Investigation of The Effect of Intersection Spacing On the Efficiency of Queue-Length Estimation from Color-Code Traffic Data

Rattaphol Pueboobpaphan*, Krittiporn Phungsombat, Natkitta Thansap, Wanwisa Naemkhuntod, Panittha Sanusit, Natapong Suwanamuk, Krittayot Torkoksung, Pongpat Jodnok, Promporn Sornsoongnern, and Suthatip Pueboobpaphan

Abstract**—** Queue length is an important input for traffic signal priority systems for emergency vehicles. Most research uses data from detectors to estimate queue lengths. In this paper, we propose a new concept by developing a Random Forest Model to estimate queue length from the color-code traffic data. These data are now available from online map service providers such as Google Maps. The proposed method could be a cost-effective alternative to cities that have not yet installed detectors. However, spacing between signalized intersections may be one of the factors that can affect the dynamics of traffic conditions between intersections, which consequently affects the queue length. This research aims to assess the effect of intersection spacing on the queue-length estimation efficiency of the proposed method. Microscopic traffic simulation is employed to model networks with spacing between signalized intersections varying as 200, 300, 500, 800, 1000, 3000, and 5000 meters. The independent variable is the traffic color-code information similar to those provided by Google Maps, while the dependent variable is the actual queue length. The results show that the developed model can detect the changes in queue length over time and the distance between the intersections does not affect the queue length estimation.

Index Terms**—** Queue length, random forest model, microscopic traffic simulation, color-code traffic, Google Maps

I. Introduction

Traffic congestion at the intersection is one of the main problems for road users, including emergency vehicles that need speedy travel but sometimes are stuck at signalized intersection for a long time. One approach that can help in resolving this issue is using intelligent traffic signal system to release vehicles stuck in the queue at the intersection to be able to pass through the intersection before the emergency vehicle arrives. Such system requires important input data such as queue length at the intersection. Research from many countries has used detectors to count traffic and estimate queue length. The most commonly used method for determining the variation in queue length is shock wave theory and data from detectors installed at the intersection are used to calculate the shock wave speed [1],[2],[3]. But in Thailand, most intersections do not use detectors to count traffic, thus making it impossible to estimate the queue length from detector.

However, Jodnok [4] proposes an alternative method for estimating queue length without using detector data, but instead using color-code traffic data from Google Maps and machine learning technique to estimate time-varying queue lengths covering both intersections with and without previous signalized intersection in the upstream direction. It is found that the results from the direction with previous signalized intersection are not as effective as those from the direction without previous signalized intersection. In addition, Jodnok and Pueboobpaphan [5] have used color-code traffic data from Google Maps, Linear Regression Analysis, and Random Forest (RF) to estimate queue length by dividing the model according to peak and off-peak periods. The results showed that queue-length estimation method using color-code data from Google Maps is sufficient to some extent but still had a high error. Therefore, Sornsoongnern et al. [6] developed further by taking the color-code traffic data from Google Maps, which originally divided the traffic condition into 4 colors: dark red, red, orange and green and arranged them into a new format. Several model development scenarios are considered including a division by direction with or without previous signalized intersection, and by peak or off-peak periods. It is found that the new format and new scenarios resulted in a better predictive performance, although the error is still high. However, it is still found that the direction without previous signalized intersection had better queue length estimation results than that from the direction with previous signalized intersection.

The aim of this study is to investigate the effect of intersection spacing on the performance of queue length estimation based on color-code traffic data. Aimsun traffic simulation model is used to generate traffic data with varying intersection spacings. Vehicle trajectory data are then extracted and manipulated, under a specific set of assumptions, to obtain color-code traffic data similar to those obtained from Google Maps. The queue length estimation model is then developed based on the Random Forest method, and the results are compared between different intersection spacings.

Manuscript received August 9, 2023.

R. Pueboobpaphan is with Suranaree University of Technology, Nakhonratchasima, Thailand

K. Phungsombat, N. Thansap, W. Naemkhuntod, P. Sanusit, N. Suwanamuk, and K. Torkoksung were undergraduate students at Suranaree University of Technology, Nakhonratchasima, Thailand.

