Enhancing Origin-Destination Matrix Estimation Using Measurements from Unmanned Aerial Vehicles

Yiolanda Englezou, Stelios Timotheou and Christos Panayiotou

September 23, 2024

Enhancing Origin-Destination Matrix estimation using measurements from Unmanned Aerial Vehicles

Y. Englezou[a*], S. Timotheou[a] and C. G. Panayiotou[a]

[a] KIOS Research and Innovation Center of Excellence & the Department of Electrical and Computer Engineering, University of Cyprus, Nicosia, Cyprus

englezou.yiolanda@ucy.ac.cy, timotheou.stelios@ucy.ac.cy, christosp@ucy.ac.cy

[*] Corresponding author

Extended abstract submitted for presentation at the Conference in Emerging Technologies in Transportation Systems (TRC-30)

September 02-03, 2024, Crete, Greece

August 9, 2024

Keywords: origin-destination demand, path demand, cell transmission model, traffic flow dynamics, optimisation.

1 INTRODUCTION

Efficiently estimating the origin-destination (OD) matrix in transportation is essential. Two key variants of this problem are considered: the *static* and the *dynamic* (or time-varying). The static variant seeks average OD demands over a fixed period using aggregated link counts collected over a longer duration. In contrast, the dynamic variant utilizes time-varying traffic data to estimate time-dependent OD matrices.

Stationary sensors, like loop detectors and cameras, are commonly used to gather traffic flow data at monitored links and provide information for inference at unmonitored links. However, their effectiveness is hindered by sparse deployment, high installation costs (Darwish & Bakar, 2015) and the inability to capture traffic data on non-equipped links (Coifman, 2014). In contrast, mobile sensors offer broader spatiotemporal coverage but come with drawbacks such as sparse data, unpredictable routes, difficulties in correlating with traffic volumes (Kurzhanskiy & Varaiya, 2015), and privacy concerns (Llorca *et al.*, 2010).

Unmanned Aerial Vehicles (UAVs) is a promising technology for various transportation applications. Research on traffic monitoring with UAVs focuses on remote capture of traffic data from above the ground using multiple sensors (Barmpounakis *et al.*, 2016). This collected data supports tasks such as surveillance, congestion management, traffic signal optimization, highway infrastructure management, and traffic parameter inference (Pham *et al.*, 2020, Barmpounakis & Geroliminis, 2020).

In this work, we aim to estimate a static OD matrix over a specified time period using density data gathered from either stationary sensors or a swarm of UAVs. We propose efficient methodologies to address both free-flow and congested traffic conditions. Our approach integrates the path-based cell transmission model (CTM) within an optimization framework for OD estimation, leveraging disaggregated measurements to minimize discrepancies between the model and actual data. We specifically explore the use of UAV measurements for OD matrix estimation, enabling comprehensive data collection across all network links at different time intervals.

2 METHODOLOGY

2.1 Proposed OD matrix estimation problem

We split the time period of interest T into K time-steps of duration T_s [hours], and define $\mathcal{T}^+ = \{1, \ldots, K\}$ and $\mathcal{T} = \{0, \ldots, K-1\}$, such that $K = T/T_s$. T_s is used to characterize both the periodicity of measurements and the discrete time-step of the chosen traffic model. For the evolution of traffic, we consider a state-space model consisting of a nonlinear traffic and an observation model described by

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{f}(\mathbf{x}_k, \mathbf{u}) + \boldsymbol{\epsilon}_k, \\ \mathbf{y}_{k+1} &= \mathbf{H}_k \mathbf{x}_{k+1} + \boldsymbol{\omega}_k, \quad k \in \mathcal{T}, \end{aligned} \tag{1}$$

where $\mathbf{x}_k \in \mathbb{R}^{M\times 1}$ is the unobserved state vector, $\mathbf{y}_k \in \mathbb{R}^{C\times 1}$ is the measurement vector, $\mathbf{u} \in \mathbb{R}^{Q\times 1}$ is the input vector and $\mathbf{x}_0$ is the unknown initial state of the process. Vector $\mathbf{y}_k$ denotes observed traffic variables. In addition, $\mathbf{f}(\cdot,\cdot)$ is a nonlinear mapping describing the traffic dynamics, e.g. as part of the cell transmission model, and $\mathbf{H}_k \in \mathbb{R}^{C\times M}$ is the observation matrix. We assume independent Gaussian errors for the model and measurement equations, $\boldsymbol{\epsilon}_k \sim N(\mathbf{0}_{(M+L)\times 1}, \boldsymbol{\Sigma}_{\epsilon,k})$ and $\boldsymbol{\omega}_k \sim N(\mathbf{0}_{C_k\times 1}, \boldsymbol{\Sigma}_{\omega,k})$, and $\boldsymbol{\Sigma}_{\epsilon,k} \in \mathbb{R}^{(M+L)\times(M+L)}$ and $\boldsymbol{\Sigma}_{\omega,k} \in \mathbb{R}^{C_t\times(M+L)}$ are the model and measurement error matrices, respectively.

