ESTIMATION OF HIGHWAY TRAFFIC FROM SPARSE SENSORS: STOCHASTIC MODELING AND PARTICLE FILTERING

A. Pascale[1], G. Gomes[2], M. Nicoli[1]

[1]Dip. Elettronica e Informazione, Politecnico di Milano, Italy, {pascale,nicoli}@elet.polimi.it
[2]PATH, UC Berkeley, CA, USA, gomes@path.berkeley.edu

ABSTRACT

Traffic control is essential for the achievement of a sustainable and safe mobility. Monitoring systems deployed over the roads collect a great amount of traffic data that must be efficiently processed by statistical methods to draw traffic macro-parameters that are needed for control operations. In this paper we propose a particle filtering approach to estimate the density over a road network starting from noisy and sparse measurements provided by road-embedded sensors. We propose a new Bayesian framework based on the link-node cell transmission model to take into account the stochastic behavior of traffic and the hysteresis phenomenon that are typically observed in real data. Numerical tests show that the estimation method is able to reliably reconstruct the traffic field even in case of very sparse sensor deployments.

Index Terms— ITS, Statistical Modeling, Particle Filtering, Traffic Densities Reconstruction.

1. INTRODUCTION

Intelligent transport systems (ITS) are expected to improve quality, safety and sustainability of mobility by integrating information and communication technologies with transport engineering. ITS rely on a capillary network of sensors, either road-embedded or mobile probes, that are on roads providing measurements of traffic macro-parameters, such as speed, flow, and density [1]. Modern traffic management systems rely on estimates and predictions of the overall state of traffic based on sparse and often noisy measurements.

The objective of this paper is to develop a statistical algorithm able to accurately estimate the evolution of traffic variables. Several approaches have been presented in the literature to reach this goal, from historical trends [2] to non-parametric methods, auto-regressive and moving average models [3, 2], Kalman filtering, and neural networks [4, 5, 6]. Recent studies have proposed Bayesian filtering to track the evolution of traffic on a real-time basis. In particular, non linear Bayesian models [7, 8], Bayesian networks (BN), graph theory [9, 10] and particle filtering (PF) [11, 8] have been proved to be promising solutions for traffic estimation.

The **original contribution** of this paper is a Bayesian framework for the estimation of the densities in highways using PF [12]. Our effort is devoted to the proposal of a new Bayesian model that extends previous approaches so as to account for the stochastic behavior of traffic, and also for phenomena of hysteresis and capacity drop that are typically observed in flow-density scatter plots [13, 14]. Although microscopic models have long incorporated random driver behavior [14, 15, 16], stochastic extensions to the cell transmission model (CTM) are relatively recent, in particular [17] contains an overview of the most relevant contributions. In this paper we add randomness to the link-node CTM (LN-CTM [18]) an application of the Godunov scheme to network topologies. Based on the modified LN-CTM, we develop a Bayesian method for the estimation of the space-time traffic evolution from sparse sensor observations. Since the traffic model is non-linear and non-Gaussian, we propose a PF approach. A numerical analysis is carried out on a realistic highway scenario where only a subset of road links is monitored by loop sensors while the rest are not observed (e.g., due to sparse sensor deployments or loop failures). The results show that the estimation method is able to provide a good reconstruction of the traffic field over all road links, as for a virtual sensor deployment with higher spatial density (increased by a factor ~ 7 in the specific scenario). The estimate accuracy outperforms the loops' accuracy up to 40% on monitored links.

2. PROBLEM FORMULATION AND MODELING

We model the road as a set of N_L links (road segments) interconnected by nodes (road junctions), as depicted in Fig. 1[1]. A subset $S \subseteq \{1, 2, \ldots, N_L\}$ of $N_S \leq N_L$ links is monitored by traffic sensors, e.g., loops installed on the links. The length of the n-th link is denoted as l_n, time evolution is sampled with time interval Δt. The variable that describes the traffic state in link n at time k is the density $\rho_n(k)$, defined as the number of vehicles per space unit [veh/mile]. The $N_L \times 1$ state vector for the whole road is defined as $\mathbf{x}_k = [\rho_1(k) \ldots \rho_{N_L}(k)]^T$ and the $N_S \times 1$ measurement vector as $\mathbf{y}_k = [\breve{\rho}_{s_1}(k) \ldots \breve{\rho}_{s_{N_S}}(k)]^T$ where $\breve{\rho}_s(k)$ is the density measured by sensor $s \in S$ at time k. Traffic is modelled as a hidden Markov model ruled by the following equations:

