http://journal.rmutp.ac.th/

Real-time Travel Time Prediction Algorithm Using Spatiotemporal Speed Interval Patterns Matching

Krit Jedwanna* and Weerathep Chanintornthep

Faculty of Engineering, Rajamangala University of Technology Phra Nakhon
1381 Pracharat 1 Road, Wong Sawang, Bang Sue, Bangkok 10800

Received 28 October 2020; Revised 30 December 2020; Accepted 6 January 2021

Abstract

Accurate, efficient and robust travel time prediction is crucial to the development of advanced traveler information systems for providing route guidance information. To achieve this goal, this paper proposed a travel time prediction through the matching of the current speed interval pattern to that in a historical database. Speed intervals, instead of speeds, are considered in this study to simplify the structure of matching patterns for improving matching efficiency. In this study, speed interval patterns are defined by sets of link speed intervals that are either spatially or temporarily correlated with the link considered. With the speed interval patterns, the algorithm is developed for searching the historical pattern(s) that is/are the closest match with the current one. Then, link speeds from these matched patterns are combined for travel time prediction. By using the GPS probe taxi data, which the collected speeds are aggregated in every 5 minutes, the proposed travel time prediction system is implemented in Bangkok. With the speed data from probe taxi, this paper has chosen four links/paths with different geometric and flow characteristics for testing the performance of the proposed travel time prediction system. From these tests, it is found that the optimal speed interval pattern should include: 1) speeds of the studied link within three preceding time intervals and; 2) speeds of links in the first connection level of the studied link. Also, while the computational time is capable of real-time application, the proposed prediction algorithm is more accurate under uninterrupted flow conditions.

Keywords: Traffic Pattern Matching; Travel Time Prediction; Time-dependent Network; Spatiotemporal Correlation

*** Corresponding Author. Tel.: +662 665 3777, E-mail Address: krit.j@rmutp.ac.th**

1. Introduction

In the Development of the Intelligent Transportation System (ITS), Advanced Travelers Information System (ATIS) plays an important role in providing real-time traffic conditions and/or traffic management measures to travelers for avoiding unnecessary delay and ensuring the reliability of road network. Among all the information provided by ATIS, travel time is one of the key information used by the route guidance system to evaluate real-time shortest paths that help in the trip planning of drivers. As route guidance system is usually used before the trip is made, an accurate prediction of future link/path travel times is necessary to ensure good performance of the system. Travel time prediction is a challenging problem as the prediction will be affected by i) complex and non-linear interactions of hetero-geneous groups of vehicles/drivers; ii) infrastructures and/or traffic management schemes that interrupt the traffic flows, and; iii) availability and types of information for travel time prediction. Owing to the complexity of the travel time prediction problem, various methods/models have been proposed in the literature and could be classified into two categories: parametric and non-parametric prediction models (H.-E. Lin et al.). [1]

Table 1 summarized some of the travel time prediction studies, with the corresponding data sources and prediction method, in the literature.

N. K. Chowdhury et al. [2] have proposed a modified moving average approach, which is another category of time series model, for travel time prediction. By eliminating unwanted fluctuations in the data set, the model proposed in N. K. Chowdhury et al. [2] outperforms the conventional moving average methods. Owing to its nature, the time series model could only provide an accurate prediction for a short forecast period (i.e., 5 ~ 10 minutes) or under stable traffic conditions. Apart from the parametric models, W.H.K. Lam et al. [3] have adopted a traffic flow simulator (TFS), in which the original-destination demands are calibrated based on the historical link counts and prior demands by an upper-level model, for travel time predictions.

Table 1 Comparison of required data and prediction methods for travel time prediction

Studies	Application	Data	Prediction Method	Properties
N.K.Chowhhury et al. (2009) [2]	Short-term prediction	GPS	Moving average	Specific route
W.H.K.Lam et al. (2005) [3]	Short-term prediction	AVI	Traffic simulation	Urban road network
S.I.Bajwa et al. (2004) [4]	Short-term prediction	Point sensor	Pattern matching By Genetic Algorithms	Expressway
T. Kim et al. (2005) [5]	Short-term prediction	Point sensor	Traffic pattern recognition	Highway
Z.-P. Li et al. (2008) [6]	Short-term prediction	AVI	Exponential Smoothing	Urban road network
W.-H. Lee et al. (2009) [7]	Traffic Classification	GPS	Fuzzy C-mean	Specific route
Y. Zhang and Y. Liu (2009) [8]	Traffic state prediction	Point sensor	Non-linear least square	Freeway
A. Khosravi et al. (2011) [9]	Accuracy interval	GPS	Bayesian updated Neural network	Specific route
A. Simroth and H. Zähle (2011) [10]	Long-term prediction	GPS	Nonparam etric Distribution-free Regression model	Nationwide Road network
W. Qiao et al. (2013) [11]	Short-term prediction	Blueto oth data	- Historical average - Auto regressive integreated moving average - Kalman filter - K-nearest neighbors	Freeway
Y. Zou et al. (2014) [12]	Short-term prediction	Point sensor	Space-time diurnal	Freeway
H. Jiang et al. (2016) [13]	Short-term prediction	Point sensor	- Neual Network - Multilinear Regression - Statistical Model	Highway
Our Proposed method	Short/ medium Term prediction	GPS	Pattern matching	Urban road network

For many cases in actual implementation, the relationships between travel time and traffic conditions are too complicated to be represented by any type of model. Thus, non-parametric models are becoming more attractive, especially for the application in a large and complex urban road network. In the literature, non-parametric travel time prediction models could be categorized into k-nearest neighbor (k-NN) algorithm (T. Kim et al. [5], W.H.K. Lam et al. [14]), expert system (W.-H. Lee et al. [7]) and artificial neural network (J.W.C. Van Lint et al. [15], A. Khosravi et al. [9]). T. Kim et al. [5]) developed pattern recognition algorithm, which based on link volumes in current and preceding time intervals, for short-term link volume/travel time prediction of a section of expressway in Washington DC. compare to the other k-NN approaches that use only current link flows, pattern recognition algorithm proposed in T. Kim et al. [5] gives a smaller prediction error in different neighborhood sizes. Apart from the prediction of traffic conditions, W.H.K. Lam et al [14] put a step forward in adopting the k-NN-based travel time prediction in incident detection. In their study, a modified k-NN approach, which depends on the estimated travel times and the corresponding temporal variance-covariance relationships, is adopted in travel time prediction. By comparing the predicted travel times and the corresponding estimated travel times from automatic vehicle identification (AVI) data, the incident could be detected if the difference exceeds the certain threshold value. In addition to the data mining approaches, expert systems improve the accuracy of predictions by the introduction of prior-knowledge rules. W.-H. Lee et al. [7] have proposed an expert system for travel time prediction in Taipei urban road network. In their study, travel times are predicted based on a weighted sum of the current (based on current speeds and flows) and historical (based on traffic patterns in the historical database) travel time predictions. Rules, which are responsive to real-time events, are introduced by the expert of the test area for the automatic determination of weights. W.-H. Lee et al. [7] show that such an expert system in travel time prediction could achieve a root mean square error (RMSE) as low as 11%. J.W.C. Van Lint et al. [15], on the other hand, has adopted a state-space neural network (SSNN) to ensure the accuracy and robustness of travel time predictions in the presence of missing data. In their study, missing data (e.g., speeds from loop detectors) are estimated using simple imputation (i.e., exponential moving average and/or spatial interpolation) and the proposed method is tested on a simulated model with different patterns of missing data. J.W.C. Van Lint et al. [15] shows that even with 40% of data are missing, their SSNN could achieve a prediction error similar to the case with no missing data. Apart from the aforementioned models, there are other non-parametric models, such as the least square support vector machines discussed in Y. Zhang and Y. Liu [8] and the nonparametric distribution-free regression model introduced in A. Simroth and H. Zähle [10]), for travel time prediction. In developing the aforementioned parametric and non-parametric models, traffic data plays a crucial role in model calibration/training that affects their performance. In practices, traffic data are commonly collected by global-positioning system (GPS), automatic vehicle identification (AVI) system and point sensor (e.g., loop detectors), while each of these methods will have their pros and cons under different types of implementation (e.g., freeway, arterial, etc). **Table 1**, summarized some of the travel time prediction studies and their required types of data. From the reviewed studies, the majority of data is collected using GPS (**Table 1**). It is because GPS data, which could be collected by inexpensive onboard GPS sensors, does not require a substantial investment in infrastructure and maintenance as in the other two methods (point sensors and AVI). Apart from the inexpensiveness, GPS data could be used to provide the precise locations of the tracked vehicle (e.g., location for every 30 seconds) for further use in the path-related analysis. Despite the high implementation and maintenance cost, point sensors are still the most common data collection method in practices for its continuous collection of relatively reliable traffic data (e.g., speeds and flows). Also, unlike the GPS data that passively depend on the installed vehicles, point sensors could be strategically placed for maximizing the coverage of the monitored network. AVI data, as another data source, provides travel time between two points through the identification and match of vehicles. Unlike the GPS data, AVI data is not able to provide the chosen path of the matched vehicles and, thus, models/algorithms should be developed to estimate the path choice from the AVI data (T. Siripirote et al. [16]). Owing to the different characteristics of data sources, data fusion techniques have recently be considered for providing an accurate and robust travel time prediction (K.P. Hwang et al. [17])