P. Jodnok and P. Sornsoongnern were graduate students at Suranaree University of Technology, Nakhonratchasima, Thailand.

S. Pueboobpaphan is with Suranaree University of Technology, Nakhonratchasima, Thailand

II. METHODOLOGY

A. *Traffic simulation*

In order to investigate the effect of intersection spacing on the performance of queue length estimation based on color-code traffic data, the intersection spacing has been altered ranging from short to long distances, divided into 7 cases: 200 meters, 300 meters, 500 meters, 800 meters, 1000 meters, 3000 meters and 5000 meters. The other factors remain constant as follows.

- Road network: the road network studied in [7] is modified to control for other external factors and simplified as a 2-lane road with one direction as shown in Figure 1. There are 2 intersections, with spacings as described above.
- Traffic volume: traffic demands during peak and off-peak from [7] are used as guideline in setting up traffic demand in this study as shown in Figure 2.
- Traffic signal: both intersections are equipped with traffic signals. Cycle length, green, yellow, and red periods are fixed based on field observation as shown in Figure 1. Small alteration of traffic signal at intersection 2 is also considered to allow for some changes in traffic conditions.
- Driving behavior parameters: we used the same values as those calibrated by [7].

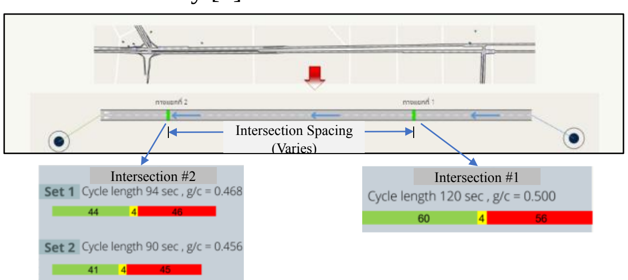

Fig. 1. Road network and traffic signal.

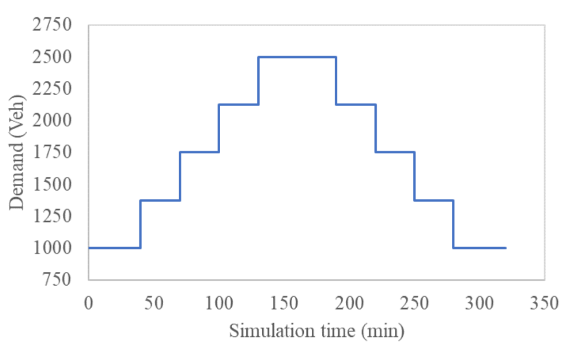

Fig. 2. Varying traffic demand used in the simulation.

B. *Data manipulation*

We develop traffic simulation model using Aimsun, in which each run takes a total time of 320 minutes. Data of vehicle trajectories are collected every 1 second. The simulation is repeated by changing the intersection spacing and the traffic signal as described above. This results in a total of 14 data sets. Each data set are processed to generate traffic color-code data. However, because there is no clear information about determining the threshold for each color code and the only known information is that Google Maps displays color according to the average speed of vehicles driving on that particular segment, therefore, the color-code assumption is set as follows.

- Dark red: 0 km/hr <= average speed <= 10 km/hr
- Red: 10 km/hr < average speed <= 40 km/hr
- Orange: 40 km/hr < average speed <= 60 km/hr
- Green: 60 km/hr < average speed

In addition, road network is divided into small segments of every 100 meters. For every 1 minute, the average speed for each segment is determined and each segment displays only one color according to its average speed. Note that data from the first 10 minutes of the simulation are not used in the analysis to allow for the warm-up periods as well as the last 10 minutes. Therefore, in total there are 300 observations for each run.

