Estimating Link-Dependent Origin-Destination Matrices from Sample Trajectories and Traffic Counts

Gabriel MICHAU[1,2], Pierre BORGNAT[1], Nelly PUSTELNIK[1], Patrice ABRY[1], Alfredo NANTES[2], Edward CHUNG[2]

[1]Laboratoire de Physique, CNRS, ENS de Lyon, Univ. de Lyon, France

[2]Smart Transport Research Centre, Queensland University of Technology, Brisbane, Australia

{prenom.nom}@ens-lyon.fr, a.nantes@qut.edu.au, edward.chung@qut.edu.au

Résumé – L'estimation des matrices Origines-Destination connait un renouveau par l'apparition de nouvelles technologies donnant accès à un échantillon de trajectoires. Ce travail se concentre sur le cas des réseaux où les décomptes de trafics sont mesurés par des boucles magnétiques et un échantillon de trajectoires est disponible, comme par exemple la ville de Brisbane où des détecteurs Bluetooth ont été installés. Cette source d'information additionnelle permet l'extension des méthodes tradionnelles d'estimation des matrices OD aux matrices OD par liens (LODM). Nous utilisons pour cela un algorithme d'optimisation convexe dont nous testons la validité sur un réseau simulé.

Abstract – Origin-Destination matrices (ODM) estimation can benefits of the availability of sample trajectories which can be measured thanks to recent technologies. This paper focus on the case of transport networks where traffic counts are measured by magnetic loops and sample trajectories available. An example of such network is the city of Brisbane, where Bluetooth detectors are now operating. This additional data source is used to extend the classical ODM estimation to a link-specific ODM (LODM) one using a convex optimisation resolution that incorporates networks constraints as well. The proposed algorithm is assessed on a simulated network.

1 Introduction

Link-dependent Origin-Destination matrices (LODM) is an extension of the Origin-Destination Matrice (ODM) concept, which is one of the mostly widespread tool used to represent demand on networks such as for Internet traffics [1, 2, 3] or transport networks [4]. LODM gives at the same time an insight of the traffic assignment and of the traffic demand. An operating example is provided by the Bluetooth scanners operated in Brisbane by the Brisbane City Council [5], whose study constitute our long-term target. These data are composed with probe trajectories (obtained from bluetooth cars) and with traffic count data at each link. The challenge is to estimate the trajectories for the whole traffic, not only the cars with bluetooth, *i.e.,* to estimate the LODM. Thus, the objective of the present work is to propose a method for recovering the LODM, by combining probe trajectories and the traffic count data.

Related Work. The generic challenge of estimating LODM from traffic counts and probe trajectories, using convex program, has been developed recently in [6] for Internet data and in [7] for car traffic. While the work in [6] has many common points with our problem (availability of link counts, sample of OD flows) it assumes the knowledge of the *Routing Matrix*, assigning traffic flows to the set of links.

Objectives. This paper improves the method proposed in [7], studying how the estimation of LODM (and ODM as a by-product) can be formulated as an inference problem on the transport networks, and then be solved by convex optimization techniques [8, 9]. Further development of the constraints induced by the topology of the transport network and their impact on the estimation process will be presented. After introducing notations, the direct problem is defined in Section 2 and the estimation as an inverse problem is formulated in Section 3. In Section 4, the proposed LODM estimation performance are explored using traffic simulations.

Notations: $\underline{A}$, $\underline{\underline{A}}$ and $\underline{\underline{\underline{A}}}$ respectively refer to vectors, matrices and tensors. The Hadamard product (element-wise product) of $\underline{\underline{\underline{A}}}$ and $\underline{\underline{\underline{B}}}$ is denoted $\underline{\underline{\underline{A}}} \circ \underline{\underline{\underline{B}}}$. For LODM, the two first indices are labelled i (origins) and j (destinations); the third index l stands for the links in the network. The symbol • is used to denote the dimension that does not contribute to a sum: e.g., the sum over first and third dimensions is written $\sum_{i,\bullet,l} \underline{\underline{\underline{A}}}_{ijl}$.