Given Model (1) and the obtained measurements, we aim to select $\mathbf{u}$, such that the states of the model are as close as possible to the measurements. This can be formulated as a nonconvex constrained least squares problem of the form

$$\begin{aligned} \min \sum_{k\in\mathcal{T}} &||\mathbf{y}_{k+1} - \mathbf{H}_{k+1}\mathbf{x}_{k+1}||^2_{[\boldsymbol{\Sigma}_{\omega,k}]^{-1}} \\ \text{s.t. } &\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k,\mathbf{u}) + \boldsymbol{\epsilon}_k,\ k\in\mathcal{T} \\ &\mathbf{x}_k \geq \mathbf{0},\ k\in\mathcal{T},\ \mathbf{u}\geq \mathbf{0}, \end{aligned} \tag{2}$$

where $\mathbf{x}_k$ and $\mathbf{u}$ are the optimisation variables. By obtaining $\mathbf{u}$ through the solution of Problem (2), the demand of each OD pair can be simply obtained by

$$\sum_{p\in\mathcal{S}_w} u_p = d_w,\ \ \forall w\in\mathcal{W}. \tag{3}$$

We assume two types of data may be available from two different sensing technologies; data collected from fixed location sensors (stationary data) and data collected by UAVs. The main difference is that when only fixed location sensor data are collected the set of measured links remains fixed $\forall k\in\mathcal{T}$, $\mathcal{C}_k = \mathcal{C}$ and $|\mathcal{C}_k| = |\mathcal{C}|$ with observed traffic counts $y_i(k)$, $i\in\mathcal{C}$, whereas if only UAV data are collected the set of measured links, $\mathcal{C}_k$, differs at each time-step, as well as the observed traffic counts, $y_i(k)$, with $i\in\mathcal{C}_k$, $k\in\mathcal{T}$.

Consider U UAVs flying in the network under study, with $\mathcal{U}$ the set of all UAVs such that $U = |\mathcal{U}|$. For simplicity we assume that all UAVs have identical flying and sensing capabilities. Each UAV traverses between neighbouring nodes with constant speed v^{UAV} [m/s] and the required time to traverse between neighbouring nodes is given by $T^U = D_{nn'}/v^{UAV}$ [s], where $D_{nn'}$ denotes the distance between nodes $n\in\mathcal{N}$ and $n'\in\mathcal{N}$. In addition assume that each UAV records measurements only for the hovering time, t^H [s], above a node, whereas no measurements are recorded during T^U.

2.2 Macroscopic Traffic Model

A path-based CTM allows us to keep track of path-based flows and cell densities and any strong assumptions on split ratios used in previous works in the OD matrix estimation research, are no longer needed. The use of the path-based CTM enables the estimation of the path demands and in consequence the estimation of OD demands, $d_w(k)$ through Equation (3).

The per path density of cell i, is:

$$\rho_{i,p}(k+1) = \rho_{i,p}(k) + \frac{T_s}{l_i}\left[\varphi^{in}_{i,p}(k) - h_i(k)\varphi^{out}_{i,p}(k)\right], \tag{4}$$

$\forall p\in\mathcal{P}_i, \forall i\in\mathcal{L}$. It follows from the definitions that if $p\notin\mathcal{P}_i$ then $\rho_{i,p}(k) = 0$, $\forall k\in\mathcal{T}^+$. Based on the characterisation of the boundary connection of the cells we can define the inflow and outflow, which are calculated according to the demand and supply paradigm. For further information regarding the path-based CTM, see Englezou *et al.* (2024).

2.3 Problem-specific dynamics

For the path-based CTM we have the state vector $\mathbf{x}_k = [\boldsymbol{\rho}(k), \bar{\boldsymbol{\rho}}(k)]^{\mathrm{T}}$, such that $\boldsymbol{\rho}(k) = [\rho_1(k),\ldots,\rho_M(k)]^{\mathrm{T}}$ is the per path density of each cell $i\in\mathcal{R}$, and every path $p\in\mathcal{P}_i$ at time-step k. In addition, $\bar{\boldsymbol{\rho}}(k) = [\bar{\rho}_1(k),\ldots,\bar{\rho}_L(k)]^{\mathrm{T}}$ is the density of each cell i in the network at time-step k. The input vector $\mathbf{u} = [u_1,\ldots,u_Q]^{\mathrm{T}}$, denotes the inflow from path $p\in\mathcal{P}_i$ in cell i and time-step k (Englezou *et al.*, 2024). The C_k-vector of observations $\mathbf{y}_k = [\bar{\rho}_{b_1}(k),\ldots,\bar{\rho}_{b_{C_t}}(k)]^{\mathrm{T}}$, $C_k \leq L$, $\mathcal{C}_k = \{b_1,\ldots,b_{C_k}\}$, denotes the density taken on measured links at time-step k, related to the state vector through $\mathbf{H}_k$.

