Traffic state estimation from partial Bluetooth and taxi data: Case study in the Brisbane Metropolitan Area

D. Ngoduy [a], A. Nantes [b], M. Miska [b] and E. Chung [b]
[a] *Institute for Transport Studies,*
The University of Leeds, United Kingdom
Email: d.ngoduy@leeds.ac.uk
[b] *Department of Civil Engineering,*
Queensland University of Technology, Australia

1. BACKGROUND

Recent rapid advances in information technology have led to various data collection systems which enrich the sources of empirical data for the traffic state estimation problems. In practice, traffic data are collected from loop detectors, floating cars, cell-phones, video cameras, remote sensing, etc. Particularly, the application of the Bluetooth (BT) technology to transportation has been enabling researchers to make accurate travel time observations, in freeway and arterial roads. These travel times are often presented through sufficient statistics, computed from the set of per-vehicle recordings at both upstream and downstream intersections. It is commonly reported that the BT enables fairly reliable travel time estimations but BT data neglect the dynamics of traffic flow within link so few studies carried out to reconstruct traffic states (i.e. transitions between free flow and congested situations) within links and network-wide.

The Bluetooth traffic data are generally incomplete, for they only relate to those vehicles that are equipped with Bluetooth devices, and that are detected by the Bluetooth sensors of the road network. The fraction of detected vehicles versus the total number of transiting vehicles is often referred to as Bluetooth Penetration Rate (BTPR). Not only is the BTPR unknown, but is also a function of time and space. Nevertheless, the detected vehicles will still flow according to the same macroscopic laws that describe the flow of all vehicles, observable and non-observable. The aim of this study is thus to precisely define the spatio-temporal relationship between the quantities that become available through the partial, noisy BT observations; and the hidden variables that describe the actual dynamics of vehicular traffic. The estimated results will be validated using taxi data, which are used as ground truth.

In this work, we propose to study the traffic dynamics through a three-class continuum model; in which BT-equipped vehicles are seen as belonging to one class and taxi and other non-equipped vehicles are considered the other class. We will incorporate our model into a Sequential Montecarlo Estimation (SME) algorithm, in order to unveil the current and future distributions of total flow and BTPRs. This is a so-called model-based traffic state estimator, which is an optimization problem of combining model predictions from a traffic model and traffic measurements from sensors. Our framework will be tested in the Brisbane Metropolitan region.

2. METHODOLOGY

In order to estimate the complete state of a system, through an observer-based estimator, a complete model of the system is needed. In this paper, the dynamics of the arterial traffic is described through a single-pipe *three-class* kinematic wave model. The three classes selected are the vehicles detectable by the Bluetooth scanners; taxi and the other undetected vehicles. As we shall see, the model proposed will enable the estimation of the traffic state; upon incomplete and noisy Bluetooth and volume observations. Our system is based on the LWR model:

$$\frac{\partial k_u}{\partial t}+\frac{\partial}{\partial x}q_u(k_u)=0 \tag{1}$$

where k_u and q_u are the density and flow, respectively, of the generic class u, for example, $u=1$ for the *Bluetooth* and $u=2$ for taxi class. The LWR model for the dynamics of the total traffic flow reads:

$$\frac{\partial k}{\partial t}+\frac{\partial}{\partial x}q(k)=0, k \equiv \sum_u k_u, q \equiv \sum_u q_u \tag{2}$$

Since the Bluetooth and taxi vehicles represent a fraction, α and β, respectively, of the entire flow we can derive the following equations for the dynamics of the fraction:

$$\frac{\partial \alpha}{\partial t} + V(k)\frac{\partial \alpha}{\partial x} = 0, \frac{\partial \beta}{\partial t} + V(k)\frac{\partial \beta}{\partial x} = 0 \tag{3}$$

Where *V(k)* denotes the mean speed of the entire traffic flow as we assume that all Bluetooth and taxi vehicles are moving with the same mean speed in the flow (i.e. no overtaking is considered). By definition

$$\alpha = \frac{k_1^t}{k^t} = \frac{q_1^t}{q^t}, \beta = \frac{k_2^t}{k^t} = \frac{q_2^t}{q^t} \tag{4}$$

The state of a road network is typically only partially observable. That is, the flow, density and speed of each cell are not directly available. Instead, only limited information is available at the intersections. For instance, if an intersection has stop-line detectors installed, in all lanes, on all approaches, one can measure the flows of the approaches at the intersections. With the taxi data available, it becomes possible to estimate the average speed that the individual (taxi) vehicles have maintained between any two scanners or intersections, within the observation time. The partial outflow and inflow as well as the travel time can also be extracted from the Bluetooth recording. Note that all observations contain some level of uncertainty or noise. This noise is to attribute, among other factors, to the location of the stop-line detectors and the aggregation time used to collect the vehicle counts, the variable scanning area, and the finite scanning frequency of the Bluetooth sensors. In this work, we shall assume that all uncertainties of the system are additive, Gaussian and zero-mean. Under this assumption, if we let $\mathbf{x}^{t+1}$ be the current state (vector) of the dynamical system, and $\mathbf{y}^{0,\dots,t}$ the set of measurements available to date, the complete model of the system becomes

$$\begin{cases} \mathbf{x}^{t+1} = f(\mathbf{x}^t, \mathbf{u}^t) + \xi^t & \xi^t \propto \aleph(0, W^t) \\ \mathbf{y}^t = h(\mathbf{x}^t) + \vartheta^t & \vartheta^t \propto \aleph(0, R^t) \end{cases} \tag{5}$$

where $\mathbf{u}$ is an external input, which contains the known parameters that are input to the system. In the context of our work, these are the free-flow speed, the jam density, the maximum flow, the green-split, and the priority ratio for all approaches of the intersections. ξ and ϑ are independent Gaussian noise terms of zero-mean and covariance matrices W and R, respectively. The update function, f, encompasses all models presented earlier for the update of the state variablesfor each cell when we discretize equations (2) and (3). The state vector is mapped to the measurement vector, $\mathbf{y}$, through a function h, whcih contains the quantities that can be observed: partial flow and travel time of Bluetooth vehicles, mean speed of taxi. System (4) is determined using a so-called Bayesian estimation technique (Arulampalam et al. 2002), which will be detailed in the full paper.

Our case study concerns a section of an arterial road (Coronation Drive) of the Brisbane metropolitan area, linking four signalized intersections. The three road segments, connecting the four intersections, have length 435m, 330m, and 710m. Travel time and flow of Bluetooth vehicles between intersections, mean speed of taxi vehicles at some locations in links on Wednesday 3 October 2012 were used in our simulation. The full results will be presented at the conference if the abstract is accepted

REFERENCES

Arulampalam, M. S., S. Maskell and N. Gordon. 2002. "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking." *IEEE Transactions on Signal Processing* 50, pp. 174--188.

Lighthill, M. J. and G. B. Whitham. 1955. "On Kinematic Waves. II. A Theory of Traffic Flow on Long Crowded Roads." *Proc. R.oy Soc. Lond. A* 229, pp. 317--345.

Richards, P. I. 1956. "Shock Waves on the Highway." *Operations Research* 4, pp.42--51.