

Demonstration of intelligent transport applications using freight transport GPS data

Can Yang^{1*}, Xiaoliang Ma², Yifang Ban¹

¹Division of Geoinformatics
Department of Urban Planning & Environment
KTH - Royal Institute of Technology
Drottning Kristinas väg 30, Stockholm 100 44, Sweden.

* Corresponding author: cyang@kth.se

²Traffic Simulation & Control (TSCLab),
Division of Transport Planning, Economics and Engineering,
Department of Transport Science,
KTH - Royal Institute of Technology,
Teknikringen 10, Stockholm 100 44, Sweden.

Submitted in July 2015 for presentation at the Transportation Research Board 95th Annual Meeting in 2016

4856 words + 9 Figures (9 X 250) = 7106 words

Abstract

Most of the existing studies on floating vehicle data are focused on traffic information retrieval. There are still few studies that demonstrate how this type of data can be, directly or indirectly, applied for ITS systems and other traffic applications. This paper presents a web-based information system which integrates database technology, information estimation methods and traffic visualization approaches. Floating vehicle data is collected from freight truck and used to estimate traffic information including link travel time, route travel time and various congestion indices. Based on the information, two applications, trip planning and traffic visualization, are designed with simple and interactive user interface. The trip planning application supports user-selected origin-destination routing and visualization of temporal variation of the route travel time. In traffic visualization application, congestion indices are computed and an animated map is designed which enables the user to explore the congestion on a network in different periods of day together with link travel time distribution. A demonstration of the system shows that it can be used for dynamic trip planning and provides an efficient way for visualizing spatial and temporal patterns of congestion on a network as well as patterns of link travel time.

1. INTRODUCTION

1.1 Background

Over the past decades there has been fast development and deployment of Intelligent Transport Systems (ITS) because of the benefits that these systems can provide for sustainable transport development with respect to traffic mobility, driving safety and general environmental impacts. Advanced traffic information and management systems can estimate and predict traffic network states, including travel times, speeds etc., based on sensor data collected by different sensing technologies. Such system aims at providing road users with pre-trip or en-route real-time traffic information so that travellers can identify optimal routes and their trip schedule to maximize their travel efficiency. In addition, and more importantly, traffic administration could predict the network-wide congestion levels and accordingly affect traffic distribution across network using proper real-time information guide and other operations so that congestion or accident induced traffic delay and vehicle emission can be minimized at a systems level.

Even the most advanced systems cannot be effective and reliable unless sufficient data are collected to characterize the overall state of the network at any given time. Traditionally, real-time data from fix locations by loop detecting, camera or local wireless sensing plays major role for traffic information system. The ITS field has recently seen an explosion in the types and number of sensors deployed for traffic data collection. Among the emerging data collection technologies, floating vehicle (FV) sensors provide rich and reliable real-time travel times information. For example, Taxi fleets equipped with GPS transmitters are commonly used to collect FV data. In Stockholm taxi FV data is currently used to implement traffic prediction methods (the Mobile Millennium Stockholm project). Automatic Vehicle Identification (AVI) based on e.g. license plate recognition is used for travel time analysis (1). Floating phone data is being used to observe path flows for long distance trips. Similarly, big companies such as Google and TomTom are collecting anonymously positions from millions of mobile phone users on the road. This data creates new opportunities in intelligent transport applications as well as traffic prediction. The quality of the data collected varies by technology and local conditions (e.g. weather). Traffic prediction accuracy and effective traffic management greatly depends on the quality of the collected data. Compared to data from fixed locations, FV data collected from GPS equipped vehicles are widely available and therefore provide information on a wide spatial coverage with relatively low cost.

1.2 Relevant studies

Previous studies have demonstrated that GPS data collected from vehicles can be used to estimate traffic information like travel time and speed and several application areas have been investigated. Combining the GPS probe data with underlying road network, link travel time (LTT) can be estimated by allocating travel time between two consecutive probes to the network that the vehicle traverses (2; 3). Based on LTT, several models have been proposed to estimate route travel time considering time delays at intersections and signals (2), correlation between LTTs (4). Dynamic route travel time could be applied for city logistics planning (5; 6). Speed information estimated from GPS data consists of spot speed and link travel speed. Spot speed is the instantaneous speed of a vehicle at a specified location. In some cases it is reported directly from GPS device but can also be approximated by dividing travel distance by time for

consecutive probes. Spot speed information has been frequently used for detecting stop points including delivery points (7) and trip pattern analysis(8). Link travel speed represents the traffic state on a road link, which can be estimated by dividing length of link by LTT or combining track average speed with spot speed (9). By classifying the link traffic speed, congestion levels on a network can be monitored and analyzed (10).

Most of the existing studies on FV data are focused on traffic information retrieval. There are still few studies that demonstrate how FV-based data can be, directly or indirectly, applied for ITS systems and other traffic applications. Web-based system has the natural advantage in wide user availability and thus has been frequently employed for ITS applications. In (11) a web based freeway performance measurement system (PeMS) is designed based on loop detector data, though providing only static maps. In (12), a lightweight, web-based visual analytics application called Fervor is designed with interactive map, histograms and two-dimension plot for transport incident datasets. In (13), a congestion and incident scanner tool was developed where a contour line approach was employed for congestion visualization, which, however, is applicable only for single road segment. An online interdisciplinary data integration and analysis platform called Digital Roadway Interactive Visualization and Evaluation Network (DRIVE Net) was proposed in (14), and its sub system RADAR NET (15) provides real-time traffic flow map and dynamic routing. A web-based freight module was developed also for the platform in (16), which allows statewide freight performance measures such as numbers of trips, average travel time and travel time index.

While web- and mobile-based system becomes the current trend in traffic information and ITS applications, integration of databases, information processing and data mining functions with application is an important procedure. For FV data, the applications, especially with interactive and efficient visual analytics, are still immature or require further development. In addition, there are other potential commercial FV data sources from private sectors such as vehicle manufacturers and fleet owners. For example, Scania AB, a top heavy-duty vehicle (HDV) manufacturer in the world, receives information from thousands of HDVs through fleet management system. Such data could potentially be used not only for private business but also by public sector in traffic management and other applications. There are also several commercial applications where the vehicles' locations are collected and analyzed to provide historical or real-time traffic status, e.g., INRIX(17), Waze(18) and Google Maps(19). However, the methodologies used in these applications are private, and only aggregate traffic state is presented to user while the detailed travel time distribution is not shown, which is also valuable especially in research related to big traffic data. Besides, there is lack of efficient way to explore and visualize the travel time distribution spatially and temporally.

