $p < 0.05$) as detailed in Box and Jenkins (1994). The parameters ϕ and θ were fitted using MLE estimation equations (6) to (9).

4.3. Weight matrix Generation

Generating the weight matrix involves several steps. First, a default $N \times N$ neighbour matrix has to be generated that records immediate neighbours. The entries in the neighbour matrix is then updated with individual link travel time from the collected data for training purposes. A Breadth First Traversal (BFT) algorithm is then applied to build traversal trees for all N links. These trees along with the matrix with individual travel time are then used to create a matrix $\mathbf{A}^{timefit}$that records the total time needed to reach any reachable destination from each origin link and which constitutes the base for generating weight matrices. $\mathbf{W}^{l}$ Matrices were generated depending on the maximum travel time for all originating links to all destination links, defined by the size of time step l. For example, if the maximum recorded travel time is 15 minutes and one timestep is 5 minutes then a total of 15/5=3 $\mathbf{W}^{l} = \mathbf{W}^{1}, \mathbf{W}^{2}, \mathbf{W}^{3}$ matrices has to be generated.

4.4. Model testing and prediction

The final step in preparing the model for travel time prediction lies in ascertaining that the derived parameters fit the data. After estimating the parameters, the predictions of the model therefore has to be tested on collected data and the accuracy evaluated before the derived parameters are being deemed suitable. In our experiments described in the section below, we therefore warmed up the complete model with one hour of collected data before starting the experiments. Predictions were performed by substituting the $(t+1)$st prediction as observed value to predict travel time at $(t+2)$, up until the desired prediction horizon (i.e. up to $t+15$). Then the predicted data for different links were used to generate a predicted travel-time matrix $\mathbf{A}_{l}^{timepredict}$ and compared with the observed travel-time matrix $\mathbf{A}_{l}^{timeobserved}$. This prediction comparison was used in order to estimate the error in the total travel time estimation, which is presented in the experiments section 5.

5. Experiments

Experiments were carried out for the two different kind of road networks. First we set up experiments for the Sydney M4 motorway network in order to recreate the type of scenario reported on in Min and Wynter (2011); Ding et al. (2010). We used these experiments to validate that our model and implementation conformed with previous work so that our results could be used to generalise previous results to cover the urban scenario. The $\mathbf{W}^l$ matrix categorisation based on LOS average speed as in Ding et al. (2010) was not used for these experiments as the test network was straightforward. The second experiment was set up for the urban scenario with a road network considerably more complex than in Lin et al. (2009); Min et al. (2009, 2010); Cheng et al. (2011), where the impact of variability (both long and short-term) of urban traffic on the STARIMA model was left unexplored. We opted to train the models based on the individual LOS levels as proposed in Min and Wynter (2011) rather than defining $\mathbf{W}^l$ based on the LOS sub category as was done in Ding et al. (2010). This allowed us to compare our results with both these previous approaches.

5.1. Experiments on the freeflow Motorway Network

5.1.1. Motorway network and Experimental set up

The first traffic network modelled the M4-Western Motorway in Sydney, Australia (fig 1) and was provided by the NSW Roads and Traffic Authorities (RTA) in Australia. The model represented a 46Km, 6 lane stretch of motorway that connects the western suburbs of Sydney at speeds ranging from 90Km/h to 110Km/h. The RTA provided travel time measurements on this stretch of road collected over the period from March 1, 2010 to Aug03, 2010 (155 days). In the data provided, the RTA divided the motorway into 64 different segments, 32 for uplink and 32 for downlink. Travel-time was measured in centi seconds (with some measurement error) and the data was collected at 3 minute intervals using loop detectors placed every 500m (kaggle.com (2010)). We calculated the mean travel time per 2 data points and used this as one data point, thus making one time step equal to 6 minutes in order to reduce the number of $\mathbf{W}$ matrices needed. The network model was divided into two separate networks without interconnection for uplink and downlink traffic with 32 points of observation each in accordance with the data collection points in the supplied data. Because of the size and unidirectional nature of the network, we modelled the segments with two separate STARIMA models, one each for the uplink and downlink. The observed travel time was used as input to the STARIMA model $\mathbf{x}(t)$. The $\mathbf{W}^l$ estimation from $\mathbf{A}^{timefit}$ was relatively straight forward for the motorway network. Model parameters were estimated using data from 24 hours of observed travel times (14,400 data points for 32 locations), and finalised after F and T tests.