C. *Development of queue-length estimation models*

In this research, the dependent variable is the average of the observed queue lengths over 2 traffic lanes and the independent variables are extracted from the color-code data according to (1) – (4). Road segments are numbered in order from 1st as the segment next to stop bar to the nth as the last segments further away in the upstream direction.

$$X_{i1} = \begin{cases} 1 \; if\ color\ of\ i^{th}\ segment\ is\ dark\ red \\ 0,\ otherwise \end{cases} \tag{1}$$

$$X_{i2} = \begin{cases} 1 \; if\ color\ of\ i^{th}\ segment\ is\ red \\ 0,\ otherwise \end{cases} \tag{2}$$

$$X_{i3} = \begin{cases} 1 \; if\ color\ of\ i^{th}\ segment\ is\ orange \\ 0,\ otherwise \end{cases} \tag{3}$$

$$If\ color\ of\ i^{th}\ segment\ is\ green, X_{i1}, X_{i2}, X_{i3} = 0 \tag{4}$$

In this paper, the queue-length estimation models are constructed using RF from the package in R, namely, caret. The method of “ranger” is used and the tuning parameter of this method, mtry, which is the number of randomly selected predictors, is tuned in the 2–9 range. Another hyperparameter of this method is Min.Node.Size, which is tuned in the range 5–9. This range setting is comprehensive and reasonable for each parameter. The 5-fold cross-validation approach is employed.

We have divided the data set separately amounting to 10% of the total data (test set) to be used to evaluate the performance of the model. By using the color-code data from the testing data set to predict queue length based on the developed RF model, the estimation results are compared with the observed queue length from the test data set. In the analysis, queue length are compared only during the red signal period. The performance indicators considered in this study are Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE).

III. RESULTS AND DISCUSSION

The result of the error analysis is shown in Figure 3. According to the study results, there is no clear trend of the relationship between the error and intersection spacing.

In addition, we made a comparison between the actual queue length values (Queue_Length) and the predicted queue length (pred.RFM) in both line plots and scatter plots as shown in Figure 4–10 and 11–17, respectively.

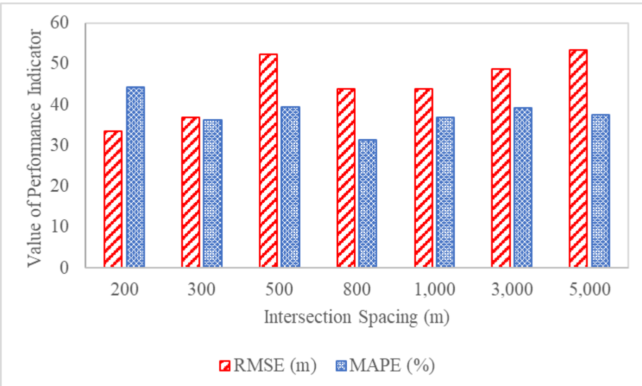

Fig. 3. Performance indicators of the developed RF models

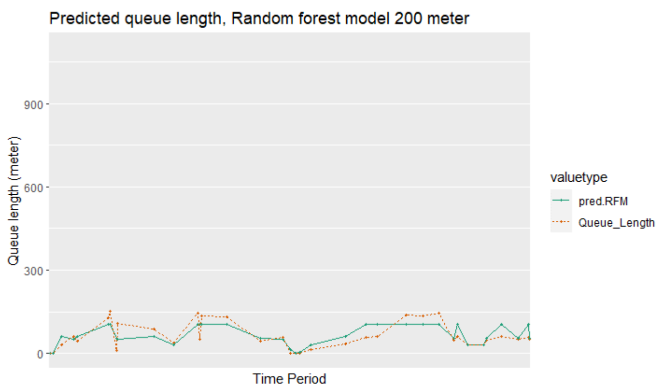

Fig. 4. Line plot for intersection spacing = 200 m.

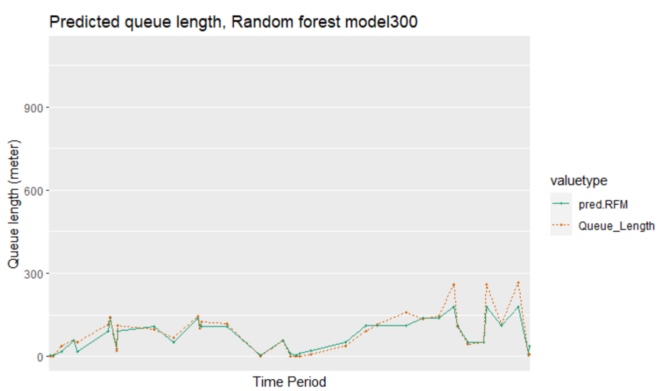

Fig. 5. Line plot for intersection spacing = 300 m.

