EXPLORING THE USE OF TRAFFIC DATA COLLECTED FROM NEW ICT BASED SENSORS TO ESTIMATE TIME DEPENDENT OD MATRICES

L. Marqués

CENIT (Center for Innovation in Transport) Technical University of Catalonia

L. Montero

Department of Statistics and Operations Research, Technical University of Catalonia

C. Carmona

CENIT (Center for Innovation in Transport) Technical University of Catalonia

J. Barceló

Department of Statistics and Operations Research and

CENIT (Center for Innovation in Transport) Technical University of Catalonia

Email: jaume.barcelo@upc.edu

Time dependent origin to destination, OD, matrices are the key input to dynamic traffic models, mainly to simulation models, micro as well as mesoscopic. Dynamic Traffic Models, DTM, are one of the major components of the Advanced Traffic Management Systems and Advanced Traffic Information Systems. DTM play a crucial role in estimating the current traffic state and forecasting its short term evolution. The quality of the results that they provide depends, not only on the quality of the models, but also on the accuracy and reliability of the inputs and, therefore, on the quality of the time dependent OD matrices as part of that input. These matrices have usually been estimated by procedures exogenous to the traffic simulation model, based typically on heuristic procedures adapted from static matrix adjustments from link flow counts. Recently dynamic approach based on Kalman Filtering, [1], [2], have been proposed, they explicitly assume that a dynamic assignment procedure is available. The quality and reliability of the measurements produced by inductive loop detectors, is not usually the one required by real-time applications, therefore one wonders what could be expected from the new ICT technologies as for example Automatic Vehicle Location, License Plate Recognition, detection of mobile devices and so on, in particular those equipped with Bluetooth technology, that are becoming pervasive data sources. Once the privacy concerns are overcome, tracking mobile devices associated uniquely to vehicles becomes a rich source of new traffic related data from which infer time dependent mobility patterns. Better results should be expected when V2I technologies are taken into account making possible paths reconstructions and from them the estimation of origin to destination matrices for each time period. The research reported in this paper explores two complementary issues for

estimating OD matrices: the exploitation of travel time measurements provided by sensors detecting Bluetooth devices equipping vehicles (Tom-Tom, Parrot, hands free...) combined with input-output flow measurement at entry and exit ramps on a motorway; and the use of data supplied by V2I technologies (i.e. positions and speeds) that allow tracking vehicles and estimate direct samples which combined with a path reconstruction process allow to estimate the OD. The first approach, suited only to Motorways or Freeways, is based in an ad hoc adaptation of Kalman-Filter, combining elements from [5] and [6]; while the second, more appropriate for networks, uses elements from [3] generalizing methods for OD estimation based on license plate recognition.

In the first case a simulation experiment has been conducted, prior to the deployment of the technology in a forthcoming pilot project. The simulation emulates the logging and time stamping of a sample of equipped vehicles. Since data from equipped vehicles constitute a random sample of traffic data of significant size, the measured travel times can be used as real-time estimates of travel times for the whole population of vehicles. The availability of real-time travel time estimates makes possible a more efficient use of Kalman Filtering for OD estimates, simplifying the equations and replacing state variables by real-time measurements. We focused our attention on dynamic OD estimation in linear congested corridors where no route choice strategy is considered since there exists a unique path connecting each OD pair, but the travel time between each OD pair is considered and affected by congestion. We propose a space-state formulation for dynamic OD matrix estimation in corridors considering congestion that combines elements of Chang and Wu [6] and Van Der Zijpp and Hamerslag [5] proposals. A linear Kalman-based filter approach is implemented for recursive state variables estimation. Tracking the vehicles is assumed by processing Bluetoooth and WiFi signals by sensors located at the entry ramps (mandatory), in the main section (as many as possible) and the off-ramps (as many as possible). Traffic counts for every sensor and OD travel time from each entry ramp to the other sensors (main section and ramps) are available for any selected interval of length higher than 1 second. Then travel time delays between OD pairs or between each entry and sensor location are directly provided by the detection layout and should no longer be state variables but measurements simplifying the approach and making it more reliable. A basic hypothesis that requires a statistic contrast for real test site applications is that equipped and non equipped vehicles follow a common OD pattern. The state variables $b_{ij}(k)$, defined in terms of proportions of trips between OD pairs (i,j), are assumed to be stochastic in nature and evolve according with an independent random walk process whose state equation is: $\mathbf{b}(\mathbf{k}+\mathbf{1}) = \mathbf{D}\mathbf{b}(\mathbf{k}) + \mathbf{w}(\mathbf{k})$ $\mathrm{b}(\mathrm{k})$ is the column vector of all feasible OD pairs (i,j), ordered by entry ramp, and $w_{ij}(k)$'s are independent Gaussian white noise sequences with zero mean and covariance matrix $\mathbf{Q}$. The state variables should additionally satisfy the structural constraints

