HAL
open science

Joint Flow and Density Reconstruction in Large Traffic Networks UsingPartial Turning Ratio Information

Martin Rodriguez-Vega, Carlos Canudas de Wit, Hassen Fourati

To cite this version:

Martin Rodriguez-Vega, Carlos Canudas de Wit, Hassen Fourati. Joint Flow and Density Reconstruction in Large Traffic Networks UsingPartial Turning Ratio Information. CDC 2018 - 57th IEEE Conference on Decision and Control, Dec 2018, Miami, FL, United States. pp.205-210, 10.1109/CDC.2018.8619321 . hal-01928699

HAL Id: hal-01928699

https://inria.hal.science/hal-01928699v1

Submitted on 21 Nov 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Joint Flow and Density Reconstruction in Large Traffic Networks Using Partial Turning Ratio Information

Martin Rodriguez-Vega*, Carlos Canudas-de-Wit* and Hassen Fourati**

***Abstract*— We address the recent problem of state reconstruction in large scale traffic networks using heterogeneous sensor data. First, we deal with the conditions imposed on the number and location of fixed sensors such that all flows in the network can be uniquely reconstructed. We determine the minimum number of sensors needed to solve the problem given partial information of turning ratios, and then we propose a linear time algorithm for their allocation in a network. Using these results in addition to floating car data, we propose a method to reconstruct all traffic density and flow. Finally, the algorithms are tested in a simulated Manhattan-like network.**

I. Introduction

The estimation of vehicle occupancy and vehicular flow in traffic networks is crucial for traffic management operators [1]. It allows to monitor the state of different urban areas and to detect traffic congestion. This information can be used to apply control strategies to reduce the impact of high traffic demands.

Network state information can be recovered from heterogeneous data sources. Floating car data (FCD) is available due to the expanded use of GPS-capable devices [1]. This provides the traces of some vehicles which can be used to estimate the mean speed over the entire network. On the other hand, fixed sensors (e.g. magnetic loops) give a reliable measure of the occupancy and traffic flow over a specific location [2]. Furthermore, vehicle re-identification sensors (e.g. Bluetooth and RFID identifiers) located in road intersections can give information about the routing preferences of vehicles [3], which can be used to estimate the turning ratios in monitored places. These could also be inferred from traffic cameras or from previous surveys [4].

Budget optimization, which traffic authorities often seek, is related with the problem of finding minimum number of sensors and their locations such that the traffic state is identifiable. This problem has been a recent topic of high interest [5], [6]. In some works it has been assumed that turning ratio information is fully available. For instance, in [7], sensors are associated with influence regions whose links are calculated by means of the measured data to provide heuristics on the upper and the lower bounds on the minimum number of sensors. In [8], the sensor location problem is solved via an optimization setup by reducing the propagation of noise when inferring unmonitored links. However, the assumption that all turning ratios are known may not be realistic in real world traffic applications.

This has motivated researchers to look at cases where no turning ratios are known. In [9], [10] only vehicular flow conservation equations are used to determine bounds on the number of sensors and their locations. The results were improved by [10]. They compute the minimum number of sensors and their locations by first expressing the flow relations as under-determined set of linear equations, and then, by operating via the Gaussian elimination method. Unfortunately, the required number of sensors range between 50% and 60% of all links for real traffic networks, which is not possible to achieve under budget constraints.

In parallel, density estimation has also been a field of active research [11]. However, most of the research in this topic has been centered in highways. For urban traffic networks, [8], [12] and [13] have provided methods to recover both flow and density by using turning ratio information in all intersections.

To the best of our knowledge, no results have been provided for density and flow reconstruction in urban traffic networks when turning ratios are only partially known. In this paper, we incorporate partial information of turning ratios to calculate the minimum number of sensors. Additionally, we present a modified version of the algorithms developed in [14] to use this partial information for the location of sensors. We use the data collected from all different sources to jointly reconstruct flow and density everywhere. We evaluate the proposed algorithms in a simulated traffic network. This work is carried out within the context of the ERC, Scale-FreeBack (see, [15]), which aims at developing holistic scale-free modeling, estimation and control methods for controlling complex network systems with application to large scale traffic systems.

II. Problem formulation

Traffic networks are modeled as a directed graph $\mathcal{G} = \{\mathcal{N}, \mathcal{E}\}$ where nodes $\mathcal{N} = \{1, 2, \ldots, n_{\mathcal{N}}\}$ represent traffic intersections, and links $\mathcal{E} = \{1, 2, \ldots, n_{\mathcal{E}}\}$ correspond to roads which are partitioned in three sets: incoming boundary roads $\mathcal{E}^{\text{in}}$, internal network roads $\mathcal{E}^{\text{net}}$, and outgoing boundary roads $\mathcal{E}^{\text{out}}$, with cardinalities n_{in}, n_{net}, n_{out}, respectively. For any $k \in \mathcal{N}$, the sets $\mathcal{I}(k) \subset \mathcal{E}$ and $\mathcal{O}(k) \subset \mathcal{E}$ are the links that go in or come out of k, respectively.