The main objective of this paper is to demonstrate the ITS applications of FV data collected from freight transport. The study develops a web-based information system by integration of database technology, information estimation methods and traffic visualization approaches. Based on information retrieved from freight GPS data, the system provides a simple and interactive user interface for two applications. The first is a trip planning case with user-selected origin-destination (OD) whereas the visualization of network congestion level is the other application, with detailed travel time distribution shown to users.

The remainder of the paper is organized as follows: in Section 2 system design and implementations including information estimation approaches are introduced. A demonstration study of the system is presented in Section 3. Finally, a conclusion is drawn in Section 4.

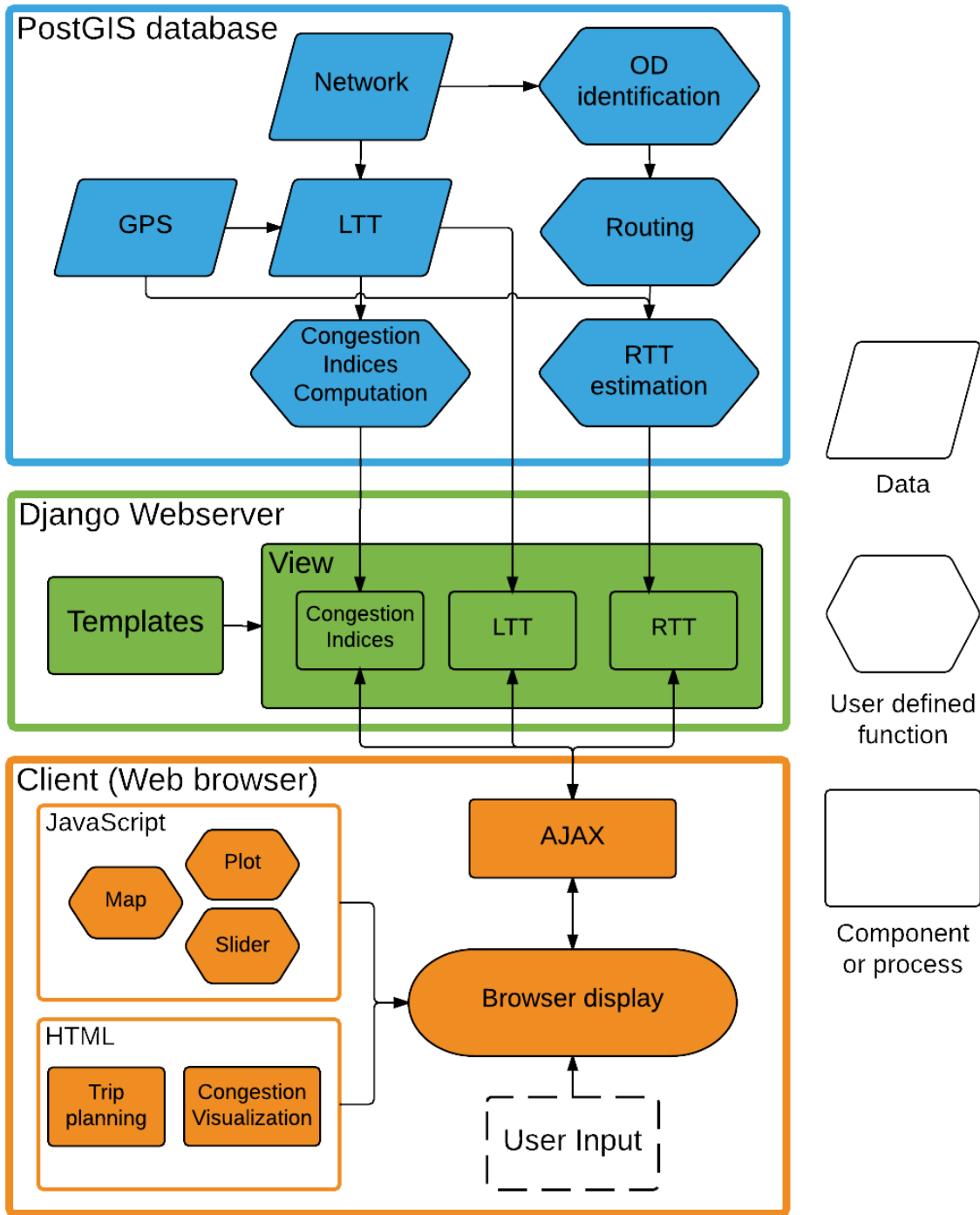


Figure 1: System architecture: components and communications between database, webserver and client

2. METHODOLOGY FOR SYSTEM DESIGN AND IMPLEMENTATION

2.1 System concept framework

The whole system is designed with the objectives of providing an efficient visualization and analysis tool based on traffic information retrieved from freight GPS data. Two application scenarios have been considered. The first scenario is about trip planning. Since the freight GPS data provides detailed information on major highways that are frequently used by HDV. The historical and real time traffic information on the roads can be retrieved from the GPS data to support fleet management and other ITS applications.

Based on the objective, three-tier client-server architecture is adopted, which is illustrated in Figure 1. The architecture is frequently used in software engineering and can be easily maintained. The details of roles for each tier are explained below.

Database

PostGIS database is employed to store the GPS probes and underlying network as it provides rich spatial functionalities in spatial data storage, indexing and query. In this system, most of the time consuming computations are performed in the database and the information is exchanged between webserver and database in JSON format. Totally three types of information are generated consisting of link travel time, route travel time and congestion indices, as shown in Figure 1. Detailed methods of estimating travel time and congestion indices are introduced in Section 2.2 and Section 2.3. Apart from traffic information estimation, two functions are defined for routing. When user specifies OD of a trip, an OD identification function can be called to find the nodes on the network, which are sent to a routing function to find the path between the two nodes by applying Dijkstra's shortest path algorithm (20).

Web server

Django is employed for web application development, which is Python-based web framework. It supports web design in a model-template-view (MTV) pattern, from which data presentation, user interface and model logics can be decoupled. The model component defines the field and behaviors of data stored in the database. Since most of the queries are implemented directly in the database, the model part is not used in this system. The template component is used to dynamically generate html files by sharing static and common content between files and inserting dynamic content separately. The view component directs user's requests to various data resources. It is also the place where connection to database is set up and queries are performed.

