Model-adaptive Event-triggering for Efficient Public Transportation Tracking

Panayiotis Kolios and Georgios Ellinas

Abstract—Accurate arrival-time predictions in public transportation systems can improve the perceived quality-of-service offered and increase usage of these systems. To date, predictions were primarily based on periodic updates of mobility information that however exhibit a tradeoff between deterministic performance and system efficiency. To further improve on this tradeoff, event-triggering is emerging as a promising operations paradigm. This work describes an innovative design of a public transportation tracking system within which arrival-time predictions are made utilizing an event-triggering framework. As a first step towards this direction, behavior models are derived through extensive analysis of real mobility data. Thereafter, an event-triggering algorithm is developed to detect changes in the model in an online fashion. The efficiency and applicability of the proposed data-driven event-triggering paradigm is demonstrated through a real-world transit scenario that compares event-triggering updating techniques.

I. INTRODUCTION

It is well understood that the ever increasing mobility needs of people and goods causes traffic congestion and accidents, that translate into substantial social and economic impacts. Further, service unpredictability in public transportation systems curbs the utilization of public transport, as shown in [1]-[5]. Most often, unpredictability is caused by the lack of information, unexpected delays, unanticipated events, and traffic jams along the traveled routes. For these reasons, a significant effort has concentrated on developing architectures and techniques to improve the public transportation service utilizing more accurate information systems.

To address this problem, electronics that provide access to various useful sensors (including positioning, speed, etc.) are employed. Commonly, GPS sensors are used to collect data for both location and timing, while cellular connectivity is predominantly used to transfer periodic updates to monitoring stations. Moreover, tracking systems are implemented using proprietary hardware or open-source platforms like Arduino [6] and Android [7],[8]. Collected data is then used to monitor, manage, and optimize various operations [9]-[12] and provide real-time information and arrival time predictions to the users [13]-[16].

Thus, emerging intelligent transportation systems (ITS) capitalize on the massive growth of affordable sensors, the rapid increase of the Internet infrastructure, and the ample computing power of embedded devices, to improve situational awareness, and ultimately offer decision support for the users

P. Kolios and G. Ellinas are members of the KIOS Research and Innovation Center of Excellence and the Department of Electrical and Computer Engineering, University of Cyprus, {pkolios, gellinas}@ucy.ac.cy

of the transportation network. As previously mentioned, usually, these systems employ sensors that sample and record data from the environment in a periodic manner to inform a centralized management/monitoring host of the system's evolution. However, periodic triggering may result in excess information communication and processing for those cases where recurrent patterns exist in the system. In these cases, clearly, an event triggering approach will be more effective and more efficient, as it minimizes processing and communication by taking action only when specific events occur. By analyzing collected data, behavior models can emerge to provide real-time information to commuters, through accurate mobility estimates. To date, regression analysis, machine learning, and Kalman filtering have been extensively used to analyze mobility data [17]-[22]. These estimates are used as a baseline to detect and trigger events when unanticipated patterns are detected. Doing so, minimizes the communication to only when deemed necessary. At the same time, processing at the local hosts is limited to comparing the actual behavior to the extracted mobility model.

This work examines the benefits of shifting from periodic to event-based triggering in monitoring the mobility of public transportation vehicles in road networks. Hereafter, events are defined as changes to the predictable system behavior extracted from analyzing recurrent mobility patterns. A novel event triggering approach is proposed where data obtained from a real-life public transportation bus fleet is first used to define models of recurrent mobility patterns. These behavior models are subsequently used to predict arrival-time estimates across predefined routes and events are triggered whenever vehicles deviate from one of the models. Both local and remote hosts adapt to mobility changes by switching to the most accurate mobility model that best represents the underlying mobility settings. The accuracy of the arrival-time estimates and the effectiveness and efficiency of the proposed model-adaptive event triggering approach is examined through extensive simulations with real mobility data for a fleet of buses operating in Nicosia, Cyprus.

The rest of the paper is structured as follows. The proposed system architecture is described in Section II, while the proposed data-driven event triggering paradigm and mobility modeling are outlined in Section III. Section IV describes the simulation set-up and the obtained performance results. Finally, Section V offers some concluding remarks.