In this paper, we consider stationary traffic networks so that flow dynamics are in steady-state. This is formally true only when external boundary flows are constant and the network dynamics reach equilibrium, or it is approximately

* CNRS, GIPSA-Lab, 38000 Grenoble, France { martin.rodriguez-vega, carlos.canudas-de-wit }@gipsa-lab.fr

** Univ. Grenoble Alpes, CNRS, INRIA, Grenoble INP, GIPSA-Lab, 38000 Grenoble, France hassen.fourati@gipsa-lab.fr

true when the external boundary flows slowly evolve in time. The reason of this assumption is based on the idea that our density/flow reconstruction can be performed in a fast enough time-scale when compared with the evolution of the natural network dynamics, and therefore the analysis can be carried out using only flow/density equilibrium equations. This assumption substantially simplifies the reconstruction algorithm design, and is typically considered in the literature [5]. Nevertheless, our proposed solution will be tested in a dynamic network simulation.

For every road $e \in \mathcal{E}$, we assign scalar variables ρ_e, φ_e, v_e representing the equilibrium density, flow, and speed, respectively. We define the turning ratio $r_{e,f} \in (0,1]$ as the proportion of flow from e that is transfered to f if they are connected to the same intersection for which e is incoming and f is outgoing. Otherwise, $r_{e,f} = 0$.

We assume that traffic flow is a conserved quantity, where the sources are only the elements of $\mathcal{E}^{\text{in}}$, and sinks are only the elements of $\mathcal{E}^{\text{out}}$. In equilibrium, flow is constant along each road, and there is no accumulation of vehicles. Thus, each intersection k must obey an equation of the form:

$$\sum_{e \in \mathcal{O}(k)} \varphi_e - \sum_{e \in \mathcal{I}(k)} \varphi_e = 0. \tag{1}$$

The collection of equations for all $k \in \mathcal{N}$ can be rewritten in matrix notation as $L^0 \boldsymbol{\varphi} = \mathbf{0}$, where, $\forall k \in \mathcal{N}$ and $\forall e \in \mathcal{E}$, we have

$$L^0_{k,e} = \begin{cases} 1 & \text{if} \quad e \in \mathcal{O}(k) \\ -1 & \text{if} \quad e \in \mathcal{I}(k) \\ 0 & \text{else} \end{cases} .$$

This matrix has dimensions $n_{\mathcal{N}} \times n_{\mathcal{E}}$.

If knowledge of turning ratios is available, the conservation equations can be rewritten to include this additional information: the outflows of every intersection can be expressed by the sum of inflow contributions weighted by its turning ratios,

$$\varphi_f - \sum_{e \in \mathcal{I}(k)} r_{e,f} \varphi_e = 0 \ : \ \forall f \in \mathcal{O}(k) \tag{2}$$

where the turning ratios satisfy the condition $\sum_f r_{e,f} = 1$. Expression (2) is only valid for the links that go out from a node in $\mathcal{N}$, so the flow of links belonging to $\mathcal{E}^{\text{in}}$ cannot be written in this way. Similarly, the flow of links belonging to $\mathcal{E}^{\text{out}}$ does not contribute to the sum in the second term of the equation.

The ensemble of all equations following (2) is expressed in matrix form as $\begin{bmatrix} \boldsymbol{\varphi}_{\text{net}} \\ \boldsymbol{\varphi}_{\text{out}} \end{bmatrix} - R^T \begin{bmatrix} \boldsymbol{\varphi}_{\text{in}} \\ \boldsymbol{\varphi}_{\text{net}} \end{bmatrix} = \mathbf{0}$, where R is the matrix of turning ratios from the roads in $\mathcal{E}^{\text{in}} \cup \mathcal{E}^{\text{net}}$ to road in $\mathcal{E}^{\text{net}} \cup \mathcal{E}^{\text{out}}$, and $\boldsymbol{\varphi}_{\text{in}}$, $\boldsymbol{\varphi}_{\text{net}}$, $\boldsymbol{\varphi}_{\text{out}}$ are the flows of roads $\mathcal{E}^{\text{in}}$, $\mathcal{E}^{\text{net}}$, $\mathcal{E}^{\text{out}}$, respectively.

To simplify, let F^- and F^+ be selection matrices such that $F^- \boldsymbol{\varphi} = \begin{bmatrix} \boldsymbol{\varphi}_{\text{in}} \\ \boldsymbol{\varphi}_{\text{net}} \end{bmatrix}$ and $F^+ \boldsymbol{\varphi} = \begin{bmatrix} \boldsymbol{\varphi}_{\text{net}} \\ \boldsymbol{\varphi}_{\text{out}} \end{bmatrix}$. With this notation, the system of equations is $L^R \varphi = \mathbf{0}$, where $L^R = F^+ - R^T F^-$.

It is clear that matrices L^0 and L^R represent the flow conservation equations, and that L^R contains also contain information about the turning ratios. Note that both matrices have $n_{\mathcal{E}}$ columns, but L^0 has only $n_{\mathcal{N}}$ rows, while L^R has $n_{\mathcal{E}} - n_{\text{in}}$ rows. Thus a unique solution does not exist. To uniquely reconstruct network flows it is thus required to include additional sensor data such that these systems can be solved.