2 Estimating Link-Dependent ODM

2.1 Problem Presentation

The road network is represented as a graph $\mathcal{G} = (V, L)$. The finite set of nodes V models the major intersections of the road network; each node is also a possible origin or destination. L is the set of directed edges, each corresponding to a direct itinerary (or road) linking two nodes (*i.e.,* not going through another node in V). The structure of the graph is then given by two matrices $\underline{\underline{I}}$ and $\underline{\underline{E}}$ called the *incidence* and *excidence* matrices respectively. These matrices describe the relations between the nodes and the edges: for every $(v, l) \in$

$\{1,\dots,|V|\}\times\{1,\dots,|L|\}$,

$$\underline{\underline{I}}_{vl} \;=\begin{cases} +1 & \text{if the edge } l \text{ is arriving to the node } v,\\ 0 & \text{otherwise,}\end{cases}$$

$$\underline{\underline{E}}_{vl} \;=\begin{cases} +1 & \text{if the edge } l \text{ is starting from the node } v,\\ 0 & \text{otherwise.}\end{cases}$$

Note that in graph theory, it is the difference $\underline{\underline{I}}-\underline{\underline{E}}$ that would be named as "incidence matrix".

The measurements assumed available on this graph are $\underline{\underline{\underline{B}}}$ and $\underline{q}$. The tensor $\underline{\underline{\underline{B}}}$, of dimension $|V|^2\times|L|$ gathers information from Bluetooth trajectories. Each trajectory adds a count of 1 into the elements of $\underline{\underline{\underline{B}}}$, denoted B_{ijl} corresponding to the origin (i), the destination (j) and the links (l) (*i.e.,* edges in $\mathcal{G}$). The vector $\underline{q}$ of dimension $|L|$, is the traffic flow measured on each edge l. These counts can be obtained by magnetic loops. Such measurements are subject to count errors modelled here by a noise $\underline{\epsilon}$.

The quantity to be recovered is the count of trajectories for all cars over the OD and links, denoted by the tensor $\underline{\underline{\underline{Q}}}$. To achieve this estimation, a variational approach has been presented in [7] that consisted in solving

$$\widehat{\underline{\underline{\underline{\alpha}}}} \in \underset{\underline{\underline{\underline{\alpha}}}}{\operatorname{Argmin}} \sum_{k=1}^{K} G_k(\underline{\underline{\underline{\alpha}}}) \tag{1}$$

where the functions $G_k\colon \mathbb{R}^{|V|\times|V|\times|L|}\to\,]-\infty,+\infty]$, for every $k\in\{1,\dots,K\}$, model several network properties and $\underline{\underline{\underline{\alpha}}}$ satisfies $\underline{\underline{\underline{\alpha}}}\circ\underline{\underline{\underline{B}}}=\underline{\underline{\underline{Q}}}$. The present contribution proposes improvements to this approach. The first improvement is to question the relevance of using the variable $\underline{\underline{\underline{\alpha}}}$ for the optimisation process; through minor changes, we can estimate the variable $\underline{\underline{\underline{Q}}}$ directly:

$$\widehat{\underline{\underline{\underline{Q}}}} \in \underset{\underline{\underline{\underline{Q}}}}{\operatorname{Argmin}} \sum_{k=1}^{K} F_k(\underline{\underline{\underline{Q}}}). \tag{2}$$

where the functions $F_k\colon \mathbb{R}^{|V|\times|V|\times|L|}\to\,]-\infty,+\infty]$, for every $k\in\{1,\dots,K\}$, model $\underline{\underline{\underline{Q}}}$ and several network properties.