Considering the problems of parametric travel time prediction models for long-term prediction under unstable traffic conditions and the computational efficiency of various data mining approaches, this study proposed a travel time prediction algorithm through the matching of link speed interval patterns, which consist of spatiotemporally correlated link speed intervals. In this study, speed intervals, instead of speeds, are considered to simplify the structure of matching patterns for improving matching efficiency. This paper is organized as follows. Section 2 will describe the variables used in this study. The formulation and solution algorithm of the proposed travel time prediction model will also be given in Section 2. Section 3 will then carried out various empirical tests based on the data and models introduced in the previous section. Lastly, the paper will be concluded in Section 4.

2. Travel time prediction through speed intervals matching

In this study, link travel times are predicted by the corresponding historical value(s) with speed interval pattern(s) that is/are the most similar to the current one. In defining speed interval patterns, this study will not only consider the speed interval of the link under concern, instead, speed intervals of the surrounding links (spatial) and previous time intervals (temporal) will also be considered. In this section, the collection and processing of speed data used in this study will be discussed in Section 2.1. Then, Section 2.2 will define the speed interval patterns used in the travel time prediction. With the speed interval patterns, Section 2.3 introduces a matching algorithm for finding the most similar speed interval pattern in the historical database. Finally, Section 2.4 will give the architecture of travel time prediction in Bangkok road network based on the collected data and developed algorithm in Section 2.1~2.3.

2.1 Collection and processing of speed data

In Bangkok, taxi companies will usually set up an IP wireless communication (Taxi radio) for normal voice-based communication with their taxi drivers. Even though taxis in Bangkok are not mandatory to equip with GPS devices, these established communication channels could also be used to transmit GPS data to enhance real-time taxi dispatching. In this research, traffic data is obtained from around 10,000 taxis, which are equipped with GPS devices, in Bangkok. The GPS devices will wirelessly transmit the latitude, longitude, altitude, traveling speed, heading, and timestamp of the equipped taxi to a designated computer server for every 45 seconds.

In this study, the urban and suburban road networks in Bangkok, which consist of 3,000 links and cover more than 3,500 origin-destination (O-D) pairs, are considered and the corresponding GPS data is collected. In the Bangkok network model, new links are usually defined wherever there is a change of road geometry (e.g., increase in several lanes), intersection with other roads (e.g., signalized junction, priority junction), and land use pattern (e.g. exit of a parking lot). The increasing detail of the modeled network (e.g., number of parking lot exits included) will substantially decrease the length of each link and, thus, cause the same network to have a larger number of links. As the number of links increases, there will be a higher chance that the real-time GPS data, especially the speed data, from probe taxi is not available for a certain link at a certain time period. With insufficient, or missing, link speed data at different traffic conditions, it is not possible to provide a reliable travel time prediction for the Bangkok network due to its highly varied traffic conditions. On the other hand, if unreasonably long links are defined, the average of collected speeds could not precisely represent the actual speed profile of the links. With such inaccurate average speeds for long links, it is not possible to achieve a reliable prediction of travel time. Thus, in this study, data availability and data accuracy are traded off in choosing the length of the links in the network. Apart from longitude and latitude, the altitude from GPS data will also be used in this study to separate speeds collected from links that are vertically overlapped. With this consideration, traffic links in this study will be categorized into three types: road (at grade), elevated (flyover), and toll road.

These link types and the other link characteristics (e.g., location of U-turn and road junction) will be stored as the local parameters of traffic link in the database.

Table 2 Samples of GPS data obtained from probe taxi

GPS_ID	Latitude	Longitude	Timestamp	Speed (km/hr)
1311042	13.581726	100.867721	2011-12-15 10:00:13	84
1310992	13.581680	100.867248	2011-12-15 10:00:14	114
1311751	13.581893	100.867095	2011-12-15 10:00:22	85
65679	13.582565	100.861358	2011-12-15 10:00:26	77
1311401	13.583236	100.860351	2011-12-15 10:01:07	93

For each set of GPS data collected from probe taxis, which some samples are shown in **Table 2**, the map-matching algorithm is adopted to check whether the GPS device (i.e. taxi) is located on road segments and screen out those data that are not on the road segment (e.g., in the parking lot). The location information will also be used to locate the collected GPS data into different links for determining the average speeds that are used for travel time prediction. Apart from the location information, the collected vehicle speeds will also be used in this study and will be considered as the spot speeds of the vehicles at the corresponding location (link) and time. Before these spot speeds could be used to estimate the historical link speeds for travel time prediction, filtering procedure will be adopted to filter out outlying and/or erroneous speed data. For instance, if a GPS device transmitted back a zero vehicle speed while most of the other speed data on the same link and in the time period are greater than 80 kph, this speed data will be considered as erroneous and filtered out.

After filtering out the outlying and erroneous spot speeds, a stratified sampling technique is applied to the remaining spot speeds for sampling a balanced number of spot speeds from each speed interval within each time period. Such stratified sampling of spot speeds is crucial for this study as the spot speeds within Bangkok network usually have a large variation (even they are from the same link and in the same time period) due to the frequent interruptions of traffic flows. For example, considering a link with a signalized intersection located at its downstream end, vehicles will queue up in front of the stop line of the intersection (i.e. at the downstream end of the link) during the red times. As the queue is building up, the number of vehicles (probed taxis) within the queues will increase and become larger than those that have not joined the queue. As a result, there will be a large proportion of spot speeds collected from the probed taxis that are in the queue as compared to those that are not. Thus, if a simple average of spot speeds is used to define the speed of that link, this estimated speed will be underestimated. To overcome such an issue, the stratified sampling technique divides the spot speeds into several speed intervals and randomly chooses the same number of data from each interval for estimating the average speed of the link (**Table 3**).

Table 3 Probe vehicle data stratified by speed intervals

Speed interval	Number of spot speed data	Randomly chosen spot speeds
0 – 12	12	1, 6, 11
13 – 25	3	15, 22, 25
26 – 38	6	28, 30, 33

The average link speed will be estimated for every 5 minutes and classified by different day of the week for further uses in travel time prediction.

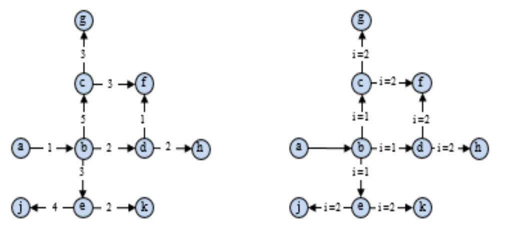
2.2 Definition of speed interval pattern

In this study, a traffic pattern, which is used to match the current traffic condition with the historical database for travel time prediction, is defined by speed interval pattern. For each of the spot speeds and historical average speeds, the corresponding speed interval is defined as:

$$\tilde{v}^{k}_{ab,t} = \left\lfloor v^{k}_{ab,t} / s \right\rfloor \quad (1)$$

where $\tilde{v}^{k}_{ab,t}$ is the speed interval of link *ab* (or the link from *a* to *b*) in time interval *t* and day *k*, $v^{k}_{ab,t}$ is the spot speed (historical average speed) of link *ab* in time interval *t*, *s* is the width of speed interval, and *x* denotes the largest integer that is less than *x*. For example, if the spot speed of link *ab* ($v^{k}_{ab,t}$) equals to 21 kph and the width of speed interval (*s*) is taken as 5 kph, the speed interval ($\tilde{v}^{k}_{ab,t}$) for this spot speed will be equal to 4. In this study, for each of the time interval (say 5 minutes duration), the speed interval of the average speed, which is determined by the stratified sampling technique described in the previous section, is determined for each of the links (**Fig 1a**).

