2015 IEEE 18th International Conference on Intelligent Transportation Systems

DynaMIT2.0: Architecture Design and Preliminary Results on Real-time Data Fusion for Traffic Prediction and Crisis Management

Yang Lu
Singapore-MIT Alliance for Research and Technology, Future Urban Mobility
1 CREATE Way, #09-02 CREATE Tower, Singapore 138602
Email address: luyang@smart.mit.edu

Francisco Câmara Pereira
Technical University of Denmark (DTU), Bygningstorvet 116,
2800 Kgs. Lyngby, Denmark
Massachusetts Institute of Technology (MIT)
Room 1-280, 77 Massachusetts Avenue, Cambridge, MA, 02139
Email address: camara@mit.edu

Ravi Seshadri
Singapore-MIT Alliance for Research and Technology, Future Urban Mobility
1 CREATE Way, #09-02 CREATE Tower, Singapore 138602
Email address: ravi@smart.mit.edu

Aidan O'Sullivan
Singapore-MIT Alliance for Research and Technology, Future Urban Mobility
1 CREATE Way, #09-02 CREATE Tower, Singapore 138602
Email address: aidanosull@gmail.com

Constantinos Antoniou
National Technical University of Athens
9, Iroon Polytechniou St., Zografou Campus, 15780
Email address: antoniou@central.ntua.gr

Moshe Ben-Akiva
Massachusetts Institute of Technology
Room 1-181, 77 Massachusetts Avenue, Cambridge, MA, 02139
Email address: mba@mit.edu

July 1, 2015

978-1-4673-6596-3/15 $31.00 © 2015 IEEE
DOI 10.1109/ITSC.2015.363

I. ABSTRACT

The ability to monitor and predict in real-time the state of the transportation network is a valuable tool for both transportation administrators and travelers. While many solutions exist for this task, they are generally much more successful in recurrent scenarios than in non-recurrent ones. Paradoxically, it is in the latter case that such tools can make the difference. Therefore, the dynamic traffic assignment and simulation based prediction system such as DynaMIT (1) demonstrates high effectiveness in the context of sudden network disturbance or demand pattern changes. This paper presents the design, development and implementation of new components and modules of DynaMIT 2.0 which is an extension of its predecessor with recent enhancements on online calibration, context mining, scenario analyzer and strategy simulation capability. Also,some preliminary results are presented using Singapore expressway to show the actual benefit of the system.

II. INTRODUCTION

Real-time traffic prediction systems can be generally divided into two types: data driven and model-based. The data-driven approach relies on fitting the current scenario to previous observations, often from a regression problem perspective where the target variable is a continuous quantity (e.g. speed, travel time, queue length) and the independent variables are obtained from sensors and spatial and temporal characteristics. Common algorithms include time series (2), artificial neural networks (3), Kalman filters (4) or combinations of techniques (e.g. fuzzy sets and genetic algorithms (5)). On the other hand, the model-based approach represents the traffic phenomenon itself, essentially through some form of simulation where supply (infrastructure, services) interacts with demand (travellers). By incorporating rather sophisticated models of traffic flow and travel behaviour, this approach allows the simulation of complex situations that may have little historical precedence (i.e. poor past data). This is the approach of dynamic traffic assignment (DTA) models such as DynaMIT (1) or DynaSMART (6).

The value proposition of both data-driven and model-based approaches is quite clear: the former offers generally efficient models that will work well under all circumstances that have relevant data precedent; the latter offers a holistic solution that can deal with any circumstance as long it is representable in the simulation and is computable in reasonable time. The limitations are also clear: data-driven approaches are too dependent on generalising from historical data and on representing the spatial temporal correlation structure of the network; the model-based approach is computationally intensive and demands accurate representation and calibration of all relevant real world phenomena.

Some research has attempted to overcome these limitations, particularly on the data-driven side. For example, Chen et al (7) implemented a Gaussian Processes model where the prediction for a certain point of the network considers not only its local characteristics but also its correlation with the rest of the network (and their predicted/estimated values). In fact, the usage of more complex models that account for spatial-temporal correlations across the network has become a natural option, for example with time series models (8) or support vector machines (9). On the model-based side, progress has also been made in several directions including computational efficiency improvement (10), joint calibration of supply and demand parameters (11, 12), representing incident information in detail (13).