This change is useful for two reasons: First, it enables the estimation of the elements of $\underline{\underline{\underline{Q}}}$ where the corresponding element in $\underline{\underline{\underline{B}}}$ is zeros. This, accounts for the possibility of a trajectory not being represented by a Bluetooth sample. Second, it gives the possibility of assuming that the variables $\underline{\underline{\underline{Q}}}$ and $\underline{\underline{\underline{B}}}$ are related by a Poisson distribution of type $\underline{\underline{\underline{B}}}\sim\mathcal{P}\left(\eta\underline{\underline{\underline{Q}}}\right)$ where η can be related to the Bluetooth penetration rate.

The second axis of improvement concerns the choice of the regularization terms within the objective function for a better modeling of the network properties. In the next section, we describe specifically the choices done for the functions $(F_k)_{1\le k\le K}$.

2.2 Criterion to minimize

Bluetooth to total flows. The first objective is to relate the known Bluetooth flows ($\underline{\underline{\underline{B}}}$) with the total flows ($\underline{\underline{\underline{Q}}}$). To do so we assume a Poisson noise, typically involved in counting processes. This Poisson noise is characterised by the scale parameter $\underline{\eta}$ that is linked to the proportion of Bluetooth vehicles on each road of the network. To ensure consistency between the estimated $\underline{\underline{\underline{Q}}}$ and a Poisson model, the function F_1 models the minus log-likelihood associated with the Poisson model, that is also known as the Kullback-Leibler divergence [10],

$$F_1(\underline{\underline{\underline{Q}}}) = \sum_{ijl} \psi_{\text{DKL}}\left(B_{ijl},\eta_l Q_{ijl}\right) \tag{3}$$

where, for every $(u,v)\in\mathbb{R}^2$, $\eta>0$,

$$\psi_{\text{DKL}}(u,\eta v)=\begin{cases} -u\log v+\eta v & \text{if } \; v>0 \;\text{ and }\; u>0,\\ \eta v & \text{if } \; v\ge 0 \;\text{ and }\; u=0,\\ +\infty & \text{otherwise.}\end{cases}$$

For this work, $\underline{\eta}=\underline{q}/\sum_{i,j,\bullet}\underline{\underline{\underline{B}}}$

Traffic Counts and total flows: The variable $\underline{\underline{\underline{Q}}}$ is linked to the known traffic counts on links $\underline{q}$ by the relationship:

$$\underline{q}=\sum_{i,j,\bullet}\underline{\underline{\underline{Q}}}+\underline{\epsilon} \tag{4}$$

where $\underline{\epsilon}$ represents the error on counts measured by magnetic loops. Accordingly to (4), a usual choice for a function ensuring the consistency of the solution on the edges is:

$$F_2(\underline{\underline{\underline{Q}}}) = \|\underline{q}-\sum_{i,j,\bullet}\underline{\underline{\underline{Q}}}\|^2.$$

Flow continuity at the nodes: An additional constraint comes from the balance of the flows on each node. It can be written using the classical ODM, denoted $\underline{\underline{T}}$:

$$\underline{\underline{T}}=\sum_{\bullet,\bullet,l}\underline{\underline{\underline{I}}}\circ\underline{\underline{\underline{Q}}}=\sum_{\bullet,\bullet,l}\underline{\underline{\underline{E}}}\circ\underline{\underline{\underline{Q}}} \tag{5}$$

where $\underline{\underline{\underline{I}}}$ and $\underline{\underline{\underline{E}}}$ are respectively the $|V|$-replication of the previous incidence and excidence matrices, defined as follows:

$$(\forall k\in V)\quad \underline{\underline{\underline{I}}}_{kjl}=\underline{\underline{I}}_{jl}\quad\text{and}\quad \underline{\underline{\underline{E}}}_{ikl}=\underline{\underline{E}}_{il} \tag{6}$$

The balance requires that, at every node n, the flow having for destination n, computed as $\underline{D}_n=\sum_{i,\bullet}T_{i,n}$, minus the flow originating from n, computed as $\underline{O}_n=\sum_{\bullet,j}T_{n,j}$, should equal the flow going through the node n. This is written as:

$$\underline{D}_n-\underline{O}_n=\sum_{\bullet,l}(\underline{\underline{I}}-\underline{\underline{E}})_{n,l}\sum_{i,j,\bullet}\underline{\underline{\underline{Q}}} \tag{7}$$