Client

On the client side, a uniform layout is designed for user interface containing four parts header, footer, map panel and a collapsible side panel, as displayed in Figure 2. The header contains title and links to various modules while the footer contains control elements such as buttons and sliders. In the middle of web page, an interactive map is displayed. The collapsible side panel is used for displaying non-spatial information such as distribution of link travel time.

The interaction between user and application is implemented with JavaScript. An interactive map is integrated to the application to support basic zoom, pan operation. Besides, user can click on the map to select OD or get detailed information for a single link. A time slider is designed for the map. When it is dragged the map will be updated simultaneously.

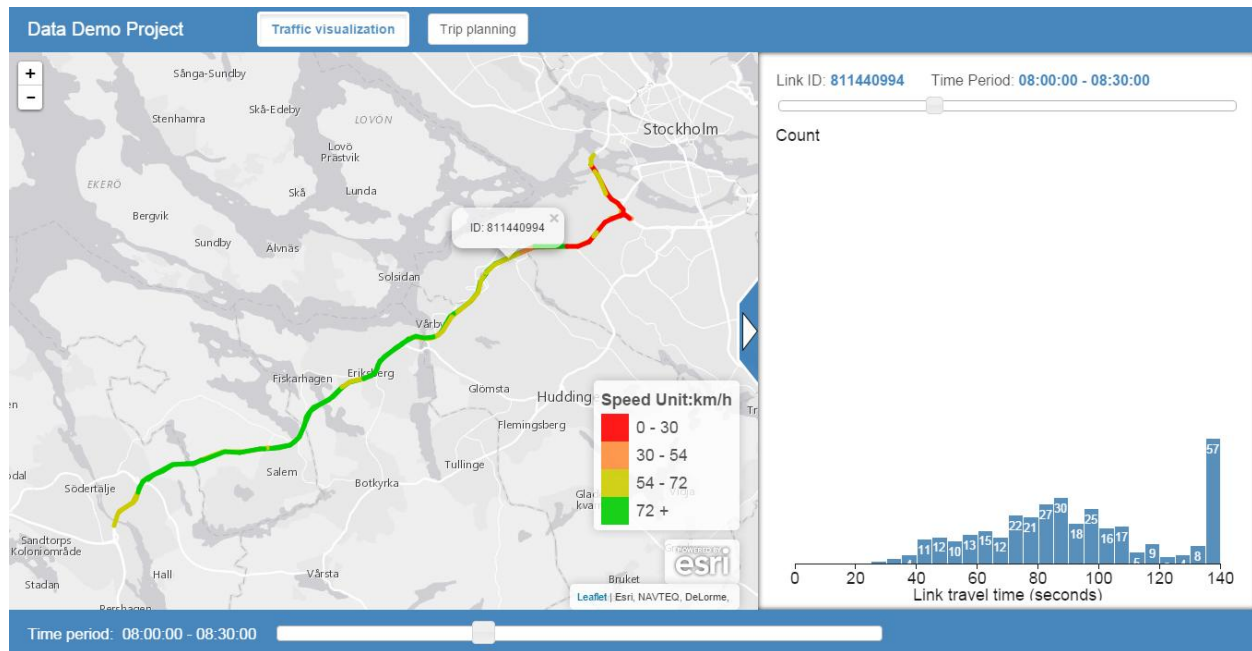


Figure 2 User interface for traffic visualization with a map showing the traffic status on the left and a histogram showing link travel time distribution on the right. The map and histogram can be updated with time controlled by the two sliders respectively.

2.2 Traffic information estimation

In traffic information estimation, an existing map matching and path inference approach (3) is adopted to estimate path travel time from GPS probe data. The path used here is defined as a sequence of links that the vehicle traverses between two consecutive probes. The principles are explained below. In map matching, a group of candidate nodes on the network can be found for each GPS probe by defining a search region. For consecutive probes, a candidate graph can be constructed by connecting all combinations of candidate nodes with path found by the shortest path algorithm. The path between two consecutive probes can be inferred as the most likely path in the candidate graph.

In this study, GPS path with travel time and constituted links is assumed already obtained. Link travel time (LTT) can be estimated for various periods of time by allocating the observed path travel time in corresponding periods to individual links based on length. In congestion visualization, the LTT can be used to derive traffic status on a network. In trip planning, user specified route is also represented as a sequence of links. The observed travel time of a path is first allocated to the overlapped section between the path and route then scaled up to the entire route to estimate route travel time. To measure the reliability of the estimation, a weight is calculated for the overlapped proportion and a threshold is defined to filter out these unreliable estimations. Since sparse GPS data is used in this study, a uniform threshold of 0.1 is adopted.

2.3 Congestion indices and visualization

The traffic information visualization module enables the user to identify and explore of spatial and temporal patterns of congestion and temporal visualization of link travel time distribution.

The congestion can be measured in various ways such as average speed and congestion indices. Average speed is calculated by dividing the length by the average travel time for each link and time period. Apart from average speed, previous studies have proposed some useful measures for congestion and travel time (21). Some of them require inputs that are unavailable in this study such as traffic volume and vehicle miles traveled and thus cannot be used in the current study.

Based on the travel time estimated from GPS probe data, four indices are calculated for each link and time period: speed index (SI), travel time index (TTI), planning time index (PTI) and buffer time index (BI), which are summarized as below

$$SI = \frac{\text{Average speed}}{\text{Speed limit}} \quad (1)$$

$$TTI = \frac{\text{Average travel time}}{\text{Freeflow travel time}} \quad (2)$$

$$PTI = \frac{95^{\text{th}} \text{ percentile travel time}}{\text{Freeflow travel time}} \quad (3)$$

$$BI = \frac{95^{\text{th}} \text{ percentile travel time} - \text{average travel time}}{\text{Average travel time}} \quad (4)$$

SI is the ratio of average speed divided by speed limit of a link. TTI is defined as the ratio of average travel time over free flow travel time for a section of freeway (21). In this study, the free flow travel time is calculated by dividing length by speed limit of each link. TTI measures the difference between the actual travel time and free flow travel time. PTI and BI measure the reliability of travel time and has been used to supplement other congestion measures (22). PTI is defined as ratio of 95% percentile of travel time over average free flow travel time and BI is defined as the difference between 95% percentile travel time and average travel time divided by average travel time.