Let $\mathcal{S} \subseteq \mathcal{E}$ be the set of roads where flow sensors are located. To represent flow measurements, the sensing matrix $C(\mathcal{S})$ is defined such that the entry (i, e) is 1 if the i-th sensor is present at road e , and 0 otherwise. Flow measurements from sensors are denoted as $\boldsymbol{\varphi}_m$. The selection of a suitable set of measured links $\mathcal{S}$ such that a unique reconstruction of the vehicular flow vector $\boldsymbol{\varphi}$ can be obtained is known as the Network Sensor Location Problem (NSLP) [9].

The fundamental diagram is a well-known empirical relation between traffic variables that allows to express flow solely as a function of density: $\varphi = \Phi(\rho)$. One of the most popular choices is the triangular fundamental diagram [2]:

$$\Phi(\rho) = \begin{cases} v_0 \rho & \text{if} \quad 0 \leq \rho \leq \rho_c \\ w(\rho_{max} - \rho) & \text{if} \quad \rho_c \leq \rho \leq \rho_{max} \end{cases} , \tag{3}$$

where v_0 is the free-flow speed, w is the congested wave propagation speed, ρ_{max} is the jam density, and the maximal flow φ_{max} occurs at the critical density ρ_c. Vehicle speed is calculated as the slope of the line that connects the origin to any point in the diagram [2]. The first linear portion of the function (density below the critical value) receives the name of free-flow regime and is characterized by all vehicles having speed v_0. The region $\rho > \rho_c$ is called the congested regime as vehicles start to interact with each other reducing the average speed of the road.

It is not possible to uniquely reconstruct density from either flow or speed alone due to the non-unicity of the inverse of the map $\Phi^{-1}(\cdot)$. For instance, for a measured flow, two values of densities can be found. Likewise, if we only measure speeds, we cannot reconstruct density below its critical value. Thus, to recover the full information it is required to use both speed and flow data. Due to widespread mobile devices, we assume that the mean speed data is available over all links of the network.

Problem statement. Given any traffic network $\mathcal{G} = \{\mathcal{N}, \mathcal{E}\}$ where turning ratio information is available only for the set $\mathcal{R} \subset \mathcal{N}$: 1. determine the minimum number of sensors and their locations to solve the NSLP, and 2. use speed data and sensor measurements to estimate flow and density for all roads in the network.

The first part of the problem will be treated in Section III, and the second in Section IV.

III. Network sensor location problem (NSLP)

The flow conservation equations were expressed for two specific cases: either turning ratios were unknown resulting in matrix L^0, or the turning ratios were assumed to be known everywhere (matrix L^R). However, in some cases it may be possible that turning ratios are only available at a reduced set of intersections. Denote $\mathcal{R} \subseteq \mathcal{N}$ as the set of intersections where turning ratio information is known. We define $L(\mathcal{R})$

as the matrix that expresses the conservation equations by using the most information, and such that $L(\mathcal{R})\varphi = \mathbf{0}$.

To find the structure of this matrix we proceed as follows. Note that L^0 can obtained by making linear combinations of the rows of L^R: for each intersection k, there are $|\mathcal{O}(k)|$ rows in L^R but only one row in L^0. In an analogue manner, we write $L(\mathcal{R}) = J(\mathcal{R})L^R$, where $J(\mathcal{R}) = \begin{bmatrix} X \\ Y \end{bmatrix}$, such that X is a diagonal matrix defined as

$$X_{e,e} = \begin{cases} 1 & \text{if} \quad e \in \bigcup_{k\in\mathcal{R}} \mathcal{O}(k) \\ 0 & \text{else} \end{cases}, \tag{4}$$

and Y is such that $\forall k \in (\mathcal{N} \setminus \mathcal{R})$ and $\forall e \in (\mathcal{E}^{\text{net}} \cup \mathcal{E}^{\text{out}})$,

$$Y_{k,e} = \begin{cases} 1 & \text{if} \quad e \in \mathcal{O}(k) \\ 0 & \text{else} \end{cases}. \tag{5}$$

Matrix $J(\mathcal{R})$ has the effect of combining the rows of L^R for which turning ratios are unknown, thus generating an equation following (1). Rows of L^R related to outflows of nodes in $\mathcal{R}$ are unmodified.

The NSLP for partially known turning ratios can be split into two steps. First, we determine the minimum number of sensors needed to solve the problem. Then, we propose an algorithm that finds the sensor locations.

A. Minimum number of sensors

To uniquely reconstruct flows, we require the system of equations

$$\begin{bmatrix} L(\mathcal{R}) \\ C(\mathcal{S}) \end{bmatrix} \boldsymbol{\varphi} = \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\varphi}_m \end{bmatrix} \tag{6}$$

to admit only one solution. Our problem is to find the set $\mathcal{S}$ of minimal cardinality $n_{\mathcal{S}}$ such that the matrix in the left hand side of (6) is full rank.