Using variable $\underline{\underline{\underline{Q}}}$ with eq. (5), it reads as

$$\sum_{i,\bullet,l}\underline{\underline{\underline{I}}}\circ\underline{\underline{\underline{Q}}}-\sum_{\bullet,j,l}\underline{\underline{\underline{E}}}\circ\underline{\underline{\underline{Q}}}=(\underline{\underline{I}}-\underline{\underline{E}})\sum_{i,j,\bullet}\underline{\underline{\underline{Q}}} \tag{8}$$

The function resulting from this constraint is

$$F_3(\underline{\underline{\underline{Q}}})=\|\sum_{i,\bullet,l}\underline{\underline{\underline{I}}}\circ\underline{\underline{\underline{Q}}}-\sum_{\bullet,j,l}\underline{\underline{\underline{E}}}\circ\underline{\underline{\underline{Q}}}-(\underline{\underline{I}}-\underline{\underline{E}})\sum_{i,j,\bullet}\underline{\underline{\underline{Q}}}\|^2.$$

Total Flow domain of definition: As the total flow is at least greater or equal to the flow of Bluetooth enabled vehicles, it is further imposed that $\underline{\underline{\underline{Q}}}$ belongs to the following convex constraint set:

$$C = \{\underline{\underline{\underline{Q}}} = \big(Q_{ijl}\big)_{(ijl)\in V\times V\times L} \in \mathbb{R}^{|V|\times|V|\times|L|} \mid Q_{ijl} \geq B_{ijl}\}$$

In the criterion, this constraint appears through an indicator function $F_4(\underline{\underline{\underline{Q}}}) = \iota_C(\underline{\underline{\underline{Q}}})$, equals to 0 if $\underline{\underline{\underline{Q}}} \in C$ and $+\infty$ otherwise.

3 Algorithm

The criterion to obtain a relevant *transport* solution, using the topology of the networks and the data available, then reads:

$$\widehat{\underline{\underline{\underline{Q}}}} \in \operatorname*{Argmin}_{\underline{\underline{\underline{Q}}}} F_1(\underline{\underline{\underline{Q}}}) + \gamma F_2(\underline{\underline{\underline{Q}}}) + \mu F_3(\underline{\underline{\underline{Q}}}) + F_4(\underline{\underline{\underline{Q}}}) \qquad (9)$$

with $\gamma, \mu \geq 0$, the weight of each constraint.

The functions involved in criterion (9) are convex, lower semi-continuous and proper. Moreover, $\gamma F_2 + \mu F_3$ is differentiable with a β-Lipschitz gradient where the value of β depends on the norm of the matrices involved in each function. According to [11, Proposition 2.2.], the function $F_1 + F_4$ is non-differentiable but it has a closed form expression for its proximity operator that is $P_C \circ \operatorname{prox}_{F_1}$ where the closed form expression of $\operatorname{prox}_{F_1}$ is given in [10, Equation (41)] and $P_C = max(\cdot, \underline{\underline{\underline{B}}})$. To find $\widehat{\underline{\underline{\underline{Q}}}}$, we employ the forward-backward algorithm, adapted from [12], described as follow:

Algorithm 1 Forward-backward algorithm

```
Set $\tau = 1.99\beta^{-1}$
For $n = 0, 1, \ldots$ until convergence
    $\underline{\underline{\underline{Q}}}^{[n+\frac{1}{2}]} = \underline{\underline{\underline{Q}}}^{[n]} - \tau\left(\gamma\nabla F_2 + \mu\nabla F_3\right)\left(\underline{\underline{\underline{Q}}}^{[n]}\right)$
    $\underline{\underline{\underline{Q}}}^{[n+1]} = \max\left\{\operatorname{prox}_{\tau F_1}\left(\underline{\underline{\underline{Q}}}^{[n+\frac{1}{2}]}\right), \underline{\underline{\underline{B}}}\right\}$
```

