Short-Term Traffic Prediction Using Rainfall

Haiyun Lu
SAP Research and Innovation, Singapore
Email: hai.yun.lu@sap.com

***Abstract*—We propose an approach to predict short-term travel time for expressways using rainfall data. The rainfall rate is quantitatively related to the travel time and is used to perform the prediction. The proposed approach is experimented on real data collected in Singapore. Results show that it performs better compared to three baseline approaches. It is also shown that the proposed approach is applicable and effective in Singapore.**

***Index Terms*—intelligent transportation systems, traffic prediction, travel time prediction, weather correlation, impulse response.**

I. Introduction

Traffic forecasting is an important component in an intelligent transportation system in the concept of smart city. Since decades ago, research effort has been devoted to various traffic prediction models and methods.

The traffic prediction problem can be classified into two categories in terms of time scale: long-term and short-term [1]. Long-term prediction targets for monthly or even yearly information of traffic states and it is used for long-term transportation planning. Short-term prediction aims for the near future, such as 15 minutes later. It can be used by experts to guide traffic flow and manage congestion. It may also be available to common citizens of the city to help them plan their trips wisely. By definition, short-term traffic forecasting is the process to predict key traffic parameters such as speed, flow, occupancy, or travel time with a forecasting horizon typically ranging from five to thirty minutes at specific locations, given real-time and historical traffic data from relevant surveillance stations [2].

Traffic prediction methods often employ statistical methodologies and there are two major types of models: parametric model and non-parametric model. Parametric models include random walk, historical average [3], linear regression [4], Kalman filter [5], [6], autoregressive moving average [7]-[9]. These methods derive a mathematical function between history states and predicted states. Non-parametric models include neural networks [10]-[12], Bayesian networks [13], K-nearest neighbor [14], [15]. These methods analyze characteristics of traffic data and approximate the model with a certain precision using a growing dataset.

Other than historical traffic data, weather information is also an important factor of traffic forecasting. Various studies show that bad weather conditions, such as rain, snow, fog, ice, flooding, wind and high temperature, generally result in more accidents on roads [16]-[19]. It is also shown that heavy precipitation conditions have impact on traffic speed, capacity, volume, intensity, flow and travel time [20]-[27]. In particular, rain condition is reported to cause speed reduction by 6% to 12% in different locations [23], [27]. Heaving rainfall condition decreases the visibility and causes wet surface on roads, so road users will slow down their vehicles in order to drive safely. Given such correlation between rainfall and traffic, it is important to include rainfall in the traffic prediction system for rainy places.

In this paper, we propose to predict short-term travel time for expressways using rainfall data. Travel time on expressways in Singapore is used for this study. Singapore has a tropical rainforest climate where rainfall is very frequent throughout the year. Thus the proposed traffic prediction approach is ideally applicable to the data obtained in Singapore.

II. Proposed Approach

In our approach, travel time on the expressway is predicted using rainfall data. We build our solution based on one assumption, that is, there is a quantitative relationship between rainfall rate and traffic condition. As mentioned previously, when there is a heavy rain, drivers will slow down to keep it safe. This is generally considered as a causal correlation. Fig. 1 shows the plot of rainfall rate and travel time for a segment of expressway on a typical weekday. We can see that travel time increases at morning and evening peak hours, which is generally expected. However, it also increases around 15:00 (off peak hour) on that day, which is overlapping with a raining period around the same time.

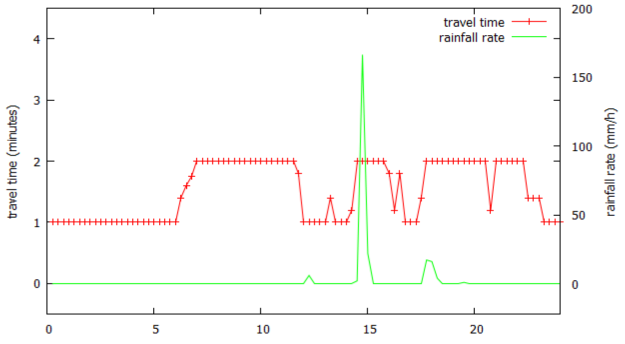

Figure 1. Rainfall rate and travel time.