Building on lessons learned from both directions, the DynaMIT 2.0 project is designed to take advantage of data driven methods to improve efficiency of the model-based approach. Each prediction cycle starts with a recognition of the current patterns (e.g. incidents, special events, weather), which provide a priori estimates for the variables of interest, both for supply and demand (e.g. capacities, speed-density parameters, Origin/Destination flows). Then, the DTA model uses these reference values to calibrate itself, considering available real-time and historical data.

DynaMIT2.0 uses a varied set of data streams, including surveillance sensors (inductive loops, cameras), floating car data (e.g. GPS probes), traffic information feeds (e.g. incident information feed) as well as data from the internet (e.g. special events websites, weather forecasts, social networks). Due to its heterogeneous nature, these sources enter DynaMIT2.0 at different stages and through different components. DynaMIT2.0 provides three relevant new functionalities:the *scenario analyzer* analyzes non-structural text based information and generates estimates of corresponding prior demand and supply parameters; *integrated online calibration* module adjusts all supply and demand parameters according to real-time traffic feeds with heterogeneous data structure;finally, *strategy simulation* allows the simultaneous evaluation of different treatment strategies for the mitigation of non-recurrent congestion.

III. BACKGROUND: REVIEW OF DYNAMIT 1.0

DynaMIT is a dynamic traffic assignment (DTA) system for traffic estimation and prediction developed at MIT Intelligent Transportation System Laboratory (1). The system receives real-time data (e.g. vehicle counts) and historical data. The core is composed of two sub-systems: 1) State estimation and 2) State Prediction. The operational structure of DynaMIT is presented in Figure 1.

State estimation combines the available surveillance with historical information to estimate the current state of the entire network. Pre-trip demand is simulated, allowing drivers with different characteristics to dynamically change their departure time, travel mode, and trip route. This is followed by OD flow estimation and network estimation. The estimated network condition is then compared to surveillance information. Based on this estimate of the current state of the network, the DTA model predicts future traffic patterns, taking into account the response of the drivers to the provided guidance and traffic information in the stage of state prediction. The outputs of the overall system are consistent forecasts of network conditions, including link density, flow speed, as well as travelers' characteristics including their travel time, route choice and departure time. The anticipated information is used to generate guidance and will be incorporated into the next round of calculations (1).

