An Adaptive Observation Site Selection Strategy for Road Traffic Data Assimilation *

Shiming Yang
Department of Computer Science and Electrical Engineering
University of Maryland, Baltimore County
Baltimore, MD
shiming1@umbc.edu

Konstantinos Kalpakis
Department of Computer Science and Electrical Engineering
University of Maryland, Baltimore County
Baltimore, MD
kalpakis@umbc.edu

Alain Biem
IBM T J Watson Research Center
New York, NY
biem@us.ibm.com

ABSTRACT

Smartphones and other vehicular sensors equipped with GPS and wireless networking capabilities, are becoming ubiquitous in transportation systems. They provide us with opportunities to gather timely information about road traffic conditions, fuse (assimilate) it with traffic flow models to improve upon the accuracy of these models, and hence supply valuable information for real-time transportation decision making. Macroscopic traffic flow models are described by systems of partial differential equations (PDEs), which are usually only solved numerically. Adaptive moving mesh methods have shown promise in handling high variability of the spatio-temporal features (e.g. shocks and discontinuities) in model's solutions.

We propose a novel low-overhead strategy to adaptively select observation sites in real time, by relying on information from the adaptive moving mesh of the numerical solver of the underlying PDEs. The idea is to place more of the limited observational resources to locations of higher variability in the numerical solution. We incorporate our strategy into a particle-filter based data assimilation framework, and compare it with the strategy of gathering and assimilating measurements from evenly spaced observation sites. We experimentally show that our strategy reduces the relative error by up to 53% in estimating vehicle density on a road during phantom jams and traffic jams due to bottlenecks.

Categories and Subject Descriptors

I.6 [**Computing Methodologies**]: Simulation and modeling; J.2.m [**Computer Applications**]: Physical sciences and engineering

*This work is partially supported by IBM.

ACM SIGSPATIAL IWCTS'12, November 6, 2012. Redondo Beach, CA, USA

General Terms

Algorithms

Keywords

traffic prediction, adaptive observation, Helbing's model, data assimilation

1. INTRODUCTION

Traffic monitoring applications systematically collect data about traffic conditions. These data are used for alerting drivers about congestion and accidents, planning new road pavements to accommodate predicted traffic loads, and so on. Many traffic models also have been developed to forecast the neat future traffic conditions in terms of certain spatio-temporal features of the system.

Unfortunately, in order to remain tractable, models of traffic phenomena are only approximations of the corresponding physical phenomena, and often require the setting of various parameters (whose exact values are dynamic and often unknown). The accuracy of these models can often be improved by fusing real observational data into the model. Data assimilation refers to methods for incorporating real observational data into the model to estimate values of the unknown parameters of the model and improve its forecasting accuracy. Assimilation is widely used in many areas, such as climate forecasting and tracking moving targets.

To collect observations, sensors are deployed in highways and secondary rural road networks. For example, the Federal and State Departments of Transportation (DOTs) carry out various programs to collect traffic data by means of inductive loop detectors, video surveillance systems, etc. The measurement devices are fixed at certain locations, and are hardly moved during their lifespan. Because of this reason, mobile observation devices attract more and more attention nowadays. For example, in modern systems, probe vehicles and camera-equipped helicopters are sent to locations where the observations are more likely to improve forecasts. Moreover, similar observations can be generated from social networks. During their trips, passengers update their information on social networks with GPS-equipped smartphones, which provide their location and possible verbal description of traffic conditions. This allows us to pull interested observation from specific locations with much less delay. The *Mobile Millennium Project* [8] at the University of California, Berkeley, conducted a series of experiments sending vehicles with GPS-enabled phones to highways and demonstrated that even a small penetration (2%∼3%) of cell phones is sufficient to accurately estimate traffic velocity on the highway. Work *et. al* [18] used ensemble Kalman filter data assimilation to fuse velocity measurements, collected from GPS phones on vehicles, with the Cell Transmission Model.

It is valuable to design observation strategies that increase the accuracy of forecast after doing data assimilation, while maintaining the cost of gathering the observations at a low level. Bishop *et. al* [3] proposed an adaptive observation sampling strategy for the ensemble transform Kalman filter, in which the new optimal observation sites are determined by the filter's predicted signal variance. Fox [5] studied the use of an adaptive number of particles for particle filters. The sampling strategy is to choose larger number of particles when the state uncertainty is high.