As can be seen from the above definition, all the indices are continuous scale and dimensionless. When directly scaling them to colors, the difference between colors is small and hard to distinguish for users. In Geographic Information System (GIS), a commonly used approach for visualization of continuous scale data on a map is to classify the data to finite number of classes and assign meaningful and distinct colors for each class. Although user can manually set the classification schemes, the visualization result could be influenced by human behavior. To minimize the influence of artifacts, several classification schemes have been proposed including equal interval, equal quantile, natural break and standard deviation (23). In order to compare all congestion indices and save computational cost in our study, a uniform equal interval classification scheme is adopted and each index is classified into four classes. For example, if speed index ranges from 0 to 1, its class interval will be 0.25.

3. DEMONSTRATION STUDY

3.1 Study area and data

The floating car data provided by Scania AB is used in our demonstration. Each GPS probe reports its vehicle id, location and timestamp of record. The study area is a freeway corridor between Södertälje and Stockholm, Sweden, where many trucks travel in between everyday. The corridor consists of road segments with three speed limits: 70 km/h, 90km/h and 100km/h, which is shown in Figure 3. Using GPS positions observed, there are about 1.2 million link travel time values derived for the freeway. Note a pair of GPS positions might lead to several link travel

times depending on the network topology. The data was collected from August in 2009 to August in 2012.

When applying the collected data for trip planning and traffic visualization, there may be several potential biases. The first one is that the traffic condition may not be stable during the three years. That is reduced partly in this study as the data are skewed towards more recent times with 7% in 2009, 25% in 2010, 41% in 2011 and 26% in 2012 (8 months). Another bias is that the freight truck usually has a low acceleration. In this paper, the biases are ignored as we mainly use the data for a demonstration of the system. With more and higher-quality data feed into the system, the estimation results can be improved.

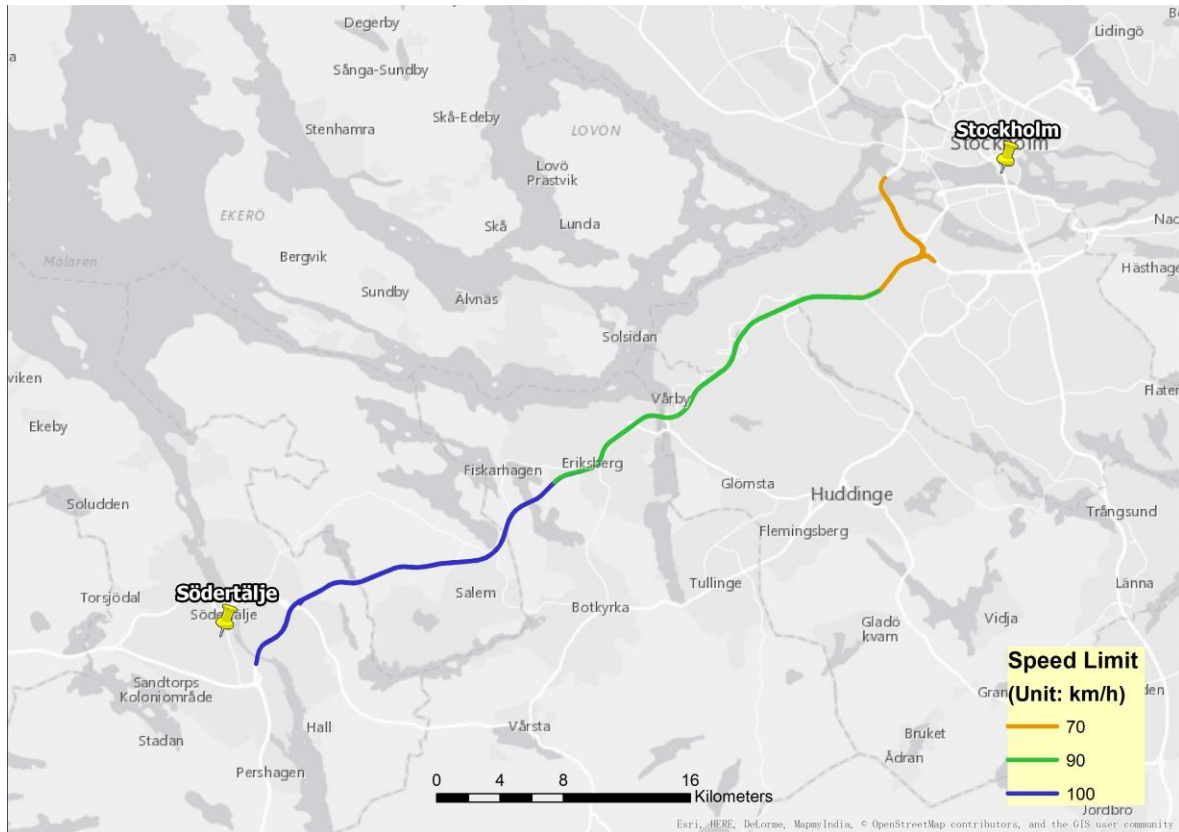


Figure 3 Study area: a corridor between Södertälje and Stockholm with three types of speed limit: 70km/h, 90km/h and 100 km/h.

3.2 Trip planning application

An example of OD selection and routing is presented in Figure 4 and the corresponding route travel time information presented in the side panel is shown in Figure 5. The side panel contains two tabs. The first tab presents length of path in text and temporal variation of average route travel time in a line chart. From that figure, it can be observed that the travel time on the route is much higher in the period 6 am to 8 am, which matched with the expectation of congestion in morning peak period. Another congested period, which is less obvious, is from 4 pm to 5 pm. The reason may be that the route direction is towards the city center and in the morning there would be more vehicles entering the city so the congestion is also heavier. The second tab

demonstrates the distribution of all historical route travel time estimated, which is shown in Figure 5 (right). It can be seen that in most cases the travel time on the route is around 3 minutes but it may exceed 6 minutes when the route is too congested.