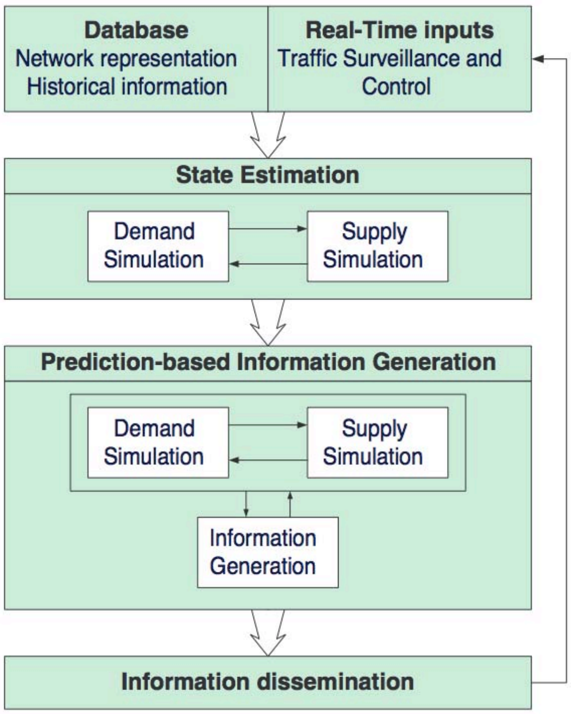

FIGURE 1 Architecture of DynaMIT1.0

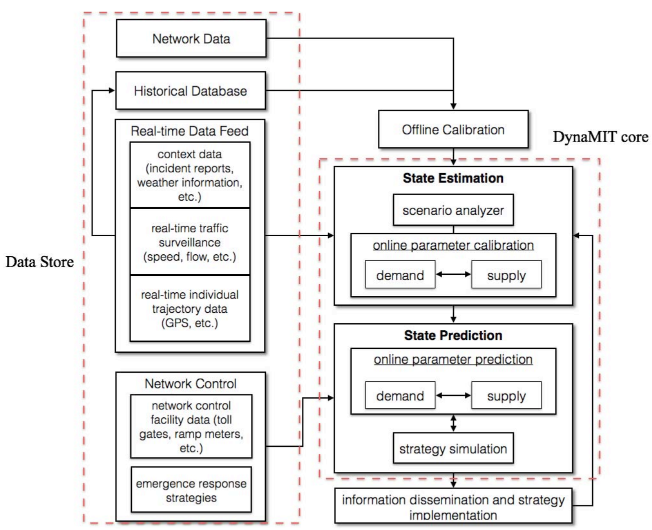

FIGURE 2 Architecture of DynaMIT2.0

IV. DYNAMIT2.0 ARCHITECTURE

Figure 2 presents the overall architecture of DynaMIT 2.0. The system is composed of two major module clusters: the data store maintains and manages all the data used by the system (both historical and real-time); the DynaMIT core is the collection of key models perform state estimation and prediction. State estimation module contains scenario analyzer and context mining, online parameter calibration; state prediction module contains online parameter prediction and strategy simulation (see right side Figure 2). The function and algorithm of each model is briefly summarized below:

A. *Scenario Analyser*

The role of the Scenario Analyser (SA) is to translate incoming data streams into a demand/supply parameter pattern that is understandable for DynaMIT2.0. All text based information are converted useful format through *context mining*. For example, special events websites provide rich spatial, temporal and semantic information that correlate with observed demand fluctuations (14, 15), social networks and micro-blogs such as Facebook and Twitter provide information in events from incidents (16) to crowd gatherings (17).

While DynaMIT core follows the model-based approach, SA builds on the data driven approach. It does so in a modular, plug-in, approach, as illustrated in Figure 3, that consists of a set of independent modules. These Scenario Analyser Modules (SAMs) need only respect the simple rule that their output match at least partially DynaMIT core's parameters. Currently, scenario analyzer contains two models: incident duration prediction and special event demand prediction. Incident prediction reads the real-time incident reports text and performs recursive duration prediction through a topic modeling technique(18). The special event SAM analyzes the web based event (social gathering, etc.) information using a Bayesian additive linear model (or BALM) and predicts the change of total arrivals at different locations. For more detail discussion, readers can refer (14).

B. *Integrated online parameter calibration*

DynaMIT 2.0 adopts an integrated online parameter calibration model to dynamically adjust all of its internal parameters based on real-time observations. The parameter set contains three types: time-dependent OD matrices, behavior model and supply parameters. The structure of online calibration module is illustrated by Figure. DynaMIT 2.0 adopts a two-stage sequential calibration method. In the first stage, OD matrices are calibrated using link flow counts through a Generalized Least Square formulation; in the second stage, route choice behavior and network supply parameters are calibrated using a SP-EKF method (simultaneous perturbation-extended Kalman filter). Also, the online calibration takes full advantage of the prior values provided by SAMs.

C. *Strategy simulation*

The function of strategy simulation is to test various control strategies in real-time to mitigate traffic congestion ranging from regular congestion to major crisis situations. By taking advantage of the parallel computing technique, strategy simulation component allows the manager to simultaneously test several options, taking into account estimated and predicted traffic state, together with travel behaviors. Optimization are performed to either minimize total travel time or maximize total traveler welfare. Currently GA (generic algorithm) is implemented in strategy simulation module to optimize network control parameters such as dynamic tolls or ramp metering rates.