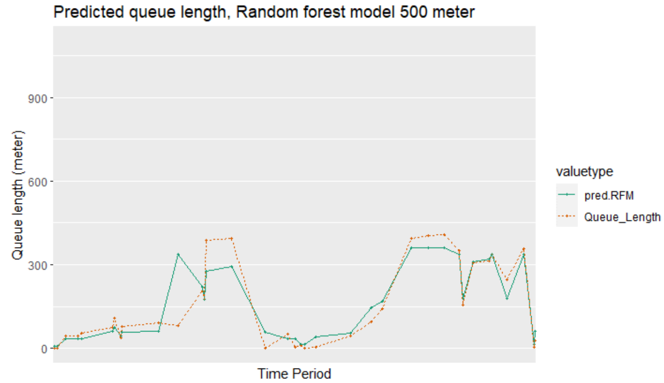

Fig. 6. Line plot for intersection spacing = 500 m.

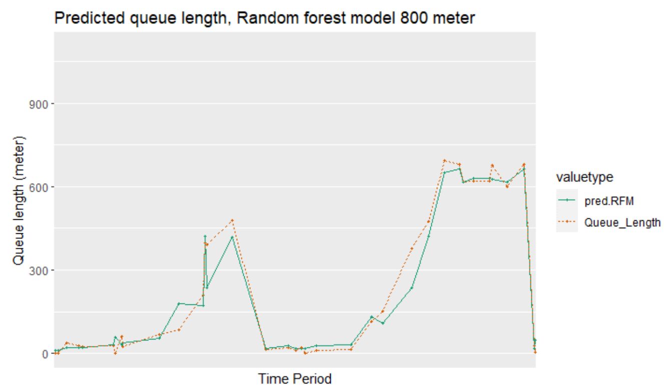

Fig. 7. Line plot for intersection spacing = 800 m.

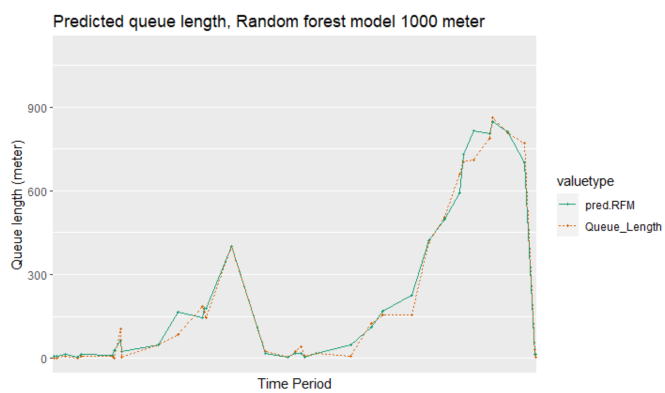

Fig. 8. Line plot for intersection spacing = 1000 m

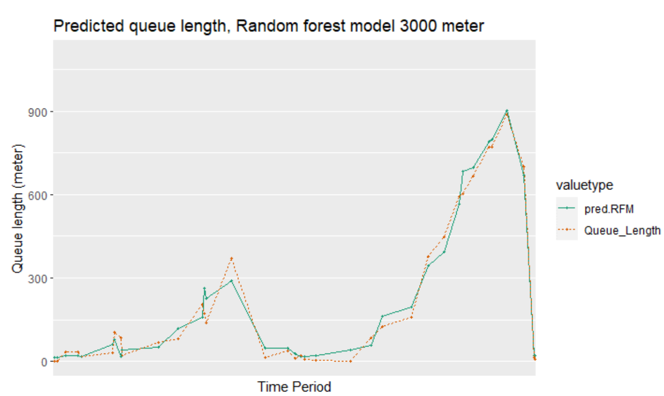

Fig. 9. Line plot for intersection spacing = 3000 m.