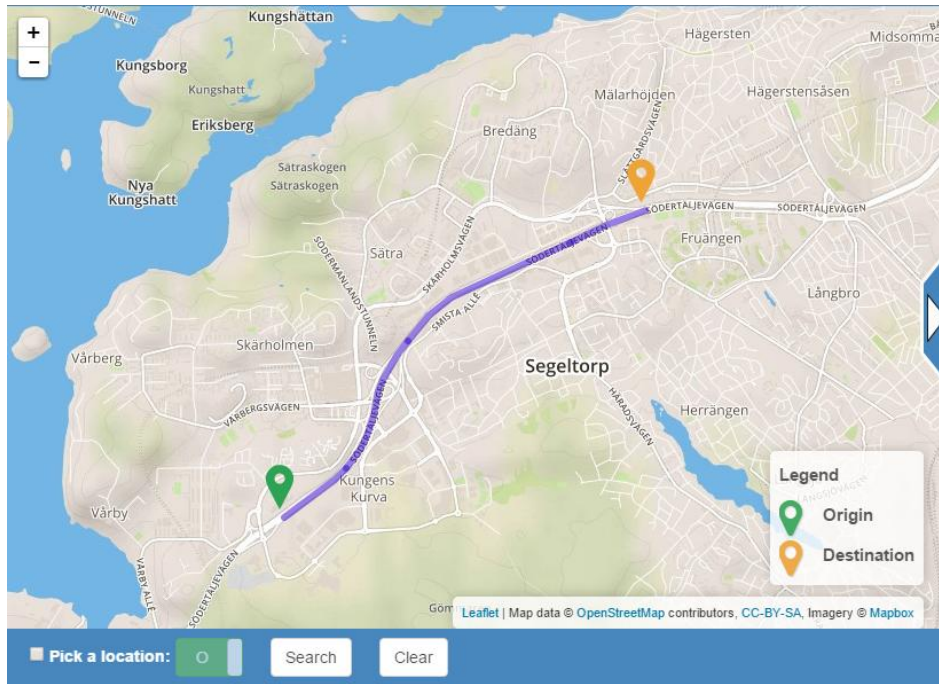


Figure 4 Origin and destination picker and routing

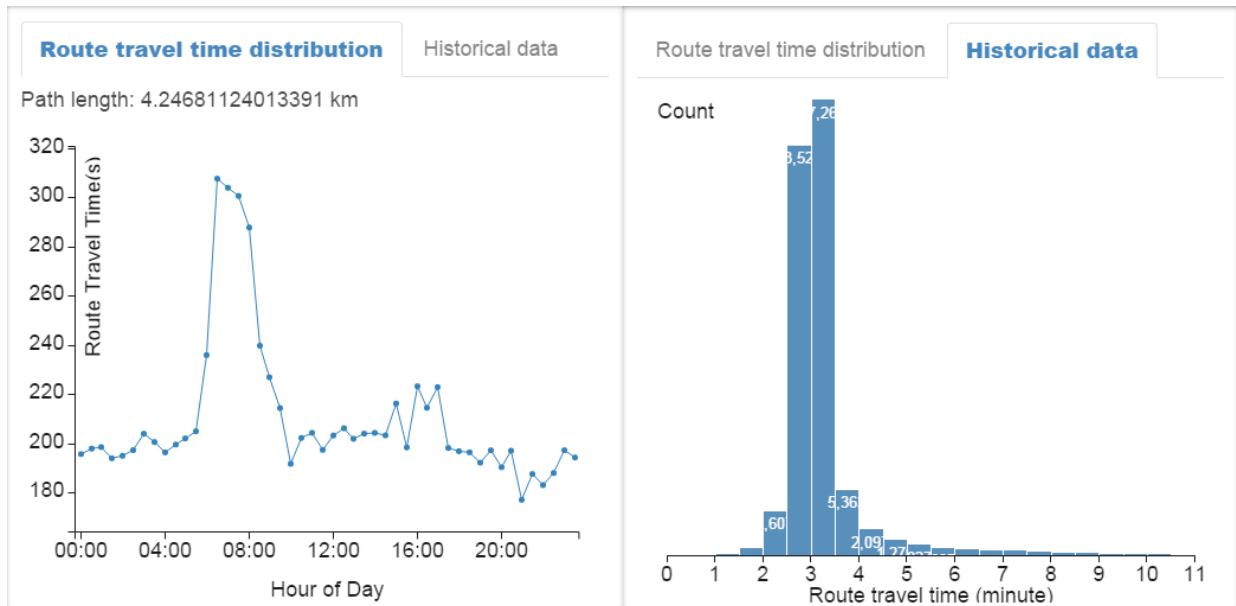


Figure 5 OD based trip planning: temporal variation of average route travel time (left) and all historical route travel time distribution (right)

