Short Term Traffic Volume Prediction for Sustainable Transportation in Urban Area

Haibo Mei, Athen Ma, Stefan Poslad, Thomas O. Oshin

Abstract- Accurate short term traffic volume prediction is essential for the realization of sustainable transportation as providing traffic information is widely known as an effective way to alleviate congestion. In practice, short term traffic predictions require a relatively low computation cost to perform calculations in a timely manner and should be tolerant to noise. Traffic measurements of variable quality also arise from sensor failures and missing data. There is no optimal prediction model so far fulfilling these challenges. This paper proposes a so-called Absorbing Markov Chain (AMC) model that utilizes historical traffic database in a single time series to carry out predictions. This model can predict the short term traffic volume of road links and determine the rate in which traffic eases once congestion has occurred. This paper uses two sets of measured traffic volume data collected from the city of Enschede, Netherlands, for the training and testing of the model respectively. The main advantages of the AMC model are its simplicity and low computational demand while maintaining accuracy. When compared with the established seasonal ARIMA and Neural Network models, the results show that the proposed model significantly outperforms these two established models.

Keywords- Absorbing Markov Chain, Fundamental Matrix, Short Term Traffic Prediction, Traffic Volume
Introduction

Congestion is a main problem found in urban areas that leads to numerous negative impacts on both the environment and society, such as pollution, elongated travel time and excess costs (Whittaker et al. 1997, Nicholoson 1974). A key to sustainable transportation is to seek effective ways in minimizing congestion so that the reliability of the transport system can be improved. Providing road users with short term traffic prediction has been proven to be an effective way to achieve this, as users would avoid congested or problematic areas so as to save time and/or money and reduce the level of pollution.

The objective of short term traffic prediction is to forecast the state of a traffic area, such as volumes, times required to travel, density, in some time interval in the future, by referring to historical and current data (Thomas et al. 2010, Rice et al. 2001, Sun et al. 2006). Common time intervals used are usually in the range of 5 minutes to half an hour (Sun et al. 2003). Different traffic prediction models have been proposed and each of them presents its own merits and drawbacks with respect to data accuracy, complexity and data quality. There is not a universal agreement on an optimal prediction model (Bolshinsky 2012) as existing methods have limitations in at least one of these areas. For example, Naïve models present a simplistic prediction approach. While Naïve models often require a low computational cost and are somewhat resilient to incomplete data, they suffer from low accuracy. Examples of Naïve models include clustering (Thomas et al. 2008, Van Grol et al. 2000, Weijermars 2005, Chrobok et al. 2004, Chung 2003) and historical average (Rakha 1995, Guo et al. 2010). A non-parametric model varies both the type and the number of parameters according to the data of interest. While parameters used in this model are not complex, defining these parameters does rely on analyzing a substantial amount of data together with an appropriate data mining strategy (Van Hinsbergen et al. 2007). Nearest neighbor regression (Smith et al. 2002), Neural network model (Yin et al. 2002, Dharia 2003) and Bayesian Network Model (Sun et al. 2006) are the well know types of non-parametric models. A parametric model, on
the other hand, predetermines the parameters and structure of the model in which a much smaller data set is required for the training process. The computational cost and complexity associated with this approach are also a lot less. However, a parametric model costs effort to calculate the complex parameters in prediction equations using historical and current traffic measurements, and the parameter training is vulnerable to incomplete data. Examples of parametric models include Autoregressive Integrated Moving Average (ARIMA) and seasonal ARIMA (Williams 2003, Kumar 1999), Support Vector Regression (SVR) (Wu et al. 2004), Kalman filter (Xie et al. 2007, Okutani 1984, Whittaker et al. 1997), linear regression model (Rice 2001), Gaussian Maximum Likelihood model (Lin 2001) and Markov Chain model (Yu et al. 2003, Hu et al. 2003).

Previous studies have provided evidence that traffic data exhibits stochastic properties and shown that the current state of the traffic is highly correlated to its previous states, where a traffic state is used to describe the different volumes of traffic (Thomas et al. 2010, Yu et al. 2003, Rakha 1995). Based on those studies, probabilistic models have been used to perform traffic prediction, and the correlation between the current and immediately preceding states is approximated using a suitable probability distribution (Yu et al. 2003, Hu et al 2003). Yu et al. (2003) and Hu et al (2003) present a Markov Chain model in the parametric category to carry out short term traffic predictions building on those probabilistic models, and in their method the Gaussian Mixture Models and Expectation Maximum algorithms are utilized to calculate the state transition probabilities of a standard Markov Chain. The method is accurate, but suffers greatly from high complexity, making it difficult to use. This paper presents an Absorbing Markov Chain (AMC) model that is accurate, relative simple without compromising on computational demand. Similar to existing Markov Chain models (Yu et al. 2003, Hu et al 2003), the AMC model also employs the probabilistic models and belongs to the parametric category of such models. In this paper, the predictions are based upon data
from a number of fixed locations (on road links) in an urban area. The time intervals in the predictions are 5 minutes each.

The AMC (Grinstead et al. 1997, Kemeny et al. 1976) is one type of Markov Chains that has at least one absorbing state. A state is considered absorbing if a transition is from another state (a non-absorbing i.e. transient state), it then cannot leave that absorbing state once transiting to it (Kemeny et al. 1976). The AMC is utilized here to model a case of congested traffic whereas the traffic volume of a given road link has reached the maximum at a given time and remained congested for a period of time. Based on this assumption, we use the AMC model and historical data to predict the short term traffic volume of the road link, and the rate the traffic eases once congestion has occurred. The proposed AMC model follows the time series theory (Thomas et al. 2010) that involves modeling a traffic volume as a function of its past observation and using an error term. The theory requires the process to be stationary (Van Hinsbergen. C.P.I et al. 2007) i.e. have recurrent variations (Thomas.T et al. 2008). Here we model stationary traffic of a road link by clustering historical traffic data with respect to different days and times of a day (Thomas et al. 2008, Van Grol et al. 2000, Weijermars 2005, Chrobok et al. 2004, Chung 2003).

In general, the key contributions of this work include:

(1) Exploiting an AMC model to predict the short term traffic volume of road links.

(2) Predicting the rate in which traffic eases once congestion has occurred.

(3) Validating the accuracy of AMC model through using real traffic traces collected from the city of Enschede, Netherlands, and by comparing the AMC model with two established methods: seasonal ARIMA and Neural Network models.