Lemma 1. *Given any traffic network $\mathcal{G} = \{\mathcal{N}, \mathcal{E}\}$ such that every link $e \in \mathcal{E}$ is part of a directed path that ends in a link belonging to $\mathcal{E}^{out}$, the minimum number of sensors needed to solve the NSLP is given by*

$$n_{\mathcal{S}} = n_{\mathcal{E}} - |\mathcal{N} \setminus \mathcal{R}| - \sum_{k\in\mathcal{R}} |\mathcal{O}(k)|. \tag{7}$$

Proof. Note that from (6), $n_{\mathcal{S}} = n_{\mathcal{E}} - \text{rank}\{L(\mathcal{R})\}$. We shall prove that

$i)$ $\text{rank}[J(\mathcal{R})] = |\mathcal{N} - \mathcal{R}| + \sum_{k\in\mathcal{R}} |O(k)|$

$ii)$ L^R is full rank.

To prove $i)$ we will first show that $\text{rank}\{J(\mathcal{R})\} = \text{rank}\{X\} + \text{rank}\{Y\}$. Then, we show that $\text{rank}\{X\} = \sum_{k\in\mathcal{R}} |O(k)|$ and that $\text{rank}\{Y\} = |\mathcal{N} \setminus \mathcal{R}|$.

By inspection of (4) and (5), a column e of X can only have a nonzero element if there exists an intersection $k \in \mathcal{R}$ such that $e \in \mathcal{O}(k)$. On the other hand, a column e of Y can only have a nonzero element if $e \in O(m)$ for some intersection $m \in (\mathcal{N}\setminus\mathcal{R})$. Due to the fact that $(\mathcal{N}\setminus\mathcal{R})\cap\mathcal{R} = \varnothing$, column e cannot be simultaneously nonzero for both X and Y. Thus, any nonzero row of Y cannot be written as a linear combination of the rows of X, and vice versa.

Now we proceed to show that $\text{rank}\{X\} = \sum_{k\in\mathcal{R}} |\mathcal{O}(k)|$. From (4), X is a diagonal matrix and hence, its rank is equal to its number of nonzero elements, which is $|\bigcup_{k\in\mathcal{R}} \mathcal{O}(k)| = \sum_{k\in\mathcal{R}} |\mathcal{O}(k)|$, because the sets $\mathcal{O}(k)$ are mutually disjoint.

To show that $\text{rank}\{Y\} = |\mathcal{N} \setminus \mathcal{R}|$ we follow a similar analysis. Elements $Y_{k,e}$ and $Y_{m,e}$ cannot be simultaneously nonzero if $k \neq m$. Thus, all rows of Y are linearly independent, as they cannot be written as linear combinations of the others. As there are $|\mathcal{N}\setminus\mathcal{R}|$ rows, $\text{rank}\{Y\} = |\mathcal{N}\setminus\mathcal{R}|$.

To prove $ii)$ we will show that L^R can be written as $L^R = [A \quad B]$ where B is an invertible matrix. Recall that $L^R = F^+ - R^T F^-$. Without loss of generality, assume that the ordering of $\mathcal{E}$ is such that the smaller indexes correspond to $\mathcal{E}^{in}$, followed by $\mathcal{E}^{net}$ and ending with $\mathcal{E}^{out}$. Thus, $F^- = [\mathbb{I}_{n_{\text{in}}+n_{\text{net}}} \; \mathbf{0}_{(n_{\text{in}}+n_{\text{net}})\times n_{\text{out}}}]$ and $F^+ = [\mathbf{0}_{(n_{\text{net}}+n_{\text{out}})\times n_{\text{in}}} \; \mathbb{I}_{n_{\text{net}}+n_{\text{out}}}]$.

Let B a square matrix by selecting the las $n_{\text{net}} + n_{\text{out}}$ columns of L^R. Thus, $B = \mathbb{I} - \tilde{R}^T$, such that

$$\tilde{R} = \begin{bmatrix} \mathbf{0}_{n_{\text{net}}\times n_{\text{in}}} & \mathbb{I}_{n_{\text{net}}} \\ \mathbf{0}_{n_{\text{out}}\times n_{\text{in}}} & \mathbf{0}_{n_{\text{out}}\times n_{\text{net}}} \end{bmatrix} R.$$

Note that all diagonal entries of $\tilde{R}$ are 0, and that for each of the first n_{net} rows, the sum of all elements is 1. Thus, $\forall e, f$,

$$|B^T_{e,e}| \geq \sum_{f\neq e} |B^T_{e,f}|, \tag{8}$$

which means that B^T is a weakly diagonally dominant matrix. The rows corresponding to the links $\mathcal{E}^{out}$ are only nonzero in the diagonal, and meet (8) with the strict inequality. As every link is part of a path that ends in an element of $\mathcal{E}^{\text{out}}$, for every row e there is a chain of nonzero coefficients $B^T_{e,a_1}, B^T_{a_1,a_2}, B^T_{a_2,a_3}, \ldots, B^T_{a_m,f}$ such that $f \in \mathcal{E}^{\text{out}}$. In [16], it was proved that matrices meeting this properties are nonsingular. Thus, B is an invertible matrix and therefore, L^R is full-row rank. □

As special cases, when $\mathcal{R} = \varnothing$, the required number of sensors is $n_{\mathcal{S}} = n_{\mathcal{E}} - n_{\mathcal{N}}$, and when $\mathcal{R} = \mathcal{N}$ the required number of sensors if $n_{\mathcal{S}} = n_{\text{in}}$. This agrees with the results presented in [8], [10].

B. Sensor location

Graph-based approaches, such as the one in [14], have been shown to be useful to solve the sensor location problem. Because of this, we devised the following algorithm that includes the partial information about turning ratios.

