Particle filter-based real-time estimation and prediction of traffic conditions

Jacques Sau1, Nour-Eddin El Faouzi2, Anis Ben Aissa2, and Olivier de Mouzon2

1 LMFA, University Claude Bernard Lyon 1
43, Boulevard du 11 novembre 1918
69622 Villeurbanne Cedex, France
(email: jacques.sau@univ-lyon1.fr)

2 Transport and Traffic Engineering Laboratory
INRETS, LICIT, laboratoire d'ingénierie circulation transports
69675 Bron Cedex, France
ENTPE, LICIT, laboratoire d'ingénierie circulation transports
69518 Vaulx-en-Velin, France
(email: elfaouzi@inrets.fr, anis.ben-aissa@inrets.fr, olivier.de-mouzon@inrets.fr)

Abstract. Real-time estimation and short-term prediction of traffic conditions is one of major concern of traffic managers and ITS-oriented systems. Model-based methods appear now as very promising ways in order to reach this purpose. Such methods are already used in process control (Kalman filtering, Luenberger observers). In the application presented in this paper, due to the high non linearity of the traffic models, particle filter (PF) approach is applied in combination with the well-known first order macroscopic traffic model. Not only shall we show that travel time prediction is successfully realized, but also that we are able to estimate, in real time, the motorway traffic conditions, even on points with no measurement facilities, having, in a way, designed a virtual sensor.

Keywords: Real-time traffic estimation, Bayesian Monte Carlo, travel time prediction.

1 Introduction

With the rapid deployment of Intelligence Transportation Systems (ITS), the users are more and more informed about traffic conditions and special events during their travel. It has been recognized that the full benefits of ITS systems cannot be realized without an ability to anticipate traffic conditions in the short-term ([Sussman, 2000]). Hence, short-term traffic prediction of traffic state could play a key role in various ITS applications, such as Advanced Traffic Management System (ATMS), Dynamic Vehicle Navigation System (DVNS) and Advanced Travel Information Systems (ATIS).

The purpose of predicting short-term traffic conditions is to forecast traffic flow variables such as traffic volume, travel speed, or travel time in the range from 5 to 30 minutes ahead. For travel time estimation and prediction, various methodologies and techniques have been explored in developing short-term prediction models. Due to the non-linearity, complexity and uncertainty of contributing factors of the traffic conditions in the traffic system, the forecasting method based on deterministic models cannot meet the accuracy needed by ITS applications. To overcome this limitation, many non-deterministic forecasting
models have largely improved the accuracy of forecasting. Such a model can be well adapted to different locations and easily transplanted, e.g. [Vlahogianni et al., 2004]. In this scope, our research effort focused on developing a stochastic traffic modeling framework that enables travel time estimation and prediction with high reliability.

The basic components, in our application, are first that the traffic can be modeled, and second that, at any time, the traffic state is completely characterized by a finite number of quantities called the state vector of the system, from which any interesting quantity can be given. The dynamic evolution of the system is modeled following the well-known macroscopic traffic model LWR (Lighthill-Whitham-Richards, [Lighthill and Whitham, 1955], [Richards, 1956]), which gives the evolution of the state vector $X_t$ at every time step. An observation equation gives the relation between measured quantities $y_t$ and state vector $X_t$. The problem then consists of estimating the state vector $X_t$, using the observations $y_t$. Since the state equation is highly non linear, this task will be carried out using the particle filter approach known also as the Bayesian Sequential Monte Carlo method ([Doucet et al., 2001], [Chen, 2003], [Doucet, 1998]).

2 Modeling framework

2.1 Traffic model in brief

A macroscopic approach is based on hydrodynamic analogy describing the behaviour of the traffic flow. The first model of traffic flow was introduced concurrently by [Lighthill and Whitham, 1955] and [Richards, 1956]. The use of this macroscopic model for simulating and predicting future traffic conditions along the roadway implies a space-time discretization. One of the best first order discretization schemes for the computation of the entropy solution (see [Velan and Florian, 2002]) of the LWR is the Godunov scheme ([Godunov, 1959], [Lebacque, 1996]). Hence, the motorway is discretized into n cells (Figure 1).

$k_i$ is the density in cell $i$ [veh/m]
$q_i$ is the exit flow of cell $i$ [veh/min]

Fig. 1. Space Discretization.

$q_{\text{max}}$

$k_{ic}$

$k_{\text{max}}$

Fig. 2. Fundamental Diagram.
Particle filter for real-time estimation & prediction of traffic conditions

The basic hypothesis, in the LWR model, is the existence of a quasi stationary relation \( q = Q_i(k) \) between flow and density in a given cell \( i \). This relation takes the general typical shape depicted in Figure 2 ([Greenshields, 1935]), where:

- \( q_{\text{max}} \) is the maximum possible flow in the cell.
- \( k_c \) is the critical density of the cell.
- \( k_{\text{max}} \) is the maximum possible density of the cell.