**Survey of Short-Term Traffic Prediction**

There are numbers of existing short term prediction models. We can categorize them into three main categories: naïve, non-parametric or parametric.
A naïve model works on the basis of historical averaging (Rakha 1995, Guo et al. 2010) and clustering (Thomas et al. 2008, Van Grol et al. 2000, Weijermars 2005, Chrobok et al. 2004, Chung 2003) on the most recent periods and makes prediction for the next period. This method is very simple and often used to provide a benchmark. Clustering and historical average (Rakha 1995, Guo et al. 2010) are examples of a naïve model, but both approaches lack of accuracy. For example, Williams (2003) examined two freeway locations, one in London, England and other in Georgia, USA, the historical average approach provides predictions on the traffic with mean absolute percentage errors of 11.43% and 12.85% respectively, compared to 8.74% and 8.97% given by a seasonal ARIMA. The main merit of this type of model is its simplicity but they are often inaccurate and susceptible to data quality.

Nearest neighbor regression is a typical non-parametric prediction model (Smith et al. 2002) which uses pattern matching (Sun et al. 2003, Bajwa et al. 2004) to find historical traffic situations that are identical to the current traffic situation so as to predict the traffic. Nearest neighbor regression is simple but requires a high computation cost and large amount of data. Good accuracy can generally be obtained in cases whereas traffic flow does not fluctuate greatly. However, if traffic flow fluctuates abruptly, prediction accuracy will be low. Another well-known non-parametric model is an approach using neural networks (Yin et al. 2002, Dharia 2003). A neural network prediction model uses adaptive learning to analyze historical traffic data so as to identify patterns in the data and predict traffic. This type of model uses complex artificial neuron network structures, together with its adaptive learning, results in high computational costs and long training times. A neural network model can be combined with other methods, such as fuzzy clustering (Yin et al. 2002), to reduce the overall training time. However, this does introduce further complexity. A Bayesian Network Model is another non-parametric model that has been proposed to utilize the traffic measurements from adjacent road links to predict a current road link under investigation (Sun et al. 2006). Both the neural and Bayesian network models can be used to carry out spatial traffic
predictions. The non-parametric prediction approach provides improvement on accuracy when compared to the naïve model; however, the accuracy depends greatly on the data quality and the approach is rather complex and computationally demanding.

The parametric category includes Autoregressive Integrated Moving Average (ARIMA) and seasonal ARIMA (S-ARIMA) which are based on linear combinations of traffic measurements from previous time intervals (Williams 2003, Kumar 1999). Both are time series methods and provide accurate predictions. The ARIMA models require a long processing time and are susceptible to incomplete data (Smith et al. 2002). A data filling technique can be used to address the latter, but this can be tricky when the traffic is complex and highly variable (Bolshinsky 2012). Similar to ARIMA that uses statistical method on time series of historical and current traffic variables, Support Vector Regression (SVR) (Wu et al. 2004) is a new and useful parametric model but has the disadvantage of a high computation cost. Kalman filter model is another parametric model (Xie et al. 2007, Okutani 1984, Whittaker et al. 1997) which can be applied to both stationary and non-stationary data analysis. It is computationally efficient, requires only very little storage, and is suitable for real-time traffic prediction. Another desirable feature of the Kalman filter model is that it can dynamically update its state variable using real-time data and better adapt to changes in traffic flows (Xie et al. 2007, Whittaker et al. 1997). However, a Kalman filter is not resilient to incomplete data. This hinders its application in practice. To simplify parametric predictions, linear regression model (Rice 2001), Gaussian Maximum Likelihood model (Lin 2001) and Markov Chain models (Yu et al. 2003, Hu et al. 2003) have often been used. These models are usually relatively complex. While they can generally provide accurate predictions, their application is limited to stationary traffic cases due to its susceptibility to data quality.

In summary, accuracy, simplicity, computational demand and susceptibility to data quality are the key factors for short-term traffic prediction. Existing approaches lack at least one of these criteria. In this paper, the proposed AMC model aims to provide an approach that
fulfills most of these criteria by providing a simple method that is accurate, computationally efficient.

**Practical applications**

Transportation information systems can be used to promote sustainable urban transportation by offering short term traffic prediction to users. For example, the EU project SUNSET (Sustainable Social Network Services for Transport) seeks to develop a system called TRIPZOOM (Holleis et al. 2012), accessed via smart phones and other web-enabled devices, to aid people to achieve more sustainable transportation choices based upon travel route, travel start time, or travel modality. The main objectives of SUNSET are congestion reduction, improved safety, greater environment protection, and greater personal wellbeing of citizens. For example, a key target for TRIPZOOM is to reduce peak time traffic in an urban area by 5%. Accurate traffic predictions can be distributed to users so that they are aware of congestion ahead and there is an incentive for them to choose an alternative route to avoid the congestion.

Short term transportation prediction can be applied to many transportation management systems such as TRIPZOOM. In these systems, the predictions can be on-line or off-line (Van Grol et al. 2000). On-line predictions are realized by models that are simple and efficient but only use the most recent traffic measurements. On-line predictions better reflect the traffic situations, but may require a higher traffic measurement quality and a lower computing cost. In contrast, off-line predictions can have a higher prediction accuracy by employing more complicated models and using a larger historical data set. However, off-line predictions are less useful if the traffic situation changes drastically. This paper proposes an AMC model as a suitable candidate model applicable for use in on-line applications through employing a simpler approach and more computational efficient algorithms.
Traffic Data

Traffic data collected are acquired from about urban road links controlled by traffic lights where the volume of traffic is represented as the count of the number of vehicles passing the traffic light every 5 minute. Traffic volumes, during the morning weekday rush hour, are used an illustrative example, see Figure 1.

Real road traffic traces were collected from the city of Enschede, Netherlands. The city of Enschede is a small city which has around 160,000 inhabitants. Inductive loop detectors are used to count the number of vehicles in all directions of a given road link in every 5 minutes. This particular length of time interval is said to be a good compromise between the amount of traffic noise and accuracy (Xie et al. 2007). The data has been screened to ensure that invalid data, e.g. errors in measurements caused by failures in the traffic sensor equipment, are identified and removed. Therefore, the training data excludes error data. This also excludes “artificial” data and other manipulated measurements that can jeopardize the prediction results (Thomas et al. 2008, Thomas et al. 2010).

The processed data are organized into traffic volume measurements per road link per measuring interval, hence the training data are represented in Eq.(1).

\[ q_{dit}^{obs} (\forall d \in D, \forall t \in T, \forall l \in S) \]  

where \( q_{dit}^{obs} \) is defined as the traffic volume measurement of a given day \( d \) and a given road link \( l \). \( t \) denote a specific 5-minute interval. \( T \) is defined as a count of sequences of 5 minutes intervals in the study, specifies the period of time such weekdays or weekends. \( S \) consists of the group of road links.