$$\begin{aligned} \mathbf{x}_k &= \mathbf{g}(\mathbf{x}_{k-1}, \mathbf{w}_k) \\ \mathbf{y}_k &= \mathbf{h}(\mathbf{x}_k, \mathbf{r}_k) \end{aligned} \tag{1}$$

where $\mathbf{g}(\cdot, \cdot)$ is the function - defined in next section - describing how densities evolve over space and time. The function

[1]Networks were built with the TOPL Network Editor [19]. Traffic data was obtained by PeMS [20].

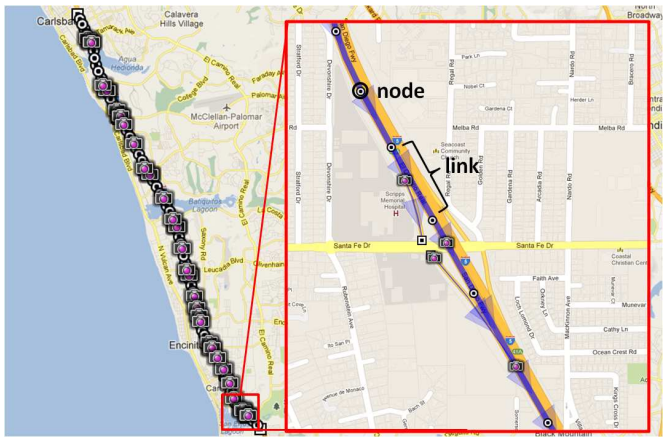

Fig. 1. Section of I-5 southbound in southern California. Camera icons indicate loop detector stations.

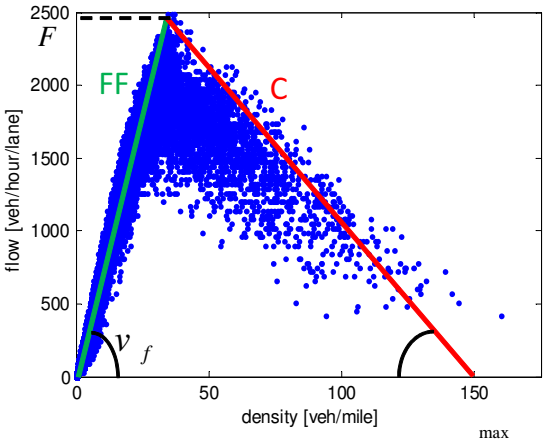

Fig. 2. 30-sec density/flow data gathered over 15 days at a single location. The red lines show the deterministic fundamental diagram computed using the technique of [18].

$\mathbf{h}(\cdot,\cdot)$ models the relation of measurements to the densities, the s-th function being defined as $h_s(\rho_s(k), r_s(k)) = \max(\rho_s(k) + r_s(k), 0)$ where $r_s(k) \sim \mathcal{N}(0, \sigma_r^2)$ is a Gaussian measurement error limited to ensure $\breve{\rho}_s(k) > 0$ on each link s. Our objective is the estimation of $\mathbf{x}_k$ given $\mathbf{y}_{1:k}$ using a Bayesian approach.

2.1. Macroscopic traffic modeling with LN-CTM

The function $\mathbf{g}(\cdot)$ is defined according to LN-CTM [18], a model derived from CTM for general road topologies. It models the evolution of traffic as the combination of two main propagation modes: an upstream wave propagating forward with velocity v_f for free flow FF (straight line with slope v_f in Fig. 2), and a downstream wave propagating backward with velocity ω for congestion C (straight line with slope ω in Fig. 2). In the fundamental diagram in Fig. 2, the maximum capacity on the road F is defined as the maximum number of vehicles that can pass from one link to the next one. The maximum density ρ_{max} represents the maximum number of vehicles that can occupy a link.