II. SYSTEM ARCHITECTURE

The proposed system architecture consists of a local host onboard the vehicle and a remote host that synchronizes a
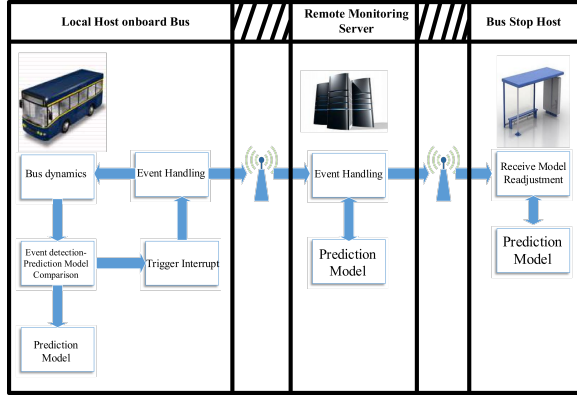
Fig. 1. Proposed system architecture.

number of terminals upon receiving mobility updates. The local host can be, for example, an embedded device onboard a bus while the remote host can be a web server (or a cloud-based server) responsible for managing the bus fleet and informing end-user terminals of arrival times, as depicted in Fig.1. The local host tracks the actual vehicle mobility and triggers events whenever a deviation is detected from the nominal mobility model used. The remote host utilizes the mobility model to estimate the vehicle mobility whenever there is no communication. Upon the arrival of a triggered event the remote host switches to the most appropriate model and re-synchronizes the monitoring terminals to reflect the most updated arrival time estimates.

More specifically, during normal operation, the local host records mobility traces, via GPS readings. The received data is compared to the expected behavior and if the actual mobility closely follows the nominal model, then no further processing or communication is required. If on the other hand, there is any delay during the traversal of the current segment of the route then an event is triggered. Upon a triggered event, the local hosts selects the most representative mobility model from a predefined list to be used and the necessary information is sent to the remote host for re-synchronization. The predicted arrival times are also adjusted accordingly to adhere to the new state of the system.

III. MOBILITY MODELING

In this section a data-driven method is proposed to extract mobility models of recurrent behavior patterns. For the use case considered, a bus route consists of a set of bus stops $\mathcal{B} = \{1, \ldots, B\}$ along a predetermined path. A trip is a sequence of traces collected along consecutive bus stops $b \in \mathcal{B}$ that can be logged with GPS samples. Since GPS traces have inherent accuracy uncertainties (as discussed in [23]), we assume that buses reach particular bus stops when found within a certain range from those bus stops. To achieve this, a circular region centered around each bus stop is defined with radii $R = \{r_1, \ldots, r_B\}$. These circular regions are used to compensate for the measurement uncertainties and are utilized to ascertain whether a measurement falls within the area of a particular bus stop. The radius of each circular region is chosen to ensure that the entire bus route is covered by consecutive

regions. This is achieved by setting $r_b = max\{\frac{d_{b,b-1}}{2}, \frac{d_{b,b+1}}{2}\}$, i.e., the entire geographical space between adjacent bus stops is covered by these circular regions.

Sequences of samples that fall within consecutive bus stop regions signify repeated trips from which statistical mobility models can be extracted. Specifically, to calculate the travel time between two arbitrary bus stops, the time duration between the GPS traces, recorded for each trip, that are closest to the designated bus stop locations are used. Given these travel times, a mobility model of a particular route can then be obtained as follows. Let t_{ln} be a random variable signifying the travel time observed for the nth trip between a pair of bus stops $\{b, b + i\} \in \mathcal{B}$, with l denoting the path segment $l = (b \mapsto b + i)$, $l \in \mathcal{L}$, and $\mathcal{L}$ is the set containing all path segments. Given these measurements, the first sample moment θ_{lN} after N sample measurements is given as follows:

$$\theta_{lN} = \frac{1}{N} \sum_{n=1}^{N} t_{ln}, \ \forall \, l \in \mathcal{L} \tag{1}$$

It is important here to also note that higher order moments (i.e., $\theta_{lN}^k = \frac{1}{N} \sum_{n=1}^{N} [t_{ln}]^k$ for the kth moment) improve the knowledge with regards to the actual probability distribution of t_{ln}.