Practically the difference in the proposed methodology when using the two different types of data is reflected in the definition of the matrix of explanatory variables $\mathbf{H}_k$. When static data

Percentage of measurements	100%	80%	60%	40%
Measured Links (C)	19	15	11	7
pCTM-OD	6.62	7.95	14.70	1.47×10^4

Table 1 – *Average RMSE for different measurement percentage using fixed location sensor data.*

t^H(s)	$T_s = 10$	$2T_s$	$3T_s$	$4T_s$	$5T_s$	$6T_s$
Scenario S1	7.26	7.71	8.16	9.01	9.23	9.32
Scenario S2	8.55	8.79	9.13	9.26	9.35	9.56
Scenario S3	7.12	7.21	7.52	8.31	8.53	8.92
Scenario S4	8.10	8.23	8.80	9.13	9.45	9.87
Scenario S5	6.74	7.32	7.35	7.39	7.38	7.69
Scenario S6	7.61	7.87	8.29	8.83	9.04	9.21

Table 2 – *Average RMSE for the six UAV scenarios using UAV-based data.*

are used to estimate the OD matrix, we have that $C_k = C\ \forall k \in \mathcal{T}$ and the $C \times (M + L)$ matrix $\mathbf{H}_k$ remains fixed for all $k \in \mathcal{T}$, such that $\mathbf{H} = \mathbf{H}_k$. However when data collected by UAVs are used for OD matrix estimation, the $C^k \times (M + L)$ matrix $\mathbf{H}_k$ differs both in its elements and its size at each time-step k. As a consequence, Problem (2) differs in terms of the objective function to be minimised. We will refer to the proposed OD matrix estimation approach as pCTM-OD.

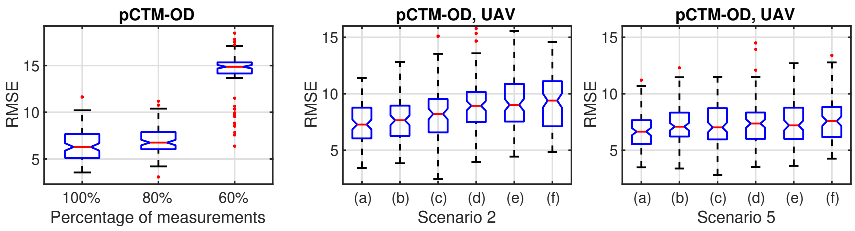
3 RESULTS

We consider an abstraction of the road network in the city of Leicester, UK which lies to the south-east of the center of Leicester and is bounded by Waterloo Way, Lancaster Road, Granville Road and London Road. The parameters considered are: $W = 9$ OD pairs, $Q = 10$ pre-defined paths and $L = 19$ directed links. In addition, we consider the following CTM parameter values: $v^f = 60$ km/h, $T_s = 10$ s, $w = 20$ km/h, $\varphi^{max} = 1500$ veh/h, $\rho^{max} = 200$ veh/km; the simulation time is consider to be $T = 1$ hour and each cell has length $l_i = 0.5$ km.

We consider: (i) *Random partial coverage* where loop detectors are randomly placed at specific links to obtain 80%, 60% and 40% coverage of the total number of links, and (ii) *Full coverage* or 100% coverage where all links are measured. In addition we consider that UAVs fly above the road network with speed $v^{\mathrm{UAV}} = 25$ m/s. For simplicity we assume that the travel time of the UAV to move to neighbouring nodes is constant and equal to $T^U = 20$ s. Each UAV has full vision of the links observed at each time-step. The following scenarios of deployed UAVs are considered: [S1.] $U = 1$ with fixed trajectory, [S2.] $U = 1$ with random trajectory, [S3.] $U = 2$ with fixed trajectories, [S4.] $U = 2$ with random trajectories, [S5.] $U = 3$ with fixed trajectories and [S6.] $U = 3$ with random trajectories. We consider 6 hovering times of each individual UAV: (a) $t^H = T_s = 10$ s, (b) $t^H = 2T_s$, (c) $t^H = 3T_s$, (d) $t^H = 4T_s$, (e) $t^H = 5T_s$, (f) $t^H = 6T_s$.