Manuscript received February 7, 2014; revised April 25, 2014.

doi: 10.12720/ijsps.2.1.70-73

To demonstrate the impact of the rainfall on traffic, we constructed the cross-correlations function between travel time and rainfall rate, as in (1), where T is 24 hours, t is time, $\delta(t)$ is the deviation from normal travel time and $r(t)$ is the rainfall rate.

$$R_{\delta r} = \frac{1}{T}\int_0^T \delta(t)\, r(t+\tau)\, dt \tag{1}$$

The deviation $\delta(t)$ is computed as in (2), where $y(t)$ is the travel time at time t and $\bar{y}(t)$ is the historical average of the travel time.

$$\delta(t) = y(t) - \bar{y}(t) \tag{2}$$

The peak value of the cross-correlation is located at the delay τ necessary to align the two time series. Fig. 2 shows a histogram of the location of the largest peak found in the cross-correlation function. The largest peak is at delay 0-15 minutes. It implies that the rainfall has a near-immediate impact on the travel time.

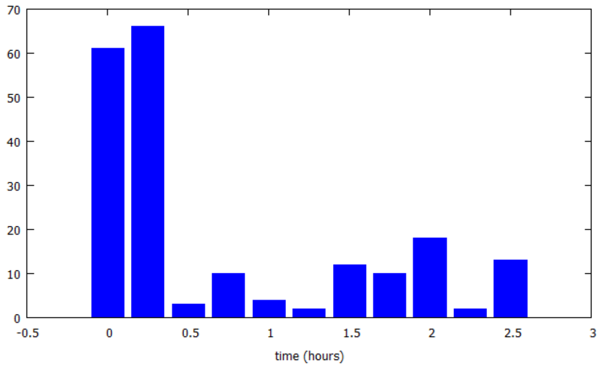

Figure 2. Histogram of location of the largest peak in cross-correlation of travel time and rainfall rate.

Given the quantitative relationship between rainfall and traffic, we model the travel time as a linear system adapted from Dailey's work [28]. Equation (3) of the system is with two components: the normal travel time and the contribution from the rainfall. The normal travel time $\bar{y}(t)$ is the historical average of travel time at time t of the day. The rainfall contribution is approximated by convolving the impulse response function $h(\tau)$ with the rainfall rate $r(t)$.

$$y(t) = \bar{y}(t) + \int h(\tau)\, r(t-\tau)\, d\tau \tag{3}$$

The impulse response function $h(\tau)$ needs to be derived from training data. To compute h, we re-write the above equation and apply Fourier transform:

$$\delta(t) = \int h(\tau)\, r(t-\tau)\, d\tau \tag{4}$$

$$\Delta(f) = H(f)R(f) \tag{5}$$

Δ is the Fourier transform of the travel time deviation δ, H is the Fourier transform of the impulse response function h, and R is the Fourier transform of the rainfall rate r. After multiplying the complex conjugate value of R, we can rewrite the equation to compute H.

$$\Delta R^* = HRR^* \tag{6}$$

$$H \approx \frac{G_{\Delta R^*}}{G_{RR^*}} \tag{7}$$

Then the impulse response function h can be approximated by the inverse Fourier transform of H, where $G_{\Delta R^*}$ and G_{RR^*} are the power spectrum. Fig. 3 shows a typical impulse response function for an expressway.

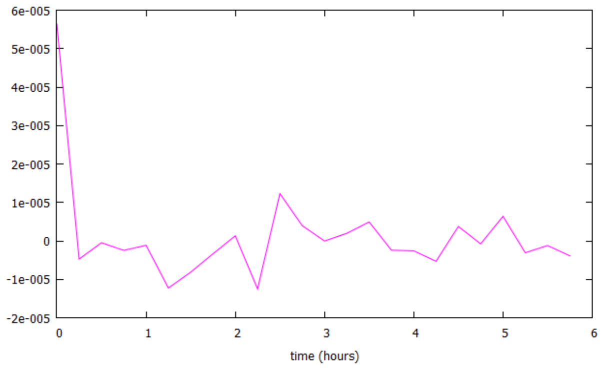

Figure 3. Impulse response function.

After obtaining the impulse response function, travel time can be predicted by the linear system model. By adding weight parameter α, the system, as in (8), is more tolerable to false contributions due to short or light rain.

$$y(t) = \bar{y}(t) + \alpha \int h(\tau)\, r(t-\tau)\, d\tau \tag{8}$$

$$\alpha = \begin{cases} 0, & \text{if } \bar{r} \leq \omega_1 \\ 0.5, & \text{if } \bar{r} > \omega_1 \text{ and } \bar{r} \leq \omega_2 \\ 1, & \text{otherwise.} \end{cases} \tag{9}$$

where $\bar{r}$ is the average of the rainfall rates used in the convolution, ω_1 and ω_2 are empirically determined thresholds.