The initialization of $\underline{\underline{\underline{Q}}}^{[0]}$ is set to zero. According to [13], the sequence $(\underline{\underline{\underline{Q}}}^{[n]})_{n\in\mathbb{N}}$ converges to $\widehat{Q}$. In practice, we consider the convergence is achieved when the relative error between two iterates is such that $\frac{\|\underline{\underline{\underline{Q}}}^{[n]} - \underline{\underline{\underline{Q}}}^{[n-1]}\|^2}{\|\underline{\underline{\underline{Q}}}^{[n]}\|^2} \leq 10^{-6}$.

4 Simulation and Results

4.1 Simulation setting

A simulation is developed to produce ground truth data. First, a schematic road network is built by locating a set of nodes randomly on a grid. The nodes are first linked by a minimum spanning tree (computed by the Kruskal's algorithm [14]). Then, links are randomly added to connect the nodes with lower degree (sum of in and out edges) provided that the added links do not cross or repeat an existing one. This is stopped when the average total degree becomes 6 per node, a value consistent with that of real road networks (notably Brisbane transport network). In practice for this simulation, the number of nodes is $|V| = 50$. This choice is driven by the tractability of the experiment and the possibility of testing varied setups easily.

Then trajectories are drawn with random origin and destination with uniform law, and use the shortest path connecting the two. Their number is set proportional to the number of links (and thus to the number of nodes).

For future comparison to Brisbane's network, the number of vehicles is set to 500 per links. Measures show that few hundreds vehicles per link are detected on average by the scanners per 15 minutes (a relevant duration in transport to estimate ODM). With a penetration rate of Bluetooth devices estimated at around 30%, 500 vehicles per link is a reasonable value.

The penetration rate is drawn for each OD pair from a Gaussian distribution of mean 30% and standard deviation of 10% (and truncated to be between 0 and 1). This choice accounts for the variability of the ownership distribution of Bluetooth devices (which is not known) from one node to another depending, as an example, on the wealth of the neighbourhoods of the node. Finally, for each trajectory, it is drawn with a probability equal to the penetration rate on its OD whether it is a sample Bluetooth trajectory, or not. This allows us to have data $\underline{\underline{\underline{B}}}$ while the full set of trajectories gives $\underline{\underline{\underline{Q}}}$ for ground truth. The measured traffic flow per link $\underline{q}$ is obtained from $\underline{\underline{\underline{Q}}}$, assuming the addition a noise $\underline{\epsilon}$, for which each independent component is drawn from a Gaussian $\mathcal{N}(0, \sigma)$ distribution where σ is proportional to $\underline{q}$; we call r the ratio of proportionality.

For some pairs of OD, the flows of a typical simulation is shown in 1(a), with a clear heterogeneity of the counts on links.

4.2 Results

Additionally to the optimal solution of (1) (see [7]) and (9), two naive solutions $\widehat{\underline{\underline{\underline{Q_0}}}}$ and $\widehat{\underline{\underline{\underline{Q_1}}}}$ are computed as the Bluetooth LODM multiplied by the averaged Bluetooth penetration rate over the whole network or over each link respectively:

$$\widehat{\underline{\underline{\underline{Q_0}}}} = \frac{\sum_l \underline{q}}{\sum_{i,j,l} \underline{\underline{\underline{B}}}} \cdot \underline{\underline{\underline{B}}}$$

and

$$\forall (i,j,l) \in \mathbb{R}^{|V|\times|V|\times|L|} \quad (\widehat{Q}_1)_{ijl} = \frac{q_l}{\sum_{i,j} B_{ijl}} \cdot B_{ijl}$$

