



International Symposium of Transport Simulation (ISTS'18) and the International Workshop on Traffic Data Collection and its Standardization (IWTDCS'18)

Identification of Road Bottlenecks on Urban Road Networks Using Crowdsourced Traffic Data

Sakitha P. Kumarage, Dimantha De Silva, J.M.S.J. Bandara*

**Department of Civil Engineering, University of Moratuwa, Katubedda 10400, Sri Lanka*

Abstract

Bottlenecks are considered as an important research area in traffic engineering due to socio economical drawbacks caused by the delays in road traffic. Identification of such bottlenecks along an urban road network can improve the traffic management of that area. Out of many static and dynamic bottleneck identification methods use of Advanced Transport Information systems is a novel approach to existing traffic management techniques. Therefore, this paper presents a method for identifying static road bottlenecks and moving road bottlenecks based on dynamic traffic data gathered from crowd sourced data and spatiotemporal analysis. Not limited to single day data analysis, the proposed method analyzes dynamic and concurrent traffic data collected for a longer period for multiple links based on google maps distance matrix application programming interface.

© 2018 The Authors. Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

“Peer-review under responsibility of the scientific committee of the International Symposium of Transport Simulation (ISTS'18) and the International Workshop on Traffic Data Collection and its Standardization (IWTDCS'18)”

Keywords: Type your keywords here, separated by semicolons ;

1. Introduction

Delays due to traffic congestion can create many socio-economic implications which can toughen day-to-day activities. Bottlenecks are major contributors for urban traffic congestion. Identification of static and dynamic road bottlenecks are carried out in different ways. Use of Advanced Transport Information systems in diagnosis process of

* Corresponding author. Tel.: +0-000-000-0000 ; fax: +0-000-000-0000 .
E-mail address: author@institute.xxx

road bottlenecks is a novel approach to existing traffic management techniques. Most of the bottleneck identification methods in the literature cannot expand for large networks simultaneously or very expensive for developing countries to prioritize. Hence, there is a need for bottleneck identification method which will be economical to implement. Identification of time and duration of congestion in bottlenecks can enable investigations and provide migratory actions (Zhang & Levinson, 2004).

The objectives of this research is identification of road bottlenecks based on traveltime variation and variation of space mean speed below a threshold value. The bottlenecks should be able to identify under spatiotemporal elements and be able to obtain an idea about their occurrence and significance. Further to explore the possibility and validity of using of Google distance matrix application programming interface traveltime for bottleneck identification.

2. Literature Review

2.1. Definitions on road bottlenecks

Many authors have taken different approaches in defining road bottlenecks (Andres L. & Alvaro, 2016; Hale et al., 2016). In general terms, a traffic bottleneck is a localized disruption of vehicular traffic on a road segment in which separates upstream queued traffic and free flowing down stream traffic. When compared to a traffic jam, a bottleneck is a result of a specific physical condition, often caused by merging and diverging traffic, lane drops, grade changes or badly timed traffic lights, intersections (Zhang & Levinson, 2004). According to the definition of (Daganzo, 1997), an active bottleneck is a restriction that separates upstream queued traffic and free flowing downstream traffic which is time dependent. (Chen, Skabardonis, & Varaiya, 2004) described freeway bottlenecks as certain freeway locations that experience congestion at nearly the same time almost every day. According to (Bertini & Myton, 2005) who followed (Daganzo, 1997), a bottleneck is a point upstream of which there is a queue and downstream of which there is freely flowing traffic. (Zhao, McCormack, Dailey, & Scharnhorst, 2013) defined a bottleneck as a poorly performing roadway segment on the basis of speed measurements and statistical predictability. Deriving from above concepts for this research a bottleneck is considered as a road location which experience recurrence congestion measured by speed reduction.