We propose a novel strategy to adaptively select locations of sensors to collect additional observations, and integrate it into a data assimilation framework. Our strategy utilizes information from the numerical solution of the system of partial differential equations (PDEs) used in modeling the macroscopic state of the traffic system, which requires very little additional calculation. In particular, we utilize adaptive moving meshes which have finer mesh cells where the numerical solution of the PDEs exhibits higher variability. The goal of our strategy is to deploy the limited observation resources at those locations that can better improve the accuracy of data assimilation, by observing more details where the solution of the PDEs may exhibit shocks, discontinuities, and high variability in general. Our strategy brings the advantages of adaptive moving mesh schemes into traffic data assimilation. We experimentally compare our strategy with the approach of assimilating observations from evenly spaced observation locations. We focus on two frequently occurring high-impact road phenomena, bottleneck and phantom jams, since during both the vehicle density has high variability at the wave fronts of the jams. We find that our strategy reduces the average relative error in estimating the vehicle density of the road during bottleneck and phantom jams by up to 53% with respect to the evenly spaced observation locations.

The rest of the paper is organized as follows. In Section 2 and Section 3, we review macroscopic traffic flow models and data assimilation methods. In Section 4, we describe adaptive moving mesh methods and our strategy of adaptively placing observation sites. Experimental results on bottleneck and phantom jams are presented in Section 5.

2. MODELS FOR TRAFFIC PHENOMENA

Macroscopic traffic flow models approach complex dynamical transportation systems by means of fundamental fluid dynamics laws, and describe traffic phenomena by certain spatio-temporal features, such as the vehicle flow rate, vehicle velocity, vehicle density, and road occupancy. These features can be of highly varying scales caused by phenomena like shock waves and contact discontinuities. Therefore, we need both models and solution schemes that are capable of capturing those phenomena.

The Lighthill-Whitham-Richard (LWR) model [7, 9] is a widely used macroscopic model. For example, the LWR model was used to study the traffic flow in the heavily congested Lincoln tunnel in New York City [15]. The LWR model assumes a fundamental relationship that describes the velocity $V(\rho)$ as a function of the density ρ, and requires the mass conservation law for vehicles, resulting in the following hyperbolic partial differential equation (PDE):

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho V(\rho)}{\partial x} = 0. \quad (1)$$

Unfortunately, studies show that the average velocity V depends not only on the density ρ at a particular location, but also on the average velocity and acceleration at nearby locations.

Payne [1] firstly suggested to take traffic acceleration into account, and extend the LWR model by including Equation 2. In this equation, P is the traffic pressure, τ is the drivers' reaction time, and V_e is the equilibrium velocity. The equation describes the behavior where drivers attempt to accelerate to match up when their vehicle speeds are less than the equilibrium speed, and decelerate otherwise. Kerner [13] proposed concrete forms for P and V_e.

$$\frac{\partial V}{\partial t} + V\frac{\partial V}{\partial x} = \frac{V_e(\rho) - V}{\tau} - \frac{1}{\rho}\frac{\partial P}{\partial x} \quad (2)$$

Helbing's model [6] extends the above models by also considering the dependence between the average velocity V and the variance of the average velocity Θ, resulting into the following system of hyperbolic PDEs

$$\frac{\partial \Theta}{\partial t} + V\frac{\partial \Theta}{\partial x} = -\frac{2P}{\rho}\frac{\partial V}{\partial x} - \frac{1}{\rho}\frac{\partial J}{\partial x} + \frac{2}{\tau}(\Theta_e(\rho) - \Theta), \quad (3)$$

where $P = (\rho\Theta - \eta_0 \partial V/\partial x)/(1 - \rho s(V))$ and

$$J = -\kappa_0(\partial\Theta/\partial x)/(1 - \rho s(V)). \quad (4)$$

In the equations above, Θ_e is the equilibrium variance of the average speed variance Θ. Eq. (2) models the pressure from the average desired velocity V_e. With such pressure, drivers adjust their speed toward V_e. Eq. (4) describes the flux of the velocity variance.