In this study, the speed interval pattern is defined by the speed intervals of a set of links that are spatially and/or temporally correlated to the link under concern (i.e., the link for travel time prediction). In defining the spatial correlation between links, this study has adopted the idea of connection level. **Fig.1b** shows the connection level of the surrounding links to the link *ab*. In this figure, link *bc*, *bd*, and *be* are the links with connection level (*i*) equals to 1 as they are directly connected to node *b*, which is the end node of the link *ab*. Other links in **Fig. 1b** (e.g., *cg*, *cf*, etc) have the connection level equals to 2 as they are connected to the end nodes of links in the previous connection level (i.e., *i* = 1). Under such a definition, the connection level of all links within the network concerning link *ab* could be established. Compare to the spatial correlation, the definition of temporal correlation of link speed intervals is simpler. In this study, link speed intervals are assumed to be temporally correlated in time interval.



(a) Speed intervals (b) Connection levels of link *ab*

Fig. 1 Speed Intervals and Spatial Connections

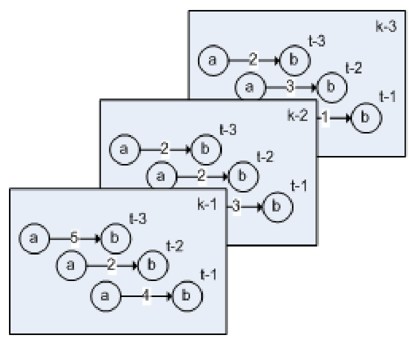


Fig. 2 Temporally correlated speed intervals for link *ab* at day *k* and time interval *t*

Fig. 2 shows the examples of temporally correlated speed intervals on different days and time intervals for the speed interval measured on link ab at day k and time interval t.

In this study, the speed interval pattern is defined by the speed interval of the set of links that are spatially and/or temporally correlated to the link under concern.

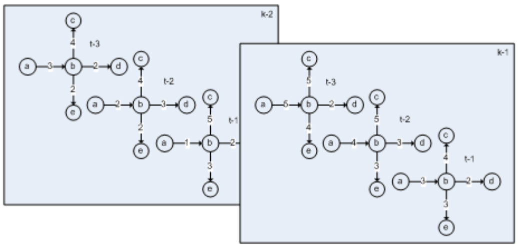


Fig. 3 Spatially and temporally correlated speed interval pattern in different days

Fig. 3 shows the example of speed interval pattern on day *k*-1 and *k*-2 for speed interval measured on link *ab* at day *k* and time interval *t*. In the speed interval patterns given in **Fig. 3**, it is assumed that: 1) speed interval of links with a connection-level equals to 1 (i.e. i = 1) is spatially correlated to the speed interval of link *ab* and, 2) the current speed intervals are temporally correlated to the 3 preceding time intervals (i.e. *t*-1, *t*-2 and *t*-3).

2.3 Speed interval pattern matching (SIPM) for travel time prediction

Based on the definition of speed interval pattern (Section 2.2), speed interval pattern for the current time interval is established and the historical database, which is built based on the data collection/processing method introduced in section 2.1, is searched for finding the most similar pattern (i.e., matching of speed interval pattern). Due to the large number of links in the Bangkok road network, an efficient and effective pattern matching algorithm should be proposed to ensure a fast matching process for real-time applications. In this study, the similarity of two-speed interval pattern (e.g., the current and historical speed interval pattern) is measured by a mismatch value, m_l^k, defined as:

$$m_l^k = \sum_{p=1}^{P}\sum_{i=1}^{I}\sum_{l'\in L_i^l}\left|\tilde{v}^k_{l',t-p} - \tilde{v}^n_{l',t-p}\right|, \qquad \forall k \in K \qquad (2)$$

where m_l^k is the mismatch of speed interval pattern derived from link l at current time t and day n as compare to the corresponding speed interval on day, P is the number of preceding time intervals that are considered to have temporal correlation with the current time interval, I is the connection levels that are considered to have spatial correlation, L_i^l is the set of links with connection-level i concerning link l, and K is the set of days in the database to be compared for determining the mismatch value. Typically, two different sets of K will be considered: 1) set of specific days on a week, $K = \{n-7, n-14, n-21, ..., n-\tilde{K}\}$, and 2) set of unclassified days, $K = \{n-1, n-2, n-3, ..., n-\tilde{K}\}$. Based on the above definition of mismatch value (m_l^k), the SIPM for travel time prediction of link l could be formulated as following constrained minimization problem:

$$\underset{\boldsymbol{\delta}}{Minimize} \quad Z_l(\boldsymbol{\delta}) = \sum_{k \in K} \delta_l^k m_l^k \tag{3a}$$

$$\text{Subject to} \quad \sum_{k \in K} \delta_l^k = 1 \tag{3b}$$

where $\delta_l^k = 1$ if speed interval pattern of day k is chosen as the match of current pattern (i.e. the corresponding speed interval pattern of day k is used to evaluate the mismatch value, m_l^k, through **equation (2)**), otherwise equal to zero. Minimizing equation (3a) guarantees that the chosen speed interval pattern, which is determined by the variable, δ_l^k, gives the smallest mismatch value (i.e. most similar pattern). Constraint (3b) is to ensure that only one historical speed interval pattern is used to match with the current one. The above minimization problem will be set up and solved for all links within the transportation network that need a speed interval matching for travel time prediction. Note that the formulation of speed interval matching problem (**equation (3)**) is general optimization framework of which additional constraint could be easily added to refine the feasible solution set. For instance, an additional constraint, $1 \le \tilde{v}_{l',t-p}^k \le 3 \quad \forall k \in K, l' \in L_i^l, p \in [1,2,...,P], i \in [1,2,...,I]$, could be added to limit the number of speed intervals, or the search space, in the database for speeding up the searching process.

As the mismatch value m_l^k is defined by speed intervals, which are integer variables, the above minimization problem (**equation (3)**) likely has multiple solutions (i.e., multiple days in the history with the speed interval patterns that give the same mismatch value as compared with the current pattern).

Thus, to handle cases with multiple matches, the procedure in **Fig. 4** is adopted to match/estimate the speed pattern that is used for travel time prediction.

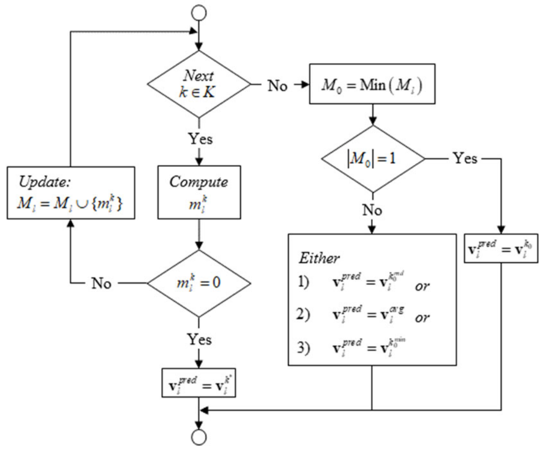


Fig. 4 Speed interval pattern search for travel time prediction

For the iterative loop on the left-hand side of **Fig. 4**, the algorithm will compute the mismatch value (m_l^k) of link l for each day $k \in K$ in a reverse chronological order (i.e., starting from the most recent day and then evaluate back in time). Within these iterative steps, the procedure will terminate whenever a speed interval pattern in the historical database ($k = k^*$) is completely matched with the current pattern (i.e., $m_l^{k^*} = 0$). Then, speeds on that day will be taken as the speed pattern used in travel time prediction ($\mathbf{v}_l^{pred}$), which is defined as:

$$\mathbf{v}_l^{pred} = \left(v_{l,t+1}^{pred}, v_{l,t+2}^{pred}, \ldots\ldots, v_{l,t+t^p}^{pred}\right)^T$$
$$= \left(v_{l,t+1}^{k^*}, v_{l,t+2}^{k^*}, \ldots\ldots, v_{l,t+t^p}^{k^*}\right)^T = \mathbf{v}_l^{k^*} \qquad (4)$$

where t^p is the number of future time intervals that travel time prediction is required. If the speed interval pattern is not a complete match with the current pattern (i.e., $m_l^{k^*} \neq 0$), the estimated mismatch value will be included in a set M_l and the mismatch value of next day (or k) will be calculated.