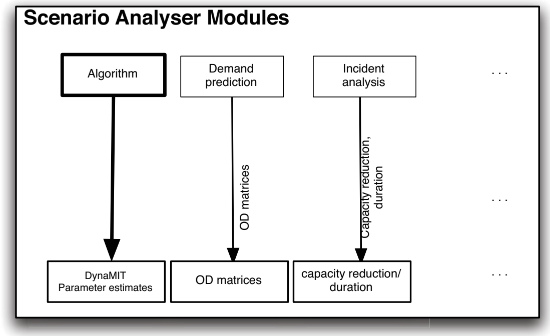

FIGURE 3 Scenario Analyser architecture

V. Preliminary results

Recognizing the ambition of the DynaMIT2.0 vision, the team has been focused on research and development of the individual modules, on a bottom-up fashion. Some results will be summarized in this section, focusing on the research achievements. We will not further detail the C++ software implementation of the system, which is in progress simultaneously to the research efforts.

For testing and model verification purpose, all the numerical results shown in this paper is conducted in a closed loop environment where the microscopic traffic simulation MITSIM is

A. Online Calibration: Closed Loop Experiments on Singapore Expressway

This section reports results from preliminary numerical experiments to investigate the performance of various online calibration algorithms. For the experiments, we make use of a closed-loop framework, interfacing DynaMIT2.0 and MITSIMLab, a microscopic simulator of the real network (19). MITSIMLab acts as a proxy to the real world, and mimics the real network, providing sensor counts for the current interval to DynaMIT. For the parameter calibration, four algorithms are compared, (1) dynamic OD estimation using a Generalized Least Squares (referred to as GLS), (2) dynamic OD estimation using an Extended Kalman Filter approach (referred to as EKF) (3) sequential calibration of demand and supply parameters using the GLS and SP-EKF method (referred to as Sequential-1), (4) sequential calibration of demand and supply parameters using EKF and SP-EKF methods (referred to as Sequential-2).

The experimental inputs for DynaMIT consist of a historical OD matrix obtained from offline calibration on the Singapore expressway network (20). The time intervals for state estimation and prediction are respectively 5 minutes and

ALGORITHM	Estimation		Prediction (RMSN)		
	MAPE	RMSN	1 Step	2 Step	3 Step
GLS	11.38%	14.98%	20.59%	21.11%	21.39%
EKF	11.25%	16.19%	18.94%	19.57%	20.53%
Sequential-1 (GLS for OD)	11.13%	14.01%	20.32%	20.82%	21.35%
Sequential-2 (EKF for OD)	11.18%	15.52%	18.23%	18.81%	19.71%
Base (Historical)	23.19%	27.69%			

TABLE I Calibration and prediction results: Errors in sensor flow counts

5/10/15 minutes. The network contains 650 sensors and 4103 origin-destination pairs.

For state estimation, the sequential-1 calibration method outperforms the remaining algorithms and yields a RMSN value of 14%, which constitutes an improvement of around 50 % over the base case (refer Table I). In contrast, for state prediction, the sequential-2 method yields the best performance with an RMSN of 18.2%, which constitutes an improvement of 34% over the base case. The estimated and predicted counts for an arbitrary sensor are plotted against the actual and historical counts in Figure 4. The preliminary experiments indicate a promising performance of the sequential calibration procedures and the next step involves comprehensive testing using real time data.

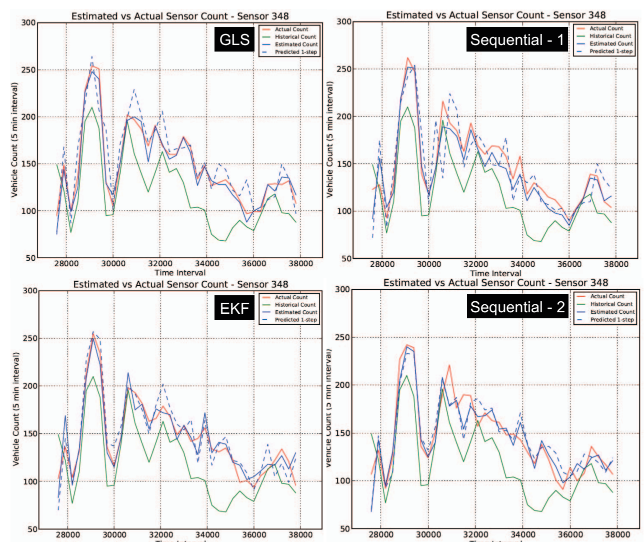

FIGURE 4 Estimated/Predicted/Actual flow counts on selected sensor of Singapore expressway network

B. Incidents SAM

In DynaMIT 2.0, an incident is represented as a capacity reduction in the affected link during a period of time. It has been demonstrated that its predictions benefit considerably from the accurate consideration of such events (13), however the task of translating incoming incident information into that representation is not a trivial one. This is the role of the incidents SAM: given an incoming feed of incident information, estimate the capacity reduction and predict the duration until clearance (18).