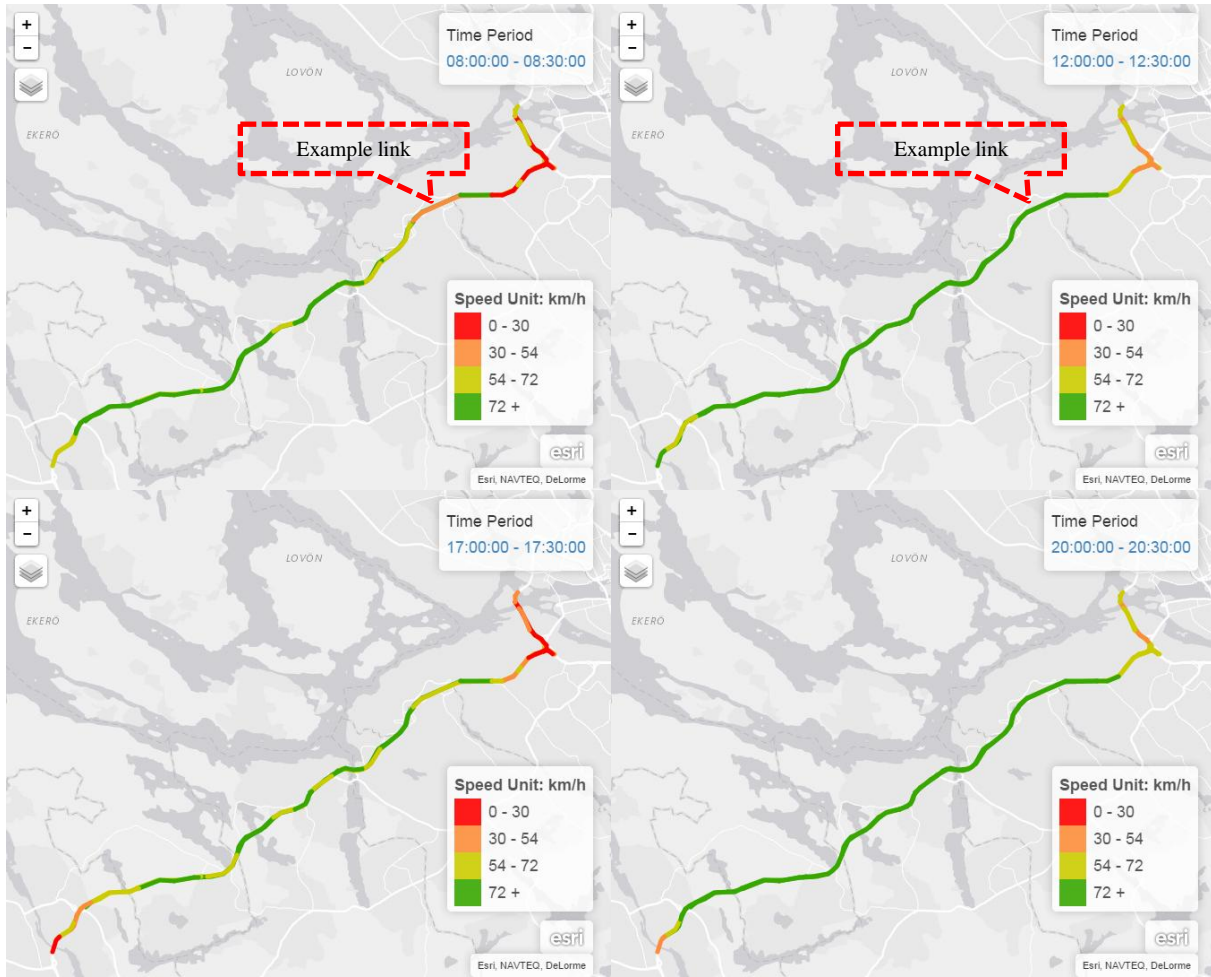


Figure 6 Visualization of link average travel speed for four periods 8:00-8:30 (upper left), 12:00-12:30 (upper right), 17:00-17:30 (under left) and 20:00-20:30 (under right)

3.3 Traffic visualization application

The traffic application enables user to explore spatial and temporal pattern of traffic on the network. An example is demonstrated in Figure 6. Four periods comprising two peak periods 8:00 – 8:30 and 17:00 - 17:30 and two non-peak periods 12:00 - 12:30 and 20:00 - 20:30, are selected to demonstrate the average link travel speed distribution on the freeway. Three thresholds 30km/h, 54km/h and 72km/h are manually defined to classify the average speed into four levels for clear presentation. It can be seen that links showing an average speed under 30 km/h can only be observed in the two peak periods. These low speed level links are majorly distributed in the region close to Stockholm City and in the evening peak period 17:00 – 17:30, some links close to Södertälje also show a low speed under 30 km/h, which means that the traffic on links close to Södertälje is heavier in the evening than that in the morning. In the two non-peak periods, all the links have an average speed over 30 km/h and the average speed on roads far away from the two city centers stays at a high level over 72 km/h. Even in the peak periods, these roads' average speed are still higher than 54km/h. As a summary, the results reveal that

the congestion on the freeway majorly occurs in peak periods on road segments close to the two city centers and for road close to Södertälje the congestion is heavier in the evening.

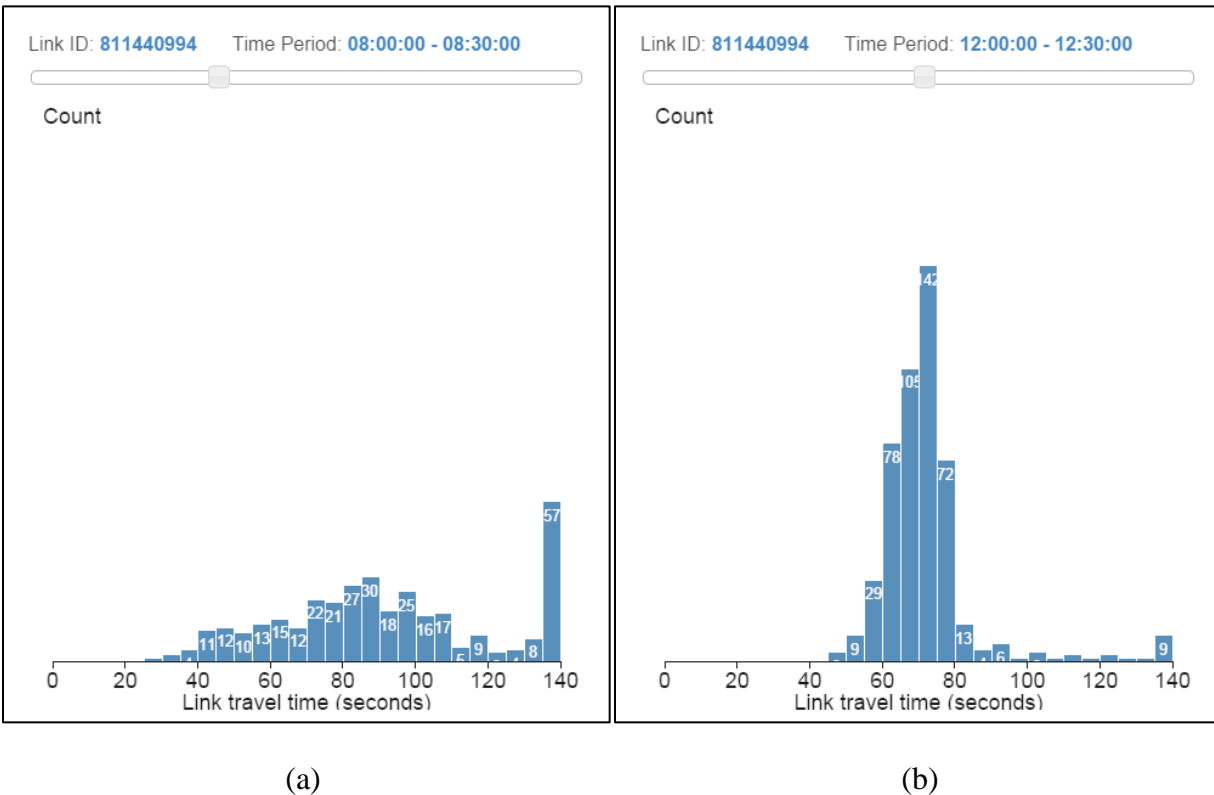


Figure 7 Example of link travel time distribution visualization: 8:00 -8:30 (a) and 12:00-12:30 (b)

Apart from exploring the spatial and temporal pattern of traffic, user can also click on links to check the detailed link travel time distribution, which is plotted as a histogram in the side panel. In Figure 7, an example of LTT distribution from two periods is presented for the “example link” marked in Figure 6. The histograms are set with fixed bins to simulate an animation for efficiently presenting the change over time. It can be seen that when the link shows a low average speed (30km/h – 54km/h) in morning peak period, LTT is distributed sparsely but in non-peak hour, its distribution is more concentrated. The link travel time visualization functionality presents a detailed knowledge of road traffic. The simultaneous visualization of road traffic on the map and link travel time on the side panel provides an interactive and efficient way for data exploration, which could be helpful for researchers studying on patterns of link travel time.