After the mismatch value of all days in the set, K is evaluated and none of these days has a speed interval pattern that is completely matched with the current pattern, the procedure in the right part of **Fig. 4** will carry out to estimate the speed pattern used for travel time prediction ($\mathbf{v}_l^{pred}$). M_0 defines a set that contains the smallest element(s) of M_l, which is the set of mismatch values for all days within K. If the size of M_0 (i.e., $|M_0|$) is equal to 1, which indicates that there is only one solution for the matching problem (3), the speeds of that day (k_0), which is

$$\mathbf{v}_l^{k_0} = \left(v_{l,t+1}^{k_0}, v_{l,t+2}^{k_0}, \ldots\ldots, v_{l,t+t^p}^{k_0}\right)^T,$$

will be taken as the speed pattern used in travel time prediction ($\mathbf{v}_l^{pred}$). The size of M_0 is larger than unity, which indicates that there exists multiple solutions for the matching problem (3), $\mathbf{v}_l^{pred}$ will be estimated by three different methods: random sample, average value, and closest speed pattern. For random sample, a day (k_0^{rnd}) is randomly chosen from the set K_0, which is the set of days within K with minimum mismatch values, for providing the speeds

$$\mathbf{v}_l^{k_0^{rnd}} = \left(v_{l,t+1}^{k_0^{rnd}}, v_{l,t+2}^{k_0^{rnd}}, \ldots\ldots, v_{l,t+t^p}^{k_0^{rnd}}\right)^T$$

used in travel time prediction. For average value, each element of $\mathbf{v}_l^{pred}$ will be taken as the average of the corresponding element of all the speed patterns within K_0 (i.e. $v_{l,t'}^{pred} = v_{l,t'}^{avg} = \frac{1}{|M_0|}\sum_{i \in K_0} v_{l,t'}^{i}, \ \ \forall t' \in \left[t+1, t+t^p\right]$).

For closest speed pattern, the day ($k_0^{\min}$) is chosen such that the speed pattern, not speed interval pattern, is the most similar (or the closest) to the current speed pattern (i.e., $k_0^{\min} = \arg\min_{k \in K} \sum_{p=1}^{P}\sum_{i=1}^{I}\sum_{l' \in L_i^l}\left|v_{l',t-p}^{k} - v_{l',t-p}^{n}\right|$). Then the speeds on that day (i.e., $\mathbf{v}_l^{k_0^{\min}} = \left(v_{l,t+1}^{k_0^{\min}}, v_{l,t+2}^{k_0^{\min}}, \ldots\ldots, v_{l,t+t^p}^{k_0^{\min}}\right)^T$), will be used for travel time prediction. With the speed vector, $\mathbf{v}_l^{pred}$, estimated by various methods in this procedure, the travel time for link l in the future time intervals (i.e., $t+1$, $t+2$, …, etc) could be predicted by dividing the link length with the corresponding speed in $\mathbf{v}_l^{pred}$.

2.4 Architecture for travel time prediction system in Bangkok

In this study, as the travel time prediction system will be implemented to the large-scale Bangkok network that consists of 3,000 links, the feasibility of real-time application will not only rely on the efficiency of prediction algorithm (Section 2.3) but also depend on the efficient flow of data between all related large-scale databases.

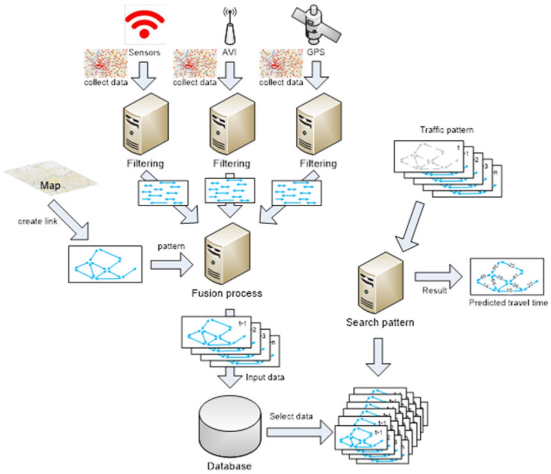


Fig. 5 Architecture for travel time prediction system in Bangkok

Fig. 5 shows the architecture of the real-time travel time prediction system for the implementation in Bangkok. In this study, data collected from GPS, AVI and various point sensors (e.g., Autoscope, loop detector, microwave sensors) are used to provide the necessary link travel time and/or link speed data (**Fig. 5**). Before these data could be used for travel time prediction, they are filtered by using the filtering procedure described in Section 2.1 to provide link speeds (link travel times) for different time intervals. Based on the transportation network extracted from the map, the link speeds (link travel times) from different data sources are combined to provide the network-wide speed pattern for each time interval and day. These filtering and combining processes will be carried out in real-time and the created speed patterns will then be stored in the database as historical speed patterns for travel time prediction. To carry out travel time prediction, speed patterns of the current and preceding time intervals are taken from the database for forming the speed interval patterns as described in Section 2.2. With the speed interval patterns, pattern search is carried out on the chosen set of speed interval patterns in the database for travel time prediction (Section 2.3).

3. Empirical studies

The proposed SIPM algorithm and the corresponding system architecture are implemented in the Bangkok road network for travel time prediction based on the collected/estimated speed data. In this section, the performance of the proposed model is evaluated in link travel time prediction (Section 3.1) and path travel time prediction (Section 3.2).

3.1 Link travel time prediction

To test the performance of the proposed model in the prediction of link travel time, four different types of links are considered: 1) link with free-flow condition on a multi-lane road without weaving, on- and off-ramp (Link A); 2) link with free-flow condition on a local side road connected to community areas (Link B); 3) main road connected to business areas with traffic signals (Link C), and; 4) link in central business district with traffic signals and U-turns (Link D). The characteristics of these links are shown in **Table 4** and their corresponding geographical locations are shown in **Fig. 6**.

Table 4 Links for travel time prediction

Link	From /To	Direction*	Distance (km)
A	Bang Khun Thian – Bang Khru	Inbound	5.7
B	Bangkok University Statium – Tollway Rangsit	Inbound	2.9
C	MRT Ratchada – The Emerald Hotel	Inbound	2.6
D	Pratumnam – Ratchatawee	Outbound	1.1

*Inbound (outbound) refers to traffic travelling towards (out of) the centre of Bangkok city

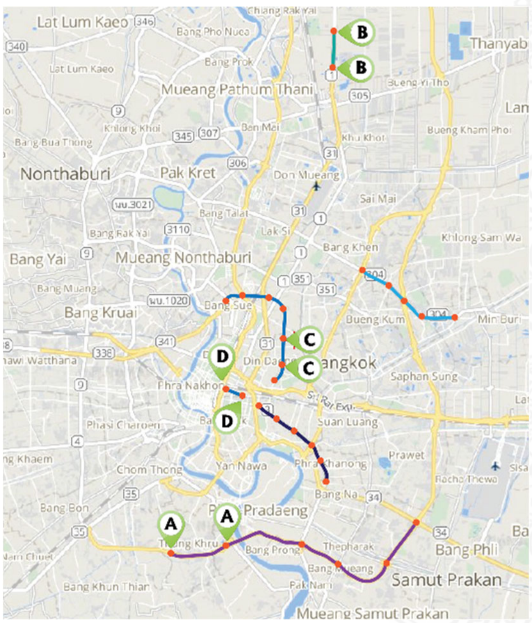


Fig. 6 Links and paths adopted in travel time prediction

The Road network in Bangkok is relatively complex as the upstream and downstream ends of links may connect to various types of roads, for example, elevated roads, toll-way ramps, and community streets. **Fig. 7** shows the set of links that are correlated to Pratunam – Ratchatawee link (Link D) at connection level $i = 1$. Among these correlated links, which the traffic flow directions are marked by the arrow in **Fig. 7**, Link *ef* represents the flyover at the intersection of Ram Inthra Road and Ramintra-At narong Expressway (Intersection *a*).