Algorithm 1. *Sensor location*

Inputs: *Traffic network graph $\mathcal{G} = \{\mathcal{N}, \mathcal{E}\}$; set of $\mathcal{R}$ intersections with turning ratio information.*

Output: *Subset of measured links $\mathcal{S}$.*

1) *Generate a virtual node k_0 and connect entering and exiting external flows to it.*
2) *For each $k \in \mathcal{R}$: arbitrarily select all but one of the outgoing links of k and remove them from the graph, such that the resulting graph is connected. Denote the removed links as $\mathcal{E}^R$.*
3) *Construct a spanning tree from the remaining graph, and denote $\mathcal{E}^u$ as the set of links contained in the tree.*
4) *Locate sensors as $\mathcal{S} = \mathcal{E} - (\mathcal{E}^R \cup \mathcal{E}^u)$: links that were not removed in Step 2 and that are not part of the tree.*

This algorithm presents an heuristic modification to the work of [14] such that turning ratios are included. In the case $\mathcal{R} = \varnothing$ the two algorithms are equal.

It is important to remark that this algorithm locates exactly the same number of sensors as indicated by the lemma in Section III-A. By adding node k_0 in Step 1, a strongly connected graph with $n_{\mathcal{N}}+1$ nodes and $n_{\mathcal{E}}$ links is obtained [17]. Then, Step 2 removes a total of $\sum_{k\in\mathcal{R}} |\mathcal{O}(k)| - |\mathcal{R}|$ links. The input to Step 3 is a connected graph with $n_{\mathcal{N}}+1$ nodes. From graph theory, any spanning tree made from this graph will have $n_{\mathcal{N}}$ links. Therefore, the number of links that do not make part of the tree is $n_{\mathcal{E}} - \sum_{k\in\mathcal{R}} |\mathcal{O}(k)| + |\mathcal{R}| - n_{\mathcal{N}}$, which is equivalent to the expression in (7).

The rationale of the algorithm is to locate sensors such that they generate linearly independent rows in (6). It has been shown in [14], [17], [18] that a basis for the null space of L^0 can be obtained from the cycle space of the graph. If sensors are located in links such that their removal causes the graph to be cycle-free, then the dimension of the null space is reduced to 0 and flow can be reconstructed. Our contribution is focused on Step 2, which is an heuristic that allows to eliminate the selection of rows that are linearly dependent due to the knowledge of the turning ratios. Putting sensors in several outgoing links of an intersection may give redundant information, as their flows can be calculated by the knowledge of incoming links. In Step 3, all the remaining cycles of the graph are removed by constructing a spanning tree, which is an optimal cycle-free structure. Note that spanning trees are not unique, but because the graph is unweighted, all of them are equally possible for the problem at hand.

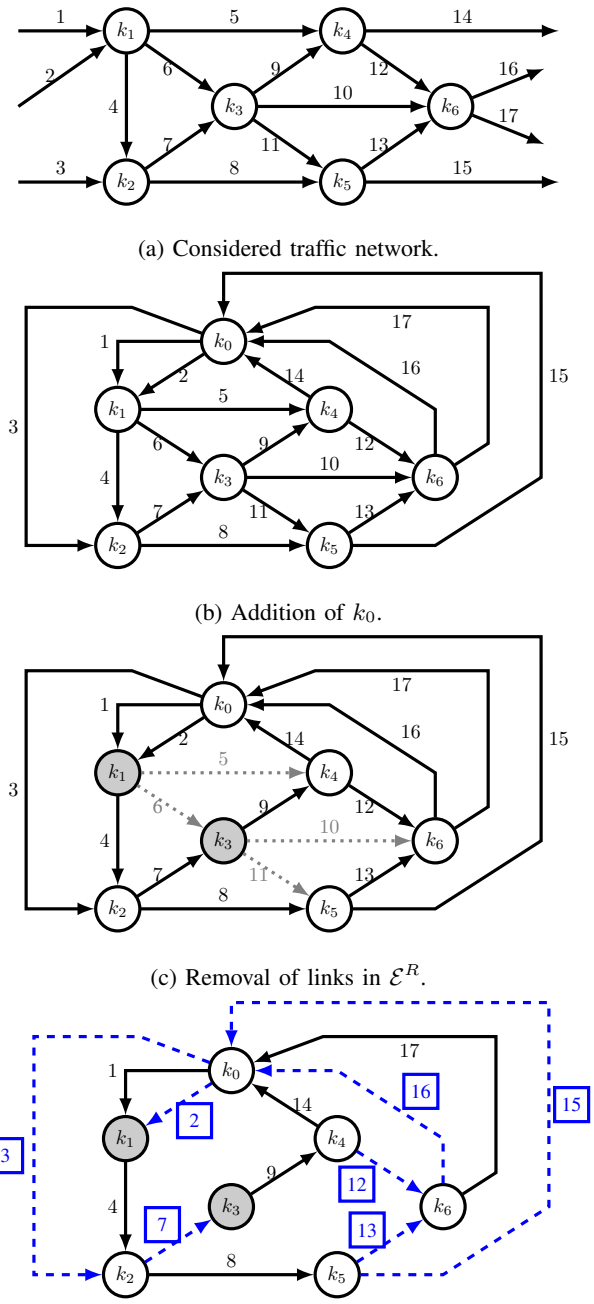
Example: Consider the traffic network shown in Fig. 1a. This graph has $n_{\mathcal{N}} = 6$ intersections and $n_{\mathcal{E}} = 17$ roads. Additionally, $\mathcal{E}^{in} = \{1,2,3\}$, $\mathcal{E}^{net} = \{4,5,\ldots,13\}$, and $\mathcal{E}^{out} = \{14,15,16,17\}$. Turning ratios are known at intersections $\mathcal{R} = \{1,3\}$. From Lemma 1, the minimum number of required sensors is $n_{\mathcal{S}} = 17-(6-2)-(3+3) = 7$. To locate the sensors, we use Algorithm 1:

- *Step 1:* add the virtual node which makes the graph strongly connected (Fig. 1b).
- *Step 2:* select links 5 and 6 from node 1, and links 10 and 11 from node 3 and remove them from the graph, as shown with dotted gray lines in Fig. 1c. Define $\mathcal{E}^R = \{5,6,10,11\}$.
- *Step 3:* construct a spanning tree from the links $(\mathcal{E}\backslash\mathcal{E}^R)$. The tree links are $\mathcal{E}^u = \{1,4,8,9,14,17\}$ shown with the black solid lines in Fig. 1d.
- *Step 4:* locate sensors in links $\mathcal{S} = \mathcal{E} \setminus (\mathcal{E}^R \cup \mathcal{E}^u) = \{2,3,7,12,13,15,16\}$. This links are shown with blue dashed lines in Fig. 1d.

Finally, note that the number of sensors is $|\mathcal{S}| = 7$, which is in agreement with the previous results.

IV. Joint density and flow reconstruction

To be able to calculate density from flow, we transform the nonlinear relation (3) into a set of linear equations using speed information. We define a selection function $S(v)$, such



(a) Considered traffic network.

(b) Addition of k_0.

(c) Removal of links in $\mathcal{E}^R$.

(d) Dashed lines contain sensors. Black lines are a tree.

Fig. 1: Example of Algorithm 1

that it identifies the current linear zone of the fundamental diagram:

$$S(v_e) = \begin{cases} 1 & \text{if} \quad v_0 - v_e < \epsilon \\ 0 & \text{else} \end{cases}, \tag{9}$$

where ϵ is a tunable parameter. Then, an expression for the density at each link is

$$\rho_e = S(v_e)\frac{\varphi_e}{v_0} + [1 - S(v_e)]\left(\rho_{max} - \frac{\varphi_e}{w}\right),$$

which can be written in matrix form as $\boldsymbol{\rho} = W(\boldsymbol{v})\boldsymbol{\varphi} + \boldsymbol{d}(\boldsymbol{v})$

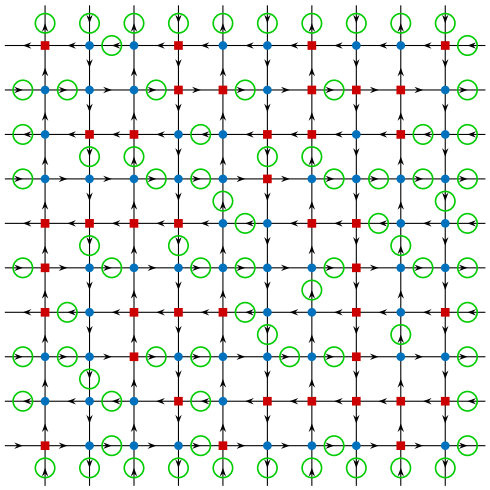


Fig. 2: Manhattan grid with 100 intersections. Square red nodes represent intersections where turning ratios are known and blue points correspond to unmeasured intersections. Circled green links represent the location of sensors.

where $\forall e \in \mathcal{E}$,

$$W_{e,e}(\boldsymbol{v}) = \begin{cases} 1/v_0 & \text{if} \quad v_0 - v_e < \epsilon \\ -1/w & \text{if} \quad v_0 - v_e \geq \epsilon \end{cases}$$

is a diagonal matrix, and

$$d_e(\boldsymbol{v}) = \begin{cases} 0 & \text{if} \quad v_0 - v_e < \epsilon \\ \rho_{max} & \text{if} \quad v_0 - v_e \geq \epsilon \end{cases}$$

is a vector.

The joint density-flow reconstruction can thus be formulated as a quadratic optimization problem subject to linear constraints:

$$\min_{\hat{\boldsymbol{\varphi}},\hat{\boldsymbol{\rho}}} \quad \left\| \begin{bmatrix} W(v) & -\mathbb{I} \\ C(\mathbb{S}) & \mathbf{0} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\varphi}} \\ \hat{\boldsymbol{\rho}} \end{bmatrix} - \begin{bmatrix} \boldsymbol{d}(v) \\ \boldsymbol{\varphi}_m \end{bmatrix} \right\|^2 ,$$

$$\text{s.t.} \quad L(\mathcal{N}_R)\hat{\boldsymbol{\varphi}} = \mathbf{0}, \; 0 \leq \hat{\boldsymbol{\varphi}} \leq \varphi_{max}, \; 0 \leq \hat{\boldsymbol{\rho}} \leq \rho_{max} \tag{10}$$

where $\hat{\boldsymbol{\rho}}$ and $\hat{\boldsymbol{\varphi}}$ are the estimated quantities.