The solutions are compared in Table 1 by looking at the relative distance D_Q between the simulated LODM $\underline{\underline{\underline{Q}}}$ and the estimated one. This is computed as a ℓ_2 norm of the difference divided by the norm of the actual LODM. Others metrics are computed the same way: D_{q_l} and D_T for, respectively, the relative distance between aggregated volumes on links and the OD matrix $\underline{\underline{T}}$. For low noise r, the naive solution $\widehat{\underline{\underline{\underline{Q_1}}}}$ is performing well compared to the other solutions for very a low amount of computation. It performs especially good on the metrics D_{q_l} as a consequence of its definition. To the opposite, the naive solution $\widehat{\underline{\underline{\underline{Q_0}}}}$ which correspond to an averaged penetration rate over the whole network, therefore less sensitive to the noise

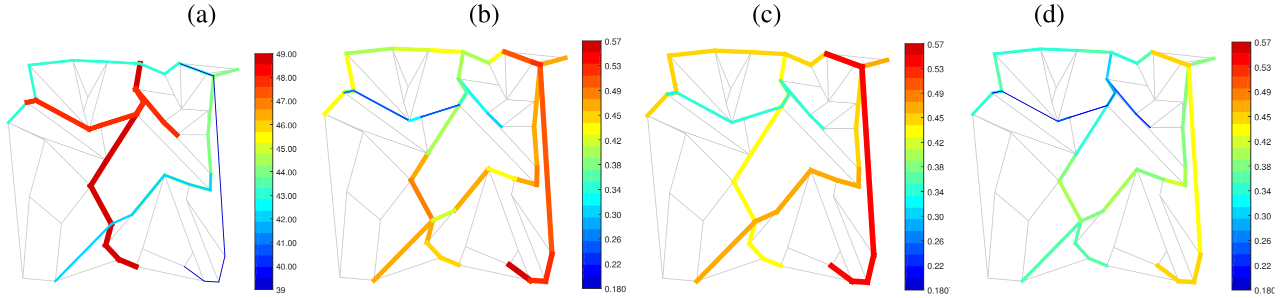

FIG. 1: Representation on the network for some OD: (a) of the number of trajectories per link for these OD; (b)-(c) of the relative errors on recovered volumes per link for the same OD, for (b) estimate $\underline{\underline{Q}} = Q_l$; (c) optimal solution of (1); (d) optimal solution of (9).

r, can give better estimates than $\widehat{\underline{\underline{Q}}}_1$ when the noise increases. The solution of (1) and (9) perform better than the naive solutions, independently of the noise which would be unknown in real traffic networks. Moreover, as shown on Figure 1(b), $\widehat{\underline{\underline{Q}}}_1$ has poor continuity of the flow along the itineraries. In fact, it doesn't satisfy Kirchhoff's law (F_3). Finally, the solution of (9) is performing better than the one of (1) both with the metrics of Table 1 and has lower relative error per OD as shown by the comparison of Figure 1(c) and (d).

r	Algo.	γ	μ	D_Q	D_{q_l}	D_T
0%	$\widehat{Q}_0$			0.399	0.03	0.400
	$\widehat{Q}_1$			0.397	**0**	0.396
	(1)	10	1*	0.397	0.01	0.395
	(9)	100	1	**0.360**	**0**	**0.358**
5%	$\widehat{Q}_0$			0.400	**0.03**	0.401
	$\widehat{Q}_1$			0.401	0.05	0.397
	(1)	0.1	1*	0.399	**0.03**	0.400
	(9)	0.1	2	**0.354**	0.04	**0.352**
10%	$\widehat{Q}_0$			0.405	**0.03**	0.406
	$\widehat{Q}_1$			0.416	0.10	0.422
	(1)	1	0.1*	0.405	**0.03**	0.406
	(9)	0.1	2	**0.364**	0.07	**0.364**

TAB. 1: Values of the metrics D_Q, D_{q_l} and D_T for optimal γ and μ. Q_0 and Q_l are the naive solutions. Note that μ for algorithm (1) is to treat with caution as G_3 in (1) was slightly different. The metrics reported are the same for all the $\widehat{\underline{\underline{Q}}}$.