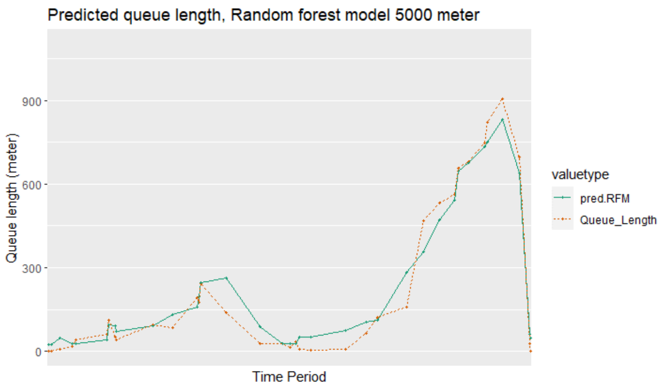

Fig. 10. Line plot for intersection spacing = 5000 m

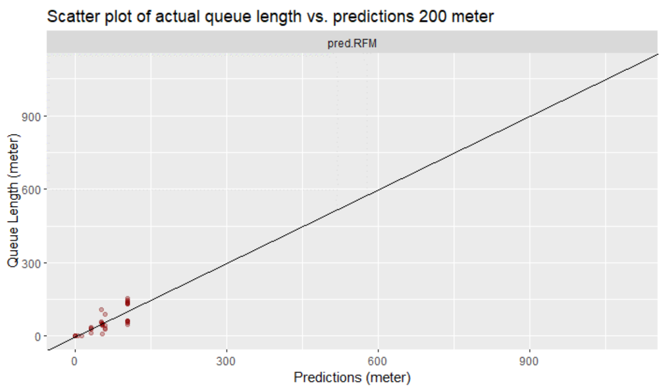

Fig. 11. Scatter plot for intersection spacing = 200 m.

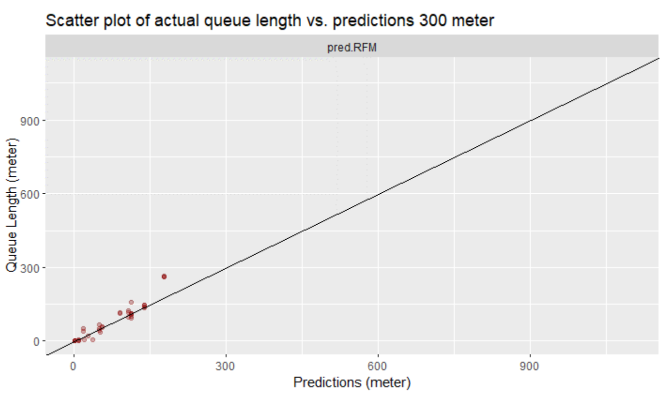

Fig. 12. Scatter plot for intersection spacing = 300 m.

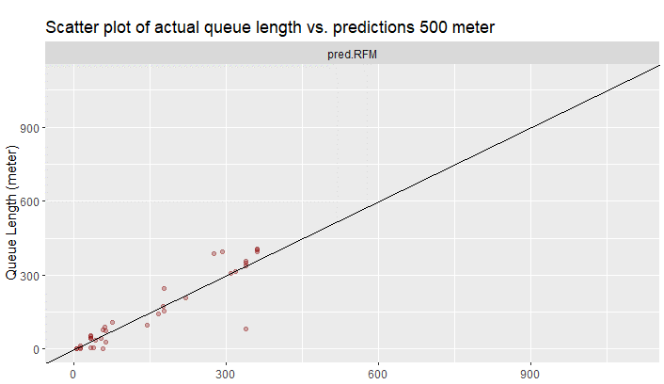

Fig. 13. Scatter plot for intersection spacing = 500 m.