In addition, sample moments can be updated recursively with each new sample. The updating for an arbitrary moment can be expressed as follows:

$$\begin{aligned} \theta_{l(N+1)}^k &= \frac{1}{N+1} \sum_{n=1}^{N+1} \left[t_{l(n+1)} \right]^k \\ &= \frac{1}{N+1} \left(\sum_{n=1}^{N} [t_{ln}]^k + \left[t_{l(N+1)} \right]^k \right) \\ &= \frac{N\theta_{lN}^k + \left[t_{l(N+1)} \right]^k}{N+1} \end{aligned} \tag{2}$$

where, as shown in eq. (2), the updating simply requires the total number of samples already fed into the calculation, N, and the latest moment value, θ_{lN}. Each moment can be updated dynamically with only these two elements. In this way, the behavior (i.e., the mobility model) can be characterized by the first $K = 1, \ldots, k$ moments for each leg of the path instead of the raw sample data, greatly reducing the processing and storage/memory load of the computing units. Of course, better accuracy can be achieved by introducing higher-order moments. In addition, the central moments can be computed from the raw moments using the binomial transform $\vartheta_{lN}^k = \sum_{i=0}^{k} \binom{k}{i} (-1)^{k-i} \theta_{lN}^k (\theta_{lN}^1)^{k-i}$.

Evidently, the distribution of travel times between bus stops is improved with an increasing number of observations made over time. This distribution can be used to set bounds on the travel times anticipated on the route and in turn trigger interrupts when deviations are observed. As expected, loose bounds can reduce the frequency of triggered deviation events, resulting however to lower tracking accuracy, whereas tight bounds will result in an unnecessarily high number of triggered events.

For simplicity, in this work the mean travel time τ_{ij}, $\{i, j\} \in$
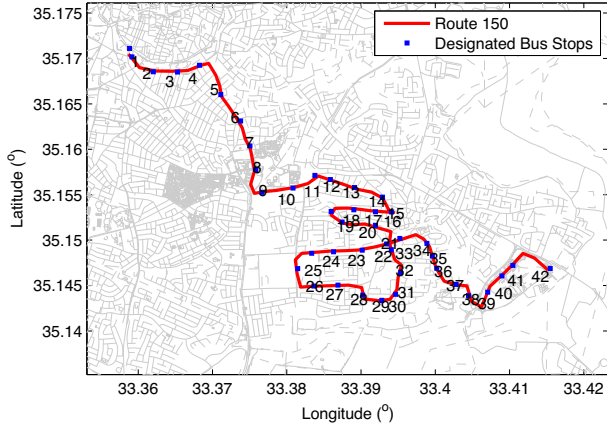
Fig. 2. Service route 150 with designated bus stops in the forward service direction (i.e., Direction 1).

$\mathcal{B}$ in segment $(i \mapsto j)$ is used to model the mobility along a route, and an arbitrary threshold α is used as a travel time bound $\tau_{ij} \pm \alpha$, to trigger deviation events. Clearly, additional models can be derived from the travel time distribution with additional corresponding thresholds. For instance, the 5th and 95th percentiles can be used to represent the left and right tail of the travel time distribution and can be used to model the behavior of the vehicle under light and heavy road utilization conditions. In the latter case, a mean and bound $\tau_{ij}^{p_5} \pm \alpha_{p_5}$ and $\tau_{ij}^{p_{95}} \pm \alpha_{p_{95}}$ for both the 5th and 95th percentiles can be used in addition to the true mean value of the distribution as additional switching models.

IV. EXPERIMENTAL SETUP AND ANALYSIS

To test the proposed framework, real traces collected from buses operating by the Transportation Organization of Nicosia District (OSEL) are considered. Approximately $50,000$ traces are collected from 3 buses serving route 150. Route 150 consists of 85 bus stops, where bus stops $1-42$ are serving the forward direction (Direction 1) and bus stops $43-85$ are serving the opposite direction (Direction 2). Naturally, separate models are used for each direction. The bus stops (Direction 1) of Route 150 are illustrated in Fig. 2. The coverage radii for each bus stop are then calculated for Direction 1 and 2 separately.