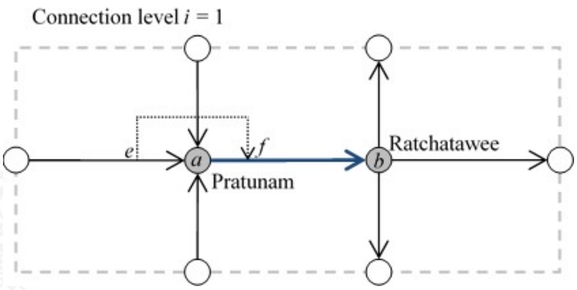


Fig. 7 Spatially correlated links of Link D (Connection level $i = 1$)

To demonstrate the variation of speed interval pattern over time, the speed interval pattern (s = 5 kph, I = 1, P = 1) of Link D during time interval 09:20 ~ 09:25 is compared with the corresponding speed interval patterns in the past 52 weeks and the mismatch values are evaluated (**Fig. 8**).

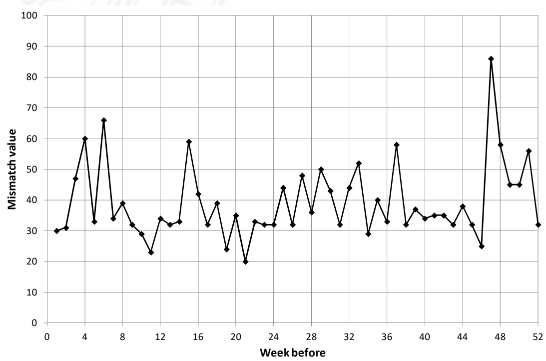


Fig. 8 Variation of mismatch values for the speed interval pattern of Link D during 09:20 ~ 09:25

In **Fig. 8**, while the majority of the mismatch values are between 30 and 50, the minimum mismatch value (i.e., 20) is found when the speed interval pattern of 21 weeks ago is considered. This indicates that traffic pattern on that day (21 weeks ago) is the most similar to the current pattern and, thus, could potentially be used in travel time prediction.

Table 5 Testing of pattern search parameters for travel time prediction on Link D

Set	P (No. of time intervals)			$\tilde{K}$ (No. of day)			I (Correlation level)		MAPE (%)	CPU time (seconds)
	1	3	5	12	24	48	1	2		
1	●			●			●		21.71	20.3
2		●		●			●		20.06	22.4
3			●	●			●		22.18	25.7
4		●			●		●		19.26	31.4
5		●				●	●		17.13	41.3
6		●				●		●	17.40	56.4

$$MAPE = \sum_{t=1}^{n} \frac{|Y_t - \hat{Y}_t|}{Y_t} \times 100$$

where Y_t is the actual travel time at time t, $\hat{Y}_t$ is the predicted travel time at time t, and n is the number of dataset.

Table 5 shows the impact of different pattern search parameters (P, I and $\tilde{K}$) on the performance of travel time prediction on Link D. In this table, the mean absolute percentage error (MAPE) is estimated by comparing the predicted and observed travel times in time interval 09:00 ~ 09:05 on 10 different Monday mornings. The algorithm was developed in Java and executed on a Linux operating system with 3.2 GHz quad-core processor. The CPU time in **Table 5** records the required time (in seconds) for carrying out the proposed algorithm in finding the predicted travel time. Comparing Set 1 and 2 in **Table 5**, it could be seen that the MAPE decreases as the number of temporally correlated time intervals (P) is increased from 1 to 3. Such a decrease suggests that the speed interval patterns of Link D are temporally correlated with the preceding 15 minutes (3-time intervals). Thus, setting P = 3 could better describe (i.e., more information) the current traffic condition for finding a closer (more similar) pattern in the historical database and, thus, resulted in a reduction in MAPE. For case P = 5 (i.e., Set 3), despite more preceding time intervals are used in forming the speed interval patterns, the MAPE is larger than that of the case P = 3. The reason for having such an increase in MAPE comes from the low correlation of the speed intervals in the two additional time intervals (i.e., t-4 and t-5) with the current one. Thus, including these two-time intervals in the speed interval pattern may result in an incorrect match that gives an inaccurate prediction of travel time. Comparing the CPU time of Set 1, 2 and 3, it could be seen that CPU time increases as the speed interval pattern are becoming more complex (i.e., a more preceding time interval is considered).

Considering the MAPE of Set 2, 4 and 5, it could be seen that as the search space increases (i.e., $\tilde{K}$ increases), prediction error reduces and CPU time increases. It is because, as the search space increases, the number of historical speed interval patterns will increase and, thus, there will be a higher chance to have a historical pattern that is similar to the current pattern (i.e., small or zero mismatch value) for improving the accuracy of travel time prediction (i.e., reduction in MAPE). Considering Set 5 and 6 in **Table 5**, despite the increase in computation time, the increase in the spatial connection level (i.e., I) could not improve the accuracy of travel time prediction on Link D (i.e., MAPE increases as I is increased from 1 to 2). Similar to the explanation for the impact of P, such the increase in MAPE could be explained by the non-correlation of speed intervals between the second-level links (L_2^D) and Link D.

Sum up, the proposed algorithm gives the best prediction of travel time for Link D when $P = 3$, $I = 1$ and $\tilde{K} = 48$. Thus, this set of searching parameters will be adopted in the remaining empirical studies of this paper. In this initial experiment, apart from the fact that $|M_0|$ is usually very small (mostly equal to 1), it is also found that speed interval patterns that give the minimum mismatch value (i.e., $m_l^k \in M_0$) will have similar accuracy on the travel time prediction. Thus, unless otherwise stated, the speed used for travel time prediction ($v_{l,t'}^{pred}$) will be taken as the average of speed patterns with minimum mismatch value (i.e., $v_{l,t'}^{pred} = v_{l,t'}^{avg}$) for remaining empirical studies. With the above set up, the proposed SIPM algorithm (Section 2.3) is adopted to predict travel times of the 4 links (Link A, B, C, and D) during 09:05 ~ 09:30 (5 intervals) on 10 different Monday mornings. Note that the travel time predictions are made during the time interval 09:00 ~ 09:05.

Table 6 shows the MAPE of the proposed SIPM algorithm and 3-interval moving average of speed (MA) for travel time prediction of Link A ~ D during 09:05 ~ 09:30. In **Table 6**, it could be seen that the performance of MA in travel time prediction decrease (i.e., MAPE increase) as the predicted time interval increases. It is because MA approach is directly depended on the speeds of the preceding time intervals for travel time predictions.

Table 6 Errors (MAPE) of link travel time predictions for 5 future time intervals (09:05 ~ 09:30)

Link	SIPM Future time interval					MA * Future time interval				
	1	2	3	4	5	1	2	3	4	5
A	13.21	12.10	13.85	14.61	14.20	12.85	14.55	16.03	21.20	22.45
B	14.50	15.22	13.92	14.68	15.32	12.90	12.17	15.46	19.25	21.10
C	16.18	16.39	16.45	17.46	15.84	15.41	20.94	23.88	25.42	29.57
D	18.20	19.23	18.69	21.40	20.22	22.45	26.40	32.10	30.56	34.20