For weekdays, the training data consists of the measurements obtained between 6:00AM-12:00AM from every weekdays between 21st of June 2010 and 21st of November 2010, giving \( |T| = (12 - 6) \times 5m + 1 = 73 \) and \( |D| = 150 \). Thus, there are \( 150 \times 73 = 10950 \) traffic measurements for each road link in \( S \). It is worth noting that Bank/Public holidays that fall on weekdays have also been included in the data. For weekends, there are

\[ \]
60 × 73 = 438 traffic measurements between 21st of June 2010 and 21st of November 2010.

In this paper, we consider weekdays and weekends separately as they present very different historical patterns. Patterns observed in the latter are also less recurrent.

Figure 1. Traffic pattern of a road link measured in the morning (weekdays) rush hour between 6:00am and noon at intervals of 5 minutes

Methodology

In urban areas, the traffic volume of a road link fluctuates with the time of day; however, it can generally be assumed that recurrent traffic patterns can be observed at specific times and on specific days. Figure 1 shows traffic recorded in the morning of 4 different weekdays between 6:00am to noon at an interval of 5 minutes - they show a similar trend particular when the traffic peaks. Figure 1 also shows the historical average of the traffic of all weekdays over 6 months. It shows that there are definitive periods of time whereas traffic volume peaks (i.e. morning or evening rush hour) and that once congestion occurs, the road will remain congested for a period of time.

A. Traffic volume prediction

Traffic volume varies depending on factors such as holidays, weather conditions, season, planned and emergency road works. According to these factors, we can classify the variations in traffic volume into systematic variations and noise. Systematic variations are caused by changes in people’s travel patterns due to seasonal changes, public holidays and general work schedule changes. Systematic variations are related to long-term correlations whereas a reoccurrence pattern is observed over time, such as heavy traffic found in the morning rush hour on weekdays. In addition, noise is introduced to the traffic volume because of unexpected events such as road accidents, temporary roadworks or changes in the weather (Thomas et al. 2008). Noise is associated with very short term correlations, reflecting
localized fluctuations in the traffic pattern. For instance, if a car has broken down and causes congestion at 10am on a given day, it is likely that the road is still congested 10 minutes later. The relationship between traffic prediction and real measurement for a road link is formulated in Eq.(2).

\[
q_{dit}^{obs} - q_{dit}^{pred} = \varepsilon_{dit} + \theta_{dit}
\]  

(2)

where, \(q_{dit}^{obs}\) and \(q_{dit}^{pred}\) are the measured and predicted traffic volume respectively, \(l\) represents the road link of interest, \(t\) is the time interval and \(d\) is the day, \(\varepsilon_{dit}\) is the systematic variation and \(\theta_{dit}\) is the amount of noise. To obtain an accurate prediction, the systematic variation and noise must be minimized.

B. Absorbing Markov Chain

A Markov Chain is a discrete random process in which a number of transient states are defined to describe the state-space of a given process. The current state of a Markov chain depends only on the current state and the transitions between two consecutive states are defined by a transition matrix as shown in Figure 2(a). An AMC introduces the notion of an absorbing state, and once this state has been entered, it will remain at that state (Kemeny et al. 1976), denoted as \(TVS_c\) in Figure 2(b).

Figure 2. State transition of (a) a Markov Chain and (b) an Absorbing Markov Chain

In the context of traffic prediction, the traffic volume of a given road link is divided into \(N\) Traffic Volume States (TVSs), see Figure 3. An absorbing state is defined to represent the traffic congestion. It is assumed that congestion occurs when the traffic volume is over a pre-defined (congestion) threshold and small fluctuations in the traffic volume during congestion are ignored. This means that the entire period of congestion is represented by the congestion absorbing state. In addition, a vacuum absorbing state is also defined to denote a clear road which occurs when the traffic volume is less than a pre-defined (clear) threshold.
Our proposed model divides the rush hour traffic volume into two halves and describes each half by an AMC see Figure 3(b). The first AMC addresses the period prior to the peak traffic. Traffic volume rises during this period and the TVS moves from one state to the next and eventually enters the congestion absorbing state prior to the time the traffic peaks. Once the model has reached this state, it will remain there until the traffic has reached its peak and then start another AMC. By examining the sojourn time in which the model spends in the congestion absorbing state, it is then feasible to estimate how long congestion will last for a given road link. Similarly, as the traffic eases, the TVS model drops from one state to the next and is gradually absorbed into a vacuum absorbing state. Once absorbed, the TVS will not transition to any other states till noon.

Figure 3. AMC absorbing state models traffic congestion absorbing state and the vacuum absorbing state (a): overall traffic situation; (b): split into two AMCs

As shown in Figure 2 and 3, in one AMC, all 8 TVSs were defined for a given road link which consists of 7 transient TVSs and 1 absorbing TVS: $TVS_{c1}$ (i.e. $TVS_{con}$ or $TVS_{vac}$ in Figure 3). The resulting transition matrix $M$, represented by Eq.(3), is $8 \times 8$ and the transition probability from one state to the next is denoted by $M_{i,j}$, where $i$ and $j$ are $TVS_i$ to $TVS_j$ respectively. This matrix is derived from the overall statistics of the historical data. The basic principle of a Markov Chain is summarized in Theorem 1 (Grinstead et al. 1997).

$$
\begin{pmatrix}
M_{1,1} & M_{1,2} & M_{1,3} & M_{1,4} & M_{1,5} & M_{1,6} & M_{1,7} & M_{1,c}
M_{2,1} & M_{2,2} & M_{2,3} & M_{2,4} & M_{2,5} & M_{2,6} & M_{2,7} & M_{2,c}
M_{3,1} & M_{3,2} & M_{3,3} & M_{3,4} & M_{3,5} & M_{3,6} & M_{3,7} & M_{3,c}
M_{4,1} & M_{4,2} & M_{4,3} & M_{4,4} & M_{4,5} & M_{4,6} & M_{4,7} & M_{4,c}
M_{5,1} & M_{5,2} & M_{5,3} & M_{5,4} & M_{5,5} & M_{5,6} & M_{5,7} & M_{5,c}
M_{6,1} & M_{6,2} & M_{6,3} & M_{6,4} & M_{6,5} & M_{6,6} & M_{6,7} & M_{6,c}
M_{7,1} & M_{7,2} & M_{7,3} & M_{7,4} & M_{7,5} & M_{7,6} & M_{7,7} & M_{7,c}
M_{c,1} & M_{c,2} & M_{c,3} & M_{c,4} & M_{c,5} & M_{c,6} & M_{c,7} & M_{c,c}
\end{pmatrix}
$$

(3)

**Theorem 1**: Let $M$ be the transition matrix of a Markov Chain, and let $u$ be the probability
vector which represents the starting distribution. Then the probability that the chain is in state $s_k$ after $n$ steps in the $k^{th}$ entry in the vector: $u^{(n)} = u \times M^n$.