The macroscopic models above are capable of simulating the formation and evolution of traffic jams. A traffic jam refers to a condition that a sequence of vehicles accumulates with small inter-vehicular distances and low moving speed. Many reasons, such as accidents or road construction, lead to reduced road capacity and contribute to the formation of *bottlenecks*, which result in jams. A *phantom jams* is a special kind of jam that happens when slight disturbances in the road result in a steady growth of vehicle density behind. For example, when the density reaches a certain level and an inexperienced driver slows down a little, it results in a steady growth of density behind. But as this driver speeds up to match the velocity of vehicles in front, the jam will disappear quickly.

3. DATA ASSIMILATION FOR TRAFFIC

Data assimilation refers to recursive computational methods that find initial conditions for a system model that results in the best short-term forecast of the system state[12]. Assimilation for state-space system models usually consists of three components: the model operator $\mathcal{M}$ that maps the current system state $\mathbf{x}_t$ to a future state, the observational operator $\mathcal{H}$ that maps the system state $\mathbf{x}_t$ to the observational data $\mathbf{y}_t$, and the assimilation algorithm. In this study, we use the Helbing's model as the model operator $\mathcal{M}$. The system state is usually hidden (not directly observable). For discrete-time systems with additive noise, we have

$$\mathbf{x}_t = \mathcal{M}(\mathbf{x}_{t-1}) + \mathbf{u}_t, \quad \text{and} \quad \mathbf{y}_t = \mathcal{H}(\mathbf{x}_t) + \mathbf{v}_t, \tag{5}$$

where $\mathbf{u}_t$ and $\mathbf{v}_t$ are model and observation noise respectively, and are often assumed to be white noise.

For a state-space model, assimilation methods find estimates of the often hidden state $\mathbf{x}_t$, given the sequence of all prior observations $\mathbf{y}_{1:t} \doteq \mathbf{y}_1, \ldots, \mathbf{y}_t$. Usually, the hidden state $\mathbf{x}_t$ represents variables that are hard to measure directly, while $\mathbf{y}_t$ represents variables that are easy to measure. In traffic problems, the road density is usually modeled as a hidden sate $\mathbf{x}_t$, since it cannot be directly measured. Other vehicle velocity and number passing a point in the road are modeled as $\mathbf{y}_t$. In Bayesian estimation, the problem becomes to recursively quantify the (*posterior*) belief (probability distribution) in $\mathbf{x}_t$ given all the prior observations $\mathbf{y}_{1:t}$. Seminal assimilation methods include Kalman filters, particle filters and so on; please refer to [12, 16] for more details.

The system state $\mathbf{x}_t$ and the model operator are determined by the discretization of the system of PDEs used to model the physical system (eg traffic system), while the observation operator is determined by the available sensors and the system state. For example, for the case of the LWR traffic model, $\mathbf{x}_t$ would correspond to the density of cars on the road at each time step t. The discretization of the PDEs is due to the fact that generally the system of PDEs does not have analytical solutions, and hence we must resort to numerical solutions. The systems of PDEs given in section 2 for modeling traffic flow do not have an analytic solution in general. Further, note that the numerical errors in solving a system of PDEs, due to the discretization and the numerical methods used, will lead into an increase of the modeling errors of the system state. It is imperative that one minimizes these additional modeling errors.

4. ADAPTIVE OBSERVATIONS

We describe a strategy for selecting at each point in time, a few observation locations where measurements when assimilated are expected to provide us with the most power in resolving the areas of high variation in the state of the traffic system. Utilizing measurements at a few locations is beneficial due to computational considerations when gathering and assimilating them, and also due to the costs (both money and time) of deploying sensors at those locations. Moreover, we consider adaptive methods since, due to the highly dynamic nature of traffic, it is unlikely that a static selection of observation locations would suffice.

4.1 Adaptive Moving Meshes

We are often interested in resolving the behavior of the solution to PDEs in those regions of their physical domain where the solution has large variations in a small area of the physical domain. For example, when solving PDEs that govern the propagation of a shock wave (e.g. due to a traffic jam), we are interested in resolving the solution around the steep front of the wave. Moreover, the PDEs most often lack analytic solutions, and hence we resort to numerical solution schemes. Numerical solutions are derived using finite difference/element solution schemes on a mesh over the physical domain. Such a mesh needs to be fine at the small areas of the physical domain where the solution is highly variable (we call these regions the high detail regions). Compared to the uniform mesh solution, a more economical and practical approach is to have more mesh points in the high detail regions of the domain. Furthermore, since the solution to the PDEs are time-varying, the mesh points would be moving over time. This gives rise to the approach of using adaptive moving meshes in finite difference/element schemes for the numerical solution of such PDEs. Please refer to [10, 11] for a detailed review of adaptive moving mesh methods.