In the next section, the proposed approach is applied on real data collected in Singapore.

III. Experiment

A. Data Collection

Travel time data with 5 minutes interval are collected for eight expressways (AYE, BKE, CTE, ECP, KJE, PIE, SLE, and TPE) from September 2013 to February 2014. Travel time is measured between consecutive exits of the expressway. In total there are 183 segments of the expressway. The travel time data are published by the Land Transport Authority of Singapore on public website (mytransport.sg). Fig. 4 shows historical average of travel time for a typical expressway segment on weekday.

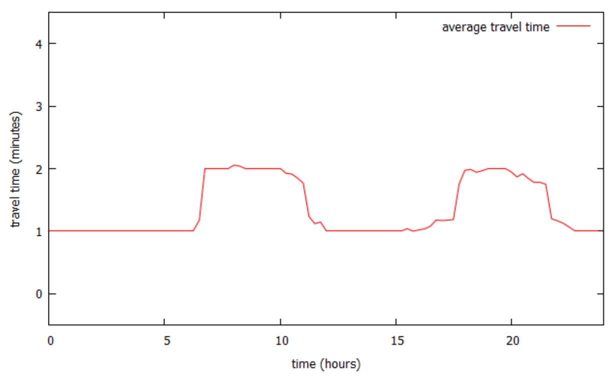

Figure 4. Average travel time on weekdays.

Rainfall data are collected for the same time period as travel time data. They are published by the National Environment Agency of Singapore on their public website (app2.nea.gov.sg). The rainfall data are obtained from weather radar, derived into image visualization (see Fig. 5) and published at every 5-10 minutes interval.

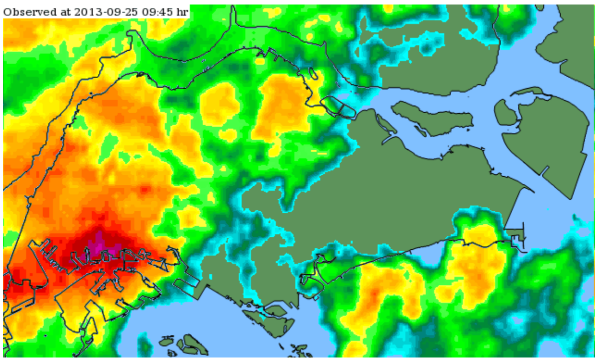

Figure 5. Rain map published on app2.nea.gov.sg.

Rainfall rate data have to be reversely derived from the image visualization using the Doppler radar reflectivity:

$$r = a\left(10^{\frac{d}{10}}\right)^{b} \tag{10}$$

where r is the rainfall rate, d is the reading taken from the image visualization, a =0.097 and b =0.997 are empirically determined parameters. The rainfall rate corresponding to a particular expressway segment is approximated by the average of rainfall rates in the area of this segment.

B. Results and Comparison

To predict the traffic, the impulse response function needs to be derived from training data. We use the data collected from September to November 2013 as training data. Impulse response function is computed for each segment of the expressways.

After learning the impulse response functions, we apply the prediction approach to predict travel time given the rainfall data from December 2013 to February 2014. The prediction interval is chosen to be 15 minutes, which is the maximum interval presented in the data collected. The results are measured by the mean absolute percentage error (MAPE) and the root mean square error (RMSE):

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\frac{|y_t - \hat{y}_t|}{y_t} \tag{11}$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{N}(y_t - \hat{y}_t)^2}{N}} \tag{12}$$

where $\hat{y}_t$ is predicted travel time at time t and y_t is actual travel time at time t.

Table I lists the final results averaged from the experiment results for all 183 expressway segments. The results are compared to three baseline prediction approaches mentioned in the studies [8], [9]. The first baseline approach models the traffic as random walk, then the forecast for next state is simply the most recent state, that is, $\hat{y}_{t+1} = y_t$. The second baseline approach predicts the next state by using previously observed historical average of the travel time, that is, $\hat{y}_t = \bar{y}_t$. The last baseline approach calculates the prediction with a smoothing parameter, that is, $\hat{y}_{t+1} = \mu y_t + (1-\mu)\hat{y}_t$ where μ=0.2. These baseline approaches are applied to the same dataset and evaluated using the same measure.