2.2. Bottleneck Identification

In the identification of bottlenecks direct and stochastic approaches has been followed mostly (Hale et al., 2016). According to (Jia et al, 2000) spatiotemporal variation should be incorporated in the identification of freeway bottlenecks. In evaluation of traditional approaches Cassidy and Bertini, Chen et al., Ban et al. and Wiecek et al used loop detector data in identifying recurrent bottlenecks. Speed differential, average delay and percentile increase in speed were methodologies used in identification.

Implementation of Intelligent Transport System (ITS) technologies such as vehicle detection by global positioning system (GPS) has enables to obtain a larger amount of information than earlier types of data collection such as sensor data and loop detectors with the. (Andrea & Marcelloni, 2016). Following the ITS concepts, (Zhao et al., 2013) suggest a methodology for identifying and ranking bottlenecks using probe data collected by commercial global positioning system fleet management devices mounted on trucks. The major drawback in this research was GPS data sample was only representing trucks which are traveling on the road. Therefore, identifying road bottlenecks generally applicable to all the vehicles moving on the road is not addressed in this method (Zhao et al., 2013).

Improving on the concept of using GPS data obtained from probe vehicles. (Andres L. & Alvaro, 2016) suggest a methodology to identify bottlenecks based on GPS data from taxis vehicles in a city limit. The GPS location reported by the GPS devices via GPRS (General Packet Radio Service) with a frequency of 10s was used to identify road bottlenecks. Finding recurrent low-speed sections, beyond expected delays in the road network was enabled by identifying road network segments that perform poorly in terms of speed, compared with upstream and downstream conditions. This could be considered as a low-cost option in identifying critical points in traffic networks in developing cities without expensive traffic-monitoring systems. Limitation in implementing this type of a methodology is the data

collected is only available for taxi vehicles in the city which is a poor sample of vehicle movement (Andres L. & Alvaro, 2016). Moreover, for countries with less number of taxis this method cannot provide a representative sample in urban traffic.

In order to get a representative sample of moving vehicles the approach taken by (Janecek et al., 2012) is commendable. The methodology suggested by (Janecek et al., 2012) for estimating vehicular travel times based on mobile cellular network. The research focuses on how vehicle travel times and road congestion can be inferred from anonymized signaling data collected from a cellular mobile network (Janecek et al., 2012). To obtain higher accuracy in estimation of travel times and a timely detection of congestions by both active users engaged in voice calls and inactive users were used. In the methodology, spatially coarse mobility data from all users (both active and inactive) is gathered to capture speed deviations in long road sections and detect congestion events. Then finer grained mobility data produced by active users is used to refine the location accuracy and classify the type of congestion event. Compared to earlier methodologies stated above this approach presents a higher accuracy and an economical approach which does not require investments in new infrastructure while utilizing mobile cellular network as a large-scale mobility sensor (Janecek et al., 2012).

Improving on using mobile network data on congestion identification (Andrea & Marcelloni, 2016) suggest a methodology to detect traffic congestion using GPS data provided by smartphone users moving in a road network. In their study traces collected from vehicles moving in the city are analyzed in real-time by means of an expert system, without the need for a learning process or historical data. The outcome of the study presented a system for detecting traffic congestion, traffic state based on the speeds of vehicles and incidents in real-time. With the study, it is possible to send a notification to users on traffic alerts indicating the affected area and traffic state whether traffic is flowing, slowed, very slowed, and blocked (Andrea & Marcelloni, 2016). This study is much more similar to the information provided by Google traffic layer (Google, 2009). A major difficulty in this research was to obtain GPS data from users moving in the city due to privacy issues (Andrea & Marcelloni, 2016). Next, the scalability of the system with limited resources was a challenge. As future work, (Andrea & Marcelloni, 2016) suggest incorporating historical data and real-time information for travel time prediction and traffic state prediction. Hence Google Distance Matrix API will be a solution for the challenges faced by authors and future work suggested by them.