Building on Theorem 1, we can predict the short term traffic volume. Specifically, the transition matrix derived from the historical traffic data of a given road link is used to calculate the state of the road link at each time interval by using Theorem 1, and the results obtained will be used to adjust the historical average of traffic volume at each interval to produce a baseline prediction. In addition, systematic variations and noise found in traffic can be addressed by taking recently measured traffic into consideration prior to prediction. This provides information with a finer granularity than historical averages, minimizing $q_{dti}^{obs} - q_{dti}^{pred}$ in Eq.(2).

The notion of an AMC is further exploited in the paper to estimate the time it takes for congestion to clear. An AMC is described by a fundamental matrix defined in definition 1:

**Definition.1:** For an Absorbing Markov Chain, the matrix $F = (I - Q)^{-1}$ is called the fundamental matrix for the Absorbing Markov Chain. The entry $f_{ij}$ of $F$ gives the expected number of times that the process is in the transient state $s_j$ when it is started in the transient state $s_i$ before reaching an absorbing state.

The actual prediction firstly depends on the likelihood the process of interest enters the absorbing state, which is stated in theorem 2.

**Theorem.2:** Let $b_{ij}$ be the probability that an absorbing chain will be absorbed in the absorbing state $s_j$ if it starts in the transient state $s_i$. Let $B$ be the matrix with entries $b_{ij}$. Then $B$ is a $t$-by-$r$ matrix, and $B = F \times R$; where $t$ is the counted number of transient states and $r$ is the number of absorbing states, $F$ is the fundamental matrix, $R$ is as in the canonical form.

$R$ is defined in the canonical form of the AMC transition matrix $M$ as shown in Eq.(4). In Eq.(4), $R$ is a nonzero $N$-by-1 matrix, and $N$ is the number of transient TVSs and 1 represents...
the only absorbing TVS in an AMC. $I$ is a 1-by-1 identity matrix, $0$ is a $N$-by-1 zero matrix and $Q$ is a $N$-by-$N$ matrix.

$$
\begin{bmatrix}
M_{1,1} & M_{1,2} & M_{1,3} & M_{1,4} & M_{1,5} & M_{1,6} & M_{1,7} \\
M_{2,1} & M_{2,2} & M_{2,3} & M_{2,4} & M_{2,5} & M_{2,6} & M_{2,7} \\
M_{3,1} & M_{3,2} & M_{3,3} & M_{3,4} & M_{3,5} & M_{3,6} & M_{3,7} \\
M_{4,1} & M_{4,2} & M_{4,3} & M_{4,4} & M_{4,5} & M_{4,6} & M_{4,7} \\
M_{5,1} & M_{5,2} & M_{5,3} & M_{5,4} & M_{5,5} & M_{5,6} & M_{5,7} \\
M_{6,1} & M_{6,2} & M_{6,3} & M_{6,4} & M_{6,5} & M_{6,6} & M_{6,7} \\
M_{7,1} & M_{7,2} & M_{7,3} & M_{7,4} & M_{7,5} & M_{7,6} & M_{7,7} \\
\end{bmatrix}
$$

We calculate the fundamental matrix $F$ and $B$ of the two AMCs in our proposed model for a given road link. An entry in $B$ represents the probability that the road link will finally come to an absorbing TVS (congestion absorbing state or vacuum absorbing state) whilst currently in a transient TVS. For example, $B_{2,1}$ is the probability that the road link will come to an absorbing TVS when it is currently in $TVS_2$. On the other hand, an entry in $F$ represents the number of times a transient TVS will transition to another transient TVS before absorbing TVS. For example, $F_{2,3}$ represents the number of times $TVS_2$ will transition to $TVS_3$ before absorbing TVS. Based on $F$, Theorem 3 can be used to estimate the total number of transient TVSs any TVS will transit to before absorbing TVS (Grinstead et al. 1997).

*Theorem 3* Let $t_i$ be the expected number of steps before the chain is absorbed, given that the chain starts in state $s_i$, and let $t$ be the column vector whose $i^{th}$ entry is $t_i$. Then

$$
t = F \times c ;
$$

where $c$ is a column vector all of whose entries are 1. $F$ is the fundamental matrix.

According to Theorem 3, $F \times c = \sum_{j \in \{1,3,4,5,6,7\}} F_{2,j}$ is the total number of transient TVSs that $TVS_2$ will transit into before going into an absorbing TVS (assuming there are 7 transient TVSs).
In summary, we can predict the short term traffic volume by using Theorem 1, and predict the rate at which traffic clears when congested, using Definition 1, Theorem 2 and Theorem 3.

The proposed AMC model

This section first describes the calculation of the transition matrixes and how they are exploited to carry out short term traffic prediction by addressing the systematic variations and noise found in the system using different strategies. Furthermore, the way the AMC model is used to predict the rate congestion clears is also described. An overview of the proposed model can be seen in Figure 4.

Figure 4. An overview of the proposed AMC model

A. Calculation of the transition matrix

We calculate AMC transition matrix of a road link of interest (say road link \( O \)) using the training data. As discussed previously, there are two AMCs for road link \( O \), leading to two transition matrices: \( M^1 \) before the traffic peak and \( M^2 \) after the peak. In \( M^1 \), \( TVS_i^1 (\forall i \in \{1..N, con\}) \) are the TVSs, and similarly, \( TVS_i^2 (\forall i \in \{1..N, vac\}) \) are the TVSs in \( M^2 \). \( N \) is the number of transient TVSs in each AMC. \( V^1_i (\forall i \in \{1..N, con\}) \) and \( V^2_i (\forall i \in \{1..N, vac\}) \) are the values to define the range of \( TVS_i^1 (\forall i \in \{1..N, con\}) \) and \( TVS_i^2 (\forall i \in \{1..N, vac\}) \). They are calculated using Eq.(5). The training data used are: \( q_{d,t}^{obs} (\forall d \in D, \forall t \in T) \).