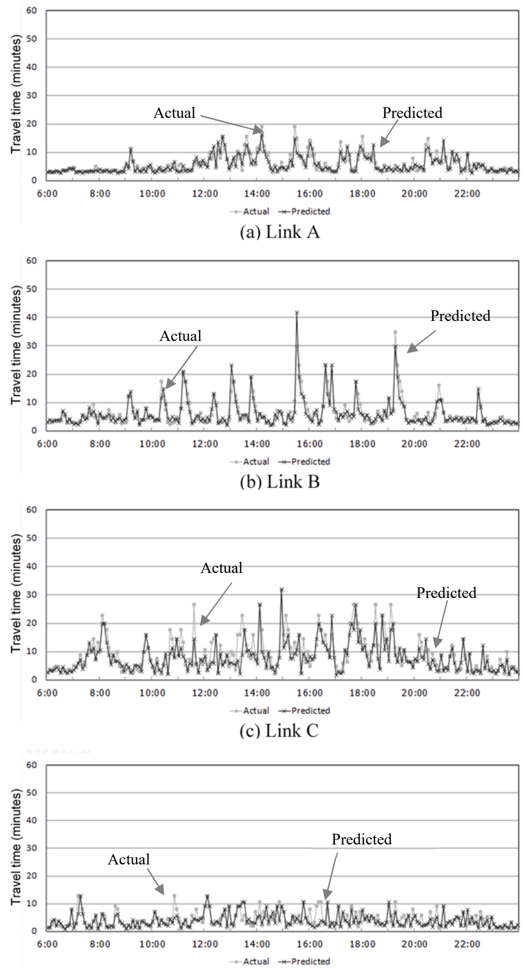
*MA = moving average of speed

Thus, error in the first few future time intervals (e.g., 09:05 ~ 09:10 in this test) will propagate to the subsequent predictions and cause a substantial increase in MAPE. For the SIPM algorithm, the MAPEs are in general smaller than that for the MA approach and they are not necessarily increased with the number of predicted future time intervals (**Table 6**). It is because in SIPM algorithm, speeds from preceding time intervals are only used in searching the historical database but not directly be used in predicting speeds and travel times. Thus, error in the first few predicted future time intervals will not directly propagate to the subsequent intervals to cause an increase in MAPE. Also, as the SIPM algorithm will independently search the historical database for travel time prediction in each of the future time intervals, it is possible to have a decrease in MAPE even when the number of future time interval increases (e.g., Link A 1st and 2nd future time interval).

Comparing the MAPEs of SIPM algorithm for travel time predictions of Link A to Link D, it could be seen that Link A has the smallest MAPEs among the four predicted links. It is because, as traffic on Link A is not interrupted (i.e., no weaving sections, on-ramp and off-ramp), the speeds along Link A will relatively stable and could be closely represented by the speeds collected at certain location on Link A. Thus, speed interval patterns that are similar to the current traffic condition could be found in the database to provide a more accurate prediction of future travel times. On the other hand, the MAPEs for Link C and D are higher than that of Link A and B. It is because the traffic

Signals on Link C and D cause extra delays (i.e., deceleration of vehicles and waiting at the stop lines) that could not be reflected by the collected speeds on that link. Thus, the speed interval pattern on that link may not truly reflect their actual traffic conditions and causes an increase in MAPE of the travel time predicted by SIPM algorithm.

Fig. 9 shows the variation of actual and predicted link travel times, which are calculated from the collected and predicted speeds for every 5 minutes, of the four links (Link A ~ D) on Monday, 26th December 2014, 06:00 – 24:00. In Fig. 9, it could be seen that the travel time predictions are relatively accurate in the morning and evening periods (i.e., 06:00 ~ 12:00 and 18:00 ~ 24:00), while most of the discrepancies (i.e., the differences between grey squares and black crosses) appear between 12:00 and 18:00. It is because during the morning and evening periods, as the links are highly congested with the peak period traffics, the traffic patterns (i.e., the spatial and temporal variation of flows) and the corresponding traffic characteristics (e.g., speed) are relatively stable. With such stable traffic characteristics, highly matched speed interval patterns could be found in the database for accurate prediction of travel times. On the other hand, as the travel patterns and traffic characteristics are largely varied during 12:00 ~ 18:00, prediction accuracies reduce in this period. Comparing



(a) Link A

(b) Link B

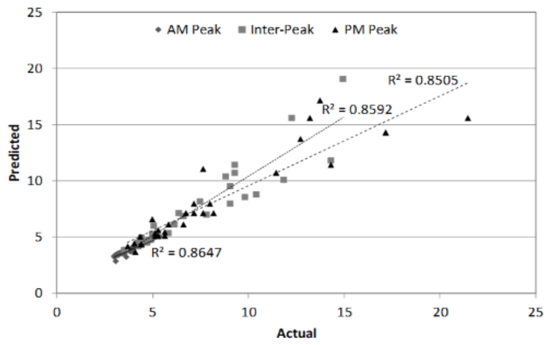
(c) Link C

(d) Link D

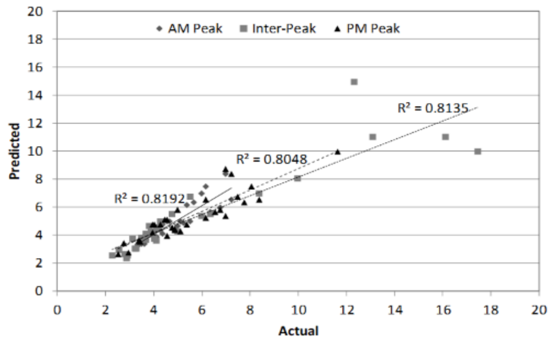
Fig. 9 Actual and predicted link (Link A ~ D) travel times on Monday, 26th December 2014, 06:00 ~ 24:00

the prediction errors in Link A/Link B with that in Link C/Link D, it could be seen that the errors are much larger for the predictions in Link C and D. It is because Link C and D are located in the central business district (CBD) with lots of disturbances to the traffic (e.g., pedestrian crossings, bus stops, on-street parking, etc). These disturbances will cause a large variety of speeds on these links and resulted in reduction in prediction accuracies.

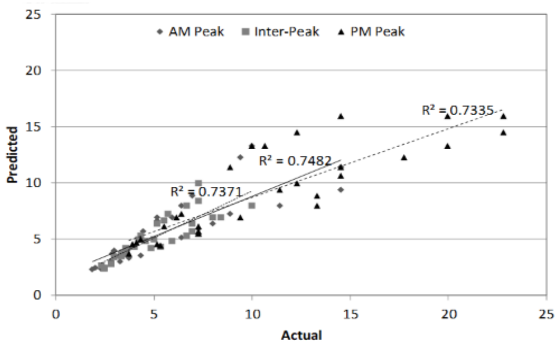
In **Fig. 10**, predicted travel times of the four links (Link A ~ D), which are predicted by the proposed SIPM algorithm, are plotted against their corresponding actual travel time. Based on the typical travel pattern in Bangkok, three representative peak periods: AM peak (09:00 ~ 09:05), inter-peak (14:00 ~ 14:05) and PM peak (18:00 ~ 18:05) are considered in this empirical study. In this figure, 30 sets of data are predicted for (collected from) each period as AM peak, inter-peak and PM peak, and the corresponding coefficient of determination (R^2) is estimated. In this study, a large R^2 value between the predicted and actual travel times indicates a high forecasting accuracy of the proposed SIPM algorithm. Considering the R^2s in Fig. 10, it could be seen that links with stable flow conditions (i.e., Link A and Link B with R^2 between 0.80 and 0.86) have a higher forecasting accuracy than links under interrupted flow conditions (i.e., Link C and Link D with R^2 between 0.64 and 0.75). Comparing among the AM peak and PM peak, it could be seen that the forecasting accuracy for PM peak is, in general, slightly lower than that for the AM peak.



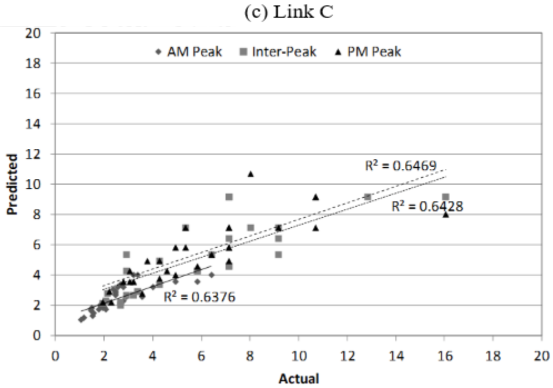
(a) Link A



(b) Link B



(c) Link C



(d) Link D

Fig. 10 Performance of link (Link A ~ D, respectively), travel time prediction by SIPM algorithm

As the start time of work is quite similar (e.g., 09:00) for all travellers, the link volumes in AM peak, which are mainly home-based work trips, are relatively stable and give a better forecasting accuracy. On the other hand, as the PM travel demands are spread out for a longer period of time (e.g., people try to avoid the traffic by shifting their time to go home or go to other places first), the link volumes in PM peak is less stable than that in the AM peak and causes a reduction in forecasting accuracy. For the inter-peak, as the travel demand is mainly based on the non-recurrent daily events (e.g., delivery of packages by courier), the link volumes and, thus, forecasting accuracies for this period are substantially varied. To sum up, the proposed SIPM algorithm gives a high forecasting accuracy on link travel time under stable traffic conditions that could either be achieved by uninterrupted traffic flow and/or congested travel conditions.