This is done in a sequential and reactive manner. Incidents SAM sequentially updates its estimates taking into account the latest received text information. Topic modelling techniques (Latent Dirichlet Allocation, LDA (21)) is applied to extract a numeric combination of latent topics which is in turn used for duration prediction using a radial basis function (RBF) algorithm.

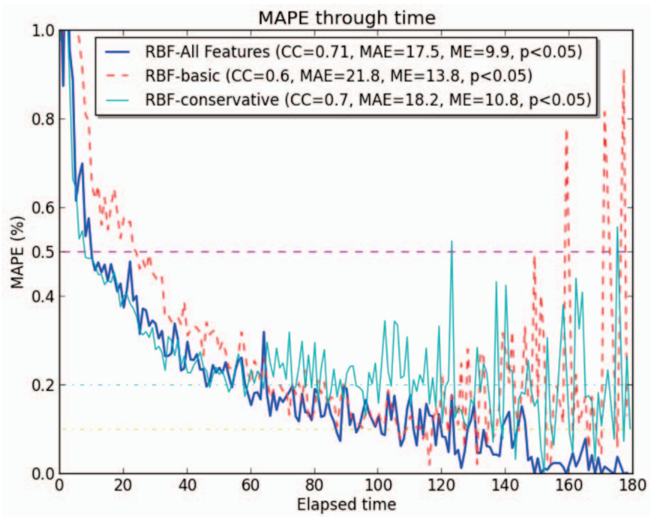

FIGURE 5 MAPE performance of incident SAM's duration prediction algorithms (CC=Correlation Coefficient; MAE=Mean Absolute Error; ME=Mean Absolute Error).

In Figure 5, we show the performance of our sequential incident duration prediction algorithm, in 3 flavours: basic, which has the report creation information together with elapsed time; conservative, which as textual information together with elapsed time; all features, which has all possible information. The graph shows the Mean Absolute Percentage Error (MAPE) through time, and it can be seen that the complete model shows the best and more resilient performance. Interestingly, the text only (conservative) model is able to have comparable performance until 90 minutes. The reader will find more details in (18).

C. Incident Response Case Study

This section demonstrates the benefits of the scenario analyzer module of Dynamit 2.0 in the context of crisis management through a simple case study on the Singapore Expressway network using the closed loop framework. Specifically, the study investigates the role of providing predictive information and real time guidance in mitigating the non-recurrent congestion caused by traffic accidents. A major incident is imposed on the east-bound lanes of the Marina Coastal expressway involving the complete blockage of three out of five lanes (Figure 6) for a duration of 45 minutes between 8:00 AM and 8:45 AM. Two scenarios are then simulated using MITSIMLab, with and without predictive information from Dynamit, for a period between 7:30 AM and 10:00 AM. It is assumed that 75% of the drivers receive the real-time predictions with a compliance rate of 70%, which implies that around 50% of the vehicles are guided.

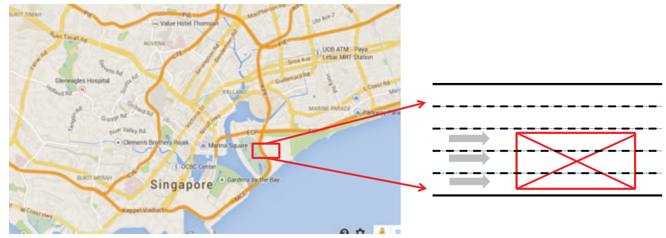

FIGURE 6 Incident Configuration

A comparison of the queue lengths on the affected segment of the MCE for the two scenarios reveals a significant reduction in the maximum queue length which reduces by almost 50% from 3000 to 1500 meters (Figure 7). The plots of queue length from the two scenarios further indicate a significantly smaller queue dissipation time in the scenario with Dynamit (see Figure 7).