A. Fault-tolerance to Measurement Errors

As noted above, localization errors are inherent in GPS measurements, with errors ranging from a few meters to a few tens of meters. Such errors can falsely indicate the arrival time at particular bus stops. This is due to the fact that, at crossroads, the coverage of adjacent bus stops can overlap (e.g., bus stop 20 in Fig. 2) causing the triggering of erroneous events as seen in Figure 3(a). To address the aforementioned challenge, a hysteresis function is introduced so that a sudden switch in the indicated bus-stop index (denoting the vehicle position) does not result in an immediate triggering of an (erroneous) event. This hysteresis function employs a window of measurements to compute the likelihood of change in position. Figures 3(b), 3(c), and 3(d) illustrate the bus stop index assigned to the received measurements for different measurement windows when a maximum likelihood estimator is used.

As shown in these plots, it is evident that the best threshold for the hysteresis window is of length 3. An additional test was also performed to ascertain whether the hysteresis window of length 3 significantly impacts true-positive event triggers. Figure 4 depicts the number of events per trip, per day of the week, triggered when no hysteresis is employed and when a hysteresis window of length 3, 4 and 5 is used. With a window of length 4 and 5, the results are similar, while with a window of length 3, the aggregated number of changes matches those of the actual schedule and thus all erroneous events are suppressed.

B. Performance Evaluation

For the results presented, four separate behavior models are used for the arrival time prediction, namely the mean travel time, and the 80th, 95th, and 99th percentiles of traveling times. In the performance evaluation that follows, results were produced for the case where only a single model is considered (the mean travel time model) and for the case where all four models are considered for mobility tracking.

In the performance evaluation, 70% of the traces were used for building the models and the rest were used for testing. For all the results depicted hereafter, a hysteresis window of length 3 is employed as detailed in Section IV-A, so as to suppress erroneous triggering events. Overall, the depicted results include 322 round-trip traces.

Two scenarios are simulated in this work. In the first scenario, the mean travel model is utilized for predicting the travel times, and delay events are triggered when the actual travel time to the next bus stop deviates by 120 seconds from the predicted one. Additionally, any deviation events are only used to update the travel times of both the remote host and the terminals without any model switching taking place.

In contrast, the second scenario involves interchanging among the four behavior models. In this case, travel time deviations between predicted and actual arrival times result in switching to the model that best matches the travel time reported in the received interrupt. The state flow diagram corresponding to the model switching process is shown in Figure 5. When a state transition occurs, a deviation event is triggered and the remote host is informed about the transition. The remote host, in turn, uses the same state flow diagram to switch its current model and to send the appropriate travel time estimates to the terminals.

