



# A METHOD TO QUANTIFY NETWORK OBSERVABILITY USING LINK AND ROUTE INFORMATION



## ABSTRACT

This study proposes a method to measure the estimation reliability of traffic flows in a network on the basis of a limited number of sensors. The proposed metric enables one to consider any type of information, i.e. link, route and OD-based in an integrated way, and to relate this information to the unobserved flows. This can be of great advantage in data fusion applications.

## MOTIVATION

Observing traffic flows in a network is important to develop both short and long term strategies. However, only a subset of flows is normally observed in a network (e.g., near intersections, on motorways). These flows are collected using different data formats (traffic counts, probe vehicles, passage at tolled gantries, etc.) and are used to provide information also of flows that are not observed. It is therefore important to identify which traffic flows are not sufficiently captured, how the observed flows can contribute to the estimation and where to place extra sensors to improve these estimates.

## METHODOLOGY

In networks we distinguish three categories: OD flows, route flows and link flows. The (static) relations between these flows are:

$$\sum_{r \in R} \delta_{ar} f_r = v_a, \quad \forall a \in A, \quad \sum_{r \in R} \rho_{wr} f_r = h_w, \quad \forall w \in W$$

The matrix elements are 1 or 0, depending on whether a route contains a link (or an OD) or not. Since these are linear relations, it is possible to write them in a general incidence matrix as follows:

$$\begin{bmatrix} v \\ h \end{bmatrix} - \begin{bmatrix} \delta \\ \rho \end{bmatrix} \vec{f} = \vec{f}$$

Matrix  $I$  completely describes the network flow variables. However, it also contains duplicate information which is reflected as linearly-dependent rows in the matrix. By removing the dependent rows, only independent flows remain. Once all remaining rows of the matrix are independent, it is possible to determine which flows are to be observed to explain other flows by applying iterations of Castillo's pivoting procedure (Castillo, 2009). From the pivoting procedure, a matrix is obtained, which describes the relations between observed flows and unobserved flows. The unobserved true flows  $f$  and the unobserved estimated flows  $f^*$  should both satisfy the relation:

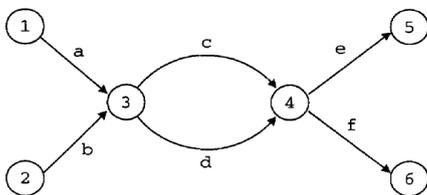
$$\sum \omega_{ij} f_j = \sum \omega_{ij} f_j^* = \vec{f}_i, \quad \forall i \in F$$

where  $\vec{f}$  denotes observed flows and  $F$  is the set of measured flows. By subtraction, this can be rewritten as  $\sum \omega_{ij} (f_j - f_j^*) = 0, \quad \forall i \in F$ .

Now define a function  $\lambda(f_p, f_j^*) = \lambda_j$  and matrix of elements  $\omega_{ij} = \frac{\omega_{ij} f_j^*}{\lambda(f_p, f_j^*)}$

Introducing these in the above equation gives:  $\sum \omega_{ij} \lambda(f_p, f_j^*) = 0$ .

The **new performance reliability function** is defined as the maximal norm of  $\lambda$ :  $\max G(\lambda) = \sqrt{\sum \lambda_j^2 / m}$



An illustrative case where even observing all link flows does not guarantee full reliability of the OD flows

## PROOF OF CONCEPT

Using the proposed function we can distinguish the maximal error in cases of partial network observability. If we have to choose three links to monitor in the toy network on the left, we would be better off by choosing e.g. couples (a,b,c), rather than e.g. (a,c,e), which exhibits a higher maximum value for  $G$ . Moreover, using the above performance measure, we find that to obtain full network observability (i.e. the observed flow variables for which  $\max G(\lambda)=0$ ) we need to monitor, for instance, 4 links (a,b,c,e) and 4 routes (a-c-e, a-d-e, a-c-f and b-c-e). This method can therefore be implemented in an optimization algorithm, i.e. given a certain budget (or number of sensors), find the flows to observe to maximize the reliability of information. This will be subject for future research.