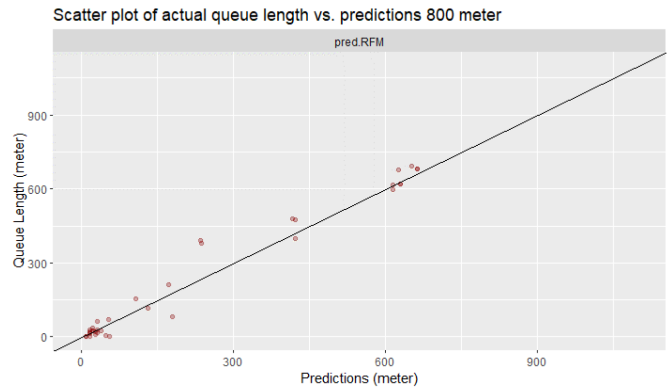

Fig. 14. Scatter plot for intersection spacing = 800 m.

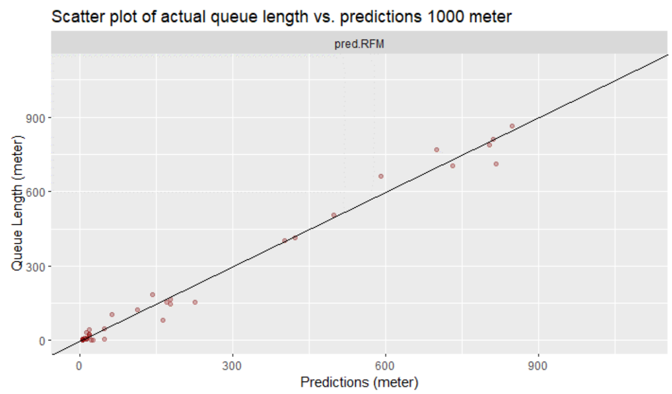

Fig. 15. Scatter plot for intersection spacing = 1000 m.

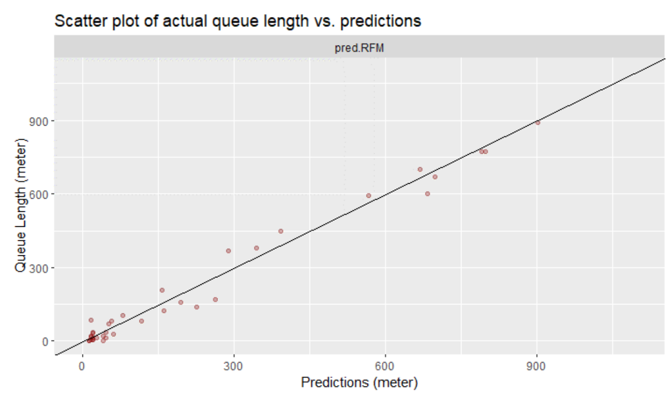

Fig. 16. Scatter plot for intersection spacing = 3000 m.

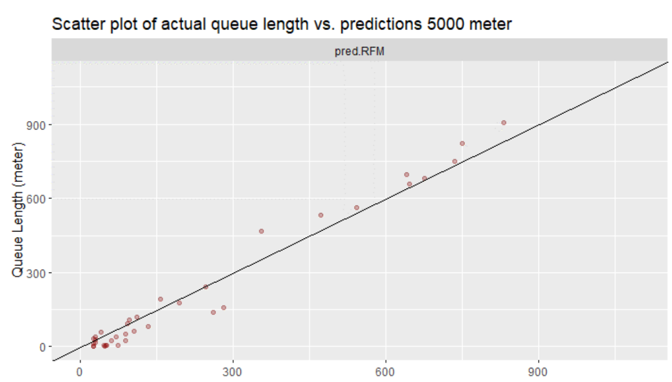

Fig. 17. Scatter plot for intersection spacing = 5000 m.

It can be seen that the predicted values and the actual queue lengths are positively correlated. It is found from the line plots that the RF model can detect the average tendency of queue length variation to some extent. It can follow the increasing and decreasing tendency according to the actual surveyed value. However, the RF method is not as efficient as it should be in detecting periods where the queuing is abruptly high or low. The reason may be that the queue is constantly changing depending on traffic volume and traffic lights. The actual queue length value obtained from Aimsun is recorded every 1 second, but the color-code traffic data are determined every 1 minute, so the variation of the color-code data with a coarser time frame may not be able to reflect real-time changes of queue length. According to the results from a study of [6], it is found that if the color-code information showing traffic conditions is timely, it would improve the forecasting results.