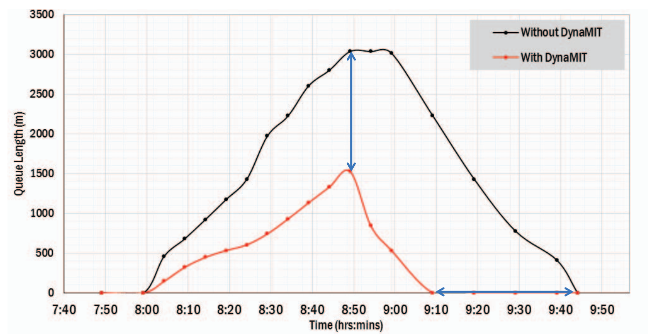

FIGURE 7 Comparison of Queue Lengths on affected corridor

An analysis of the route choice fractions for a subset of relevant O-D pairs further indicates that the provision of predictive real time information results in a significant proportion of users (18%) re-routing from the incident affected MCE route to the alternate CTE-PIE route (refer Figure 8). This results in an overall reduction in network travel times by 12.8% and a total travel time reduction of 31.3% for the relevant vehicles using the incident affected links.The changes in route choice proportions are summarized in Figure 8

VI. Conclusion

This paper presented the architecture and preliminary results of the DynaMIT 2.0 system which aims to combine the data-driven/statistical approaches and traffic assignment/simulation model-based approaches. Building on an extensive experience with its parent project, the development of new functionality is introduced and preliminary results on the Singapore expressway network suggest a promising future for the system. One important future work of DynaMIT 2.0 is to expand its prediction capability into a multi-modal network where both private traffic, public transport services and freight vehicles are taken into account. The online calibration of the model in a multi-modal environment is one of our ongoing studies.

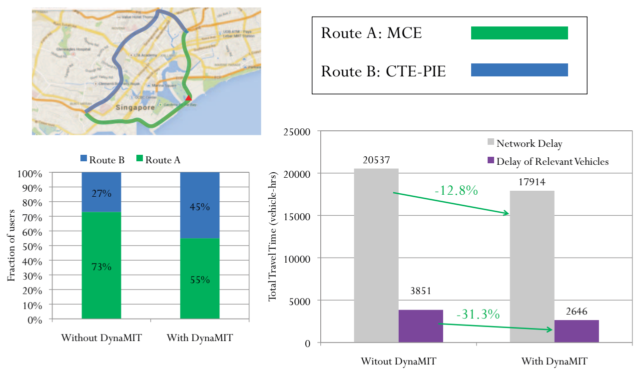

FIGURE 8 Driver Behavior and Network Performance

REFERENCES

[1] Ben-Akiva, M., H. N. Koutsopoulos, C. Antoniou, and R. Balakrishna, Traffic Simulation with DynaMIT. In *Fundamentals of Traffic Simulation* (J. Barceló, ed.), Springer, New York, NY, USA, 2010, pp. 363–398.

[2] Zhou, D., S. Chen, and S. Dong, Network traffic prediction based on ARFIMA model. *IJCSI International Journal of Computer Science*, Vol. 9, No. 3, 2013.

[3] Park, J., D. Li, Y. L. Murphey, J. Kristinsson, R. McGee, M. Kuang, and T. Phillips, Real time vehicle speed prediction using a neural network traffic model. In *Neural Networks (IJCNN), The 2011 International Joint Conference on*, IEEE, 2011, pp. 2991–2996.

[4] Zhang, L., J. Ma, and J. Sun, Examples of Validating an Adaptive Kalman Filter Model for Short-Term Traffic Flow Prediction. *Bridges*, Vol. 10, 2014, pp. 9780784412442–094.

[5] Zhang, X., E. Onieva, A. Perallos, E. Osaba, and V. Lee, Hierarchical fuzzy rule-based system optimized with genetic algorithms for short term traffic congestion prediction. *Transportation Research Part C: Emerging Technologies*, 2014.