$$b_{ij}(k) \geq 0 \qquad i = 1 \ldots I, \quad j = 1 \ldots J$$

$$\sum_{j=1}^{J} b_{ij}(k) = 1 \qquad i = 1 \ldots I$$

Let's denote by: $q_i(k)$ number of equipped vehicles entering the freeway from on-ramp i during interval k and i=1,...I; $s_j(k)$ number of equipped vehicles leaving the freeway by off-ramp j during interval k and j=1,...,J; $y_p(k)$ number of equipped vehicles crossing main section sensor p and p=1,...,P; $G_{ij}(k)$ number of vehicles entering the freeway at on-ramp i during interval k with destination to off-ramp j; $g_{ij}(k)$ number of equipped vehicles entering the freeway from ramp i during interval k that are headed towards off-ramp j; IJ = IxJ, number of feasible OD pairs depending on entry/exit ramp topology in the corridor; $t_{ij}(k)$ average measured travel time for equipped vehicles entering from entry i and leaving by off-ramp j during interval k; $t_{ip}(k)$ average measured travel time for equipped vehicles entering from entry i and crossing sensor p during interval k; $b_{ij}(k)=g_{ij}(k)/q_i(k)$ the proportion of equipped vehicles entering the freeway from ramp i during interval k that are destined to off-ramp j; e(k)=e a colum vector of dimension I containing ones; z(k) vector of observation variables during interval k; i.e. a column vector of dimension I+J+P, whose structure is $z(k)^T=[s(k), y(k), e(k)]^T$. Let's define $U_{ijq}^{h}(k)= 1$ if the average measured time-varying travel time during interval k to traverse the freeway section from entry i to sensor q takes h time intervals, h = 1,...,M and q = 1,...Q and Q = J+P (the total number of main section and off-ramp sensors), and M the maximum number of time intervals required by vehicles to traverse the entire freeway section considering a high congestion scenario; the value is 0 otherwise. Let's also define:

E	:	Matrix of row dimension I containing 0 for columns related to state variables in time intervals $k-1,\ldots,k-M$ and **B** for time interval k.
B	:	Matrix of dimension I x IJ defining equality constraints (sum to 1 in OD proportions for each entry) for state variable in time interval *k*.
F(k)	:	Matrix of dimensions (1+M)IJ x (1+M)IJ consisting on diagonal matrices $f(k),\ldots,f(k-M)$ containing input on-ramp volumes. This applies to each OD pair and time interval. Each $f(.)$ is a squared diagonal matrix of dimension IJ.
$\mathbf{g(k)}$	:	*Column vector* of OD flows of equipped vehicles for time intervals $k, k-1,\ldots,k-M$
A	:	Matrix of dimensions (J+P) x (1+M)(J+P) that adds up for a given sensor *q* (main section or off-ramp) traffic flows from any previous on-ramps arriving to sensor at interval *k* assuming their travel times are $t_{iq}(k)$

Defining the measurement equation as $z(k) = \begin{pmatrix} \mathbf{H(k)} \\ \mathbf{E} \end{pmatrix} \mathbf{b(k)} + \begin{pmatrix} v'(k) \\ 0 \end{pmatrix} = \mathbf{R(k)b(k)} + v(k)$ where $v'_{ij}(k)$'s are independent Gaussian white noise sequences with zero mean and covariance matrix $\mathbf{R'}$.

The Kalman-Filter algorithm for the dynamic estimation of OD matrices in Motorway is:

KF Algorithm	**:**	Let K be the total number of time intervals for estimation purposes and M maximum number of time intervals for the longest trip
Initialization	**:**	$\mathbf{b}_k^k = \mathbf{b}(0)$ k=0; Build constant matrices and vectors: e, **A, B, D, E, R, W** where each time interval and each row is set to the maximum indetermination proportion $1/J_i$