\[
\begin{align*}
A_t &= \frac{\sum_{(d,t)D} q_{d,t}^{obs}}{|D|}; \quad (\forall t \in T) \\
T_{max} &= \arg_t (\max_{\forall t \in T} (A_t)) \\
V^1_{con} &= \max_{\forall t \in T} (A_t) \\
V^2_{vac} &= \frac{\sum_{(d,t)T} A_t}{|T|} \\
V^1_j &= \frac{V^1_{con}}{N} \times (j - 1) \quad (\forall j \in \{1..N\}) \\
V^2_j &= \frac{V^1_{con} - V^2_{vac}}{N} \times (j - 1) \quad (\forall j \in \{1..N\})
\end{align*}
\]
\( V_i^1 (\forall i \in \{1..N, con\}) \), \( V_i^2 (\forall i \in \{1..N, vac\}) \), \( TVS_i^1 (\forall i \in \{1..N, con\}) \) and \( TVS_i^2 (\forall i \in \{1..N, vac\}) \) can be specified in a simple way. We first consider \( M^1 \). If a road link \( O \) has a traffic volume \( q \) that \( V_k^1 \leq q < V_{k+1}^1 \) \((1 \leq k \leq (N - 1))\), then the traffic volume of road link \( O \) is in \( TVS_k^1 \). If \( V_N^1 \leq q < V_{con} \), road link \( O \) is in \( TVS_N^1 \), and if \( V_{con}^1 \leq q \), road link \( O \) is in \( TVS_{con}^1 \). As for \( M^2 \), if a road link \( O \) has a traffic volume \( q \) such that \( V_{k+1}^2 \leq q \leq V_k^2 \) \((1 \leq k \leq (N - 1))\), then the traffic volume of road link \( O \) is in \( TVS_k^2 \). If \( V_{vac}^2 \leq q \), road link \( O \) is in \( TVS_N^2 \), and if \( q \leq V_{vac}^2 \), road link \( O \) is in \( TVS_{vac}^2 \). With reference to Eq. (5), \( V_{vac}^2 \) is the total average of the training data and \( V_{con}^1 \) is the highest average traffic volume. Therefore, we define road link \( O \) being in a vacuum absorbing state when its traffic volume is lower than \( V_{vac}^2 \) and being in a congested absorbing state when its traffic volume are higher than \( V_{con}^1 \). \( T_{max}^1 \) is defined as the peak time that separates two AMCs for a specific road link, i.e., \( O \).

Finally, with \( TVS_i^1 (\forall i \in \{1..N, con\}) \) and \( TVS_i^2 (\forall i \in \{1..N, vac\}) \), we can calculate the probability of each TVS transiting to other TVSs using simple training data statistics (Grinstead 1997, Kemeny 1976). This is a simpler way than the transition matrix generation in (Yu.G 2003 and Hu.J 2003). This method uses training data to calculate how many times a road link is in a TVS and the number of times the road link transits from one TVS to another TVS in sequential intervals. Afterwards, the probability of the former TVS transiting to the latter TVS is the ratio of these two calculations. After processing all the training data, \( M^1 \) and \( M^2 \) can be obtained. One thing important is that transition matrices can be updated when new traffic data is read in. This enables AMC transition Matrix to meet the time varying nature of traffic condition.

B. Short term prediction-AMC(short term)

As shown in Figure 4, short term prediction in AMC model includes 3 intermediate steps: baseline prediction, 24-hour prediction and finally short term prediction.

The baseline prediction uses the average of the historical data as the training data. Historical average never comes out best compared to more advanced models, although sometimes it can
outperform some prediction techniques on longer horizons (Van Hinsbergen et al. 2007). We use TVS transition probabilities in $M^1$ and $M^2$ to augment the historical average of traffic volume at each time interval: $A_{Ot}(\forall t \in T)$ in Eq.(5) to generate the baseline prediction for road link $O$.

Table 1 Algorithm to Generate the Baseline Prediction of road link $O$

The algorithm for generating the baseline prediction is shown in Table 1. The algorithm uses Theorem 1 and $A_{Ot}(\forall t \in T)$ to predict which TVS (say $S_{kMC}^O$) road link $O$ will on average be in at any time interval (say $k$). In the prediction, $u$ is initialized as a $(N+1)$ size probability vector representing which TVS road link $O$ is in, at a previous time interval $(k - 1)$ (step 9 or step 13). In $u$, the entry to the TVS that road link $O$ is in, at interval $(k - 1)$, is set to 1; while all other entries are set to 0. Afterwards, $S_{kMC}^O$ is calculated out by Theorem 1 as the most probable TVS that road link $O$ will be in at time interval $k$ (step 10 or step 14). After the prediction, $S_{kMC}^O$ is compared to TVS: $S_{kbase}^O$ to decide whether $A_{Ok}$ needs a correction or not (Step 11 or 15). $S_{kbase}^O$ is defined by $A_{Ok}$ in step 3 or step 5. In comparison, if ($S_{kMC}^O <> S_{kbase}^O$), $A_{Ok}$ is corrected to be $B_{Ok}$. The correction uses the TVS discrepancies and traffic volume variation: $dev$ between two sequential TVSs. This correction leads to a higher prediction accuracy than using a simple historical average alone. When the algorithm finishes, $B_{Ot}(\forall t \in T)$ will be obtained as the baseline prediction of road link $O$.

The drawback of the baseline prediction is that it cannot predict the system variations and the noise of traffic volume of road link $O$ on different days. For a given time interval, the 24-hour predication method refers to the baseline prediction for the same period 24 hours ahead to calculate the difference in the traffic volume between the two consecutive days, see Eq.(6):
where $q_{d,t}^{24}$ is the 24-hour traffic prediction for road link $O$ at a given time interval $t$. $w$ is the number of time intervals to be considered before and after $t$ and the time window used for calculation is between timeslot $(t - w)$ to $(t + w)$. $\alpha$ is a revision factor which is the ratio between the measured traffic volume and the baseline prediction over the time window. $f(d, O, t, w)$ is the exponent of the revision factor.