TABLE I. EXPERIMENT RESULTS

Approach	MAPE	RMSE
Random walk forecast	5.372%	0.383
Historical average forecast	7.357%	0.435
Smoothed historical average forecast	5.347%	0.394
Our proposed approach	**4.669%**	**0.339**

From the comparison, it is clear that the proposed approach achieves better experiment results. Fig. 6 shows two examples of the predicted travel time versus the actual travel time plotted for one day. It shows that the increase of travel time can be effectively predicted for raining periods at off peak hours.

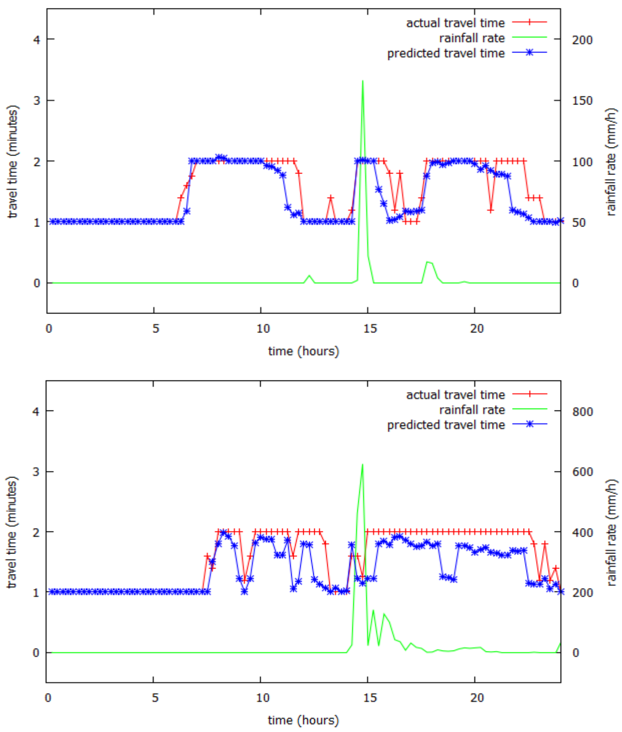

Figure 6. Predicted travel time versus actual travel time.

IV. CONCLUSION

In this paper, we presented an approach to predict short-term travel time for expressways using rainfall data. We use an impulse response function to quantitatively relate the rainfall rate to travel time and use a weighted linear system to perform the prediction. The experiment is conducted on real data collected in Singapore. The results show that the proposed approach achieves lower errors compared to three baseline approaches. It is also shown that the proposed approach is applicable and effective in Singapore.

However, there is still room for further research. For example, some parameters are empirically determined. It could be a challenging task to find suitable values for those parameters in practice. Another issue is that the proposed model may be only applicable to rainy places. The prospect is to combine this model with others to enhance the flexibility and robustness.

REFERENCES

[1] S. M. Hashemi, M. Almasi, R. Ebrazi, and M. Jahanshahi, "Predicting the next state of traffic by data mining classification techniques," *International Journal of Smart Electrical Engineering*, vol. 1, no. 3, pp. 180-192, 2012.

[2] J. Zhu, "A self-learning short-term traffic forecasting system through dynamic hybrid approach," Ph.D. dissertation, The University of Hong Kong, 2007.

[3] Y. J. Stephanedes, P. G. Michalopoulos, and R. A. Plun, "Improved estimation of traffic flow for real-time control," *Transportation Research Record*, vol. 795, pp. 28-39, 1981.

[4] X. Zhang and J. Rice, "Short-term travel time prediction," *Transportation Research Part C*, vol. 11, no. 3, pp. 187-210, 2003.

[5] M. Chen and S. I. J. Chien, "Dynamic freeway travel-time prediction with probe vehicle data - link based versus path based," *Transportation Data and Information Technology*, vol. 1768, pp. 157-161, 2001.

[6] A. Stathopoulos and M. G. Karlaftis, "A multivariate state space approach for urban traffic flow modeling and prediction," *Transportation Research Part C*, vol. 11, no. 2, pp. 121-135, 2003.

[7] M. van der Voort, M. Dougherty, and S. Watson, "Combining Kohonen maps with ARIMA time series models to forecast traffic flow," *Transportation Research Part C*, vol. 4, no. 5, pp. 307-318, 1996.

[8] B. Williams and L. Hoel, "Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results," *Journal of Transportation Engineering*, vol. 129, no. 6, pp. 664-672, 2003.