In LN-CTM the density on link n evolves according to a conservation principle:

$$\rho_n(k) = \rho_n(k-1) + \frac{\Delta t}{l_n}\left(f_n^{in}(k) - f_n^{out}(k)\right), \qquad (2)$$

where $\rho_n(k)$ is the density on the n-th link in the k-th time interval, $f_n^{in}(k)$ and $f_n^{out}(k)$ are respectively the flows that enter and exit the link during the k-th interval. Each node may in general have M input links and N output links. The flows through the node from each input m to each output n are computed as follows. First, the demand from each input link is computed using the sending function of the CTM (i.e., using the left side of the fundamental diagram in Fig. 2):

$$d_m(k) = \min\left[v_{f,m}\rho_m(k-1), F_m^d\right], \qquad (3)$$

where $v_{f,m}$ and F_m^d refer to link m. The supply of space in each exiting link is then computed with the receiving function of the CTM (i.e., the right side of the fundamental diagram):

$$c_n(k) = \min\left[\omega_n(\rho_{\max} - \rho_n(k-1)), F_n\right], \qquad (4)$$

where the parameters ω_n and F_n are related to link n. Note that LN-CTM assumes $F_n^d = F_n$. Here we set $F_n^d > F_n$ in order to account for capacity drop. The third step is to transfer the demands from input links to output links. This is done using a given split ratio matrix with element $B_{m,n}$ which specifies the portion of the flow from input link m that goes to output link n. The demand on output link n is $d_n^{out}(k) = \sum_{m=1}^{M} B_{m,n} d_m(k)$. If the output demand exceeds output supply, then the flows must be scaled by a factor:

$$\gamma_n(k) = \min\left[d_n^{out}(k), c_n(k)\right] / d_n^{out}(k). \qquad (5)$$

$\gamma_n(k)$ determines the state of congestion of link n. If $\gamma_n(k) < 1$ then link n is said to be "congested", otherwise it is in "free flow". Input links are congested if any of the output links that they feed are congested. Finally the flows entering the node are computed by taking the most restrictive scaling factor among the exits fed by each input link m:

$$f_m^{out}(k) = \min_{n|B_{m,n}>0}\left[\gamma_n(k)\right] \cdot d_m(k). \qquad (6)$$

The flows entering the downstream links are found using the split ratio matrix: $f_n^{in}(k) = \sum_{m=1}^{M} B_{m,n} f_m^{out}(k)$.

2.2. Extension of LN-CTM for stochastic traffic modeling

Observations of real measurements have shown traffic to have a significant stochastic component, linked principally to the granular characteristic of flows and to driver behavior. Figure 2 shows a typical plot of 30-seconds aggregates. It can be seen that data in the congested regime tend to spread widely from the fundamental diagram.

Previous authors have noted that continuous and concave fundamental diagrams fail to capture some important features of traffic behavior, namely the "capacity drop" and "traffic

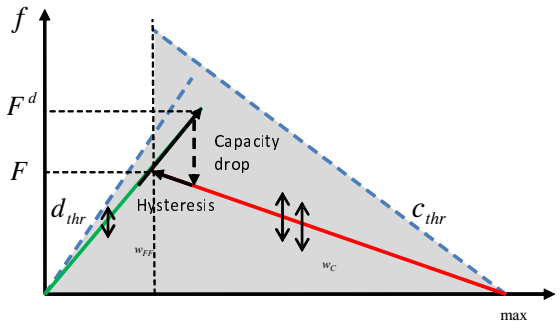

Fig. 3. Fundamental diagram with limits in (9) and hysteresis.

hysteresis". Capacity drop occurs when a section that is operating at capacity becomes an active bottleneck. This transition is usually accompanied by a sharp decrease in the flow exiting the section. Hysteresis occurs when the recovery from congestion follows a different path from the breakdown (see Fig 3). Here we propose to capture both capacity drop and hysteresis by setting $F^d > F$ and replacing the "min" function in (5) with a probabilistic variant defined below:

$$\min_p [a,b] = \begin{cases} a, & \text{if } a < b \\ b, & \text{with probability } p \text{ if } b < a \\ a, & \text{with probability } 1-p \text{ if } b < a \end{cases} \quad (7)$$