We estimate $\mathbf{u}$ for the different loop detector scenarios and the six scenarios of deployed UAVs described above, and calculate the root mean squared error: $\mathrm{RMSE} = \sqrt{1/Q \sum_{p=1}^{Q} (u_p - \hat{u}_p)^2}$. Here, u_p is the true value of the pth unknown path demand, $\hat{u}_p$ is the pth estimated path demand using the output of the optimisation method and Q is the total number of paths in the network under study. Once the estimated path demand $\hat{\mathbf{u}}$ is obtained the OD matrix can then be easily derived using Equation (3). We run the optimisation procedure $\tilde{B} = 100$ times and present the average RMSE for each loop detector layout (Table 1) and UAV scenario (Table 2).

As shown in Table 1 decreasing the percentage of the measurements, i.e decreasing the measured links from $C = 19$ to $C = 7$, the average RMSE is increased on average three orders of magnitude. As shown in Table 2 the estimation results remain very similar for all UAV deployment scenarios. Scenario 5 (a) appears to have the lower average RMSE and Scenario 4 (f) yields the higher average RMSE. As we increase t^H the average RMSE is increased probably due to the fact that when observing the same links for a larger time duration the UAV is not able to visit all nodes multiple times in the total time horizon under study affecting the estimation results. In addition as we increase the number of UAVs that hover above the road network the ARMSE is decreased. Furthermore, assuming random trajectories for the UAVs results in higher average RMSE compared to the results assuming fixed trajectories. A further investigation of



(a) (b) (c)

Figure 1 – *Boxplots of the residuals of the pCTM-OD using (a) static data, and (b) and (c) UAV-based measurements for Scenario 2 and 5, respectively.*

the RMSE obtained using pCTM-OD for different sensor layouts and deployed UAVs scenarios is presented in Figure 1, in order to evaluate the estimation variance of each method. We select to present results for two UAV scenarios; Scenario 2 and 5 which result in the worst and best estimation results, respectively, compared to all six scenarios. The RMSE obtained using the different sensor layout is rapidly increased as we decrease the number of sensors, as also shown in Table 1 (we omit results obtained using 40% percentage of measurements as they exhibit very large RMSE). In addition, for 100% measurements the variance of the RMSE is lower than the results obtained using UAV measurements. As we increase the number of deployed UAVs the mean and variance of the RMSE for the 100 different estimation procedures is decreased, as also shown in Table 2.

4 DISCUSSION

This work investigates OD matrix estimation in the context of the path-based CTM and noisy measurements obtained from (i) fixed location sensors and (ii) UAVs flying above the road network. Towards this direction a model-based estimation approach has been formulated and has been demonstrated the estimation advantage of using measurements obtained from UAVs compared to measurements collected by fixed location sensors layouts. Estimation results show that UAV data provide significantly better results even when the number of measurements per time-step with the UAV swarm is smaller.

Acknowledgments. This work is supported by the European Union (i. ERC, URANUS, No. 101088124 and, ii. Horizon 2020 Teaming, KIOS CoE, No. 739551), and the Government of the Republic of Cyprus through the Deputy Ministry of Research, Innovation, and Digital Strategy. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

References

Barmpounakis, E., & Geroliminis, N. 2020. On the new era of urban traffic monitoring with massive drone data: the pNEUMA large-scale field experiment. *Transportation Research Part C*, 50–71.

Barmpounakis, E. N., Vlahogianni, E. I., & Golias, J. C. 2016. Unmanned Aerial Aircraft Systems for transportation engineering: current practise and future challenges. *International Journal of Transportation Science and Technology*, **5**, 111–122.

Coifman, B. 2014. Revisiting the empirical fundamental relationship. *Transportation Research Part B*, **68**, 173–184.

Darwish, T., & Bakar, K. A. 2015. Trafc density estimation in vehicular ad hoc networks: a review. *Ad Hoc Networks*, 337–351.

Englezou, Y., Timotheou, S., & Panayiotou, C. G. 2024. Path-Based Origin-Destination Matrix Estimation Utilizing Macroscopic Traffic Dynamics. *IEEE Transactions on Intelligent Transportation Systems*.

Kurzhanskiy, A. A., & Varaiya, P. 2015. Traffic management: An outlook. *Economics of Transportation*, **4**, 135–146.

Llorca, D. F., Satelo, M. A., Sánchez, S., na, M. Oca Rodríguez-Ascariz, J. M., & Garcìa-Garrido, M. A. 2010. Traffic Data Collection for Floating Car Data Enhancement in V2I Networks. *EURASIP Journal of Advanced Signal Processing*, 425–464.

Pham, H. Q., Camey, M., Pham, K. D., Pham, K. V., & Rirett, L. R. 2020. *GIGOS 2019, Innovation for Sustainable Infrastructure*. Springer. Chap. Review of Unmanned Aerial Vehicles (UAVs) Operation and Data collection for driving behaviour analysis, pages 1111–1116.