[9] G. Yu and C. Zhang, "Switching ARIMA model based forecasting for traffic flow," in *Proc. Acoustics, Speech, and Signal Processing*, 2004, pp. 429-432.

[10] D. Park and L. R. Rilett, "Forecasting freeway link travel times with a multilayer feed forward neural network," *Computer-Aided Civil and Infrastructure Engineering*, vol. 14, pp. 357-367, 1999.

[11] H. M. Zhang, "Recursive prediction of traffic conditions with neural network models," *Journal of Transportation Engineering*, vol. 126, no. 6, pp. 472-481, 2000.

[12] M. Zhong, S. Sharma, and P. Lingras, "Analyzing the performance of genetically designed short-term traffic prediction models based on road types and functional classes," *Lecture Notes in Computer Science*, vol. 3029, pp. 1133-1145, 2004.

[13] W. Zheng, D. Lee, and Q. Shi, "Short-term freeway traffic flow prediction: Bayesian combined neural network approach," *Journal of Transportation Engineering*, vol. 132, no. 2, pp. 114-121, 2006.

[14] R. K. Oswald, W. T. Scherer, and B. L. Smith, "Traffic flow forecasting using approximate nearest neighbor nonparametric regression," Center for Transportation Studies, University of Virginia, 2001.

[15] R. E. Turochy, "Enhancing short-term traffic forecasting with traffic condition information," *Journal of Transportation Engineering*, vol. 132, no. 6, pp. 469-474, 2006.

[16] P. J. Codling, "Weather and road accidents," in *Climatic Resources and Economic Activity*, J. A. Taylor, Ed. Newton Abbot: David and Charles Holdings, 1974, pp. 205-222.

[17] L. Fridstrøm, J. Ifver, S. Ingebrigtsen, R. Kulmala, and L. K. Thomsen, "Measuring the contribution of randomness, exposure, weather, and daylight to the variation in road accident counts," *Accident Analysis & Prevention*, vol. 27, no. 1, pp. 1-20, 1995.

[18] J. B. Edwards, "The temporal distribution of road accidents in adverse weather," *Meteorological Applications*, vol. 6, pp. 59-68, 1999.

[19] D. Eisenberg, "The mixed effects of precipitation on traffic crashes," *Accident Analysis and Prevention*, vol. 36, no. 4, pp. 637-647, 2004.

[20] K. K. Knapp, and L. D. Smithson, "Winter storm event volume impact analysis using multiple-source archived monitoring," *Transportation Research Record*, vol. 1700, pp. 10-16, 2000.

[21] K. Keay, and I. Simmonds, "The association of rainfall and other weather variables with road traffic volume in Melbourne, Australia," *Accident Analysis & Prevention*, vol. 37, no. 1, pp. 109-124, 2005.

[22] T. H. Maze, M. Agarwai and G. Burchett, "Whether weather matters to traffic demand, traffic safety, and traffic operations and flow," *Transportation Research Record*, vol. 1948, pp. 170-176, 2006.

[23] H. Rakha, M. Farzaneh, M. Arafeh and E. Sterzin, "Inclement weather impacts on freeway traffic stream behavior," *Transportation Research Record*, vol. 2071, pp. 8-18, 2008.

[24] M. J. Koetse, and P. Rietveld, "The impact of climate change and weather on transport: An overview of empirical findings," *Transportation Research Part D: Transport and Environment*, vol. 14, no. 3, pp. 205-221, 2009.

[25] F. J. Camacho, A. Garcia, and E. Belda, "Analysis of impact of adverse weather on freeway free-flow speed in Spain," *Transportation Research Record*, vol. 2169, pp. 150-159, 2010.

[26] M. Cools, E. Moons and G. Wets, "Assessing the impact of weather on traffic intensity," *Weather, Climate, and Society*, vol. 2, no. 1, pp. 60-68, 2010.

[27] D. Akin, V. P. Sisiopiku, and A. Skabardonis, "Impacts of weather on traffic flow characteristics of urban freeways in Istanbul," *Procedia-Social and Behavioral Sciences*, vol. 16, pp. 89-99, 2011.

[28] D. J. Dailey, "The Use of Weather Data to Predict Non-recurring Traffic Congestion," Technical Report, Washington State Transportation Center, University of Washington, Sep. 2006.

Haiyun Lu obtained PhD in Computer Science from National University of Singapore in 2012, and Bachelor of Computing from the same university.
She is currently a Researcher with SAP Research & Innovation in Singapore. Her research interests include computer vision and machine learning.