5 Conclusion

A new methodology for estimating LODM from traffic counts and sample trajectories has been presented. Compared to previous method, the assumption of a Poisson law underlying the Bluetooth distribution together with more extensive constraints on network topology lead to better solutions. These results are an encouragement for developing additional constraints, especially on Origin-Destinations. Results have shown that metrics on aggregated data per links have very good values while improvements have to be done on the OD one. Trails of improvement might rely on testing other distribution relating $\underline{\underline{Q}}$ to $\underline{\underline{B}}$. Indeed, the combination of F_1 and F_4 corresponds to a truncated Poisson distribution and may be replaced by a binomial law. Otherwise, implementing a norm of smoothness of the penetration rate along the graph as per [15] or assuming a long term spatio-temporal correlation assumption.

References

[1] M Coates, A. O Hero, R Nowak, and B Yu, "Internet tomography," *Signal Processing Magazine, IEEE*, vol. 19, no. 3, pp. 47–65, 2002.

[2] A Girard, B Sansò, and F Vazquez-Abad, *Performance Evaluation and Planning Methods for the Next Generation Internet*, vol. 6, Springer, 2006.

[3] M Roughan, Y Zhang, W Willinger, and L Qiu, "Spatio-Temporal compressive sensing and internet traffic matrices (extended version)," *Networking, IEEE/ACM Transactions on*, vol. 20, pp. 662–676, 2012.

[4] E Castillo, J. M Menéndez, S Sánchez-Cambronero, A Calviño, and J. M Sarabia, "A hierarchical optimization problem: Estimating traffic flow using gamma random variables in a bayesian context," *Computers & Operations Research*, vol. 41, pp. 240–251, Jan. 2014.

[5] A Bhaskar and E Chung, "Fundamental understanding on the use of bluetooth scanner as a complementary transport data," *Transportation Research Part C: Emerging Technologies*, vol. 37, pp. 42–72, Dec. 2013.

[6] M Mardani and G. B Giannakis, "Estimating traffic and anomaly maps via network tomography," *arXiv preprint arXiv:1407.1660*, 2014.

[7] G Michau, P Borgnat, N Pustelnik, P Abry, A Nantes, and E Chung, "Estimating Link-Dependent Origin-Destination matrices from sample trajectories and traffic counts," Ecole Normale Supérieure de Lyon, FRANCE, Sept. 2015.

[8] H. H Bauschke and P. L Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, CMS Books in Mathematics Ser. Springer, Apr. 2011.

[9] N Parikh and S Boyd, "Proximal algorithms," *Foundations and Trends in optimization*, vol. 1, no. 3, pp. 123–231, 2013.

[10] P. L Combettes and J Pesquet, "A Douglas–Rachford splitting approach to nonsmooth convex variational signal recovery," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 1, no. 4, pp. 564–574, 2007.

[11] C Chaux, J Pesquet, and N Pustelnik, "Nested iterative algorithms for convex constrained image recovery problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 33, 2009.

[12] P. L Combettes and V. R Wajs, "Signal recovery by proximal Forward-Backward splitting," *Multiscale Modeling & Simulation*, vol. 4, no. 4, pp. 1168–1200, 2005.

[13] P. L Combettes and J. C Pesquet, "Proximal splitting methods in signal processing," in *Fixed-point algorithms for inverse problems in science and engineering*, p. 185–212. Springer, 2011.

[14] T. H Cormen, C. E Leiserson, R. L Rivest, and C Stein, "The algorithms of kruskal and prim," *Introduction to Algorithms*, p. 631–638, 2009.

[15] V Kalofolias, X Bresson, M Bronstein, and P Vandergheynst, "Matrix completion on graphs," *arXiv preprint arXiv:1408.1717*, vol. abs/1408.1717, 2014.