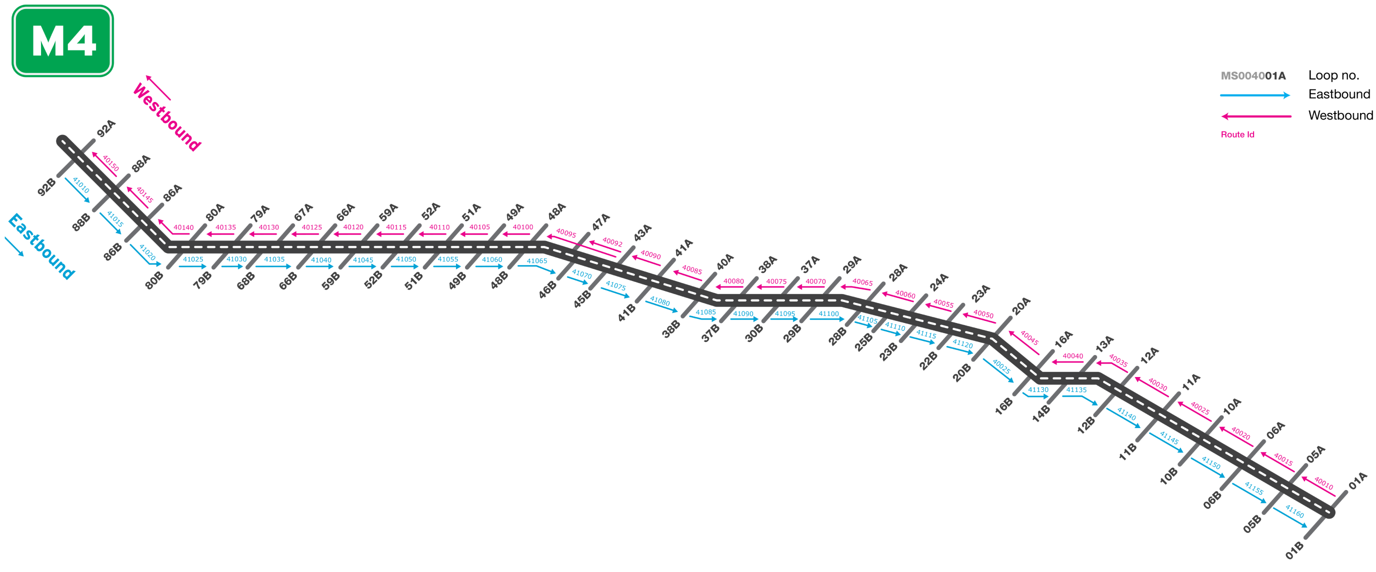

Figure 1: M4 motorway network

5.1.2. Experiments

We used two separate STARIMA models, one for each direction along the motorway in our experiments. The first model was built using data from traffic during weekdays and the second model was built for weekend traffic. Traffic data from holidays such as Easter were excluded from testing. Each model was then tested with 24 hour data during a period of fifteen weeks and the prediction performance was measured. For weekdays, the prediction accuracy reached a maximum of 83%- 85.26% up to 14 time steps (84 minutes) ahead in time and for weekends the result was 84.44%-86% up to 14 time steps ahead in time (fig2),(fig3) . These results conform with those presented in Min and Wynter (2011) and are somewhat better than the results reported in Ding et al. (2010). However, this was expected as the authors in Ding et al. (2010) trained the model on weekday and weekend traffic alike which introduces more noise in the system than our approach.

5.2. Experiments on Urban Network

5.2.1. Urban Network Experimental set up

To test the model in urban mixed network scenarios a part of the Sydney Canterbury region (referred to as the urban network) constituting 8 suburbs with an area of approximately 5.2 Km×2.5km was recreated in Quadstone Paramics (fig4). The total length of the road segments in the urban network was 286km with 486 intersections and 52 traffic signals. Timing of traffic signals, roundabouts, turning priority and behaviour of intersections were collected from open street map (OpenStreetMap.org