Congestion indices are also integrated in the traffic visualization application. The distributions of SI, TTI, PTI and BI in morning peak period 8:00- 8:30 and evening peak periods 17:00- 17:30 are demonstrated in Figure 8 and Figure 9 respectively. For clear presentation, the background base map is hidden. From Figure 8 it can be observed that the pattern of congestion on the freeway measured by TTI is similar with PTI and SI is similar with BI. From both indices, the congested links are identified to be close to the Stockholm City. Although the ranges of indices are different, with equal interval classification scheme used, TTI and PTI identified similar road links. SI has the most diverse visualization outcome. Comparing the two figures, it

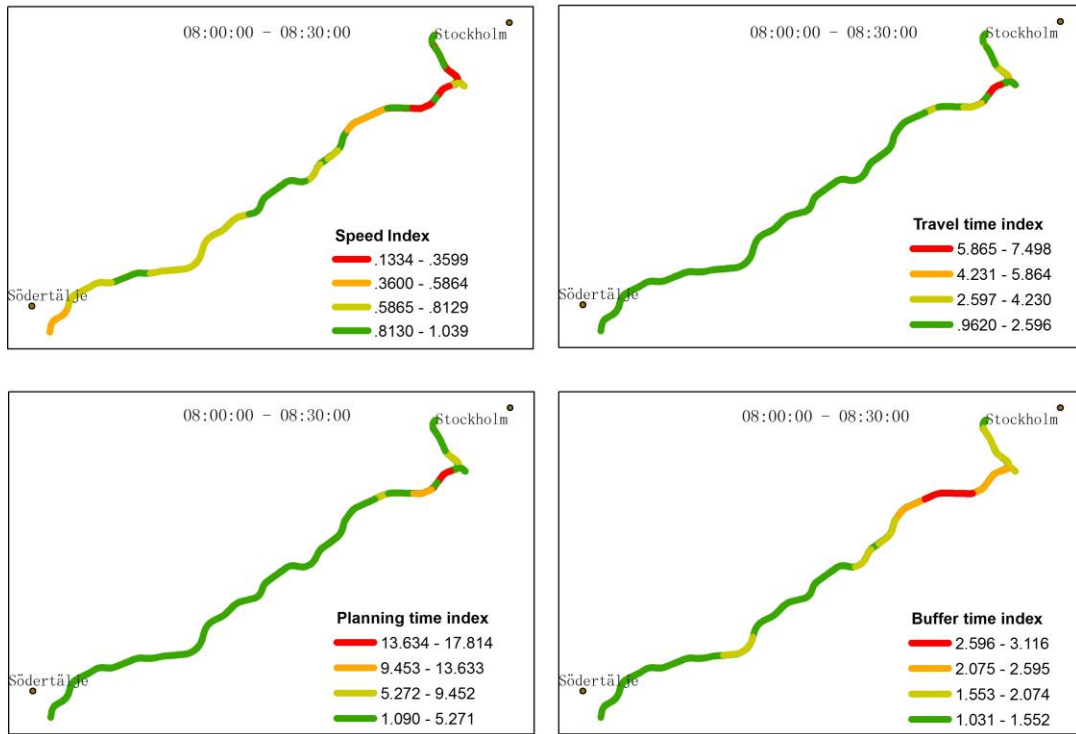


Figure 8 Congestion indices for morning peak period

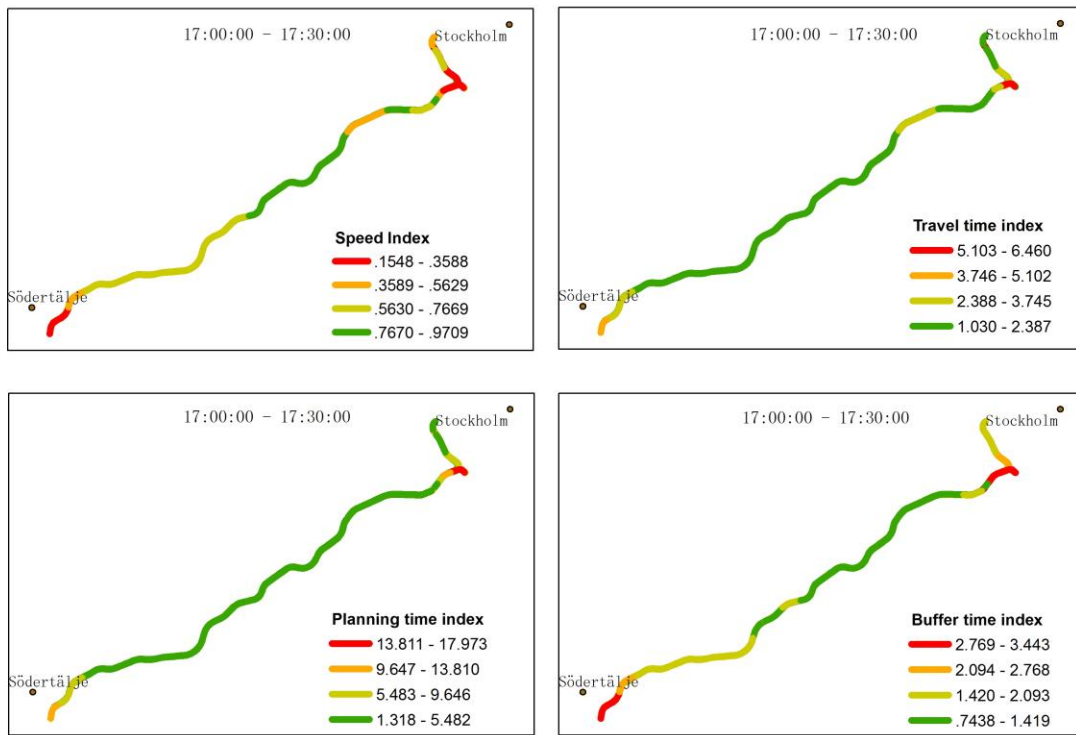


Figure 9 Congestion indices for evening peak period

also demonstrates that in the evening peak hour, the congestion close to the Södertälje is heavier than the morning peak hour. Both TTI and PTI are homogeneously distributed on the road while SI and BI are distributed in a heterogeneous way. The reason is that PTI and TTI, has a larger range than SI and BI so that equal interval classification tends to generate class with large interval. When only a few links depict large index value, most of left links will be divided into the same class. At the same time, both the travel time index and planning time index highlighted the most time consuming road links are those close to Stockholm.