2.3. Ranking and Reliability of Bottlenecks

When a road bottleneck is identified, the next consideration is to measure the reliability of the bottleneck. Travel time index used by FHWA include the 90th or 95th percentile travel time and the buffer index, which is the extra time needed to allow the traveler to arrive on time (FHWA 2011). This index is computed as the difference between the 95th percentile travel time and the mean travel time, divided by mean travel time. Hale et al suggest three evaluation criteria for bottleneck ranking based on spatiotemporal traffic state matrix (Hale et al., 2016). Impact Factor for ranking of bottlenecks based on the duration of congestion and length of congestion over a longer time period was proposed to rank the bottlenecks (Hale et al., 2016). Annual reliability matrix (ARM) is suggested by authors to understand the annual variability of a bottleneck impact factor (Hale et al., 2016). Further improving the ARM a bottleneck intensity index is proposed to capture the size and shape of ARM to a single number (Hale et al., 2016). Emanm and Al-Deek developed a methodology to estimating travel time reliability based on stochastic approach (Emam & Al-Deek, 2006). Their research confirms lognormal distribution is best-fit to compute the travel time reliability as the probability that a trip between a given origin-destination pair could be made within a specified time interval (Emam & Al-Deek, 2006). Zhao et al classified the travel reliability of roadway segments into three categories as unreliable, reliably slow, and reliably fast (Zhao et al., 2013). It was based on the hypothesis that roadway reliability is stochastically represented by a unimodal or a bimodal probability density function over a certain time period (Zhao et al., 2013). (Zheng & Chang, 2017) suggested that Weibull and Log logistic statistical distribution best fits to evaluate capacity and congestion duration respectively. Hence confirming to estimate the probability of congestion occur and the time it will last using those estimates.

3. Methodology

The methodology for identification of road bottlenecks in urban road networks was carried out in two stages. The first the methodology on acquiring data and secondly methodology on analysis the data collected for identification of road bottlenecks. In this paper, we gather data from Google Distance Matrix API. Using the API, it is possible to collect traveltime and distance values of a given route between start and end points. When a user enables Google services on a smart phone it sends anonymous bits of data back to Google describing how fast the user is moving. The traveltime prediction algorithms on google servers combine users speed with the speed of other phones on the road, across thousands of phones moving around a city at any given time in which a significantly reliable estimate on traffic conditions could be gained(Google, 2009).

Collecting traveltime data for identification of bottlenecks was carried out using a cloud server and a program written in Hypertext Preprocessor(PHP) language. As shown in FIGURE 2 the cloud server contains a file with origin and destination GPS coordinates, a PHP script and a file to store traveltime and distance vales. For a predefined frequency and time duration, the script calls the Google Distance matrix API and stores the collected data to the could server. Spanning this process for a longer period of time can gather data for analysis purposes.