The state equation is then built as follows:

- **Conservation equation:**
  \[
  k'_i = k'_{i+1} + \frac{\Delta t}{\Delta x} (q'_{i+1} - q'_{i+1}).
  \]

- **Supply-demand equation:**
  - Supply of cell \( i+1 \) (to its upstream):
    \[
    \Omega\left(k'^{i+1}\right) = Q_{i+1}\left(\max\left(k'^{i+1}, k'_{i+1}\right)\right).
    \]
  - Demand of cell \( i \) (to its downstream):
    \[
    \Gamma\left(k'_{i}\right) = Q\left(\min\left(k', k'_{i}\right)\right).
    \]
  - Resulting flow leaving cell \( i \):
    \[
    q'_i = \min\left(\Omega\left(k'^{i+1}\right), \Gamma\left(k'_{i}\right)\right).
    \]

The measured variables are flows and/or densities on some specific cells. The observation equation is then written:

\[
y_i = CX_i,
\]

where \( C \) is the observation matrix, with as many lines as measured variables.

However, for a numerical application, one must distinguish the observation time step \( \Delta t_o \) from the numerical time step \( \Delta t_N \) used for the numerical solution of the state equation. The observation time step is an on-field constraint, typically 6 min in our application. The numerical time step is constrained by the Courant-Friedrichs-Lewy (CFL) condition in order to obtain a stable numerical scheme ([Godunov, 1959], [Lebacque, 1996]). The numerical time step will be always such that \( \Delta t_o \) is multiple of \( \Delta t_N \), i.e. \( \Delta t_o = N_o \Delta t_N \), with \( N_o \) chosen so that the CFL condition will be always satisfied everywhere.

With state vector defined as \( X_i = (k_1, k_2, \cdots k_n; q_0, q_1, \cdots q_n)^T \), the dynamic system will be written in term of observation time step:

\[
\begin{align*}
X_{i+1} &= F(X_i, u_i) \\
y_i &= CX_i
\end{align*}
\]

Therefore, \( N_o \) numerical time steps are needed for an observation time step. The function \( F(X_i, u_i) \) contains then implicitly \( N_o \) numerical Godunov time steps.
In this system, the inputs are the traffic conditions at the boundaries: 
\[ u_t = (u_1, u_2)^T \], where \( u_1 \) (resp. \( u_2 \)) is the upstream demand (resp. downstream supply).

### 2.2 State vector estimation

In the previous section, a traffic model has been described. The aim is then to estimate the state vector using the observations. The traffic model is highly non-linear, therefore Kalman filter can hardly be used.

Sequential Monte-Carlo or particle filter approach provides the solution to deal with such a case. Monte Carlo methods are widely used to simulate the dynamic of complex systems. This approach will be used in the present work (see e.g. [Ben Aissa et al., 2006], [Mihaylova and Boel, 2004] and [Doucet et al., 2001] for a detailed description of the application of SMC to the traffic application, or [Arulampalam et al., 2002] and [Doucet et al., 2001] for the theory of the SMC).

In general, importance sampling method is used, as it is in fact difficult to sample directly from the posterior density \( P(X_{0:t} | y_{0:t}, u_{0:t-1}) \). However, in this first application, since the measurements directly concern components of the state vector, we considered that partial Gaussian state space case, see [Doucet et al., 2001] and [Doucet, 1998] for a detailed description of this case.

### 3 Application to Travel Time Prediction

#### 3.1 Data and Network Site

The application has been performed on a French motorway section, owned by the “APRR” private company, from ‘Mâcon Sud’ to ‘Belleville’. This route is approximately 21 km long. Two traffic detectors are located in this route. The section between the detectors is 15 km long. A road accident occurred in between the traffic detectors, causing congestion. This site is very interesting due to available data: Additionally to detector traffic data, data from toll collection system was also available for some vehicles which both enter the motorway before the first detector and leave it after the last one. This data provides individual travel time (magnetic toll stamps). Once filtered, it constitutes a reference experienced travel time. Therefore, predicted travel time coming from the Bayesian Monte Carlo state vector reconstruction can be validated by these independent measurements issued from what we can call probe vehicles.