This method takes into consideration the correlation between two consecutive days of a road link which helps minimize system variations that arise from seasonal changes, providing more accurate predictions. Specifically, considering road link $O$, if the measured traffic volumes at day $(d - 1)$ are on average higher (or lower) than the baseline prediction within the time window, this increase (or decrease) will be taken into account to adjust the prediction for day $d$. The size of the time window affects the overall accuracy as a small window size can result in too much local influence from noise but a large window size averages out these variations. We employ a time window with $w = 6$. $f(d, O, t, w)$ is used to calculate the correlation level between day $d$ and $(d - 1)$. For a given $t$ of a given day $d$, we refer to the AMC transition matrix for the same period of time the day before to define $f(d, O, t, w)$. Considering a time interval $t'$ between $(t - w)$ and $(t + w)$ the measured traffic volume defines a TVS, then $P_{(d-1)t'}^{obs}$ is the historical probability that road link $O$ is in this TVS at time interval $t'$. In parallel, the baseline prediction $B_{ot'}$ defines a TVS, $P_{(d-1)t'}^{base}$ is the historical probability that road link $O$ was in this TVS at time interval $t'$. $f(d, O, t, w)$ is then related to the average ratio between the historical probabilities derived from the measured traffic volume and baseline prediction. $f(d, O, t, w)$ should not be higher than the highest possible correlation between two continuous weekdays, $e_{24}$, which is set to be 0.8. In practice, $e_{24}$ should neither be too high nor too low to correctly reflect traffic correlations.
between two weekdays. The \( f(d,O,t,w) \) definition follows the principle that, if the measured traffic volume taken 24 hours ago reflects the real traffic situation of road link \( O \) better (or worse) than the baseline prediction, then the discrepancy between the measured traffic volume and baseline prediction on the previous day \((d-1)\) could have a stronger correlation in traffic situations of day \( d \). Thus \( f(d,O,t,w) \) should be higher.

Finally, short term prediction addresses the issues related to noise by taking short term correlations into account. For a given time interval \( t \), measured traffic volume in short period of time (10 minutes in our case) and previously obtained 24-hour prediction to the specified interval are exploited to provide an accurate short term prediction. The approach is summarized in Eq.(7).

\[
\begin{align*}
q_{d,t}^{24} &= q_{d,t}^{24} \times \beta^{h(d,O,t,r)} \\
\beta &= \left( \frac{\sum_{t' = t-r}^{t} q_{d,t'}^{\text{obs}}}{\sum_{t' = t-r}^{t} q_{d,t'}^{24}} \right) ^\frac{1}{r} \\
h(d,O,t,r) &= e_{st}
\end{align*}
\]

where \( q_{d,t}^{24} \) is the short term traffic volume prediction for road link \( O \) at time interval \( t \) of day \( d \); \( r \) is the number of time intervals prior to \( t \) to be used for comparison; \( \beta \) is the ratio between the measured traffic volume \( q_{d,t}^{\text{obs}} \) and 24-hour prediction \( q_{d,t}^{24} \) over a time period between \((t-r)\) and \( t \). This revision follows the principle that the traffic situation of road link \( O \) shortly before \( t \) would be highly correlated to current traffic state. Similarly, the size of time window needs to be carefully chosen to ensure that results are not greatly affected by noise or weak correlations. \( h(d,O,t,r) \) is related to \( r \). The higher \( r \) is, the weaker the correlation. In this paper, \( r \) is 2, and as a result, a high correlation is expected within the given time window, leading to \( e_{st} \) with a value of 1.

**C. Prediction on the rate in which traffic clears upon congestion – AMC (Clear)**

A novel way to predict the rate at which traffic congestion clears is derived from theorem 2, 3 and Definition 1. We use the canonical form of the transition matrices \( M^1 \) and \( M^2 \) to
calculate the fundamental matrices $F_1$ and $F_2$ respectively. The rate in which congestion clears is determined by theorem 3 and is given by Eq.(8),

$$T_{vac} = \begin{cases} ((F_1 \times c) + ((F_2 \times c)[N])) \times Int & (if \ t < T_{max}) \\ (F_2 \times c) \times Int & (if \ t \geq T_{max}) \end{cases}$$

(8)

where $T_{vac}$ is the prediction and $N$ is the size of the vector. $(F_1 \times c)$ and $(F_2 \times c)$ are each a vector of size $N$, $c$ is a $N \times 1$ vector and $Int$ is a fixed value of 5 minutes. If time interval $t < T_{max}$, we add $N^{th}$ entry of $(F_2 \times c)$ to each entry of $(F_1 \times c)$ as the road link $O$ has to go through two AMCs to be absorbed into $TVS_{vac}^2$. If road link $O$ is in $TVS_{k}^1$ at the time interval $t$ ($t < T_{max}$) in the morning of a weekday, road link $O$ will take $(T_{vac}[k])$ minutes to reach $TVS_{vac}^2$. Similarly, if road link $O$ is in $TVS_{k}^2$ ($t \geq T_{max}$) at time interval $t$ ($t \geq T_{max}$), it will take $(T_{vac}[h])$ minutes before being absorbed into $TVS_{vac}^2$.

**Results**

A separate set of test data is used to validate the AMC model. This data set was collected between 22nd November 2010 and 22nd April 2011 in Enschede and is recorded using the same method and time intervals as the training data discussed in the training data section. We consider both weekdays and weekends in the study. It is worth noting that traffic patterns observed in weekends are much less recurrent. Two groups of road links are considered in this paper. Group 1 consists of road links that have a highly recurrent morning rush hour, and road links in group 2 have a lowly recurrent morning rush hour. The recurrent level is decided by the deviation between the measured traffic volume and the historical average. The higher the deviation is, the lower the recurrent level is, and vice versa. In each group, we consider 10 road links.

We compare the short term prediction in AMC model with the S-ARIMA (Williams 2003, Kumar 1999) and Neural network (NN) models(Yin et al. 2002, Dharia 2003). S-ARIMA and NN here are implemented by using Econometrics Toolbox in Matlab R2012b. For a single road link, we apply the previous 10 days’ traffic data that have daily traffic patterns as the
training data to predict traffic volumes of a target day in S-ARIMA, and 30 days traffic data as training data in NN. The Root Mean Square Error (RMSE) of the predictions in each time interval is used as the quality indicator for prediction accuracy. For a time interval \( t \), RMSE is formulated as Eq. (9), and considers the road links in group \( G \). In addition, we also refer to the Mean Absolute Percentage Error (MAPE), given by Eq. (10).

\[
RMSE_t = \frac{\sum_{\forall e \in G} (q_{\text{obs}}_{t,e} - q_{\text{pred}}_{t,e})^2}{|D| \times |G|} \quad (\forall t \in T) \quad (9)
\]

\[
MAPE = \frac{\sum_{\forall e \in G} |q_{\text{obs}}_{t,e} - q_{\text{pred}}_{t,e}|}{|D| \times |G| \times |T|} \quad (10)
\]

The value of \( N \) in the AMC(short term) is set to 7, resulting in a transition matrix of size \( N+1 \), \( 8 \times 8 \). Figure 5 and table 2 show the prediction accuracy for AMC(short term), S-ARIMA and NN, for both groups of road links during weekdays. The results for RMSE and MAPE show that AMC(short term) outperforms both the S-ARIMA and NN models. The error difference is more pronounced during peak traffic flow compared to the equilibrium following peak traffic flow. While the data does contain noise due to Bank/Public holidays that have been included, the results also show that the proposed method is resilient to noise.