For $p = 1$, this function is equivalent to the usual min function. For $p = 0$, it equals a always. To model capacity drop and hysteresis we replace the min in equation (5) with $\min_p$:

$$\gamma_n(k) = \min_p \left[d_n^{out}(k), c_n(k) \right] / d_n^{out}(k), \quad (8)$$

The value of p is selected depending on the current state of congestion of the link: $p = p_h \in [0,1]$ if FF and $p = 1$ if C. If the link is in free flow and $p_h = 0$, the link will reach the higher capacity F^d before transitioning into congestion. Once in congestion, because $p = 1$ and $F < F^d$, the link will remain on the lower congested branch, as shown in Fig. 3.

Furthermore we propose a statistical representation of inflows/outflows that enter/exit a node. In contrast to a previous approach in [11], the stochastic behavior is added to both demands and supplies. We do not add noise to the fundamental diagram parameters directly, as done in [17]. We start our description from the statistical characterization of supplies and demands in (3)-(4). Those equations are modified as:

$$\begin{aligned} \widetilde{d}_m(k) &= \min\left(\max\left(d_m(k) + w_F(k), 0\right), d_{\text{thr},m}\right), \\ \widetilde{c}_n(k) &= \min\left(\max\left(c_n(k) + w_C(k), 0\right), c_{\text{thr},n}\right), \end{aligned} \quad (9)$$

where $d_m(k)$ and $c_n(k)$ are taken from (3)-(4), $d_{\text{thr},m} = \rho_m(k-1)\frac{l_m}{\Delta t}$ and $c_{\text{thr},n} = (\rho_{max} - \rho_n(k-1))\frac{l_n}{\Delta t}$. The two independent random terms $w_F \propto \mathcal{N}\left(0, \sigma^2_{w_F}\right)$ and $w_C \propto \mathcal{N}\left(0, \sigma^2_{w_C}\right)$ describe the modeling uncertainty of flow entering or exiting the node. The terms $\widetilde{d}_m(k)$ and $\widetilde{c}_n(k)$ obtained by (9) are then substituted to $d_m(k)$ and $c_n(k)$ after the computation of the min in (8). The effects of the modeling noises on the fundamental diagram are presented in Fig. 3 with arrows. It is important to point out that adding statistical behavior brings to unpredictable values of flows that should be maintained under reasonable limits. Particular care should thus be paid to avoid that density on any link becomes negative or larger than the maximum. In particular the feasibility constraint of the CTM $v \leq \frac{l}{\Delta t}$ should be always maintained (i.e., a stream of vehicle cannot pass two links in one time-step). These conditions are represented in (9) by $d_{\text{thr},m}$ and $c_{\text{thr},n}$ as blue lines in Fig. 3. Grey areas represent acceptable values of flows. They model the relation between flow and density when the maximum value of speed $v_f^{\max} = \frac{l_m}{\Delta t}$ and $\omega^{\max} = \frac{l_n}{\Delta t}$ are reached. Once these conditions are satisfied, the density on each link is checked so as to assure that it is restricted to the range $[0, \rho_{\max}]$, otherwise input and output flows are coherently adjusted. This process ensures that the results of the simulation are feasible.

3. PARTICLE FILTERING

We propose a recursive Bayesian approach for the estimation of the traffic densities over the links, using the measurement model (1) and the traffic density model (2)-(9). At each time step we update the a-posteriori probability density function (pdf) $p(\mathbf{x}_k|\mathbf{y}_{1:k})$ according to the Bayesian rule [21]:

$$p(\mathbf{x}_k|\mathbf{y}_{1:k}) = \frac{1}{p(\mathbf{y}_k|\mathbf{y}_{1:k-1})} p(\mathbf{y}_k|\mathbf{x}_k)\, p(\mathbf{x}_k|\mathbf{y}_{1:k-1}), \quad (10)$$

using the current measurement likelihood $p(\mathbf{y}_k|\mathbf{x}_k)$ and the a-priori pdf $p(\mathbf{x}_k|\mathbf{y}_{1:k-1})$. The a-priori pdf is computed from the a-posteriori pdf of previous time step using the Chapman-Kolmogorov equation as

