Vulnerability analysis of network observability in link flow inference problems

Marco Rinaldi[1], Francesco Viti[1]*

[1]MobiLab Transport Research Group, University of Luxembourg, Luxembourg
[*]Corresponding author: marco.rinaldi@uni.lu

Extended abstract submitted for presentation at the
7th International Conference on Transport Network Reliability (Sydney, 17-19 January, 2018)

1. Introduction & Literature review

Traffic information and management applications strongly rely on how traffic flows are monitored. Locating traffic sensors in a network is therefore considered a problem of paramount importance in transportation engineering, in particular within estimation problems (e.g. real time traffic state estimation, OD flows estimation, link flow inference, travel time estimation and path flow reconstruction (Ahmed et al., 2014; Castillo et al., 2008b; Viti et al., 2008; Zhou and List, 2010). From a technological and practical perspective, a few researchers have been focusing on establishing the impact that missing / incomplete data due to sensor failure in networks might have on the aforementioned problems, leading to schemes and techniques aimed at directly deploying sensors bearing robustness characteristics wrt. sensor failure (Li and Ouyang, 2011).

In this work, we specifically consider network flow observability problems, in which the main goal is to exploit algebraic and topological relationships on a set of observed link, route or OD flows in order to infer information on the remaining non-observed link (or route, or OD) flows. In these problems, through appropriate algebraic transformations, candidate locations for sensors in the network are subdivided into two sets: a topologically independent set (which is candidate for sensor deployment) and a topologically dependent set, for which measurements can be analytically derived through the algebraic relationship between the two sets.

Assessing the impact of sensor failure in these problems is not trivial: a sensor whose information acts as a keystone for several algebraic dependencies might represent a higher risk / vulnerability compared to a sensor whose information is entirely independent from any other source. We hypothesize that a well-conditioned observability solution should exhibit considerable robustness to sensor failure, therefore featuring both a low mean error due to sensor failure and a small variance, while less refined solutions will suffer from larger errors.

To empirically verify this hypothesis, in this work we perform a series of experiments to specifically measure how different network flow observability approaches – varying in terms of how densely packed the algebraic relationship information is – are vulnerable against randomly distributed sensor failures on a real-life sized network.

2. Methodology

In order to quantify the impact of stochastic sensor failure on network observability problems, we begin by deriving four different possible full observability solutions for the proposed network. This is achieved by firstly generating four different candidate route sets for the given network.

We compare two ad-hoc route set generation strategies, whose objective is specifically that of maximizing available dependent-independent information, namely Castillo's Algorithm 1 (Castillo et al., 2014) and our own Maximum Independent route set generation technique (Rinaldi and Viti, 2017); we furthermore compare these to two purely enumerative techniques, namely the well-known

K-Shortest Path (KSP) algorithm (Yen, 1971) and a variant proposed in (Rinaldi et al., 2015) which includes lax independence constraints (KISP). As we have recently shown in (Rinaldi and Viti, 2017), these four different route set generation techniques yield very different levels of information in terms of exploitable algebraic relationships. In this work we aim to further assess whether the denser information characteristics of our recently proposed approach yield robustness properties when dealing with sensor failure.

Based on the obtained route sets, expressed in terms of link-route incidence matrices $[A_{MI}, A_{KSP}, A_{KISP}, A_C] \in {}^{|L|\times|R|}$, where L is the link set and R the route set, the corresponding four full observability solutions $[\Omega_{MI}, \Omega_{KSP}, \Omega_{KISP}, \Omega_C] \in {}^{|indep|\times|dep|}$ are obtained using Castillo's pivoting method (Castillo et al., 2008a), which operates in a rows-by-columns fashion on matrix A to isolate independent-dependent variable relationships. From these four matrices, we can generate partial observability solutions, i.e. optimal solutions in which only a budget of sensors (in this paper, 50) has been installed on the network. To generate these solutions, we base ourselves on our own metric (shown in Eq. 1) and greedy addition algorithm, introduced in (Viti et al., 2014).

$$NSP = \frac{\| \Omega^T \cdot B' \|_F}{\| \Omega \|_F} \tag{1}$$

The same metric is then used to evaluate how the failure of a subset of the installed sensors affects the overall observability of the network. We perform 100 random draws (uniformly distributed) in each of which a fixed amount of sensors (10% - i.e. 5 sensors) will fail on the network. For each of these random draws, the corresponding partial observability is then computed.

As we discussed extensively in (Viti et al., 2014), metric (1) is a normalized quantity between 0 and 1, representing the amount of topological error introduced by partially observing a network. Sudden sensor failures will therefore increase this value, and we expect a considerable relationship will exist between how the full observability solution has been generated and how large the impact of sudden information loss will be. In the next Section, results obtained on the Rotterdam network are shown.

3. Results and conclusions

We test our approach on the city network of Rotterdam, shown in Figure 1.