Adaptive moving mesh methods are highly desired in the numerical solution schemes for traffic models, since we will see that some traffic phenomena, such as phantom jams, are accompanied with high variability in certain regions, while the locations with free flow are less variable.

An adaptive moving mesh approach typically consists of three components: a strategy to move the mesh points, a method to discretize the physical domain, and an approach to solve the coupled system of physical and mesh equations. The strategy to move the mesh requires finding the new locations of the mesh points. The new locations are computed with the help of a mesh density function and using the equidistribution principle. The mesh density function is a user-defined function that is a proxy for the error density of the solution on each mesh cell. The equidistribution principle states that the integral of the mesh density function at each mesh cell should be the same for all cells. The idea is that, when viewing the numerical solution method as a function interpolation method, the total interpolation error at each mesh cell is the same for all cells. Methods for solving the mesh equations are given in [11, 14], while various mesh density functions are given in [4, 11].

4.2 Adaptive Selection of Observation Sites

Accurate, real-time observations with sufficient temporal-spatial resolution improve accuracy of traffic condition real-time forecasts, especially with data assimilation methods. Traditional observation networks, consisting of point detectors, video surveillance systems, and microwave radars, are expensive to build, maintain, and hard to adapt to new transportation infrastructure or dynamic travel demands. Probe vehicles have advantages in that they are mobile and easy to be deployed in different locations based on real demands with low cost. However, it is impractical to deploy a large number of probe vehicles. Also, it takes time to send probe vehicles to target locations. Therefore, computationally efficient methods for determining accurate 'hotspots' for observation is important for efficient and economical operation of a fleet of probe vehicles. In addition, smartphones equipped with GPS have recently attracted lots of interest in collecting real-time traffic conditions. Passengers on highway send timestamped messages with geo-location tags, which could provide information on the vehicle speed, density, travel time, and so on. The expected large number of smartphone and social network users provides better observation coverage and flexibility.

Mobile observation platforms, such as probe vehicles and smartphones, allow us to dynamically deploy observation sites to select locations that have high impact on correcting system prediction. Especially useful would be low-latency observations gathered from social networks. Once a set of "good" locations is determined, a query can be immediately executed to gather the most recent data from observation platforms in the vicinity of these locations.

Because the adaptive moving mesh strategy is desired especially for traffic models with shocks, discontinuities, and large variations in their solution, the newly calculated mesh provides us information about the regions of high detail on the domain. When the observation resources are limited, it is natural that we place more observation devices at the places where the numerical solution has difficulty to capture the details. Hence, the estimation of system state at those locations can be better corrected with more information from the real system.

Current mesh shapes are good for deciding current observation locations, which is feasible for observing from social networks. Observations from probe vehicles have higher latency, since it takes some time to send them to the desired locations. Nevertheless, the strategy of making observations based on an adaptive moving mesh can still work for these devices, by simply forwarding the solution (prediction) of the PDE model to a long enough future time; the new locations of the mesh points indicate locations in the road where high details are likely to happen. In this paper, we focus only on the determination of observation locations, and leave the problem of scheduling the probe vehicles to future work.

Algorithm 1 describes in detail our strategy of allocating limited observation points to those mesh cells of the adaptive moving mesh of higher detail and interest. Let M, M_1, N, and s be user-specified parameters. Given a fixed number of available observation points M and a mesh with N cells, a large proportion M_1/M of observation points are allocated to cells selected from the s cells with the smallest volume. The remaining observation points are assigned to cells selected from the remaining $N - s$ cells. There are many different strategies to select cells from the two groups of cells. One simple yet effective strategy is to uniformly sample from each group of cells.

Different preferences can be adopted to choose values for M_1 and s. For example, one extreme case is to assign all available M observation sites to the first M least-volume cells (i.e. set $M_1 = s = M$). Although this spirit agrees with the idea of the algorithm, it is better to reserve some observation locations for the low detail regions of the physical domain, because both the PDE model and the mesh (error) density function are only approximations.