$$p(\mathbf{x}_k|\mathbf{y}_{1:k-1}) = \int p(\mathbf{x}_k|\mathbf{x}_{1:k-1}, \mathbf{y}_{1:k-1})\, p(\mathbf{x}_{k-1}|\mathbf{y}_{1:k-1})\, d\mathbf{x}_{k-1}, \quad (11)$$

where $p(\mathbf{x}_k|\mathbf{x}_{1:k-1}, \mathbf{y}_{1:k-1})$ is the traffic density evolution obtained from the model equations (2)-(9) while $p(\mathbf{x}_{k-1}|\mathbf{y}_{1:k-1})$ is the a-posteriori pdf at time $k-1$.

We point out that the pdf's in the update and prediction steps (10)-(11) cannot be written in closed form as the traffic density model is highly non-Gaussian and non-linear. Namely, the model equations (2)-(9) are non linear, and the model noises w_F and w_C are not Gaussian because flows are limited as explained in the previous section. For these reasons, we use a PF approach for implementation of the Bayesian procedure. PF allows for the description of the traffic density distributions by a set of random samples (particles) that are updated at each time-step using the importance sampling principle. The PF algorithm is carried out using the procedure in [21] with model parameters calibrated on a real highway scenario as described in the following section.

4. ESTIMATION RESULTS

We simulate the traffic behavior on the real highway scenario in Fig. 1. It is composed by 94 links, of which 71 are mainline

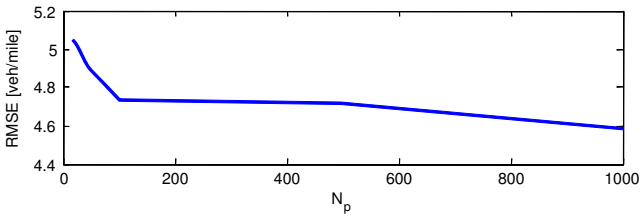

Fig. 4. Density estimate accuracy vs the number of particles.

links and the remainder on/off ramps. A subset of 12 links is monitored by sensors. The simulation step is $\Delta t = 5$s. Parameters v_f and F^d are calibrated for each link on real data following the procedure in [18]. Parameters ρ_{max}, ω and F are calibrated fitting the data with a straight line on the right part of the fundamental diagram for each link. We use real historical demands as boundary conditions obtaining a realistic scenario of traffic evolution. For testing the algorithm we select a time interval T of 350 min starting from 5:50 am. Data are collected by sensors every 30 seconds (each observed density is an average over $6 \cdot \triangle t$), while density estimation is carried out every 5 seconds.

Fig. 5 (top) shows the traffic density simulated using $\sigma_r = 10$/mile/lane, $\sigma_{w_C} = 400$veh/hour/lane, $\sigma_{w_F} = 100$veh/hour/lane, and $p_h = 0.4$. Grey areas indicate the links that are monitored by sensors. In Fig. 5 (bottom) the estimate of the traffic density is plotted for the PF algorithm with $N_p = 1000$. The picture shows that the algorithm is able to reconstruct with high accuracy the evolution of traffic.

As performance metric we consider the root mean square error of the estimate: RMSE$= \sqrt{E\left[(\hat{\rho}_n(k) - \rho_n(k))^2\right]}$. Averaging is carried out over all the N_L links - both monitored and non monitored - observed over the time window T. We numerically verified that an equivalent σ_r can be considered because of the function $\max(\cdot,\cdot)$ when applying $\mathbf{h}(\cdot,\cdot)$ in (1), that is $\sigma_r^{eq} \simeq 9$. In Fig. 5 this equivalent threshold is represented. The RMSE results are plotted on the bottom of Fig. 5, where the red line represents σ_r^{eq}. The figure shows that the estimate error on links covered by sensors is always under this value, and the same behavior is observed for non-monitored links underlying the ability of the algorithm to accurately reconstruct the density field.