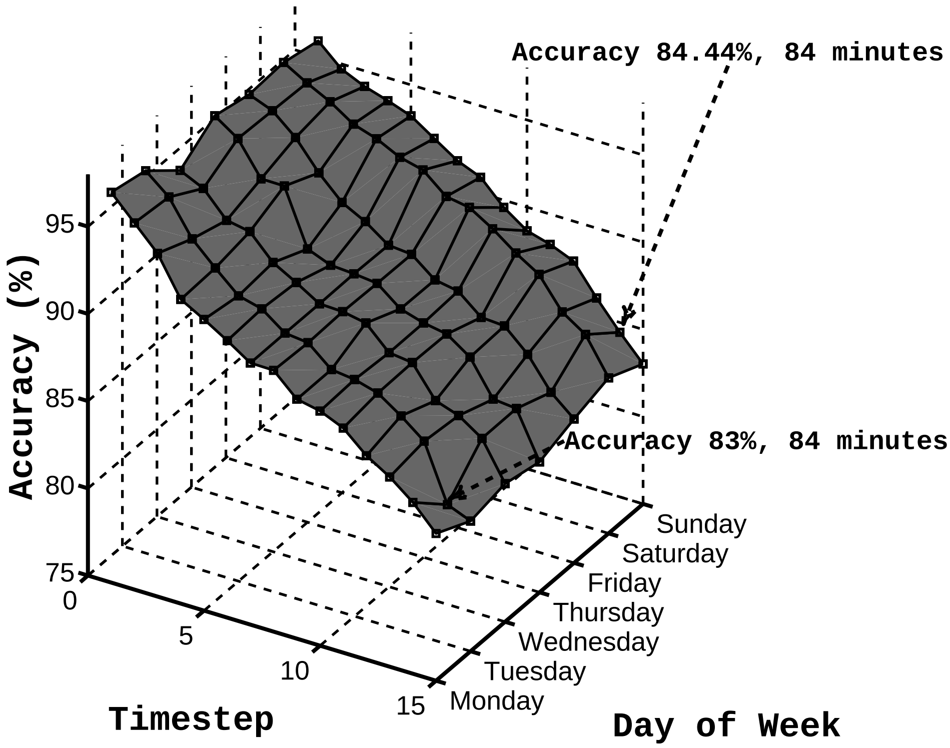

Figure 2: M4 motorway eastbound prediction

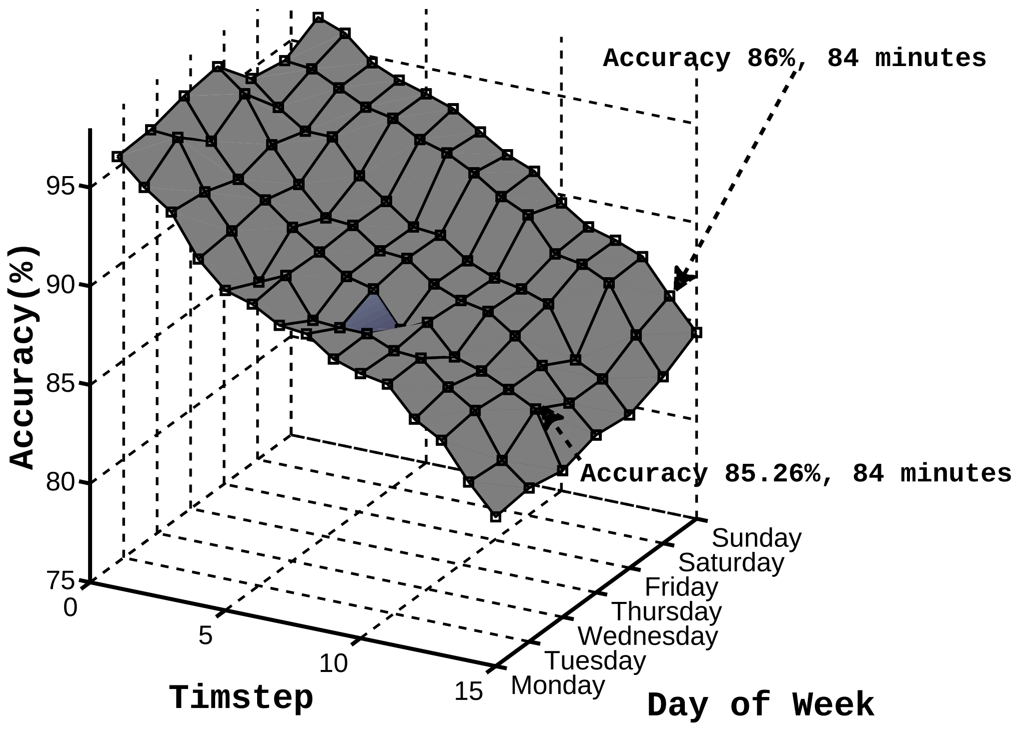

Figure 3: M4 motorway westbound prediction

(2010)). The network includes a total of 5 arterial links which constitute gateways to surrounding regions. Vehicles use these arterial links to reach other regions including the 8 suburbs within the urban network. Network include road segments varying from 1-3 lanes in each direction with a speed limit of 50km/h-70km/h. Various traffic control methods were included in the model, such as turning restriction, giveway rules, stop signs, roundabouts, one way streets etc; however pedestrians and pedestrian crossings were not included in the model. The network was divided into 29 zones comprising 8 zones for suburbs and 21 zones for road links in and out of the network. An initial demand matrix was constructed that generated a base traffic load of approximately 30% and later adjusted using the load shift parameter to vary the traffic intensity. A total of 572 loop detectors were placed to collect data (speed and volume) every 5 minutes. In Cheng et al. (2011), the authors studied the STACF and STPACF and found that a global STARIMA model would fail to accurately capture the movement of traffic between the road links, and they therefore proposed to partition the global road network into smaller regions. Our model differs from the one used in this work since we have defined the $\mathbf{W}$-matrix differently, but when testing for this property, we saw that the partitioning was necessary in our case also. Based on this finding, the road network was divided into a 8×24 grid, where each grid comprise four links (both directions) on average depending on link connectivity. 12 clusters were formed by combining 16 neighbouring grids with no overlapping road links and a unique STARIMA model was developed for each of these clusters. For each experiment scenario a different $\mathbf{W}$ matrix was then calculated with 1 time-step of 5 minutes for each model. Outputs of all the models were later combined for the entire network. Depending on the number of parameters and the offered load of a candidate model, 10-30 simulation runs of 3 hr (plus 1 hr for warming up) each were used. A candidate model was finalised after performing F and T testing.

5.2.2. Experiments

The main goal of the experiments on the urban network was to gain insight into the STARIMA model's capability to predict travel-time in an urban/mixed urban network.

In our first experiment the prediction performance was tested for various STARIMA models trained for specific LOS levels and the prediction performance for each time-step was compared with the real-time $\mathbf{A}_l^{timeobserved}$ matrix. The results show that the STARIMA model prediction accuracy started deviating (<80%) with increased load and further time-steps in future as expected. The model performed well (accuracy≥80%) for up to midpoint loads of LOS level C (loads up to 45%) and 13 time steps into the future. When the model was trained with higher LOS levels (loads above 50%) the accuracy degraded