Algorithm 1: Adaptively placing observation according to cell size.

input : mesh $\mathcal{J} : J_1, J_2, \ldots, J_N$, possible number of observations M
output: Observation locations and their observed value

calculate the size of each cell $L = (l_1, l_2, \ldots, l_N)$;
sort(L);
Find the first s smallest cells;
Randomly select M_1 from the s cells for observation;
From the rest $N - s$ cells, randomly select M_2 for observation, where $M = M_1 + M_2$ and $M_1 > M_2$;
Query observation values at these M locations. Feed the sampled observation into a data assimilation algorithm;

The running-time (overhead) of Algorithm 1 is dominated by the time to find the s mesh cells out of N with the least volume, which can be done in $O(N)$ time using standard order statistics algorithms.In most situations, the mesh size N is not large (e.g. 100s of cells for a 10km road).

5. EXPERIMENTAL RESULTS

We analyze our proposed strategy of adaptive observations using two experiments with traffic jams due to bottlenecks and phantom traffic jams.

Both experiments share the same circular road of 10 Km length with a periodic boundary condition. We use meshes with $N = 100$ cells and always start with a uniform mesh. Each timestep is 1 sec. Vehicles are evenly placed on the cells, and the initial average velocity in each cell is equal to $110Km/hr$ with a small added perturbation to mimic the diversity of driving speed. We use $M = 20$ observation points at each time step. For the data assimilation, we use a particle filter with 200 particles and a Gaussian proposal distribution. We also use adaptive moving mesh with the mesh density function proposed by Beckett and Mackenzie [2], and the monotone upstream-centered schemes for conservation laws (MUSCL) scheme [17] to find numerical solutions to the systems of PDEs.

We use relative errors in the system state to compare our proposed strategy with the approach of evenly spaced observation locations. In particular, every 50 seconds, we compute the average relative error in estimating the system state over the whose domain during the last 50 secs.

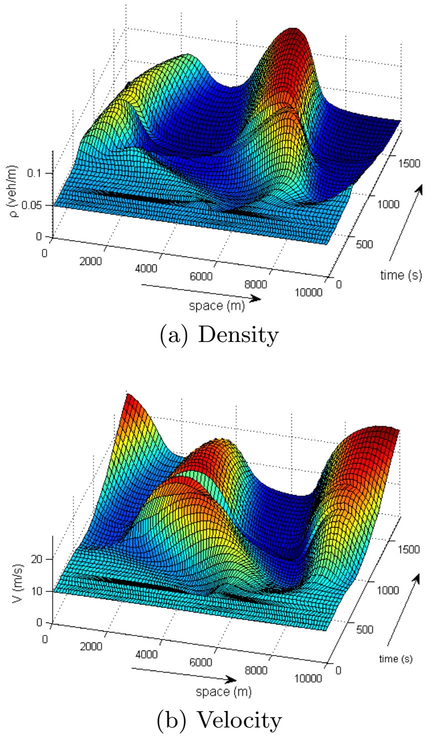
5.1 Jams due to bottlenecks

We setup an identical-twin experiment to generate observations from the true system state, and to test the data assimilation algorithm by assimilating these observations with the state of a similar model with different parameters as the one used to generate the true system state. First, Helbing's model is used to generate the true system state by reducing the road capacity from $200veh/Km$ to 20% of the initial capacity during the two time periods $200 \leq t \leq 240$ and $1010 \leq t \leq 1040$ at the locations $5000m - 5500m$ and $2000m - 7500m$ respectively. A twin Helbing's model, which is the same as the true model above except that it has no knowledge of the changes of road capacity, attempts to reproduce the true state by assimilating partial observations from the true system state.

In setting observation sites, at each time step, we uniformly allocate (sample) $M_1 = 12, 13, 14, 15$ observations to mesh cells among the $s = 25$ cells of least volume to observe at each iteration, corresponding to 60%, 65%, 70% and 75% of the $M = 20$ available observation points. The rest $M - M_1 = 8$, 7, 6, 5 observation sites are uniformly sampled without replacement from the remaining 75 mesh cells. Figure 2 shows the relative forecasting error with a uniform mesh (blue curve) and our adaptive observation strategy (red curve). The relative error for the evenly spaced observations approach has two obvious peaks during the periods the bottleneck occurs; the adaptive observation strategy does not exhibit such a sharp deterioration of forecasting ability during the bottleneck occurrences. Furthermore, our proposed method reduces the relative estimation by 47%∼53% with respect to using the evenly spaced observations approach.