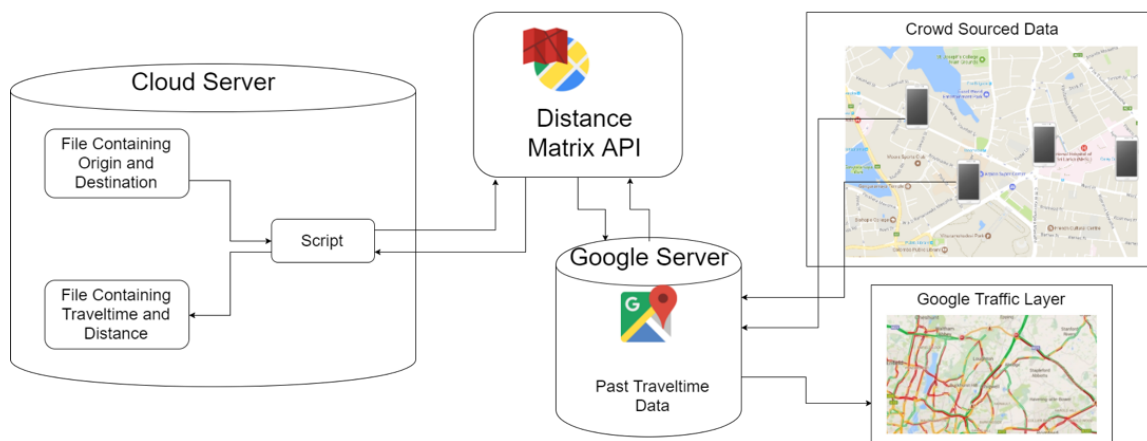


Fig 1: Data collection process from Google Distance Matrix API

In obtaining data for bottleneck identification, spatiotemporal variation of speed in between two adjacent road segments is evaluated. Average traveltime and space mean speed are the governing variables which will be considered in this research for evaluation. Both parameters could be obtained using the Google Distance Matrix API. Initially, the API calls are utilized to gather distance value in meters and traveltime estimates in seconds to the considered segments. then traveltime value and distance value obtained from the API could be used to calculate the space mean speed of each considered segment.

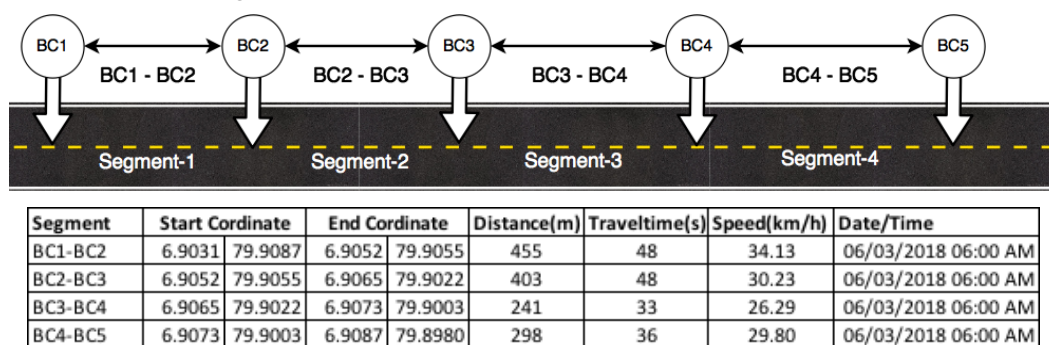


Fig 2 : Illustration of data collection database

Figure 3 shows how a sample of data being collected to server by using the Google distance matrix API. Iteration of the process at an frequency (every 5 min, 10min, 15min) will enable to collect traffic data continuously for desired time period.

3.1. Bottleneck identification

In this paper, a spatiotemporal variation of speed data is recorded in the Spatio-temporal matrix (STM) for each corridor or road segment considered. On identification of bottlenecks, the average space mean speed of each segment over a period of a week is evaluated. For identification of bottleneck initially, the segment need to be evaluated for congestion. In that aspect, any segment with space mean speeds less than 21km/h was considered (Transportation Research Board's, 2010). The criteria was to consider average speeds of arterial roads at the level of service (LOS) class E and Class F as congested scenarios (Transportation Research Board's, 2010).

When a road segment is identified as congested the next step is to check for its recurrence. In ranking bottlenecks recurrence is a significant factor to be considered as mitigation of such occurrences is important. In the research the methodology adopted was to prepare STM with the weekly average the speed data and identify congested segments along with spatiotemporal identities. In this way, recurring bottlenecks could be identified. Further on validating the bottleneck visual observing the spatiotemporal traffic state matrix could be used.