V. Numerical example

To evaluate the proposed method, we considered the simulated traffic network shown in Fig. 2, which is a Manhattan-grid with 100 intersections and 220 roads. Turning ratios are selected to be known at 40% of the intersections chosen randomly.

As discussed in Section II, the proposed method considers steady-state networks. However, in real traffic applications this is generally very hard to achieve. For this purpose, we tested the performance of the algorithms under time-varying conditions. To generate synthetic dynamic traffic data we use the well-known Cell Transmission Model (CTM) [19], which is a discretization of the Lighthill-Williams-Richards (LWR) first order traffic model:

$$\rho_e(t+1) = \rho_e(t) + \frac{\Delta t}{L}\left[\varphi_e^{up}(t) - \varphi_e^{down}(t)\right], \tag{11}$$

where t is the discrete time index, Δt is the simulation step time, L is the road length, and $\varphi_e^{up}, \varphi_e^{down}$ are the upstream and downstream flows in link $e \in \mathcal{E}$, respectively. At intersections, flows are determined using the junction model presented in [20], [21] such that the intersection's throughput (sum of all flows) is maximized while also respecting the turning ratios. The traffic network parameters are shown in Table I. The time step was chosen such that $L/\Delta t > v_0$ to satisfy the CFL condition [20].

TABLE I: Parameters used for the simulations

Variable	Symbol	Value	Units
Road length	L	500	m
Time step	Δt	15	s
Free-flow speed	v_0	50	km/h
Max density	ρ_{max}	125	veh/km
Max flow	φ_{max}	1980	veh/h
Critical density	ρ_c	39.6	veh/km
Congested wave speed	w	23.18	km/h

To initialize the traffic network, a first simulation with constant input flows was run until equilibrium was obtained. Afterwards, each of the boundary input flows was subject to sinusoidal variations:

$$\varphi_e^{ext} = \varphi_{ss} + A\sin(\omega t) \quad , \quad \forall e \in \mathcal{E}^{in},$$

where φ_{ss} was the flow value obtained for the steady-state initial simulation, the amplitude $A = 0.25\varphi_{max}$ was chosen such that the input flows would reach φ_{max}, and ω is an arbitrary frequency. Several runs of the algorithm were made by modifying the value of ω, such that it followed different scenarios of real world traffic patterns, where the associated periods vary between 5 minutes and 6 hours.

To evaluate the algorithm's performance over time, we calculate the normalized root mean square error (NRMSE) of the density along all links in the network relative to the maximum values:

$$NRMSE_\rho(t) = \frac{1}{\rho_{max}}\sqrt{\frac{1}{n_\mathcal{E}}\sum_{e=1}^{n_\mathcal{E}}(\rho_e(t) - \hat{\rho}_e(t))^2}.$$

An analogous expression is used to calculate $NRMSE_\varphi(t)$ relative to flow.

A. Results and discussion

With the partial knowledge of turning ratios in the considered scenarios, 80 fixed flow sensors are needed, which represents 36% of all the links. This is an improvement in comparison with the case when turning ratios are unknown, which requires 120 sensors (55% of the links). Furthermore, the case when turning ratios are fully known locates only 20 sensors (9%), albeit with the requirement of knowing turning ratios in 60 additional intersections.

Figure 3 shows the obtained RMSE for both density and flow for three different scenarios of varying input frequency, with corresponding periods of 10 minutes, 1 hour and 6 hours. Through all cases, spikes of high and low errors can be associated with the regions of maximum and minimum derivative of the input. As the traffic propagates in the network, roads will follow the input but with phase differences. The blue dashed line shows the case with period of 10 minutes, whose rate of change is faster than traffic propagation in the network. Steady-state conditions are almost never achieved and the total error ranges from 5% to 30%. By

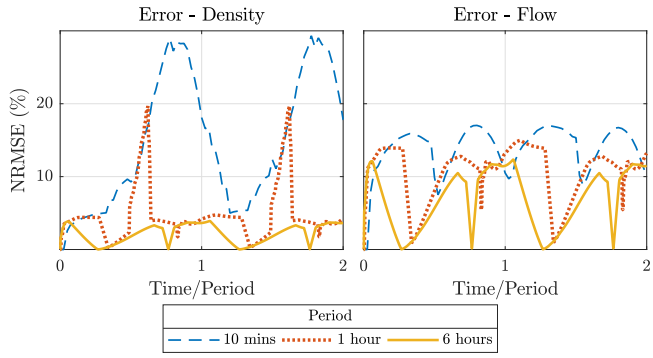


Fig. 3: NRMSE for flow and density reconstruction for different input flow variation frequencies.

increasing the period to 1 hour (red dotted line), the overall error is reduced, and the network can reach equilibrium-like state when the input's time-variation is small. Nevertheless, when the time derivative of the input is steep, error peaks of 20% are obtained. For the case with period of 6 hours shown with a continuous yellow line, errors are below 15% for flow and below 5% for density. Error minima of 0% show that input variations can be sufficiently constant to provide network equilibrium.