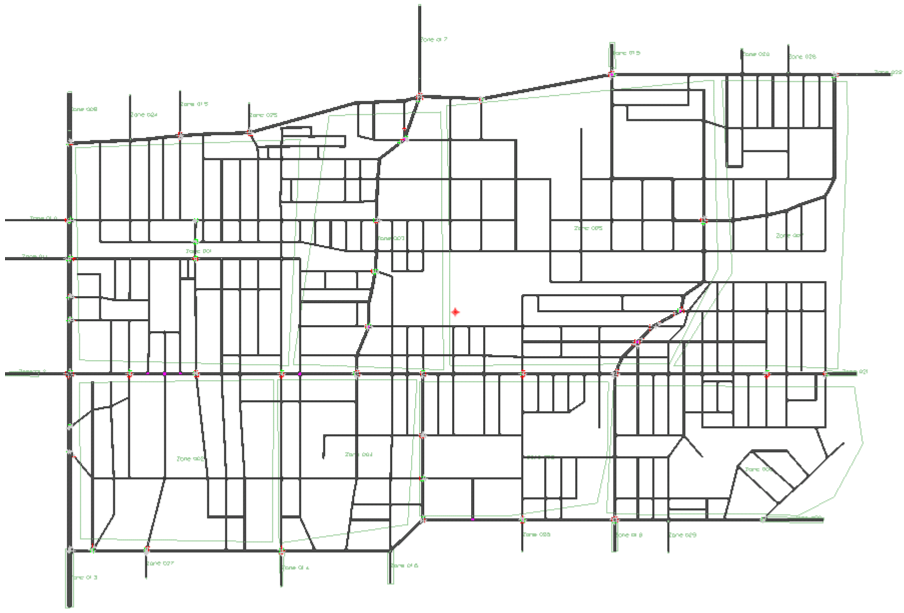

Figure 4: Urban network(Sydney, Canterbury)

significantly. In fig 5 the accuracy of prediction is plotted for loads 15%, 30%,45%,60%,70% and 80% to represent LOS A-F respectively. A cut off plane is plotted in the figure to illustrate when the time-steps prediction accuracy goes below 80%.

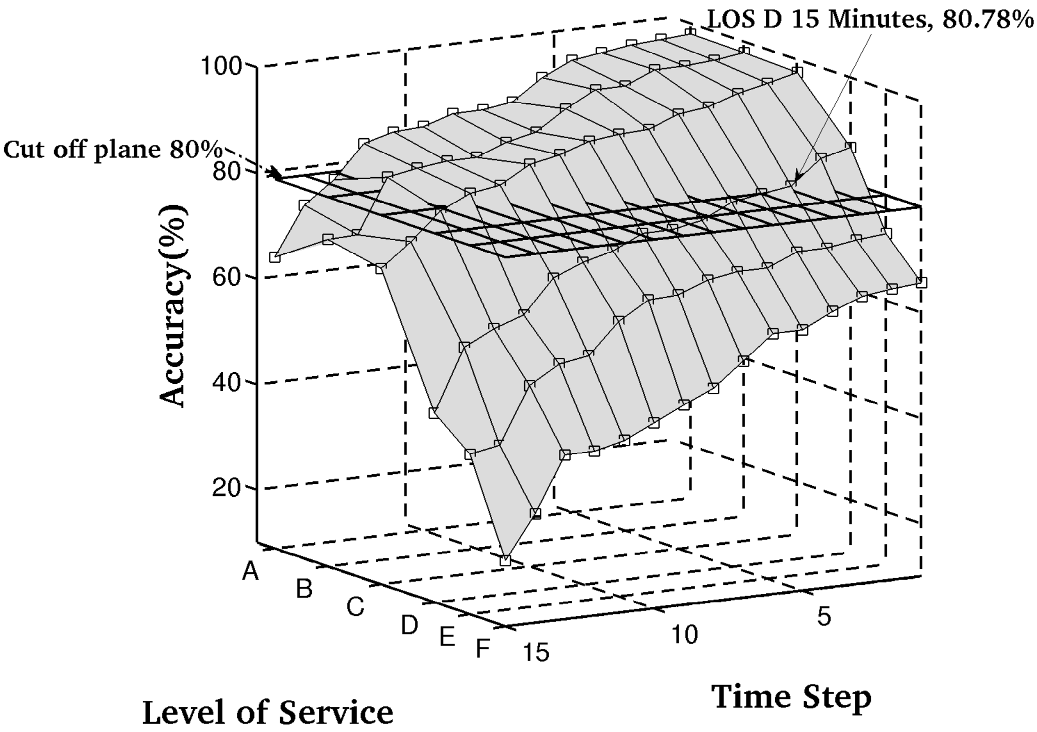

Figure 5: Urban network prediction

In the following experiment, a STARIMA model was trained and fitted for a particular load and was tested with additional load of 15% in 5% increments. The results determine under what level of load increase the model started to deviate significantly from the observed time series and the prediction performance was also measured. The results suggests that a STARIMA model can accurately predict travel time when the maximum increased load is 5% and when the total load in the network is less than 55%. When we tested a load increase of 10%, the prediction accuracy of the model became less than 80% after 8 time-steps and 6 time-steps when trained for 10% and 15% load respectively. For models trained for 25%-50% (5% increment), the prediction accuracy was around 80% up to 5-6 time-steps ahead in time for 5% increased load. Fig 6 and Fig 7 plots the travel-time prediction accuracy when additional load of 5-10% applied on models trained for loads 10% and 50% respectively.

