Fault-Adaptive Origin-Destination Matrix Estimation using the Cell Transmission Model

Yiolanda Englezou, Stelios Timotheou, Christos G. Panayiotou

KIOS Research and Innovation Center of Excellence and the Department of Electrical and Computer Engineering, University of Cyprus, Nicosia 20537, Cyprus
{englezou.yiolanda, timotheou.stelios, christosp}@ucy.ac.cy

Keywords: origin-destination matrix; state space model; sensor faults; fault identification; optimisation.

1. Introduction

The estimation of the origin destination matrix, referred to as O-D matrix, is a crucial requirement for transportation planning and traffic modelling and an essential input to traffic assignment and network equilibrium models (Cascetta and Nguyen, 1988). However, the O-D matrix cannot be observed directly hence it has to be efficiently estimated. The estimation of the O-D matrix of a particular network of interest is based on a set of observed counts on the links aiming to estimate the number of vehicles that traverse from link $i$ to $j$ for each pair of nodes $n_i$ and $n_j$ in the network.

There is considerable literature regarding the O-D matrix estimation problem, using a set of link counts and different proposed models. The O-D matrix estimation is typically divided in two problems with different formulations: static and dynamic (or time-varying). The static O-D matrix estimation problem is treated as a steady state situation (traffic flows are considered time independent) over a fixed period and an average O-D demand is estimated for a specific fixed time window $T$ (e.g. Tamin and Willumsen, 1989, Fisk, 1989, Cascetta and Nguyen, 1988, Bell, 1991, Li, 2005). The dynamic O-D matrix estimation problem is based on a sequence of traffic counts taken in small intervals (for example 10-15 minutes) and aims to estimate time varying O-D matrices (e.g. Maher, 1983, Airoldi and Blocker, 2013). Despite the extensive study of efficient O-D matrix estimation, there is no research work in our knowledge that assumes faulty traffic sensors, an assumption that in practise is true.

The aim of this work is the efficient estimation of O-D matrices for a fixed period in the presence of faulty sensors (static O-D matrix estimation), and differs from previous research in the area as follows.

- A path-based CTM (Englezou et al., 2019) results in a state space model where link densities (measurements) are associated with per path densities (state vector) and the path demand pattern. Using the state space model under free flow conditions we formulate a convex optimisation problem that aims to minimise the discrepancy between the model and the measurements and simultaneously identifies faulty sensors and their fault magnitude.
- An optimisation problem for fault-tolerant estimation is solved to identify the level of faulty behaviour of each sensor. Then each sensor is classified as faulty or non-faulty and a second optimisation problem is solved to improve the accuracy of the O-D matrix estimation based on the identification of faulty sensors.

2. Problem Formulation

We consider a mathematical representation of the physical system of interest consisting a set of input, output and state variables related by first-order difference equations. We assume traffic networks which remain in the free flow condition for $t = 1, \ldots, K$. As a result, equations describing the path-based CTM reduce to a linear state space model.

Assume the stochastic linear dynamic model

\begin{align*}
    x_{t+1} &= A_t x_t + B_t u_t + \epsilon_t, \\
    y_{t+1} &= H_t x_{t+1} + \omega_{t+1},
\end{align*}

(1)
where $x_t$ is the unobserved state vector, $y_t$ are the noisy observations, $u_t$ is the input vector and $x_0$ is the initial state of the process. Matrix $A_t$ describes the evolution in time of the unknown states and $H_t$ represents the matrix of explanatory variables. Both equations are subject to Gaussian error denoted by $e_t$ and $ω_t$, which are independent, both between them and within each of them. We assume that $e_t \sim N(0, Σ'_t)$ and $ω_t \sim N(0, Σ''_t)$, where $Σ'_t$ and $Σ''_t$ are the model and measurement matrices, respectively.

For the path-based CTM, $x_t$ denotes the density at each cell per path passing through the cell, and the density of each cell in the network at time $t$. The input vector $u_t$, denotes the path demand entering the network for each cell at time $t$, that can be fixed $u_p(t) = λ_p$ ($p \in P$ a path in the network with $p = 1, \ldots, Q$ and $P$ the set of all paths in the network under study) or assumed to follow a Poisson distribution with rate $λ_p$ such that $u_p(t) \sim \text{Pois}(λ_p)$. The main objective of this work is to estimate $λ = [λ_1, \ldots, λ_P]^T$ in the presence of faulty sensors. The demand for O-D pair $w \in W$ ($W$ is the set of all O-D pairs in the network under study), $d_w(t)$, is related to the path demand pattern, $λ_p(t)$ on each path $p \in S_w$ ($S_w$ the set of all paths connecting O-D pair $w$ and $P = \cup_{w \in W} S_w$) and time $t = 1, \ldots, K$, by $\sum_{p \in S_w} λ_p(t) = d_w(t), \forall w \in W$. We assume $λ_p(t)$ remains fixed for $t = 1, \ldots, K$ and hence drop subscript $t$. Estimating $λ$ is sufficient to obtain an estimate of the O-D matrix of interest.

In the presence of faulty sensors a more appropriate measurement model is

$$y_{t+1} = H_{t+1}x_{t+1} + ω_{t+1} + o_t,$$

where $o_t$ denotes the sensor fault residuals, with $o_{i,t}$ having a zero value if sensor $i$ is not faulty at time $t$ and non-zero otherwise (Timotheou et al., 2015).

3. Solution Approach

We propose a three-stage algorithm to provide the best of both worlds (healthy versus faulty operation) by detecting faulty sensors and compensating only for their respective identified fault residuals. This is outlined as follows.

1. The first stage of the algorithm involves the estimation of $x_t$, $o_t$, and $u_t$ using a fault-tolerant least squares formulation of (1) and (2), following Timotheou et al. (2015). The resulting problem is a convex optimisation problem that can be solved to optimality.

2. The second stage, uses measurements, $y_1, \ldots, y_{t-1}$, and maximum fault residuals $o_1, \ldots, o_{t-1}$ to decide which are the healthy and which are the faulty sensors. This is achieved by separating the considered $y_{i,t-1}$ values into two clusters, identifying whether the clusters belong to the same or different data class (healthy or faulty) and classifying the current time values $y_{i,t}$ appropriately, using a clustering technique, e.g. $K$-Means.

3. In the third stage, the faulty sensor measurements are compensated according to the $o_{i,t}$ values obtained in the first stage, and the path demand pattern estimation problem is resolved, with no fault-tolerance consideration.

The fault-tolerant O-D matrix estimation approach (2) has been demonstrated to jointly provide robust estimation of $u_t$, very close to the non-faulty formulation (1), and also successful detection, isolation, and identification of sensory faults in the context of the path-based cell transmission model. Building on this, the proposed algorithm will adaptively compensate faults leading to better estimation performance irrespective of the fault magnitudes.

Acknowledgements

This work has been supported by the European Union’s Horizon 2020 research and innovation programme under grant agreements No 739551 (Project: KIOS CoE) and No 101003435 (Project: MSCA-WF-02-19/BITS), the Government of the Republic of Cyprus through the Directorate General for European Programmes, Coordination and Development and through the Research Promotion Foundation (Project: CULTURE/BR-NE/0517/14).

References