3.2. Ranking of Bottlenecks

When a bottleneck is identified its reliability and influence is the next challenge to be addressed. As congestion is a random event and formation of bottlenecks is followed by such a random event, (Zheng & Chang, 2017) the attention is significant to the aspect of the influence of the bottleneck and its reliability (Zheng & Chang, 2017). To gather an idea on speed reduction at the bottleneck percentage variation from existing network average speed was evaluated. Hence that will decompose an idea on how worse the performance of the bottleneck to the overall network traffic.

$$\%Effect = \frac{V_{Network} - V_{Bottleneck}}{V_{Network}} * 100 \quad (3)$$

The effect of a bottleneck on the segment performance is another parameter to be evaluated when ranking bottlenecks. This will be a reliability measure of the segment considered. Traveltime Index is an optimum measure as it compares the traveltime during congestion to the time required to make the same trip at free-flow speeds. (Transportation Research Board's, 2010). With respect to the bottleneck ranking, Traveltime index was taken as the ratio of traveltime at bottleneck congestion situation to traveltime at freeflow.

$$TTI = \frac{TT_{Bottleneck}}{TT_{Freeflow}} \quad (4)$$

Although a bottleneck gives a high traveltime index it does not become significant for ranking if the bottleneck exists only for a short period of time (Hale et al., 2016). Further vice versa a bottleneck with lower TTI could be much impactful if it exists for a longer period of time of the day (Hale et al., 2016). Hence the congestion period should be taken into consideration Therefore, in order to incorporate the duration of congestion into the evaluation of the impact of a bottleneck the research defines Impact Factor incorporating traveltime index and duration of congestion.

$$IF = \frac{TT_{Bottleneck} * D_{congestion}}{TT_{Freeflow} * D_{Observed}} \quad \begin{array}{l} IF = \text{Impact Factor} \\ D_{congestion} = \text{Duration of Congestion} \\ D_{Observed} = \text{Duration Observed} \end{array} \quad (5)$$

In defining the Impact Factor, the ratio between the duration of congestion to an observed duration which the analysis was taken into consideration, was included as shown in equation (5). Hence with this approach, it is possible to distinguish the impact of traveltime index and congestion duration ratio as explained earlier.

4. Analysis

4.1. Data validation

Data obtained from Google Distance Matrix API was validated using a floating car data test and license plate survey. Data collected from floating car data is then compared with the traveltime obtained from the Google Distance Matrix API. The results observed showed a minor variation from the traveltime collected from floating car data and Google Distance Matrix API. The observed R squared value is 98% which indicates that the traveltime estimates given by Google Distance Matrix API does not have a significant difference from the observed floating car data. Fig 3(a) shows a comparison of traveltime data obtained from the above mentioned two methods for 108 events in Colombo urban road network.

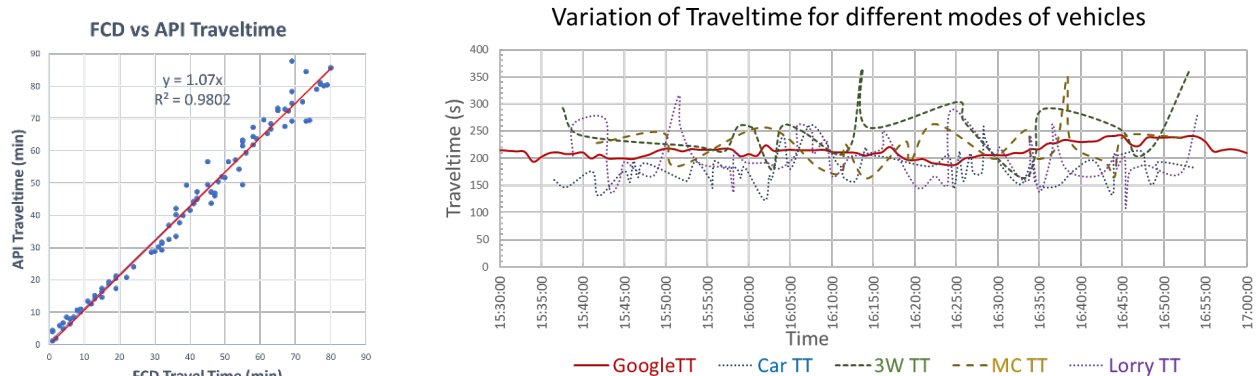


Fig 3 : (a) Graph of FCD vs API Traveltime (b) Variation of traveltime for different modes of vehicles