3.2 Path travel time prediction

With the testing of SIPM algorithm in travel time prediction on single links (Section 3.1), this section focuses on the more practical use of the algorithm on path travel time prediction. In this study, four different paths, which are all formed by a series of links, are tested: 1) Path E is a path with free-flow condition on a multi-lane road without weaving, on- and off-ramp (similar to Link A); 2) Path F is a path with free-flow condition on a local side road connected to community areas (similar to Link B); 3) Path G, which contains Link C, is arterial connected to business areas, and; 4) Path H is path in CBD with traffic signals and U-turns (similar to Link D). The characteristics of these paths are shown in **Table 7** and their geographical locations are shown in **Fig. 6.**

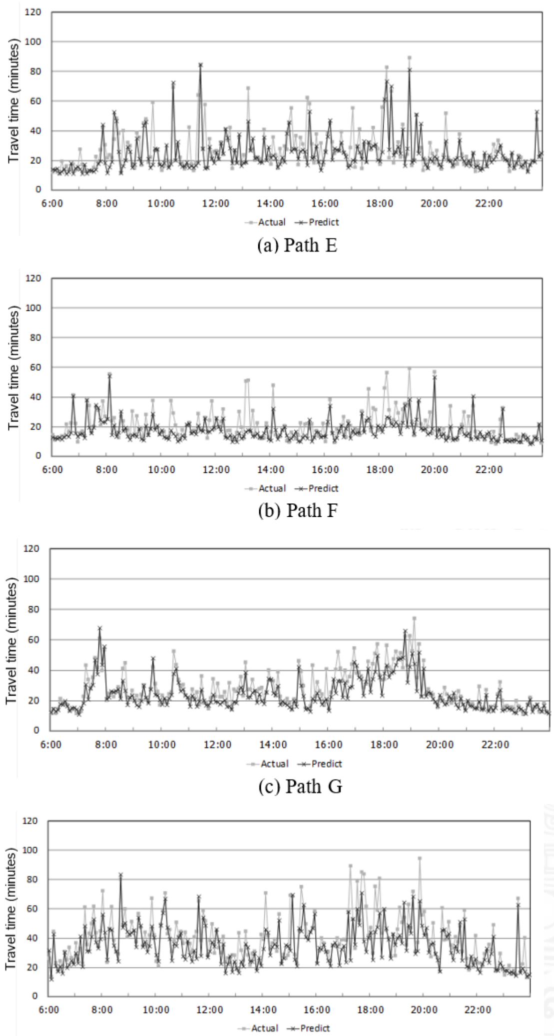
Table 7 Paths for travel time prediction

Path	From / To	Direction*	Links	Distance (km)
E	Bang Khru – Bang Pli	Inbound	4	19.7
F	Minburi – Watcharaphol	Inbound	4	9.8
G	Ratchada 4 – Wongsawang	Outbound	6	12.0
H	BTS Nana – Udom Suk	Outbound	5	9.4

*Inbound (outbound) refers to traffic travelling towards (out of) the centre of Bangkok city

To predict path travel times, the proposed algorithm will first predict the link travel time of each of the links within the network for the next time interval (e.g., 09:05 ~ 09:10). Then, based on the predicted departure time of the first links of the considered path, the speed interval pattern that is used for travel time prediction of the second link is generated. For example, if the predicted departure time of the first link is 09:11, the speed interval pattern for travel time prediction in the second link will be generated by taking 09:05 ~ 09:10 as the current time interval, t. The above steps will be repeated for all links within the considered paths and path travel times could be calculated by summing up the predicted travel time of all constituent links.

Fig. 11 shows the variation of actual and predicted path travel times, which are calculated from the collected/predicted speeds for every 5 minutes, of the four paths (Path E ~ H) on Monday, 26th December 2014, 06:00 – 24:00. Owing to the same reasons discussed for link travel time prediction, the majority of discrepancies between actual and predicted path travel times are concentrated in the period 12:00 ~ 18:00. Similarly, Path G and H, which are within CBD, tend to have lower prediction accuracies due to the frequent disturbances of traffic flows along these paths.



(a) Path E

(b) Path F

(c) Path G

(d) Path H

Fig. 11 Actual and predicted path (Path E~ H) travel times on Monday, 26th December 2014, 06:00 ~ 24:00

Comparing the prediction errors of link and path travel time (i.e., the difference between the actual and predicted travel times in **Fig. 9** and **11**), it could be seen that the prediction of path travel time is less accurate and start to have significant errors in morning/evening period. The main reason for having such reductions inaccuracy comes from accumulation of errors in link travel time prediction for each of the links within the path considered.

In order to have a more detailed analysis of errors in path travel time predictions, **Table 8** shows the mean absolute error (MAE) and MAPE of the four paths in AM period (07:00 ~ 09:00), Midday period (13:00 ~ 15:00) and PM period (17:00 ~ 19:00). Comparing the MAPE of each of the paths for different periods, it could be seen that the AM periods have the smallest error while the midday periods have the highest.

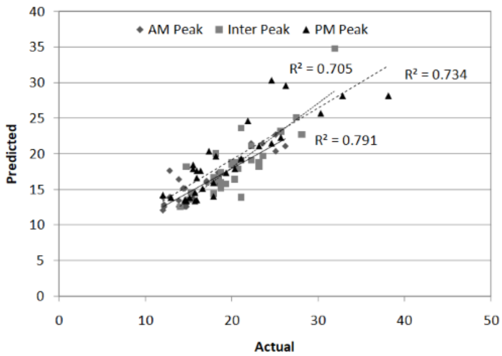
Table 8 Errors of path travel time prediction

Path	Period*	Time	MAE	MAPE
	AM period	16.26	1.75	10.53
E	Midday period	20.52	2.76	13.33
	PM period	19.77	2.63	12.62
	AM period	13.87	1.74	12.02
F	Midday period	14.02	2.47	16.81
	PM period	17.06	2.45	13.63
	AM period	19.96	3.18	15.88
G	Midday period	23.17	5.31	23.51
	PM period	40.62	8.00	19.35
	AM period	44.48	11.06	21.56
H	Midday period	34.48	10.51	30.91
	PM period	53.65	16.85	29.43

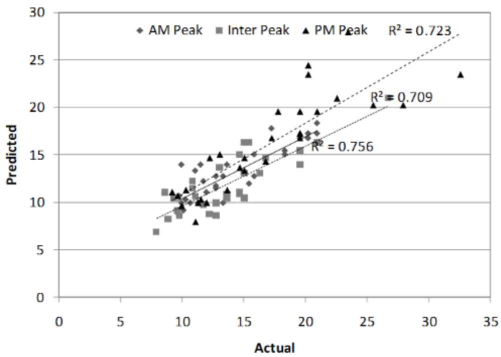
*AM period: 07:00~09:00; Midday period: 13:00~15:00; PM period: 17:00~19:00

$$MAPE = \sum_{t=1}^{n} \frac{\left|Y_t - \hat{Y}_t\right|}{Y_t} \times 100$$

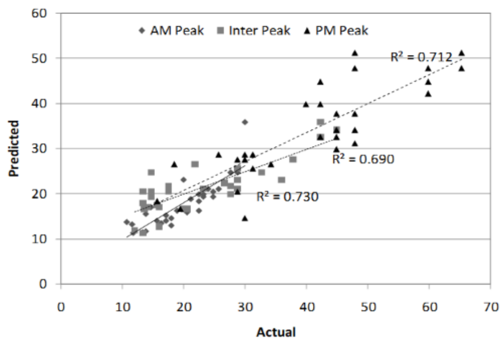
As discussed in Section 3.1, the low MAPE in AM period could be explained by the stable flows, which comes from the concentration of demand during the AM



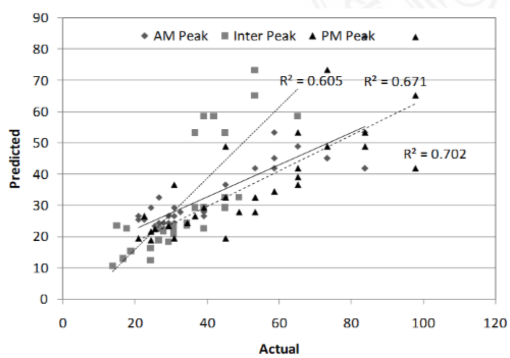
(a) Path E



(b) Path F



(c) Path G



(d) Path H

Fig. 12 Performance of path (Path E~H) travel time prediction by SIPM algorithm

period, of the predicted path. For the PM periods, as the demand is spread out for a longer period of time, path flows are less stable than the AM periods and, thus, give a slightly higher MAPE. Lastly, the high MAPEs in midday periods are resulted from the unpredictable path volumes (e.g., business trips) at this time of the day. Thirty sets of path travel time data (predicted and actual) from each of the period (AM peak, inter-peak and PM peak) are plotted in **Fig. 12** with the corresponding R^2 estimated. Comparing the R^2s for different time of the day and paths, it could be seen that performance of the proposed SIPM algorithm in path travel time prediction is similar to that in link travel time prediction (Section 3.1).