The computation cost of these models is examined by referring to the relative computation time the ratio between other prediction models computing time to the computing time of the NN model. The NN model, by far, requires the most computational time (figure 6). AMC(short term) provides the best performance in terms of speed due to its simplicity.

Figure 5. Comparison of the traffic volume prediction using RMSE with the AMC (short term), NN and S-ARIMA models, for two defined groups of road links: (a) group 1 (b) group 2.

Figure 6. Average relative computational times obtained from NN, S-ARIMA and AMC(short term) models
So far, the size of the transition matrix has been used is $8 \times 8$. We increase $N$ to 11 and 14, resulting in transition matrixes of $12 \times 12$ and $15 \times 15$ respectively. The purpose is to investigate whether or not different $N$ values leads to different predictions. These values are typical ones presenting raw, medium and high Markov Chain state granularity. The AMC (short term) model is rerun using these matrixes. Figure 7 and table 2 compare the RMSE and MAPE obtained from $N = 7, 11$ and 14; and it can be seen that changes in $N$ has a negligible effect on the overall accuracy.

In additional to weekdays, table 2 also considers AMC(short term) for weekends. The accuracy for weekend traffic prediction is significant lower than that for weekdays due to the lack of recurrent traffic patterns in the former.

Figure 7. Comparison of the traffic volume prediction results using RMSE with the AMC(short term) using transition matrices of different sizes: (a) group 1 and (b) group 2.

Table 2. Average MAPE for all the considered prediction models

Figure 8 shows the predicted average time it takes for road links to become clear from each time interval. The results in Figure 8 concern 12 TVSs. According to the results, road links in Group 1 and Group 2 will always become clear using a range of time intervals. We also carry out a basic estimation by calculating the average time it takes for traffic to clear for a given road link with reference to the TVSs as determined by the Historical Average prediction, and this estimation is compared with the AMC (Clear) model. For a clearer comparison, Figure 9 shows the ratio of how the AMC(Clear) outperforms the Historical Average prediction when predicting the time a road link clears. Specifically, the ratio is calculated by using AMC(clear) accuracy to minus Historical Average accuracy then divided by Historical Average accuracy to get a percentage. Thus, the higher the ratio the better the AMC model is in comparison to Historical Average prediction. Figure 9 considers 8, 12 and 15 TVS prediction settings for
Group 1 and Group 2. According to this comparison, the AMC model outperforms the Historical Average prediction.

Figure 8. Rate at which traffic clears upon congestion (a) group 1 and (b) group 2

Figure 9. The improved accuracy the AMC model compared to the Historical Average expressed as a percentage difference of AMC(clear) accuracy minus the Historical Average accuracy, divided by the Historical Average accuracy

Analysis

The AMC model works is designed for road links that have a recurrent traffic pattern. Results show that the model works well when predicting traffic for the morning rush hour during weekdays. Similarly, the AMC model can be applied to the evening rush hour. In the event of weekends in which recurrent in traffic patterns are found to be low, the overall accuracy is lower. In theory, if the traffic volume in a road link fluctuates frequently in continuous time intervals, the transition matrix $M^1$ and $M^2$ will have entries with low differences. In that case, $M^1$ and $M^2$ cannot predict which TVS a road link will be in at a specific time interval. This will jeopardize the traffic volume predictions and the time taken to clear a road link. Specifically, when generating baseline prediction in a road link, $S_k^{MC} = \arg\max(u \times M^1)$ or $S_k^{MC} = \arg\max(u \times M^2)$ in table 1 is used to decide whether to correct the historical average or not. Thus, if the objective road link has no recurrent traffic pattern, $S_k^{MC}$ as the most possible TVS in the time interval will not correctly reflect the real traffic situation. This is because the other TVSs almost have the same possibility as $S_k^{MC}$. The same situation happens when using TVS probabilities to generate the adjustable exponent $f(d, O, t, w)$ for a 24 hour prediction. If a road link has no clear traffic pattern, exponent $f(d, O, t, w)$ will not correctly reflect the correlation level. In addition, if the road link has no recurrent traffic pattern, the system variations and noise correlation within successive days or
time intervals will be weak. As a result, the AMC model will have a poor performance on traffic volume predictions. This was shown in the previous validation section through showing that road links in Group 2 have worse prediction results than Group 1. Finally, when predicting the time to clear road links, the model is based on the absorbing state assumption. Thus, if a traffic pattern has no recurrent traffic pattern, e.g., morning rush hour, the time prediction will not be applicable.

The number of TVSs and the amount of training data used in the AMC model are the other two factors that mostly influence the prediction results. These two factors are co-related. In theory, the number of TVSs should be neither too high nor too low and the amount of training data should be large enough but does not result in an unacceptable computing cost. When generating $M^1$ and $M^2$, if there are $M$ samples in the training data, the computing cost is: $M \times (N + 1)^2$, where $(N + 1)$ is the number TVSs. Therefore, $M$ and $N$ should not be too large. On the other hand, if the AMC model has too many TVSs, then there will not be enough sampling data to calculate a reasonable probability that a TVS transitions to another TVSs for a road link. This leads to $M^1$ and $M^2$ having entries that are similar. This problem becomes worse when the data samples are distributed non-uniformly within different TVS ranges. Using the training data in section 5 as an example, there are 150 $\times$ 73 = 10950 training data samples for road link $O$. Thus if the number of TVSs is 8 each in $M^1$ and $M^2$, there will be $((10950/73) \times T_{max})/8$ data samples used to calculate the possibility that a TVS transitions to another TVS in $M^1$, and $((10950/73) \times (73 - T_{max}))/8$ data samples used to calculate the possibility that a TVS transition to another TVS in $M^2$. On average, there are $\lfloor 10950/(8 \times 2) \rfloor = 684$ samples for a TVS. Suppose we set the number of TVSs to be 30, there will be only $\lfloor 10950/(30 \times 2) \rfloor = 182$ samples for a TVS. These are less data samples to support feasible $M^1$ and $M^2$. Hence, less TVSs can increase the sampling data used in each TVS. However, the number of TVSs should not be too low. This is because if there are too few TVSs, the AMC model will have too large granularity to generate useful
prediction. Arguably, the S-ARIMA and NN models require significant amount of input data, however, our results show that they provide less accuracy. If the (training) data set used in the AMC model were applied to the S-ARIMA and NN models, these models would fail to complete or to run.