When compared the variation of traveltime data with different modes of vehicles (see Fig 3(b)), the Traveltime obtained by google distance matrix API gives a representative value for the vehicle composition of the considered road segment. This hypothesis was confirmed by the licence plate survey conducted for 256 vehicles comprised of Motor cars, three-wheelers, Motor cycles and Lorries within a period of 2 hours at a 3.06 km distance. The deviation of observed average travel time to traveltime given by Google Distance matrix API was 12 seconds and deviation of space mean speed was recorded 3km/h. From the results obtained in two verification methods, it could be concluded that the use of Google traveltime estimates as a novel approach in collecting traveltime data for analysis purposes.

4.2. Bottleneck identification

Analysis for identification and ranking of road bottlenecks was carried out at three road networks in Colombo metropolitan area and one analysis on a major arterial road is presented. The details of the analysis on location 1 is given in Table 1. The illustrated spatiotemporal graph shows the road segments in the horizontal axis and time on the vertical axis. The average speeds over a period of three months were shown in the matrix after excluding weekends speed values and extreme outliers. The matrix was colour coded to identify the formation of bottlenecks. The threshold speed value for detection of bottlenecks was 21km/h which was suggested by the level of service criteria of HCM2010 (Transportation Research Board's, 2010). The speeds which are lower than 21km/h are colored as red, and red colour gets darker when speed reduces from a threshold level. When the speeds are greater than 21km/h the indicated colour becomes green, and when speed levels increase further from threshold level the green colour gets darker.

When the location1 was analyzed it was able to observe 4 bottleneck incidents. Further, there were two segments which did not indicate any bottlenecks. The observed bottleneck incidents are recurring bottlenecks as the weekly average speeds were considered. Hence the STM method was able to capture the effect of day to day variations.

The effect of each incident was identified under three performance measures. when percentage variation from existing network traffic is considered segment 4 has a higher change. When traveltime index is considered segment 3

has the highest drop. Segment 5 is critical under influence factor considerations, which consider both traveltime index and duration of congestion.

The segment 1 shows a bottleneck incident in the morning this is mainly due to a high school located in that segment and vehicles which arrive at school create the traffic by parking at roadsides. When observed the segment 4 long term congestion could be observed. The reason behind this incident is a long-distance bus stop and malfunctioning signal system.

Table 1 : Analysis of an arterial road in Colombo - Location 1

Location 1

The above analysis reveals STM illustration is a successful tool to identify congestion patterns and road bottlenecks. The advantage of this method is it can incorporate spatiotemporal variations and long-term impacts. Use of this method in planning and management is an effective approach. When a segment is identified as highly congested a finer experiment by analyzing smaller road segments of the congested segment will enable to identify exact locations which cause bottlenecks. Any field visit too can be helpful in this regard. A regular traveler might have an idea about the congestion in a road but for a planner, this method could be an efficient tool to identify bottlenecks and congestion patterns. In future work, the experiment could be extended to a larger network and observe congestion patterns and bottlenecks. The visual identification method could be incorporated into a pattern identification algorithm and conduct the manual work much faster. Incorporation of weather data and observing rainfall induced changes in congestion patterns is a possibility.

5. Conclusion

This paper focused on identification of road bottlenecks based on traveltime variation and variation of space mean speed below a threshold value. The methodology proposed was to gather data from Google Distance Matrix Application Programming Interface (API). The traveltime given by the API is used to calculate the space mean speed of the road segment. The methodology is justified by a stepwise literature review on suitability. Traveltime samples obtained from API were validated through floating car data and license plate survey.