4. Conclusions

In this paper, the proposed system is an extension of the real-time travel time estimation system in BAL-Labs. [18] and K. Sringiwai et al. [19] to path travel time predictions for effective route planning in Bangkok. For predicting path travel times, this paper has proposed an algorithm for the matching of speed interval pattern (i.e., SIPM algorithm), which has included speeds (or speed intervals) that are spatially and temporally correlated to the current link speed, of the current traffic conditions with those in the historical database. Historical speed interval pattern with the minimum mismatch value, which is determined by the sum of absolute difference of the speed intervals as compared to the current pattern, is considered to be the matched traffic condition and the speeds in the subsequent time intervals of that time and day will be used for travel time prediction. For cases with multiple historical speed interval patterns that have achieved the same minimum mismatch value, this paper has proposed three different methods (i.e., random sample, average value, and closest speed pattern) in determining the speed pattern used for travel time prediction.

To demonstrate the efficiency of the SIPM algorithm and the performance of corresponding travel time prediction system, empirical tests are carried out on four different links and paths in Bangkok. In these tests, it is found that the optimal speed interval pattern, which gives the minimum MAPE in link travel time prediction, should include 3 preceding time intervals (*P*) and links in the 1st connection level (*I*). Apart from the optimal speed interval pattern, these tests also show that the proposed algorithm for travel time predictions under uninterrupted flow conditions (MAPE: 12% ~ 15% for links and 11% ~ 17% for paths) are more accurate than under the interrupted flow conditions in central business areas (MAPE: 16% ~ 21% for links and 16% ~ 31% for paths).

Concerning the proposed real-time path travel time prediction algorithm, there are four directions of future work. First, the development of multi-threaded processing of the proposed algorithm should be completed to further reduce the computational time for practical implementation in real-time route planning system in Bangkok. Second, an adaptive traffic data fusion algorithm should be developed to make use of other potential sources of data (e.g., cell phone probes, GPS, traffic flow data, etc) for improving the accuracy of travel time prediction. Third, location-dependent speed interval pattern should be considered to improve the performance of travel time prediction on arterials and/or in congested areas. Lastly, this study aimed to capture parameters under uninterrupted flow conditions. The main effective variable is speed. The study of other variables and methods would make the prediction more accurate and would be included in the future study.

5. Acknowledgements

This research was supported by the Rajamangala University of Technology Phra Nakhon (RMUTP).

6. Reference

[1] H.-E. Lin, R. Zito and M. Taylor, “A review of travel-time prediction in transport and logistica,” in *Proceedings of the Eastern Asia Society For Transportation Studies*, 2005, vol. 5, pp. 1433 – 1448.

[2] N. K. Chowdhury, R. P. D. Nath, H. Lee and J. Chang, “Development of an Effective Travel Time Prediction Method Using Modified Moving Average Approach,” in *Knowledge-Based and Intelligent Information and Engineering Systems*, Berlin, Heidelberg, 2009, pp. 130–138.

[3] W.H.K. Lam, K.S. Chan, M.L. Tam and J.W.Z. Shi, “Short-term travel time forecasts for transport information system in Hong Kong,” *Journal of Advanced Transportation*, vol. 39, no. 3, pp. 289-306, 2005.

[4] S. I. Bajwa, E. Chung and M. Kuwahara, "An Adaptive Travel Time Prediction Model Based on Pattern Matching," *11th Intelligent Transportation Systems World Congress*, Nagoya, Japan, 2004.

[5] T. Kim, H. Kim and D. J. Lovell, "Traffic flow forecasting: overcoming memoryless property in nearest neighbor non-parametric regression," in *Proceedings. 2005 IEEE Intelligent Transportation Systems, 2005*, Sep. 2005, pp. 965–969.

[6] Z.-P. Li, H. Yu, Y.-C. Liu and F.-Q. Liu, "An Improved Adaptive Exponential Smoothing Model for Short-term Travel Time Forecasting of Urban Arterial Street," Acta Automatica Sinica, vol. 34, no. 11, pp. 1404–1409, Nov. 2008.

[7] W.-H. Lee, S.-S. Tseng and S.-H. Tsai, "A knowledge based real-time travel time prediction system for urban network," Expert Syst. Appl., vol. 36, pp. 4239–4247, Apr. 2009.

[8] Y. Zhang and Y. Liu, "Comparison of Parametric and Non-Parametric Techniques for Non-Peak Traffic Forecasting." *International Journal of Mathematical and Computational Sciences*, vol. 3, no. 3, pp. 8-14, 2009.

[9] A. Khosravi, E. Mazloumi, S. Nahavandi, D. Creighton and J. W. C. Van Lint, "Prediction Intervals to Account for Uncertainties in Travel Time Prediction," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 2, 2011, pp. 537-547.

[10] A. Simroth and H. Zähle, "Travel Time Prediction Using Floating Car Data Applied to Logistics Planning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 1, pp. 243–253, Mar. 2011.

[11] W. Qiao, A. Haghani and M. Hamedi, "A Nonparametric Model for Short-Term Travel Time Prediction Using Bluetooth Data," *Journal of Intelligent Transportation Systems*, vol. 17, no. 2, pp. 165–175, Apr. 2013.

[12] Y. Zou, X. Zhu, Y. Zhang and X. Zeng, "A space–time diurnal method for short-term freeway travel time prediction," *Transportation Research Part C: Emerging Technologies*, vol. 43, pp. 33–49, Jun. 2014.

[13] H. Jiang, Y. Zou, S. Zhang, J. Tang and Y. Wang, "Short-Term Speed Prediction Using Remote Microwave Sensor Data: Machine Learning versus Statistical Model," *Mathematical Problems in Engineering*, vol. 2016, p. e9236156, Feb. 2016.

[14] W.H.K. Lam, M.L. Tam, A. Sumalee, C.L. Li, W. Chen, S.K. Kwok, Z. Li and E.W.T. Ngai, " Incident detection based on short-term travel time forecasting," in *Proceedings of the 13th International Conference of Hong Kong Society for Transportation Studies: Transportation and Management Science*, 2008, p. 83-92.

[15] J.W.C. Van Lint, S. P. Hoogendoorn and H.J. van Zuylen, "Accurate Travel time Prediction with State-Space Neural Networks under Missing Data,"

Transportation Research. Part C, vol. 13, pp. 347-369, 2005.

[16] T. Siripirote, A. Sumalee, D. P. Watling and H. Shao, "Updating of Travel Behavior Model Parameters and Estimation of Vehicle Trip Chain Based on Plate Scanning," *Journal of Intelligent Transportation Systems*, vol. 18, no. 4, pp. 393–409, Oct. 2014.

[17] K. P. Hwang, W.-H. Lee and W.-B. Wu, "Travel time prediction by weighted fusion of probing vehicles and vehicle detectors data sources," in *2012 12th International Conference on ITS Telecommunications*, Nov. 2012, pp. 476–481.

[18] BAL-Labs. (2010). "Siamtraffic1.0: Real-Time Travel Time Estimation." BAL-Labs. [Online]. Available: http://bal.buu.ac.th/bal2010/?q=siamtraffic1.0

[19] K. Sringiwai, N. Indra-Payoong and A. Sumalee, "RFID-based travel time estimation : development case in Bangkok," in *Proceeding of International Conference of Hong Kong Soceity for Transportation Studies*, 2010, pp. 139-146.