4. CONCLUSIONS

The main objective of this paper is to demonstrate the ITS applications of FV data collected from freight transport. The study develops a web-based information system by integration of database technology, information estimation methods and traffic visualization approaches. Based on travel time information retrieved from freight GPS data, the system provides a simple and interactive user interface for two applications. The first one is a trip planning application supporting user-selected origin-destination (OD) routing and visualization of the temporal variation of average route travel time as well as distribution of all historical route travel time. The second one is a traffic visualization application where an animated map and a slider are designed, enabling users to explore congestion on a network in various periods of day and to check and utilize the link travel time distribution simultaneously.

The system has potential both for more complex commercial applications and demonstration on scientific research. The trip planning application could be used for applications in better logistic planning. Traffic visualization application is useful for planners in analyzing traffic or congestion patterns in an area. Besides, simultaneous visualization of traffic and link travel time has the potential for assisting researchers in studying patterns of travel time. Our future study would be focused on more integrated data mining and analysis as well as more efficient visualization of traffic patterns with different resolutions on time and space.

ACKNOWLEDGMENT

The authors would like to acknowledge the funding support by Integrated Transport Research Lab (ITRL), a joint research center among KTH, Scania and Ericsson. Scania provides the data for the application development. The first author is a PhD student financially supported by the China Scholarship Council (CSC).

REFERENCES

1. Ma, X.-L., and H. Koutsopoulos. Estimation of the automatic vehicle identification based spatial travel time information collected in Stockholm. *Intelligent Transport Systems, IET*, Vol. 4, No. 4, 2010, pp. 298-306.
2. Jenelius, E., and H. N. Koutsopoulos. Travel time estimation for urban road networks using low frequency probe vehicle data. *Transportation Research Part B: Methodological*, Vol. 53, 2013, pp. 64-81.
3. Rahmani, M., E. Jenelius, and H. N. Koutsopoulos. Non-parametric estimation of route travel time distributions from low-frequency floating car data. *Transportation Research Part C: Emerging Technologies*, Vol. 58, Part B, 2015, pp. 343 - 362.

4. Ramezani, M., and N. Geroliminis. On the estimation of arterial route travel time distribution with Markov chains. *Transportation Research Part B: Methodological*, Vol. 46, No. 10, 2012, pp. 1576-1590.
5. Ehmke, J. F., S. Meisel, and D. C. Mattfeld. Floating car based travel times for city logistics. *Transportation Research Part C: Emerging Technologies*, Vol. 21, No. 1, 2012, pp. 338-352.
6. Simroth, A., and H. Zöhle. Travel time prediction using floating car data applied to logistics planning. *Intelligent Transportation Systems, IEEE Transactions on*, Vol. 12, No. 1, 2011, pp. 243-253.
7. Yang, X., Z. Sun, X. J. Ban, and J. Holguín-Veras. Urban Freight Delivery Stop Identification with GPS Data. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2411, 2014, pp. 55-61.
8. Liu, X., and Y. Ban. Uncovering Spatio-Temporal Cluster Patterns Using Massive Floating Car Data. *ISPRS International Journal of Geo-Information*, Vol. 2, No. 2, 2013, pp. 371-384.
9. De Fabritiis, C., R. Ragona, and G. Valenti. Traffic estimation and prediction based on real time floating car data. In *Intelligent Transportation Systems, 2008. ITSC 2008. 11th International IEEE Conference on*, IEEE, 2008. pp. 197-203.
10. Liao, C.-F. Generating Reliable Freight Performance Measures with Truck GPS Data: Case Study in Twin Cities Metropolitan Area, Minnesota. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2410, 2014, pp. 21-30.
11. Chen, C., K. Petty, A. Skabardonis, P. Varaiya, and Z. Jia. Freeway performance measurement system: mining loop detector data. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1748, 2001, pp. 96-102.
12. Wongsuphasawat, K., M. L. Pack, D. Filippova, M. VanDaniker, and A. Olea. Visual Analytics for Transportation Incident Data Sets. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2138, 2009, pp. 135-145.
13. Lund, A. S., and M. L. Pack. Dynamic Wide-Area Congestion and Incident Monitoring Using Probe Data. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2174, 2010, pp. 1-9.
14. Ma, X., Y.-J. Wu, and Y. Wang. DRIVE Net: E-Science Transportation Platform for Data Sharing, Visualization, Modeling, and Analysis. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2215, 2011, pp. 37-49.
15. Wu, Y.-J., S. An, X. Ma, and Y. Wang. Development of Web-Based Analysis System for Real-Time Decision Support on Arterial Networks. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2215, 2011, pp. 24-36.
16. Ma, X., E. D. McCormack, and Y. Wang. Processing Commercial Global Positioning System Data to Develop a Web-Based Truck Performance Measures Program. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2246, 2011, pp. 92-100.
17. *INRIX Traffic App*. <http://inrix.com/inrix-traffic-app/>. Accessed November 13, 2015.
18. *Waze*. <http://www.waze.com/livemap>. Accessed November 13, 2015.
19. *Google Maps*. <http://www.google.se/maps>. Accessed November 13, 2015.
20. Dijkstra, E. W. A note on two problems in connexion with graphs. *Numerische mathematik*, Vol. 1, No. 1, 1959, pp. 269-271.

21. Schrank, D. L., and T. J. Lomax. *The 2007 urban mobility report*. Texas Transportation Institute, Texas A & M University, 2007.
22. Lyman, K., and R. Bertini. Using travel time reliability measures to improve regional transportation planning and operations. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2046, 2008, pp. 1-10.
23. Brewer, C. A., and L. Pickle. Evaluation of methods for classifying epidemiological data on choropleth maps in series. *Annals of the Association of American Geographers*, Vol. 92, No. 4, 2002, pp. 662-681.