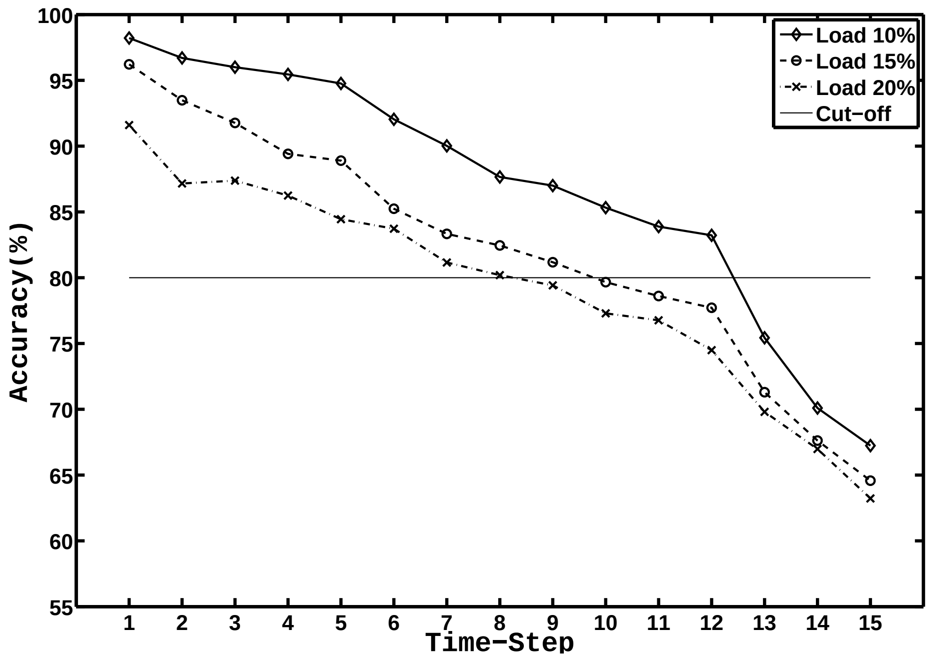

Figure 6: Load tolerance of STARIMA model, load 10%

5.3. Experimental Result Analysis

It is evident from the experimental results that travel time prediction using the STARIMA model varied greatly between the two different traffic networks. The model performed adequately in urban network as long as the traffic in the road network was at steady flow with limited load (LOS C, paramics load<54%). When the traffic built up to steady state but at high loads (LOS D-F), even after training the accuracy of the model degraded rapidly. We therefore see that for urban traffic settings, using **W** matrix categorisation based on LOS levels as proposed in Ding et al. (2010) may work up to a point

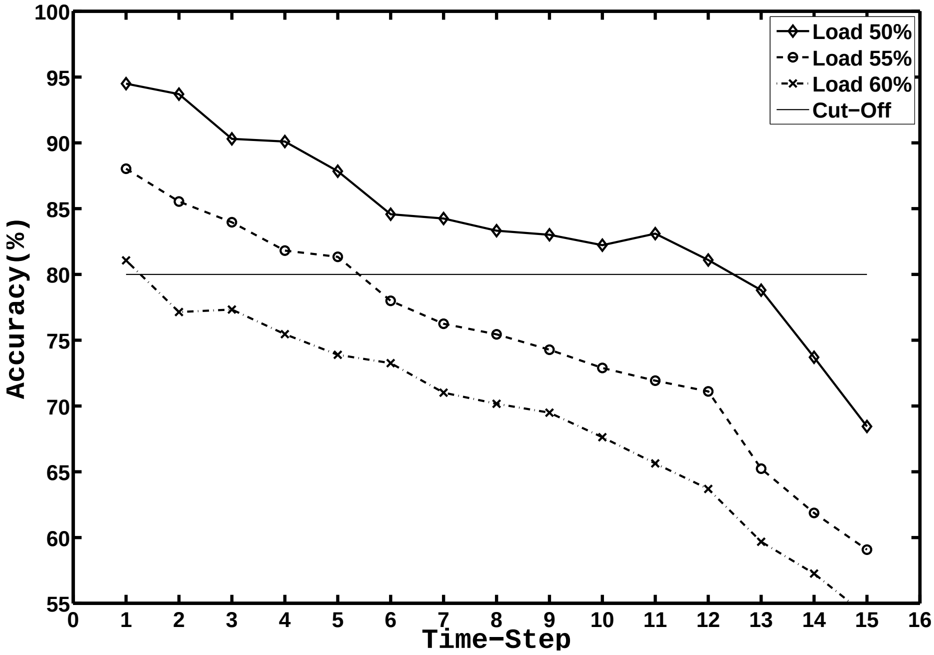

Figure 7: Load tolerance of STARIMA model, load 50%

but certainly not for traffic building up to congestion. The need for a revised or different model was evident whenever the traffic load was increased around 5% from the original training load. It can also be deduced from the results that the prediction horizon of a STARIMA model trained for a specific traffic load is reduced with increased load. The results from both our experiments on urban networks indicate a critical load limit(50%-55%) for the STARIMA model. For completeness, we illustrate the confidence in our results in fig 8, 9 and 10. The figures illustrate the confidence intervals for three LOS levels at 95% confidence. For all our presented results, the confidence interval varies between ±1.0% and ±4.2%

These results show that the STARIMA model is unsuitable for use in urban networks that reach a congestion level higher than LOS C or when there is a discrepancy between the load at which the model is trained and the actual load in the network. This indicates that the STARIMA model is not well suited to handle varying traffic loads as is the case in urban networks during peak hour - off peak cycles. On the other hand, this model performs well under free flow condition (motorway) and may be considered as a possible method for travel time prediction. For completeness, fig 11 and 12 show the travel time prediction for midweek traffic in both directions along the M4 Motorway plotted with 95% confidence intervals. For all our presented results, the confidence interval varies between ±1.0% and ±4.5%.