VI. Concluding Remarks

In this paper, we study and solve the network sensor location problem and joint flow and density reconstruction for large scale traffic networks by considering partial information of turning ratios. It is shown that this information can be used to effectively reduce the minimum number of required sensors, and a heuristic algorithm for their locations is proposed. In addition, flow sensor measurements and floating car data (speed measurements) are incorporated in an optimization problem that allows to reconstruct flows and densities for all roads in a traffic network. The proposed methods were evaluated in a simulated Manhattan grid. The results show that, even when the assumption that the network is in equilibrium state is not met, the algorithms find reasonable close results compared to the real values.

Future research on this aspect involves the reconstruction of the traffic state while taking into account network dynamics. Other future works include the development of methods to place sensors in optimal locations when budget constraints forces the use of less sensors than the minimum number determined in this paper.

Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement 694209), see [15].

References

[1] A. Mohan Rao and K. Ramachandra Rao, "Measuring Urban Traffic Congestion – a Review," *International Journal for Traffic and Transport Engineering*, vol. 2, no. 4, pp. 286–305, 2012.

[2] M. Treiber and A. Kesting, *Traffic flow dynamics: Data, models and simulation*, Springer, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013.

[3] 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, 2013.

[4] S. R. Hu and H. T. Liou, "Estimation of time-dependent intersection turning proportions for urban signal controls," *2012 12th International Conference on ITS Telecommunications, ITST 2012*, pp. 124–128, 2012.

[5] M. Gentili and P. B. Mirchandani, "Locating sensors on traffic networks: Models, challenges and research opportunities," *Transportation Research Part C: Emerging Technologies*, vol. 24, pp. 227–255, 2012.

[6] F. Viti, M. Rinaldi, F. Corman, and C. M. Tampère, "Assessing partial observability in network sensor location problems," *Transportation Research Part B: Methodological*, vol. 70, pp. 65–89, 2014.

[7] L. Bianco, G. Confessore, and P. Reverberi, "A Network Based Model for Traffic Sensor Location with Implications on O/D Matrix Estimates," *Transportation Science*, vol. 35, no. 1, pp. 50–60, 2001.

[8] E. Lovisari, C. Canudas de Wit, and A. Y. Kibangou, "Density/Flow reconstruction via heterogeneous sources and Optimal Sensor Placement in road networks," *Transportation Research Part C: Emerging Technologies*, vol. 69, pp. 451–476, 2016.

[9] S. R. Hu, S. Peeta, and C. H. Chu, "Identification of vehicle sensor locations for link-based network traffic applications," *Transportation Research Part B: Methodological*, vol. 43, no. 8-9, pp. 873–894, 2009.

[10] M. W. Ng, "Synergistic sensor location for link flow inference without path enumeration: A node-based approach," *Transportation Research Part B: Methodological*, vol. 46, no. 6, pp. 781–788, 2012.

[11] T. Seo, A. M. Bayen, T. Kusakabe, and Y. Asakura, "Traffic state estimation on highway: A comprehensive survey," *Annual Reviews in Control*, vol. 43, pp. 128–151, 2017.

[12] F. Zuurbier, J. V. Lint, and V. Knoop, "Traffic Network State Estimation Using Extended Kalman Filtering and DSMART," *IFAC, Delft*, pp. 37–42, 2006.

[13] A. Ladino, C. Canudas-de-Wit, A. Kibangou, H. Fourati, and M. Rodriguez-Vega, "Density and flow reconstruction in urban traffic networks using heterogeneous data sources," *To appear in European Control Conferenc ECC18 Limasol Cyprus*, 2018.

[14] S. X. He, "A graphical approach to identify sensor locations for link flow inference," *Transportation Research Part B: Methodological*, vol. 51, no. 516, pp. 65–76, 2013.

[15] "Scale-Free Control for Complex Physical Network systems. Scale-FreeBack," *ERC Advanced Grant*, no. 694209. [Online]. Available: http://scale-freeback.eu/

[16] P. N. Shivakumar and K. H. Chew, "A Sufficient Condition for Nonvanishing of Determinants," *Proceedings of the American Mathematical Society*, vol. 43, no. 1, p. 63, mar 1974.

[17] F. Y. Tascikaraoglu, J. Lioris, A. Muralidharan, M. Gouy, and P. Varaiya, "PointQ model of an arterial network: calibration and experiments," jul 2015.

[18] E. Castillo, A. Calviño, H. K. Lo, J. M. Menéndez, and Z. Grande, "Non-planar hole-generated networks and link flow observability based on link counters," *Transportation Research Part B: Methodological*, vol. 68, pp. 239–261, 2014.

[19] C. F. Daganzo, "The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory," *Transportation Research Part B*, vol. 28, no. 4, pp. 269–287, 1994.

[20] M. Garavello, K. Han, and B. Piccoli, *Models for Vehicular Traffic on Networks*, B. Piccoli, Ed. American Institute of Mathematical Sciences, 2016.

[21] C. F. Daganzo, "The cell transmission model, part II: Network Traffic," *Transportation Research Part B*, vol. 29B, no. 2, pp. 79–93, 1995.