Conclusions

This paper presents a traffic prediction method based upon Absorbing Markov Chains (AMC). This method combines the use of historical data and a transition matrix to provide an improvement over a baseline prediction method based upon averaging the historical data. The AMC model has been enhanced to address systemic variations and noise in traffic flows so as to provide short term predictions. Furthermore, the method is able to forecast the rate in which traffic clears upon congestion. Real traffic traces collected from the city of Enschede, The Netherlands, over a period of 6 months, are used for the AMC training process. Results presented in the form of Root-Mean-Square-Error (RMSE) and the Mean Absolute Percentage Error (MAPE) show that the AMC model is accurate, with the best performance being provided by a short term prediction approach. Compared to seasonal ARIMA and neural network models, the AMC model also shows a higher prediction accuracy. The proposed model is relatively insensitive to noisy data and is computationally efficient.

Acknowledgement

This research is supported by SUNSET (FP7/2007-2013) funded by the European Union Seventh Framework Program under grant agreement n°270228. We thank the Centre of Transport Studies, University of Twente for providing the traffic data.
References


Table 1 Algorithm to Generate the Baseline Prediction of road link $O$

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
</table>
| 1. | **Input:** $A_{Ot} (\forall t \in T)$, $M^1$ and $M^2$, $T_{max} = \arg\max_{t \in T}(A_t)$  
**Output:** $B_{Ot} (\forall t \in T)$  
**Initialization:** $B_{Ot} = A_{Ot} (\forall t \in T)$; |
| 2. | WHILE ($\exists k \in T$)  
2.1. IF ($k \leq T_{max}$) // AMC before peak time  
2.2. FOR $\forall m \in \{1..N, con\}$ { IF ($A_{Ok}$ is in $TVS^1_m$) $s^{base}_k = m$; ENDIF;}  
2.3. IF ($k > T_{max}$) // AMC after peak time  
2.4. FOR $\forall m \in \{1..N, vac\}$ { IF ($A_{Ok}$ is in $TVS^2_m$) $s^{base}_k = m$; ENDIF;}  
2.5. END WHILE |
| 3. | WHILE ($\exists k \in T$)  
3.1. IF ($1 < k \leq T_{max}$) // AMC before peak time  
3.2. $u = \text{zeros}(1,(N+1))$; $u(1, s^{base}_{k-1}) = 1$;  
3.3. $s^{MC}_k = \arg\max(x(u \times M^1))$;  
3.4. IF ($s^{MC}_k \gg s^{base}_k$) { $B_{Ok} = B_{Ok} + (s^{MC}_k - s^{base}_k) \times \text{Dev}$; }  
3.5. IF ($k > T_{max}$) // AMC after peak time  
3.6. $u = \text{zeros}(1,(N+1))$; $u(1, s^{base}_{k-1}) = 1$;  
3.7. $s^{MC}_k = \arg\max(x(u \times M^2))$;  
3.8. IF ($s^{MC}_k \gg s^{base}_k$) { $B_{Ok} = B_{Ok} + (s^{MC}_k - s^{base}_k) \times \text{Dev}$; }  
3.9. END WHILE |

Accepted Manuscript  
Not Copyedited
Table 2. Average MAPE for all the considered prediction models

<table>
<thead>
<tr>
<th>Target group</th>
<th>Method</th>
<th>MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weekdays</td>
<td>AMC (Short term)</td>
<td>8.91% (N = 7); 8.64% (N = 11); 8.87% (N = 14)</td>
</tr>
<tr>
<td>(Group 1)</td>
<td>S-ARIMA</td>
<td>19.61%</td>
</tr>
<tr>
<td></td>
<td>NN</td>
<td>10.32%</td>
</tr>
<tr>
<td>Weekdays</td>
<td>AMC (Short term)</td>
<td>15.55% (N = 7); 15.50% (N = 11); 15.40% (N = 14)</td>
</tr>
<tr>
<td>(Group 2)</td>
<td>S-ARIMA</td>
<td>25.51%</td>
</tr>
<tr>
<td></td>
<td>NN</td>
<td>18.27%</td>
</tr>
<tr>
<td>Weekends</td>
<td>Baseline prediction</td>
<td>58.72%</td>
</tr>
<tr>
<td>(Group 1)</td>
<td>AMC (Short term)</td>
<td>18.43% (N = 7)</td>
</tr>
<tr>
<td>Weekends</td>
<td>Baseline prediction</td>
<td>67.93%</td>
</tr>
<tr>
<td>(Group 2)</td>
<td>AMC (Short term)</td>
<td>21.52% (N = 7)</td>
</tr>
</tbody>
</table>
Figure 1. Traffic pattern of a road link measured in the morning (weekdays) rush hour between 6:00am and noon at intervals of 5 minutes

Figure 2. State transition of (a) a Markov Chain and (b) an Absorbing Markov Chain

Figure 3. AMC absorbing state models traffic congestion absorbing state and the vacuum absorbing state (a): overall traffic situation; (b): split into two AMCs

Figure 4. An overview of the proposed AMC model

Figure 5. Comparison of the traffic volume prediction using RMSE with the AMC (short term), NN and S-ARIMA models, for two defined groups of road links: (a) group 1 (b) group 2.

Figure 6. Average relative computational times obtained from NN, S-ARIMA and AMC(short term) models

Figure 7. Comparison of the traffic volume prediction results using RMSE with the AMC(short term) using transition matrices of different sizes: (a) group 1 and (b) group 2.

Figure 8. Rate at which traffic clears upon congestion (a) group 1 and (b) group 2
Figure 9. The improved accuracy the AMC model compared to the Historical Average expressed as a percentage difference of AMC(clear) accuracy minus the Historical Average accuracy, divided by the Historical Average accuracy.

Table 1 Algorithm to Generate the Baseline Prediction of road link $O$

Table 2. Average MAPE for all the considered prediction models.
(a)

(b)
Historical Data

Calculate AMC transition matrixes: $M^1$ and $M^2$

Input: $M^1, M^2$

Apply Theorem 1 to calculate $B_{O_t}$

Input: $B_{O_t}$

Calculate 24-hour prediction: $q^{24}_{dO_t}$

Input: $q^{24}_{dO_t}$

Calculate short term prediction $q^{0}_{dO_t}$

Output: $q^{0}_{dO_t}$

AMC(Short term)

Apply Definition 1 to calculate the fundamental matrix:

$F = (I - Q)^{-1}$

Apply Theorem 2 to find

Prob (traffic clears): $B = F \times R$

Apply Theorem 3 to calculate the time it takes to clear traffic:

$t = F \times c$

AMC(Clear)
The ratio of how the AMC model outperforms the Historical Average (%) for different group sizes.

- Group 1: N=7, N=11, N=14
- Group 2: N=7, N=11, N=14

Copyright 2013 by the American Society of Civil Engineers