Figure 6 displays the histograms of events at each bus stop per trip for the two scenarios considered. As indicated in the figures, the average number of deviation events is approximately 0.35 per trip, while 0.06 erroneous switching events per trip are observed. The second scenario yields 0.29 deviation events per trip and 0.05 erroneous switching events, resulting in a 17.62% decrease in deviation events and in a
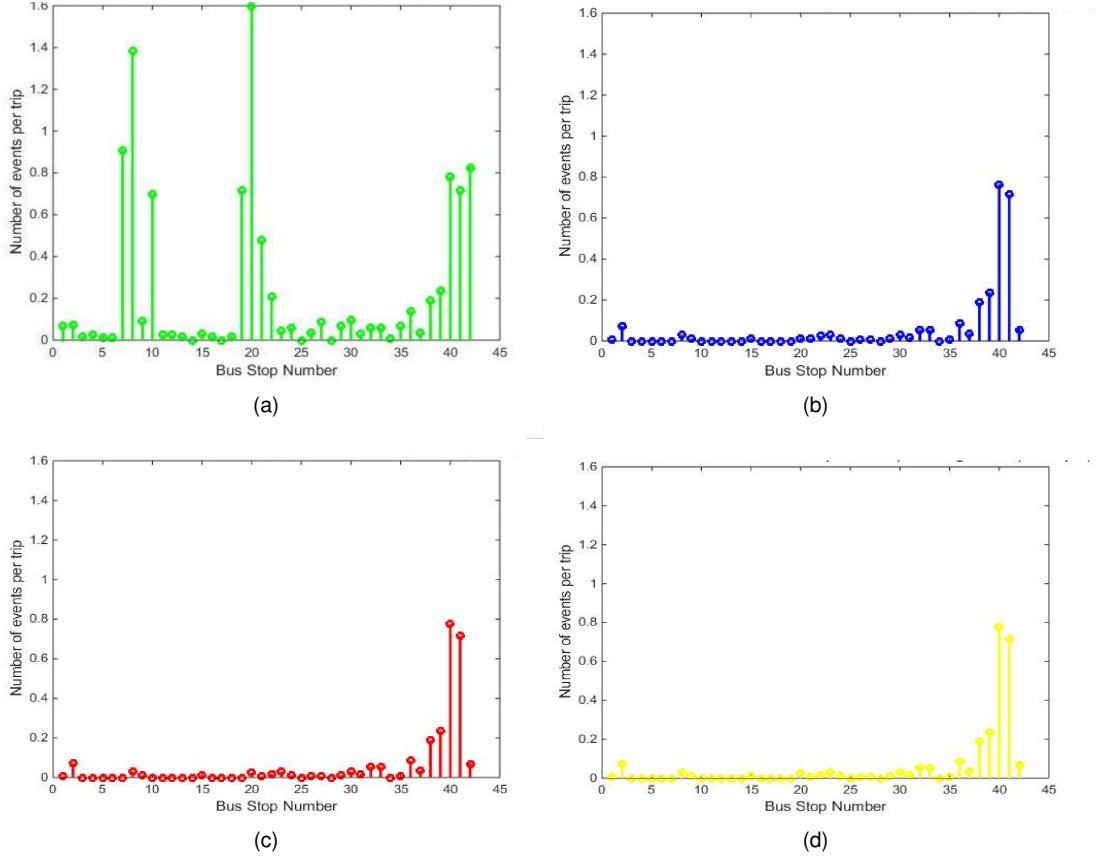
Fig. 3. (a) Number of events near each bus stop during normal assignment. (b) Number of events near each bus stop during assignment utilizing a 3-sample hysteresis. (c) Number of events near each bus stop during assignment utilizing a 4-sample hysteresis. (d) Number of events near each bus stop during assignment utilizing a 5-sample hysteresis.
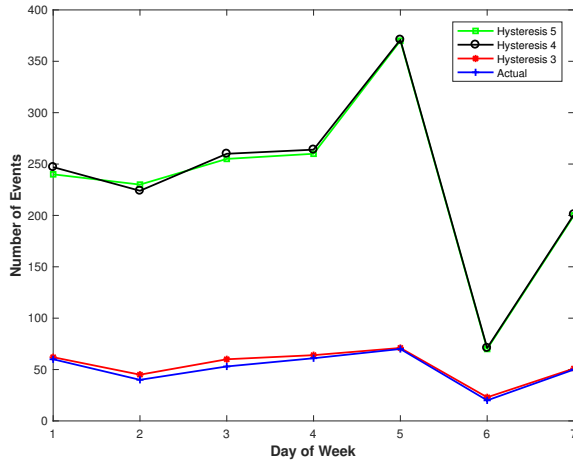

Fig. 4. Direction changes per day of the week.

Fig. 5. Behavior model transitions during normal mode operation.