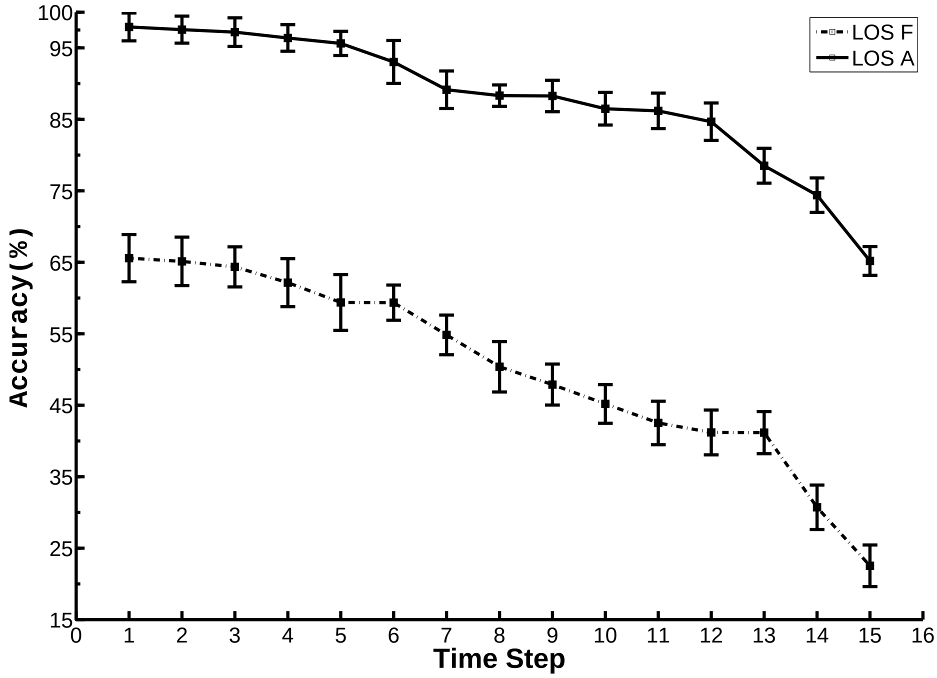

Figure 8: Confidence interval of prediction outcome LOS A and F

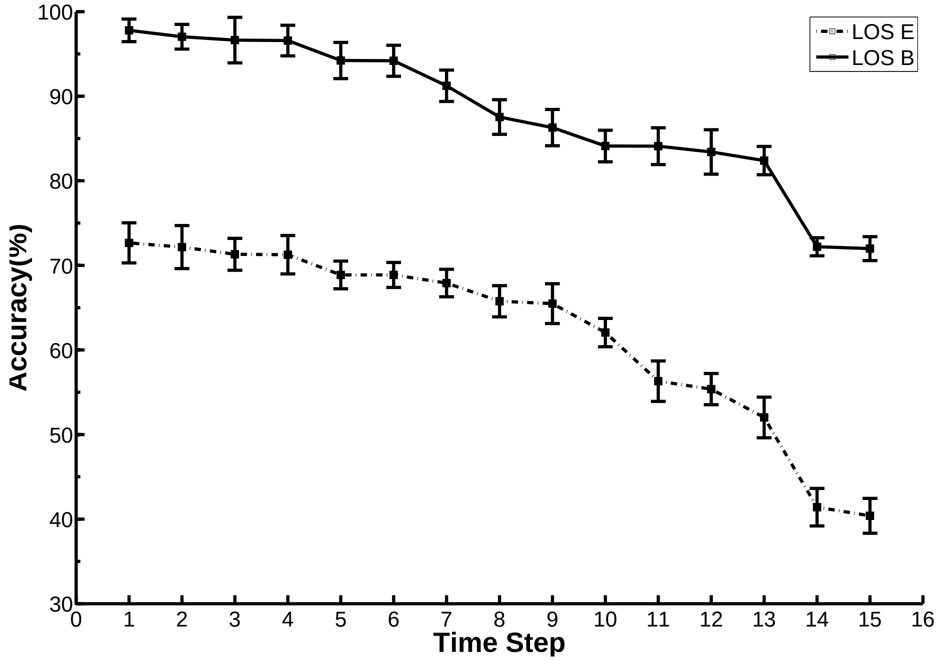

Figure 9: Confidence interval of prediction outcome LOS B and E

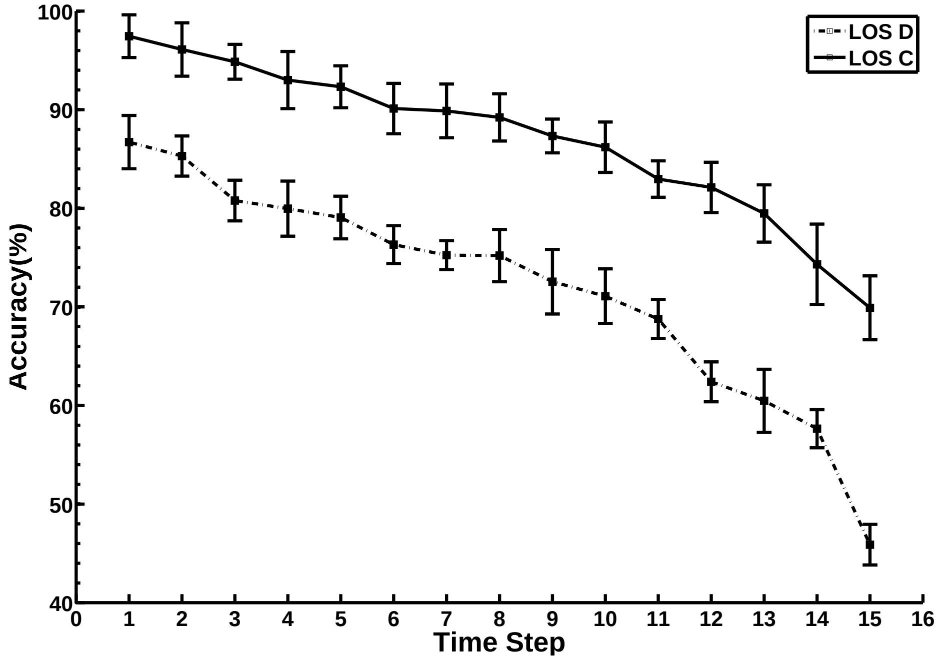

Figure 10: Confidence interval of prediction outcome LOS C and D