Finally, in Fig. 4 the RMSE is shown versus the number of particles, N_p, used by the PF algorithm. As expected, the error decreases as the number of particles increases. An error floor is reached around $N_p \approx 1000$, which proves that this value is a good choice for estimation. For $N_p \geq 100$, the algorithm allows to reach a higher accuracy compared to the one provided by sensors (σ_r^{eq}). Also, recalling that measurements are provided every 30s and the estimates every 5s, the algorithm is able to provide a traffic monitoring with improved accuracy (i.e., lower RMSE) and increased spatial-temporal sampling of the traffic field, compared to the original sensors. The result is a virtual monitoring network with a higher spatial density of sensors and temporal sampling, respectively by a factor 7 and 6. This is very useful when the back-propagation of congestion (caused by traffic or accident) needs to be accurately estimated or predicted.

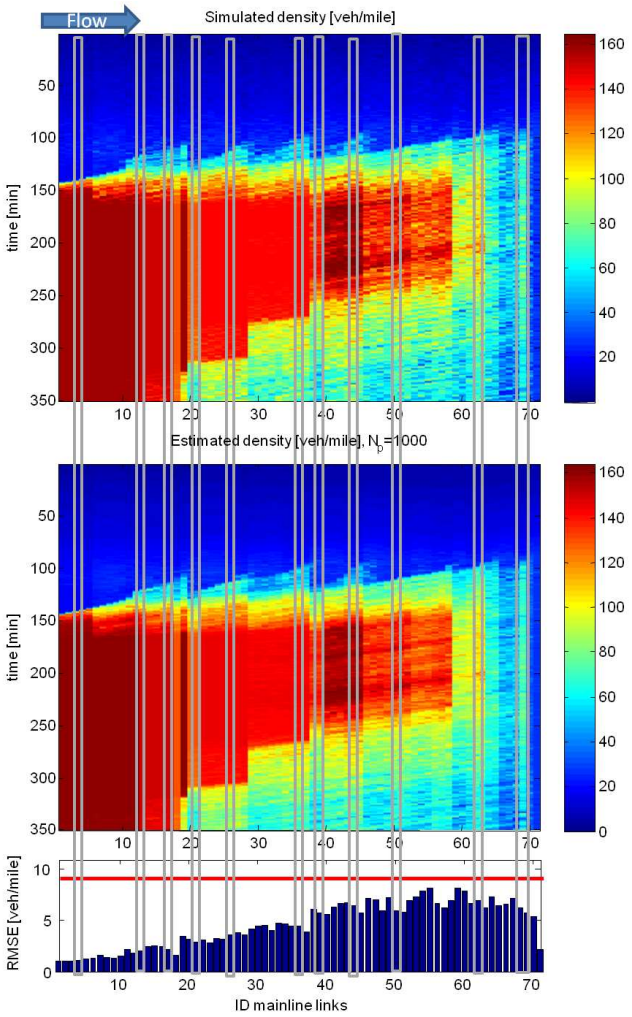

Fig. 5. Top: real density; center: estimated density for $N_p = 1000$; bottom: RMSE of the density estimate on each link.

5. CONCLUSIONS

In this paper we proposed a PF approach to traffic density estimation. In particular we improved the LN-CTM model to take into account the statistical behavior of traffic together with the hysteresis and the capacity drop phenomena usually observed in real data. Using this modified version of LN-CTM, we developed a PF estimation method for the reconstruction of traffic evolution given sparse and noisy data. We tested our approach on data simulated using a real highway scenario obtaining results with high estimate accuracy both on links with and without sensors. As on-going activities we are calibrating and validating the algorithm on real data coming from the scenario of I5S and extending to distributed processing [22]. Moreover we will compare the performance obtained with our method to existing algorithms using real data on more complex scenarios.

6. REFERENCES

[1] A. Pascale, M. Nicoli, F. Deflorio, B. Dalla Chiara, U. Spagnolini, "Wireless sensor networks for traffic management and road safety," *Intelligent Transport Systems, IET*, vol. 6, no. 1, pp. 67-77, March 2012.

[2] R. Chrobok, O. Kaumann, J. Wahle, and M. Schreckenberg, "Different methods of traffic forecast based on real data," *European Journal of Operational Research*, vol. 155, issue 3, pp. 558-568, June 2004.