V. CONCLUSIONS

In this work, a data-driven event triggering technique has been developed for accurate mobility tracking that minimizes communication overhead between remote hosts and terminal nodes, as well as the information processing and storage/memory requirements at the local hosts, compared to the case of periodic triggering. The proposed technique effectively addresses the trade-off between tracking efficiency and communication/processing overhead, thus minimizing the system's

6.29% decrease in the number of erroneous switching events. These improvements are due to the model-adaptive capabilities considered in the second scenario, demonstrating the validity and effectiveness of the proposed framework.
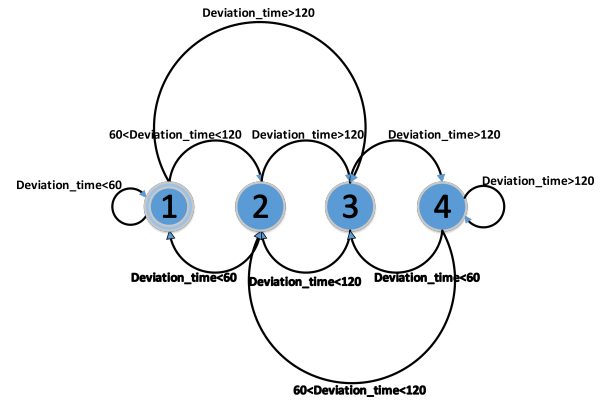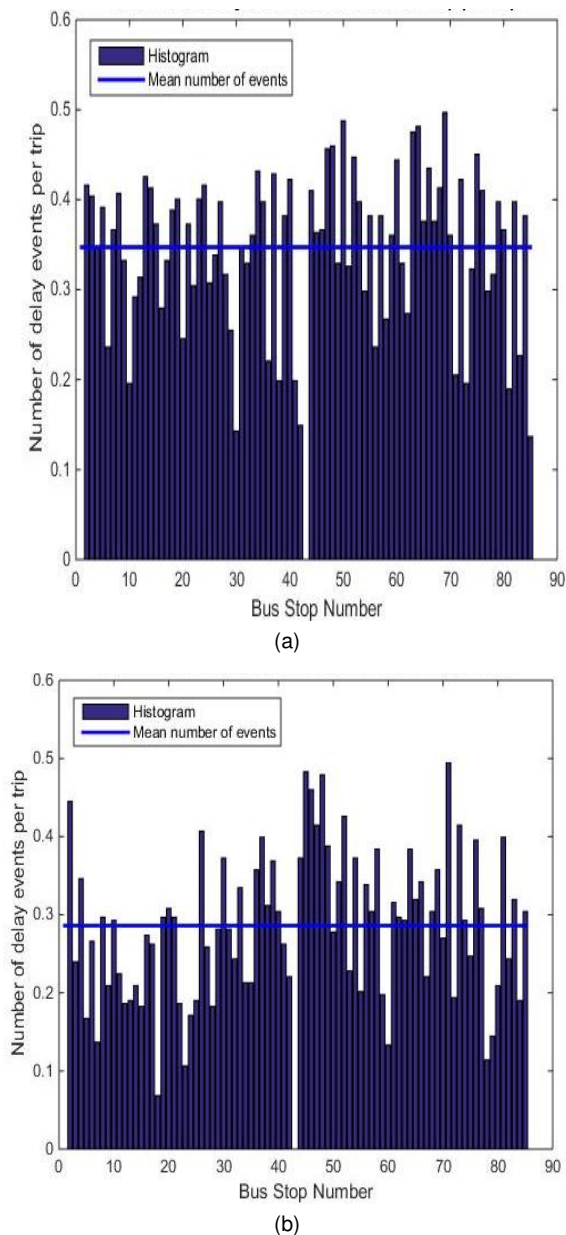
Fig. 6. Performance evaluation for a) Scenario 1, and b) Scenario 2.

operating cost, while providing accurate vehicle tracking. The proposed technique has been evaluated using real GPS traces collected from a fleet of buses serving a particular bus route, demonstrating an improvement in the system's performance when utilizing the multi-model approach compared to the use of single-model event-triggering techniques.

REFERENCES

[1] J. Brown, et al., "Fare-free Public Transit at Universities: An Evaluation", *Journal of Planning Education and Research*, 23(1):69-82, 2003.

[2] L. Tang and P. Thakuriah, "An Analysis of Behavioral Responses to Real-Time Transit Information Systems", *Proc. Association of Collegiate Schools of Planning Annual Meeting*, Fort Worth TX, 2006.