5.2 Phantom jams

On highways, phantom jam is a commonly observed phenomenon, even when there is no shutdown of lanes nor other change of road capacity. When traffic is of high density but



(a) Density

(b) Velocity

Figure 1: True density and velocity of a bottleneck jam simulated using Helbing's model.

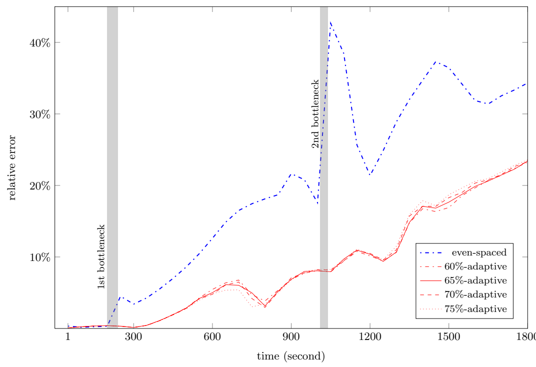
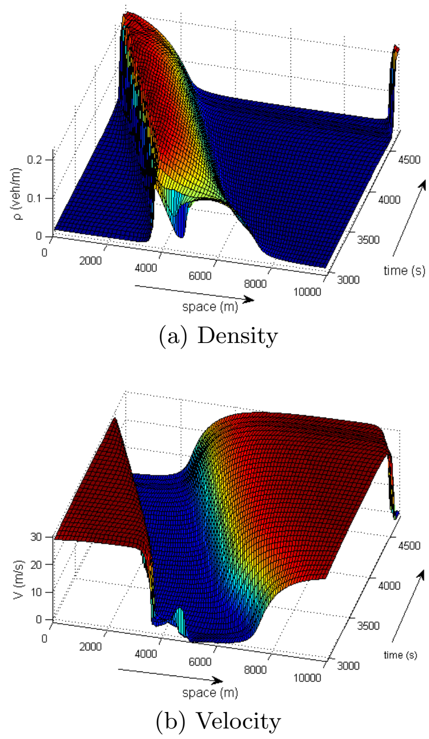


Figure 2: Relative estimating error for the evenly spaced observations approach and the adaptive observations approach for different proportions M_1/M of observations allocated to the s least volume cells.

still in free flow, sudden self-sustaining traffic jams may form after small disturbances, such as when an unskillful driver slows down. We can observe a sequence of reduced speed areas propagate backward along the road, as Figure 3(a) and 3(b) illustrate.

In this experiment, the true system state was generated with Kerner's model. A shockwave can be observed during the period 3000-4800 secs. We use Helbing's model to reproduce the true system state by assimilating M observations from the true system state at various locations. Two methods are used to decide on the observation locations at each time step: the evenly spaced observations approach and our adaptive observations strategy

The proposed adaptive observation strategy reduces the relative error by 29.7%~45.8% (for M_1/M ranging between $60\% - 74\%$) on average with respect to the evenly space observations approach.



(a) Density

(b) Velocity

Figure 3: Density and velocity of true system state of a simulated phantom jam using Kerner's model.

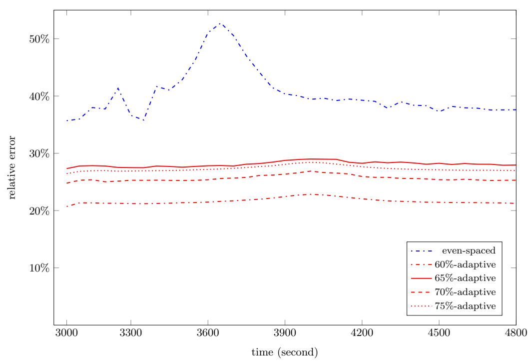


Figure 4: Relative estimation error for the phantom jam experiment for the evenly spaced observations approach and the adaptive observations strategy for different proportions M_1/M of observations allocated to the s least volume cells.

6. SUMMARY

Mathematical models sketch traffic phenomena on road networks, such as traffic jams, shock waves. When provided with the most recent traffic condition as their initial conditions, those models allow us to forecast trajectories of possible future traffic conditions, and hence support transportation management decisions. With data assimilation, forecasting can be improved by fusing real observations with the model predictions. Smartphones and social networks can supply a huge amount of flexible and rich information on traffic condition through timestamped messages with geo-location tags, which well compensate the shortcoming of fixed observation networks like loop detectors and video surveillance systems currently deployed. Because of the mobility and ubiquity of smartphones, it is possible for us to pull a small amount of observations from regions of high variability.