Identification of road bottlenecks with a spatiotemporal variation of speed data was graphically visualized on traffic state matrix and impact over time is evaluated. The reliability and significance of road bottlenecks were identified using three evaluator measures, traveltime index, bottleneck influence factor and speed variation from overall network average speed. The proposed methodology was illustrated using an arterial road in Colombo metropolitan area which showed successful results. Therefore it could be concluded that the methodology is very advantageous in planning and forecasting traffic management activates which require higher accuracy and low cost of implementation. Hence this is a promising and economical method to identify congestion patterns in large urban networks in developing countries

6. References

- Andrea, E. D., & Marcelloni, F. (2016). Detection of Traffic Congestion and Incidents from GPS Trace Analysis Eleonora. *Expert Systems With Applications*. <https://doi.org/10.1016/j.eswa.2016.12.018>
- Andres L., J., & Alvaro, R. (2016). Exploratory Methodology for Identification of Urban Bottlenecks Using GPS Data. *Journal of Traffic and Transportation Engineering*, 4(5), 280–290. <https://doi.org/10.17265/2328-2142/2016.05.007>
- Bertini, R. L., & Myton, A. M. (2005). Use of Performance Measurement System Data to Diagnose Freeway Bottleneck Locations Empirically in Orange County, California. *Transportation Research Record: Journal of the Transportation Research Board No 1925*, 48–57. <https://doi.org/10.3141/1925-06>
- Chen, C., Skabardonis, A., & Varaiya, P. (2004). Systematic identification of freeway bottlenecks. *Transportation Research Record: Journal of the Transportation Research Board*, (January). <https://doi.org/https://doi.org/10.3141/1867-06>
- Daganzo, C. F. (1997). *Fundamentals of transportation and traffic operations*. Oxford: Pergamon.
- Emam, E., & Al-Deek, H. (2006). Using Real-Life Dual-Loop Detector Data to Develop New Methodology for Estimating Freeway Travel Time Reliability. *Transportation Research Record: Journal of the Transportation Research Board*, 1959, 140–150. <https://doi.org/10.3141/1959-16>
- Google. (2009). The bright side of sitting in traffic: Crowdsourcing road congestion data. *Googleblog*. Retrieved from <https://googleblog.blogspot.de/2009/08/bright-side-of-sitting-in-traffic.html>
- Hale, D., Hajbabaie, A., Ma, J., Hu, J., Park, H., & Bared, J. (2016). Proposed Data-Driven Performance Measures for Comparing and Ranking Traffic Bottlenecks. In *Transportation Research Procedia; International Symposium on Enhancing Highway Performance Performance* (Vol. 15, pp. 483–494). Elsevier B.V. <https://doi.org/10.1016/j.trpro.2016.06.041>
- Janecek, A., Hummel, K. A., Valerio, D., Ricciato, F., & Hlavacs, H. (2012). Cellular Data Meet Vehicular Traffic Theory : Location Area Updates and Cell Transitions for Travel Time Estimation. In *ACM Conference on Ubiquitous Computing* (pp. 361–370).
- Jia, Z., Varaiya, P., Chen, C., Petty, K., & Skabardonis, A. (2000). Congestion, excess demand, and effective capacity in California freeways. Retrieved from http://people.eecs.berkeley.edu/~varaiya/papers_ps.dir/delay7-eff.pdf
- Transportation Research Board's. (2010). *Highway Capacity Manual 2010*.
- Zhang, L., & Levinson, D. (2004). Some Properties of Flows at Freeway Bottlenecks. *Transportation Research Record: Journal of the Transportation Research Board No 1883*, 122–131. <https://doi.org/https://doi.org/10.3141/1883-14>
- Zhao, W., McCormack, E., Dailey, D. J., & Scharnhorst, E. (2013). Using Truck Probe GPS Data to Identify and Rank Roadway Bottlenecks. *Journal of Transportation Engineering*, 139(1), 1–8. [https://doi.org/10.1061/\(ASCE\)TE.1943-5436.0000444](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000444)
- Zheng, L., & Chang, Y. (2017). Risk assessment model of bottlenecks for urban expressways using survival analysis approach. *Transportation Research Procedia*, 25, 1544–1555. <https://doi.org/10.1016/j.trpro.2017.05.181>