IV. CONCLUSION

This paper investigates the effect of intersection spacing on the queue length estimation performance by the Random Forest (RF) model using data from color-code traffic information. Microscopic traffic simulation is used to create different spacings between intersections to cover short distances to long distances. The results showed that the developed RF model is capable of detecting the trend of changes in queue length to a certain extent. However, the results indicated that the distance between intersections is not significantly correlated with the performance of the queue length estimation model based on color-code traffic data.

Future consideration should be given to the display resolution of the color band, both in terms of data update time and length of the band.

REFERENCES

[1] B. Bae, Y. Liu, L.D. Han, and H. Bozdogan, “Spatio-temporal Traffic Queue Detection for Uninterrupted Flows”, *Transp. Res. B Methodol.* 2019, 129, 20–34. https://doi.org/10.1016/j.trb.2019.09.001

[2] S. Yang, J. Wu, Y. Du, Y. He, and X. Chen, “Ensemble Learning for Short-Term Traffic Prediction Based on Gradient Boosting Machine”, *J. Sens*. 2017, 7074143. https://doi.org/10.1155/2017/7074143

[3] M. Noaeen, R. Mohajerpoor, B.H. Far, M. Ramezani, “Real-Time Decentralized Traffic Signal Control for Congested Urban Networks Considering Queue Spillbacks” *Transp. Res. Part. C Emerg. Technol*. 2021, 133, 103407. https://doi.org/10.1016/j.trc.2021.103407

[4] P. Jodnok, “Estimation of Queue Lengths at a Signalized Intersection Using Google Maps’ Color-Code Traffic Information”, Suranaree University of Technology: Nakhonratchasima, Thailand, 2021.

[5] P. Jodnok, R. Pueboobpaphan, “Estimating Time-Varying Queue Length at Signalized Intersection by Using Color-Coded Traffic on Google Maps”, *UBU Eng. J.* 2021, 14, 48–60.

[6] P. Sornsoongnern, S. Pueboobpaphan, R. Pueboobpaphan, “Innovative Dynamic Queue-Length Estimation Using Google Maps Color-Code Data”, *Sustainability* 2023, 15, 3466. https://doi.org/10.3390/su15043466.

[7] J. Suwanko, M. Deesiang, T. Jabthong, K. Yooma, N. Songkrut, T. Buapa, P. Kunmueang, C. Sooksai, “Traffic signal improvement on major highway network, a case study of highway 304 at the 80th Birthday Anniversary Sports Stadium intersection, Wing 1 intersection, and Suranaree University of Technology intersection” *Research Report*, Suranaree University of Technology: Nakhonratchasima, Thailand, 2021.

Rattaphol Pueboobpaphan was born in Nakhonratchasima Province, Thailand. He is an Assistant Professor at Suranaree University of Technology since 2010. He obtained PhD in Urban and Environmental Engineering (Transportation Engineering) from Hokkaido University, Japan in 2006. He spent two years during 2008 - 2010 as an Assistant Professor at Centre for Transport Studies, University of Twente, the Netherland. He has more than 20 years of professional experience in traffic engineering, transport planning, and logistics engineering. His research interests focus on Intelligent Transportation System (ITS), sustainable transportation, public transport, transit-oriented development, application of machine learning techniques in transportation and traffic analyses.

He has involved in many aspects of traffic engineering and transportation planning studies including public transit planning, transportation planning, planning and development of Intelligent Transport System (ITS), and feasibility study of several transportation projects.

Suthatip Pueboobpaphan is a lecturer at Suranaree University of Technology since 2014. She obtained Ph.D. in Logistics Management from Burapha University, Thailand in 2014. Her research interests focus on Transportation and Logistics Management.

She has involved in many aspects of Transportation Engineering and Logistics studies including transportation planning, and feasibility studies of several transportation and logistics projects.