Prediction Step	:	$\mathbf{P_k^k} = \mathbf{V}[\mathbf{b(0)}]$ $\mathbf{b_{k+1}^k} = \mathbf{Db_k^k}$ $\mathbf{P_{k+1}^k} = \mathbf{DP_k^kD^T} + \mathbf{W}$
Kalman gain computation	:	Get observations of counts and travel times: $q(k+1),\quad s(k+1),\quad y(k+1),\quad t_{ij}(k+1)\quad t_{ip}(k+1)$. Build $z(k+1)$, $\mathbf{F(k+1)}$, $\mathbf{U(k+1)}$. Build $\mathbf{R_{k+1}} = \mathbf{R(k+1)}$. Compute $\mathbf{G_{k+1}} = \mathbf{P_{k+1}^k R_{k+1}^T}\left(\mathbf{R_{k+1} P_{k+1}^k R_{k+1}^T} + \mathbf{R}\right)^-$ (where $(.)^-$ denotes the pseudoinverse)
Filtering	:	Compute $\mathbf{d_{k+1}} = \mathbf{G_{k+1}}\left(z(k+1) - \mathbf{R_{k+1} b_{k+1}^k}\right)$ filter for state variables and errors $\boldsymbol{\varepsilon}_{\mathbf{k+1}} = \left(z(k+1) - \mathbf{R_{k+1} b_{k+1}^k}\right)$ Search maximum step length $0 \le \alpha \le 1$ such that $\mathbf{b_{k+1}^{k+1}} = \mathbf{b_{k+1}^k} + \alpha\, \mathbf{d_{k+1}} \ge 0$ $\mathbf{P_{k+1}^{k+1}} = \left(\mathbf{I} - \mathbf{G_{k+1} R_{k+1}}\right)\mathbf{P_{k+1}^k}$
Iteration	:	k=k+1 if k=K **EXIT** otherwise **GOTO Prediction Step**
Exit	:	*Print results*

In a benchmark conducted at a Toll Plaza of the Motorway Site the number of Bluetooth devices associated with vehicles was in average the 27.67%, that determined the significance of the sample used in the experiments. The increasing penetration of the technologiy guarantees larger samples in the near future. A pending task planned for the near future will be to determine the influence of the sample size in the accuracy of the results. A Set of computational experiments has been conducted with time sliced OD flows with time horizon split in four time intervals of 15 minutes and the demand accordingly distributed to account for the 15%, 25%, 35% and 25% of the total demand in each interval. The results can be summarized as follows: for time intervals where traffic flow varies from free flow to dense but not yet saturation conditions the filtering approach works as expected and its performance seems not affected as traffic flows become congested. RMSE values are of a similar order of magnitude, ranging in the interval [0.63, 6.35] (x10^{-2}). The convergence to the true values is quite satisfactory as prove the computational results in the full paper.

The second approach, intended for more general networks, has been based on the use of disaggregated flows, as in the case of the license plate recognition, [3]. The procedure works as follows: given a sample of equipped vehicles in the V2I scenario, their positions are tagged along their paths, there will be various classes, trips crossing the scenario at entry and exit tagged points as well as at intermediate positions, trips starting outside by a tagged entry point and ending inside, trips starting inside, leaving by a tagged exit point and ending a a destination outside the scenario, and finally trips starting and ending within the scenario, whose pats are tagged at intermediate points.

In all cases we have assumed that vehicles interact with the infrastructure at V2I sensors located in the border tagging all entries and exits and that there is a sensor layout in the network tagging the vehicles at intermediate points in their routes. From the point of view of the observability, defined in terms of identifying if a set of available measurements is sufficient to estimate the state of a system [2], in the first approach the detection layout has been set up in such way that intercepts flows for all OD pairs in the Motorway section, and therefore satisfies the observability conditions. While in the second case it is guaranteed by a suitable design of the sensor layout [7], determined by the network topology and the identification of the most likely used paths between origins and destinations. This layout allows to collect a sample of the OD matrix, for each time interval, that can be expanded to the whole population as a function of a initial OD matrix and the rate of penetration of the technology. Computational results will be included in the full version.

References

[1] M. Ben-Akiva, M. Bierlaire, D. Burton, H.N. Kotsopoulos, R. Mishalani, (2001), Network State Estimation and Prediction for Real-Time Traffic Management, Networks and Spatial Economics, 1, pp. 293-318.

[2] E. Castillo, A. J. Conejo J. M. Menéndez, P. Jiménez, The Observability Problem in Traffic Network Models, Computer-Aided Civil and Infrastructure Engineering **23** (2008) 208–222

[3] E. Castillo, J.M. Menéndez, P. Jiménez, (2008), Trip matrix and path flow reconstruction and estimation based on plate scanning and link observations, Transportation Research Part B 42 455–481

[4] Mahmassani, H.S., and Zhou, X., (2005), Transportation System Intelligence: Performance Measurement and Real-Time Traffic Estimation and Prediction in a Day-to-Day Learning Framework, Chapter 16 in Advances in Control, Communication Networks, and Transportation Systems, In Honor of Pravin Varaiya, edited by E. Abed, Birkhauser,

[5] Van Der Zijpp, N.J. and Hamerslag, R., (1994). An improved Kalman filtering approach to estimate Origin—Destination matrices for freeway corridors. In: *Transportation Research Records 1443*, National Research Council, pp. 54–64.

[6] Chang, G.L. and Wu, J., (1994). Recursive estimation of time-varying Origin—Destination flows from traffic counts in freeway corridors. Transportation Research B Volume 28B No.2 pp. 141-160

[7] Gillieron, F. and Barceló, (2008) J.' Optimization of Road Traffic Counts Location in a Network, Conference Proceedings (Paper 2728), ITS Europe 2008.