[3] B.L Smith, B.M. Williams, R. Keith Oswald, "Comparison of parametric and nonparametric models for traffic flow forecasting," *Transportation Research Part C: Emerging Technologies*, vol. 10, issue 4, pp. 303-321, 2002.

[4] I. Okutani, Y.J. Stephanedes, "Dynamic prediction of traffic volume through Kalman filtering theory," *Transportation Research Part B: Methodological*, vol. 18, issue 1, pp. 1-11, February 1984.

[5] E.I. Vlahogianni, J.C. Golias, M.G. Karlaftis, "Short-term traffic forecasting: Overview of objectives and methods," *Transport Reviews*, vol. 24, issue 5, pp. 533-557, 2004.

[6] M. Dougherty, "A review of neural networks applied to transport," *Transportation Research Part C: Emerging Technologies*, vol. 3, issue 4, pp. 247-260, August 1995.

[7] X. Sun, L. Munoz, R. Horowitz, "Highway traffic state estimation using improved mixture Kalman filters for effective ramp metering control," Proceedings, *42nd IEEE Conference on Decision and Control, 2003*, vol. 6, pp. 6333- 6338, 9-12 Dec. 2003.

[8] K. Stanková, B. De Schutter, "On freeway traffic density estimation for a jump Markov linear model based on Daganzo's cell transmission model," *2010 13th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pp. 13-18, 19-22 Sept. 2010.

[9] A. Pascale, M. Nicoli, "Adaptive Bayesian network for traffic flow prediction," *IEEE Statistical Signal Processing Workshop (SSP)*, *2011*, pp. 177-180, 28-30 June 2011.

[10] S. Sun, C. Zhang, G. Yu, "A Bayesian Network Approach to Traffic Flow Forecasting," *IEEE Trans. on Intelligent Transportation Systems*, vol. 7, no. 1, pp. 124-132, March 2006.

[11] L. Mihaylova, R. Boel, A. Hegyi; "Freeway traffic estimation within particle filtering approach," *Automatica*, vol. 43, no. 2, pp. 290-300, 2007.

[12] M.S. Arulampalam, S. Maskell, N. Gordon, T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. on Signal Processing* , vol. 50, no. 2, pp.174-188, Feb 2002.

[13] S.P. Hoogendoorn, and P.H.L. Bovy, "State of art of vehicular traffic modelling," *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, vol. 215, no. 4, pp. 283-303, 2001.

[14] M. Fukui, Y. Sugiyama, M. Schreckenberg, and D.E. Wolf, "Traffic and granular flow '01'," *Springer*, 2001.

[15] D. Choudhury, L. Santen, and A. Schadschneider, "Statistical physics of vehicular traffic and some related systems," *Physics Reports (Elsevier)*, vol. 329, pp. 199-329, 2000.

[16] P.I. Bratanov, E. Bonek, "Mobility model of vehicle-borne terminals in urban cellular systems," *IEEE Trans. on Vehicular Technology*, vol.52, no.4, pp.947-952, July 2003.

[17] A. Sumalee, R.X. Zhong, T.L. Pan, W.Y. Szeto, "Stochastic cell transmission model (SCTM): A stochastic dynamic traffic model for traffic surveillance and assignment," *Transportation Research Part B: Methodological,* vol. 45, issue 3, pp. 507-533, 2011.

[18] A. Muralidharan, G. Dervisoglu, R. Horowitz, "Freeway traffic flow simulation using the Link Node Cell Transmission model," *2009 American Control Conference*, 10-12 June 2009.

[19] TOPL network editor - `http://vii.path.berkeley.edu/NetworkEditor`

[20] PeMS - Performance Measurement System. [Online] UC Berkeley, PATH and Caltrans. `http://pems.dot.ca.gov//`.

[21] N.J. Gordon, D.J. Salmond, and A.F.M. Smith, "A Novel Approach to Non-Linear and Non-Gaussian Bayesian State Estimation," *IEE Proceedings F, Radar and Signal Processing*, vol. 140, pp. 107-113, 1993.

[22] M. Nicoli, D. Fontanella, "Fundamental Performance Limits of TOA-Based Cooperative Localization," *IEEE International Conference on Communications Workshops, ICC Workshops 2009*, pp. 1-5, 14-18 June 2009.