The section between the detectors has been discretized in 50 cells of 300 m long. The observation time step of the detector data base is 6 min. This is the time step for the re-estimation of Monte Carlo Bayesian procedure.
3.2 Model calibration

The model parameters, i.e. critical density and maximum flow, have been determined from independent data. Data from a different - but similar - motorway section have been retrieved. These data cover traffic conditions from the free flow to heavy saturated one, but with no perturbing events like accidents or other incidents. Comparison between the measured flows and the corresponding ones given by our model in the same conditions allows the best estimation of the model parameters. A criterion is needed for this task. The relevant criterion is the minimization of the joint entropy, (or maximization of mutual information) between measures and model (see [Li and Vitany, 1997] and [Guiasu, 1977]). The minimization of this positive and parameter dependant quantity gives the optimal choice for the parameter values. As an illustration, the optimal maximum flow and critical density for a three lanes motorway stretch are founded to be:

• $k_c = 0.12 \text{ veh/m}$,
• and $q_{\text{max}} = 100 \text{ veh/min}$.

3.3 State vector estimation and Travel time prediction

Once the traffic model established, Sequential Monte Carlo Bayesian estimation of the state vector of the system is performed. After the model calibration, and in order to fulfil the CFL condition, a numerical time step of 0.143 min has been chosen, leading to $N_o = 42$ numerical time steps for an observation one.

First, we act as if the traffic takes place in normal conditions. The diagnosis comes then from the gap between on-field measures and measures rebuilt by the Bayesian procedure. Figure 3 shows the measured and the re-estimated flows at the downstream end of the section. Most of the time, the re-estimated curve is close to the measured one. However, two kinds of events can be observed on the graph. First, clear breakdowns of the sensor can be seen: In these cases, the returned value is zero. The measure is correctly rebuilt by the Bayesian procedure.

Fig. 3. Downstream flows with normal traffic conditions hypothesis.  
Fig. 4. Downstream flow taking into account the accident occurrence.
Second, in the 2:00 pm - 3:00 pm slot an anomaly is clearly occurring in the motorway traffic observation. It can be due to a sensor default like a drift, or a traffic incident. With no other information, sensor default is the hypothesis, and the measure is rebuilt. In fact an accident occurred, and was reported by the patrol data and lasted one hour. This accident is modelled as restricted values of critical density and maximum flow during the accident duration at the accident location which corresponds to the cell number 25. In this cell, free velocity and maximum flow have been respectively divided empirically by 4 and 9, which correspond to reasonable values.

With these new physical conditions the comparison between downstream measured flow and re-estimated one is shown figure 4. The agreement is now good, including the growth of flow which follows the release of vehicles at the end of the accident.

A global view of the spatio-temporal traffic concentration and flow in the section, as estimated by the Bayesian procedure, is depicted in figure 5. The concentration and flow dependence in abscissa and time are surfaces in a 3-D diagram. The traffic jam provoked by the accident is seen in the concentration graph, with a decrease of concentration downstream the jam. At the same time, flow is almost zero downstream the accident because of no vehicles travelling and in the jam because of very low speed. The sudden rise of flow just after the jam is also very well reproduced. The maximum length of the traffic jam can be evaluated which gives a value of around 2 km, in comparison to the value of 2.2 km reported in the patrol data base.

![Concentration and Flow](image)

**Fig. 5.** Global view of concentration and flow in space and time in the motorway section.

The estimated time evolution of concentration and flow in cell number 23, just upstream the accident and where no direct measurement is performed, are depicted Figure 6. The rise of concentration, which reaches almost the maximum value, and the decrease of flow during the traffic jam caused by the accident
appear distinctly. During this jam, the vehicle velocity can be calculated using concentration and flow values, and is around 3 km/h. The estimated density and flow in a non-direct measured cell behaves then like a virtual sensor.

The predicted travel time is shown Figure 7. It is the travel time predicted for vehicles entering the section. Indeed, at current time $t$ we have estimated the history of past state vectors. We can then, using the traffic model, perform a prediction of travel time for vehicles entering the section.

As we said above, for this particular motorway section, toll gates exist before and after the section. Therefore, a real-world measurement of travel time experienced by the vehicles can be calculated from toll gate data base. A comparison between this real measurement and the prediction can be performed and is depicted Figure 7. A very good agreement is clearly seen, showing that estimations of densities and flows based on a traffic model provide relevant travel time predictions.

![Fig. 6. Estimated concentration (veh/m) and flow (veh/min) in cell 23.](image)

![Fig. 7. Predicted travel time for vehicles entering the section.](image)

4 Conclusion

In this paper, we have proposed an iterative stochastic approach to capture dynamically the spatio-temporal behaviour of traffic flow for the purpose of short-term travel time prediction. The basic hypothesis is the ability to model the traffic evolution by a state equation. Therefore, the travel time estimation and prediction problem can be successfully reformulated and solved as a sequential estimation process, using Monte Carlo procedure. Second, the calibration process was performed using maximization of mutual information principle. Finally, the obtained results pointed out the predictive capabilities of the underlined estimation process, and its benefit for real-time travel time prediction.

Acknowledgements

The authors wish to thank the APRR motorway company for providing the real-world data used in this research.
References