[6] Mahmassani, H., T. Hu, and R. Jayakrishnan, Dynamic traffic assignment and simulation for advanced network informatics (DynaSMART). *Urban traffic networks: Dynamic flow modeling and control. Springer, Berlin/New York*, 1995.

[7] Chen, J., K. H. Low, C. K.-Y. Tan, A. Oran, P. Jaillet, J. M. Dolan, and G. S. Sukhatme, Decentralized data fusion and active sensing with mobile sensors for modeling and predicting spatiotemporal traffic phenomena. In *Proceedings of the 28th Conference on Uncertainty in Artificial Intelligence (UAI-12)*, Catalina Island, CA, 2012, p. 163âĂŞ173.

[8] Wynter, L. and W. Shen, Real-Time Traffic Prediction Using GPS Data with Low Sampling Rates: A Hybrid Approach. In *Transportation Research Board 91st Annual Meeting*, 2012, 12-1692.

[9] Asif, M., J. Dauwels, C. Y. Goh, A. Oran, E. Fathi, M. Xu, M. Dhanya, N. Mitrovic, and P. Jaillet, Unsupervised learning based performance analysis of n-support vector regression for speed prediction of a large road network. In *Intelligent Transportation Systems (ITSC), 2012 15th International IEEE Conference on*, IEEE, 2012, pp. 983–988.

[10] Wen, Y., *Scalability of Dynamic Traffic Assignment*. Ph.D. thesis, Cambridge, MA, USA, 2009, aAI0821759.

[11] Balakrishna, R., M. Ben-Akiva, and H. N. Koutsopoulos, Offline calibration of dynamic traffic assignment: simultaneous demand-and-supply estimation. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2003, No. 1, 2007, pp. 50–58.

[12] Antoniou, C., *On-line calibration for dynamic traffic assignment*. Ph.D. thesis, Massachusetts Institute of Technology, 2004.

[13] Lopes, J. A., *Traffic prediction for unplanned events on highways*. Ph.D. thesis, Instituto Superior Tecnico, IST, 2012.

[14] Rodrigues, F., B. Ribeiro, and F. C. Pereira, A Bayesian additive linear model for understanding public transport usage in special events. In *Submitted, under review*, 2014.

[15] Gowrishankar, S., R. Stern, and D. Work, Including the social component in smart transportation systems. In *National Workshop on Transportation Cyber-Physical Systems*, Arlington, VA, 2014.

[16] Mai, E. and R. Hranac, Twitter Interactions as a Data Source for Transportation Incidents. In *Proc. Transportation Research Board 92nd Ann. Meeting*, 2013, 13-1636.

[17] Sugitani, T., M. Shirakawa, T. Hara, and S. Nishio, Detecting Local Events by Analyzing Spatiotemporal Locality of Tweets. In *Advanced Information Networking and Applications Workshops (WAINA), 2013 27th International Conference on*, IEEE, 2013, pp. 191–196.

[18] Pereira, F. C., F. Rodrigues, and M. Ben-Akiva, Text analysis in incident duration prediction. *Transportation Research Part C: Emerging Technologies*, Vol. 37, 2013, pp. 177–192.

[19] Ben-Akiva, M., H. Koutsopoulos, T. Toledo, Q. Yang, C. Choudhury, C. Antoniou, and R. Balakrishna, Traffic Simulation with MITSIMLab. In *Fundamentals of Traffic Simulation* (J. BarcelÃş, ed.), Springer New York, Vol. 145 of *International Series in Operations Research & Management Science*, 2010, pp. 233–268.

[20] Lu, L., X. Yan, C. Antoniou, and M. Ben-Akiva, W-SPSA: An Enhanced SPSA Algorithm for the Calibration of Dynamic Traffic Assignment Models. *Transport Research Part C*, Accepted - Forthcoming.

[21] Blei, D. M., A. Y. Ng, and M. I. Jordan, Latent dirichlet allocation. *J. Mach. Learn. Res.*, Vol. 3, 2003, pp. 993–1022.