[3] S. F. G. Taylor, "Waiting for Service: Perceptions Management of the Wait Experience", *Services Marketing and Management*, Sage Publications, 2000, pp. 171-190.

[4] W. Toor and S. Havlick, *Transportation and Sustainable Campus Communities: Issues, Examples, Solutions*, Island Press, 2004.

[5] F. Zhang, et al., "An Examination of Traveler Responses to Real-time Bus Arrival Information Using Panel Data", *Proc. 87th Annual Meeting of the Transportation Research Board*, 2007.

[6] T. M. Bojan , et al., "An Internet of Things based Intelligent Transportation system", *Proc. Vehicular Electronics and Safety (ICVES)*, 2014.

[7] A. Thiagarajan, et al., "Cooperative Transit Tracking using Smartphones", *Proc. 8th ACM Conference on Embedded Networked Sensor Systems*, 2010.

[8] P. Zhou, et al., "How Long to Wait?: Predicting Bus Arrival Time with Mobile Phone based Participatory Sensing", *IEEE Transactions on Mobile Computing*, 13(6):1228-1241, 2013.

[9] L. Zhang, et al., "Multi-stage Approach to Travel-mode Segmentation and Classification of GPS Traces", *Proc. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2012.

[10] L. Xiao and Z. x. Wang, "Internet of Things: a New Application for Intelligent Traffic Monitoring System", *Journal of Networks*, 6(6):887–894, 2011.

[11] H. O. Al-Sakran, "Intelligent Traffic Information System Based on Integration of Internet of Things and Agent Technology", *International Journal of Advanced Computer Science and Applications*, 6(2):37-43, 2015.

[12] P. Papadimitratos, et al., "Vehicular Communication Systems: Enabling Technologies, Applications, and Future Outlook on Intelligent Transportation", *IEEE Communications Magazine*, 47(11):84-95, 2009.

[13] O. Cats and G. Loutos, "Real-Time Bus Arrival Information System: An Empirical Evaluation", *Journal of Intelligent , Transportation Systems: Technology, Planning, and Operations*, 20(2):138–151, 2015.

[14] Y. Wang and H. Qi, "Research of Intelligent Transportation System Based on the Internet of Things Frame", *Wireless Engineering and Technology*, 3:160-166, 2012.

[15] F.-Y. Wang, "Parallel Control and Management for Intelligent Transportation Systems: Concepts, Architectures, and Applications", *IEEE Transactions on Intelligent Transportation Systems*, 11(3):630-638, 2010.

[16] S. S. Dukare, et al., "Vehicle Tracking, Monitoring and Alerting System: A Review", *International Journal of Computer Applications*, 119(10):39-44, 2015.

[17] B. M. Williams and L. A. Hoel, "Modeling and Forecasting Vehicular Traffic Flow as a Seasonal ARIMA Process: Theoretical Basis and Empirical Results", *Journal of Transportation Engineering*, 129(6):664-672, 2003.

[18] W. H. Lin and J. Zeng, "Experimental Study of Real-Time Bus Arrival Time Prediction with GPS Data", *Transportation Research Record*, 1666(1):101-109, 1999.

[19] S. Chien, et al., "Dynamic Bus Arrival Time Prediction with Artificial Neural Networks", *Journal of Transportation Engineering*, 128(5):429-438, 2002.

[20] J. Patnaik, et al., "Estimation of Bus Arrival Times. Using APC Data", *Journal of Public Transportation*, 7(1):1-20, 2004.

[21] D. Sun, et al., "Predicting Bus Arrival Time on the Basis of Global Positioning System Data", *Transportation Research Record: Journal of the Transportation Research Board*, 2034(1):62-72, 2007.

[22] M. Chen, et al., "A Dynamic Bus-Arrival Time Prediction Model Based on APC Data", *Computer-Aided Civil and Infrastructure Engineering*, 19(5):364-376, 2004.

[23] A. Steed, "Supporting Mobile Applications with Real-Time Visualisation of GPS Availability", *Proc. Mobile Human-Computer Interaction (Mobile HCI)*, pp. 373–377, 2004.