We proposed a strategy to place observations at select locations computed using adaptive moving mesh scheme for numerically solving the underlying system of PDEs. We place a large proportion of observation sites to the most fine mesh cells of an adaptive moving mesh of the physical domain of the PDEs, which are expected to be the most beneficial for the numerical solution of the PDEs. By observing more at places where numerical solutions are of sharp variations, data assimilation methods have the opportunity to improve estimation at those locations, and hence achieve higher forecasting accuracy. We experimentally show, using simulated traffic bottleneck jams and phantom jams that our proposed adaptive observations strategy improves upon traffic estimation accuracy by up to 53% with respect to evenly space observation locations.

Our strategy brings the advantages of adaptive moving mesh numerical solution schemes into the data assimilation framework, providing an effective adaptive dynamic observation site placement. Our strategy utilizes information from the numerical solution schemes of the underlying system of PDEs in the model, with marginal running time overhead, while offering substantially higher estimation accuracy.

7. ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers for their careful review and comments. This project is partially supported by the IBM. In particular, we would like to thank Dr. Bob Szabo for his passionate support of this work.

8. REFERENCES

[1] J. Barceló, editor. *Fundamentals of Traffic Simulation*, volume 145 of *International Series in Operations Research & Management Science*. Springer, 2010.

[2] G. Beckett, J. A. Mackenzie, A. Ramage, and D. M. Sloan. On the numerical solution of one-dimensional pdes using adaptive methods based on equidistribution. *J. of Comp. Physics*, 167(2):372–392, 2001.

[3] C. H. Bishop, B. J. Etherton, and S. J. Majumdar. Adaptive sampling with the ensemble transform kalman filter. part i: Theoretical aspects. *Monthly Weather Review*, 129:420–436, March 2001.

[4] C. J. Budd, W. Huang, and R. D. Russel. Moving mesh methods for problems with blow-up. *SIAM Journal on Scientific Computing*, 17(2):305–327, March 1996.

[5] D. Fox. Kld-sampling: adaptive particle filters. *Statistics*, 14:713–720, 2001.

[6] D. Helbing. Improved fluid-dynamic model for vehicular traffic. *Physical Review E*, 51(4):3164–3169, April 1995.

[7] D. Helbing. Traffic and related self-driven many-particle systems. *Reviews of Modern Physics*, 7(3):1067–1141, October 2001.

[8] J.-C. Herrera, D. Work, X. Ban, R. Herring, Q. Jacobson, and A. Bayen. Evaluation of traffic data obtained via gps-enabled mobile phones: the mobile centuryfield experiment. *Transportation Research C*, 18:568–583, 2010.

[9] S. P. Hoogendoorn and P. H. Bovy. State-of-the-art of vehicular traffic flow modelling. *Journal of Systems and Control Engineering*, 215(4):283–303, 2001.

[10] W. Huang, Y. Ren, and R. D. Russel. Moving mesh partial differential equations (mmpdes) based on the equidistribution principle. *SIAM Journal on Numerical Analysis*, 31(3):709–730, June 1994.

[11] W. Huang and R. D. Russel. *Adaptive Moving Mesh Methods*. Springer, 2011.

[12] E. Kalnay. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 2003.

[13] B. S. Kerner and P. Konhäuser. Cluster effect in initially homogeneous traffic flow. *Physical Review E*, 48(4):R2335–R2338, 1993.

[14] B. J. Lucier. A moving mesh numerical method for hyperbolic conservation laws. *Mathematics of computation*, 46(173):59–69, January 1986.

[15] B. Piccoli and A. Tosin. *Encyclopedia of Complexity and Systems Science*, chapter A Review of Continuum Mathematical Models of Vehicular Traffic. Springer, to appear.

[16] B. Ristic, S. Arulampalam, and N. Gordon. *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Artech House, 2004.

[17] B. van Leer. Towards the ultimate conservative difference scheme. v. a second-order sequel to godunov's method. *Journal of Computational Physics*, 32:101–136, 1979.

[18] D. Work, S. Blandin, O.-P. Tossavainen, B. Piccoli, and A. Bayen. A traffic model for velocity data assimilation. *Applied Mathematics Research eXpress (AMRX)*, 1:1–35, 2010.