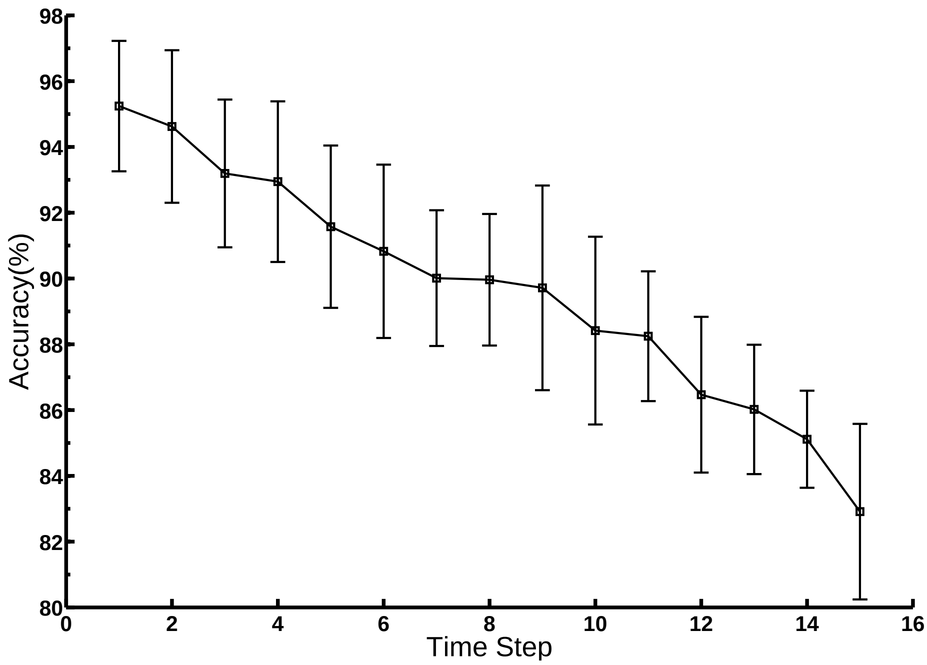

Figure 11: Confidence interval of prediction outcome using M4 eastbound Wednesday data

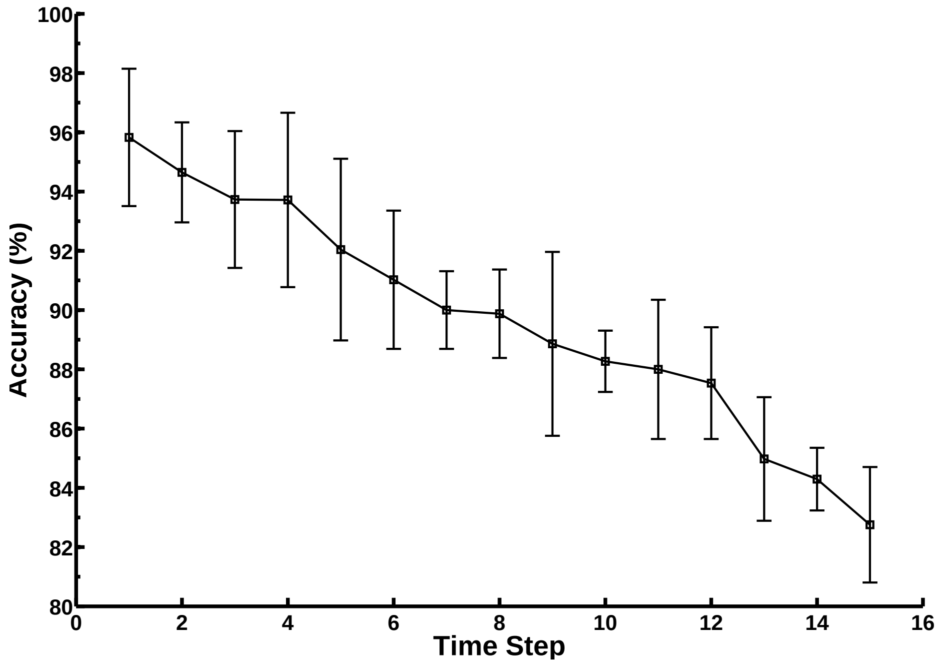

Figure 12: Confidence interval of prediction outcome using M4 westbound Friday data

6. Conclusion

In this work we explored how the STARIMA model performed when predicting travel times in two types of road traffic networks. Our study on Freeway traffic verifies previous findings that the STARIMA model is able to accurately predict travel times under steady state freeflow conditions, when the road network is of relatively low complexity. We further conducted an experimental study on a large-scale urban network and found that the prediction accuracy degrades significantly as the traffic load goes beyond LOS C and when the network moves out of steady state, as is the case during build up to peak hour traffic and slow down to off-peak traffic. Since the STARIMA model performs so well in the freeway scenario, it would be interesting to study how the model could be complemented and revised for better tolerance of traffic load variations. Building such a holistic model that can accurately predict both freeway and urban travel times would make a very important contribution to the ITS community. The nature of the STARIMA approach lends itself to control theoretic extensions; especially, a natural consideration is that of incorporated feedback loops into the model in order to enable accelerated convergence when the variability of traffic is large. We therefore intend to investigate this approach in future work.

Acknowledgement

NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program.

References

Box, G., Jenkins, G., 1994. Time Series Analysis: Forecasting and Control, 3rd Edition. Prentice Hall.

Cheng, T., Haworth, J., Wang, J., 2011. Spatio-temporal autocorrelation of road network data. Journal of Geographical Systems.

Danech-Pajouh, M., Aron, M., 1991. Athena: A method for short-term inter-urban motorway traffic forecasting. In: Recherche Transports Sécurité. Vol. 6. pp. 11–16.

Ding, Q., Wang, X., Zhang, X., Sun, Z., 2010. Forecasting traffic volume with space-time arima model. Advanced Materials Research 156-157, 979–983.

Hall, F., Persaud, B. N., 1989. Evaluation of Speed Estimates Made with Single Detector Data from Freeway Traffic management systems. Transportation Research Record.

kaggle.com, November 2010. Rta freeway travel time prediction competition. URL `http://www.kaggle.com/c/RTA/`

Kamarianakis, Y., Prastacos, P., Aug. 2002. Space-time modeling of traffic flow. Ersa conference papers, European Regional Science Association.

Lin, S., Huang, H., Zhu, D., Wang, T., 2009. The application of space-time arima model on traffic flow forecasting. In: 8th International Conference on Machine Learning and Cybernetics.

Liu, H., 2008. Travel time prediction for urban networks. Ph.D. thesis, Delft University of Technology.

Min, W., Wynter, L., 2011. Real-time road traffic prediction with spatio-temporal correlations. Transportation Research Part C: Emerging Technologies 19 (4), 606 – 616.

Min, X., Hu, J., Chen, Q., Zhang, T., Zhang, Y., oct. 2009. Short-term traffic flow forecasting of urban network based on dynamic starima model. In: Intelligent Transportation Systems, 2009. ITSC '09. 12th International IEEE Conference on. pp. 1 –6.

Min, X., Hu, J., Zhang, Z., 2010. Urban traffic network modeling and short-term traffic flow forecasting based on gstarima model. In: 13th International IEEE Conference on Intelligent Transportation Systems.

Nicholson, H., Swann, C., 1974. The prediction of traffic flow volumes based on spectral analysis. In: Transportation Research. Vol. 8. pp. 633–638.

Nookala, L., 2006. Weather impact on traffic conditions and travel time prediction. Master's thesis, University of Minnesota Duluth.

Okutani, I., Stephenades, Y., 1984. Dynamic prediction of traffic volume through kalman filtering. In: Transportation Research. Vol. 18(B). pp. 1–11.

OpenStreetMap.org, November 2010. Open street map.
URL `http://www.openstreetmap.org/`

Pfeifer, P., Deutsch, S., 1980. A three-stage iterative procedure for space-time modeling. In: Technometrics. Vol. 22. pp. 35–47.

Rad, M., 2007. Mining spatio-temporal data for travel time estimation in urban traffic networks. Master's thesis, University of Zagreb.

Regiolab-Delft, November 2010. The regiolab-delft project highway data website.
URL `http://www.regiolab-delft.nl/`

Rodrigue, J.-P., Comtois, C., Slack, B., 2009. The Geography of Transport Systems, 2nd Edition. Routledge.

Torday, A., Dumont, A., 2004. Parameters influencing probe vehicle based travel time estimation accuracy. In: The 4th Swiss Transport Research Conference.

TransportationResearchBoard, 2000. Highway capacity manual. Proceedings of The National Academy of Sciences, Washington D.C.

Zhou, J., Mitchell, K., 2005. Spatial autoregressive models for resource demand prediction in mobile wireless networks. In: International Teletraffic Congress(ITC) 19.

ISBN *978-1-74210-269-6*

School of Information Technologies
Faculty of Engineering & Information Technologies
Level 2, SIT Building, J12
The University of Sydney
NSW 2006 Australia

T +61 2 9351 3423
F +61 2 9351 3838
E sit.information@sydney.edu.au
sydney.edu.au/it

ABN 15 211 513 464
CRICOS 00026A