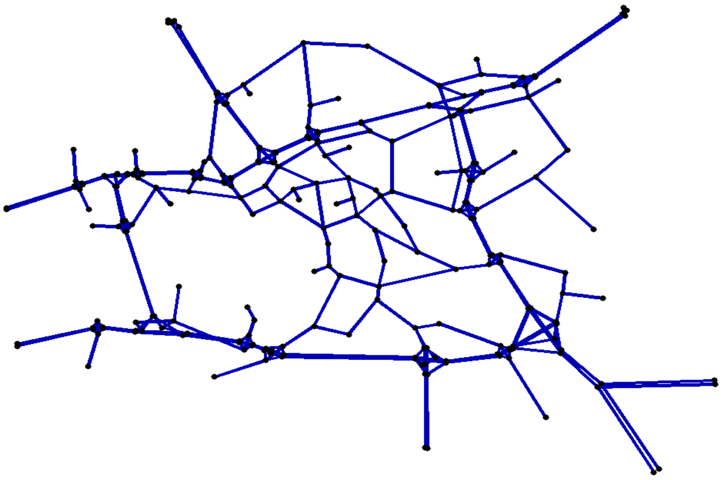

Figure 1: The Rotterdam city network.

This network consists of 476 links, 243 nodes and 1890 OD pairs.

In Figure 2 we present the four observability solutions and the relationship between their level of partial observability and the amount of sensors installed. For additional details on the performance of the four route enumeration criteria and the full observability solutions shown in the figure we suggest the interested reader to look at (Rinaldi et al., 2015) and (Rinaldi and Viti, 2017). As mentioned earlier, for all four, 50 sensors have been installed (as represented by the vertical dashed black line) and randomly 5 sensors are selected to be faulty.

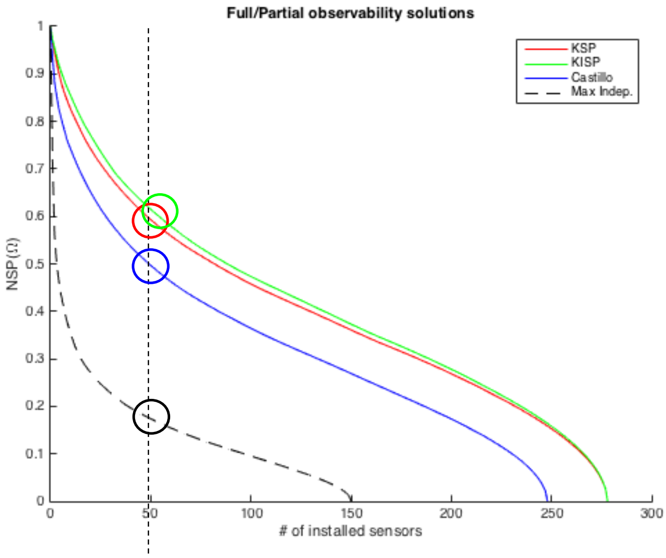

Figure 2: Full/partial observability solutions.

To showcase the different levels of robustness offered by the four partial observability solutions $[\Omega'_{MI}, \Omega'_{KSP}, \Omega'_{KISP}, \Omega'_{C}] \in {}^{50\times|dep|}$, we present the results in graphical form. In Figure 3(a), the experienced distributions of observability errors, as measured by equation (1) are shown. In Figure 3(b), instead, the relative error increase, compared to the levels obtained when all 50 sensors are installed, is shown.

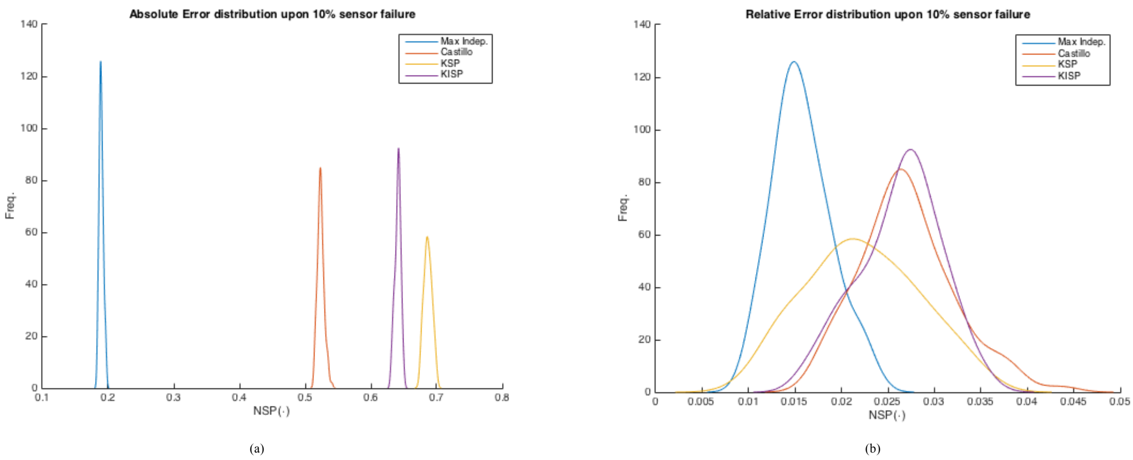

Figure 3: Distribution of observability errors upon sensor failure.

Analysing the results, it appears evident that the maximum independent route set methodology, other than minimizing the overall amount of sensors necessary to observe the network (150), also exhibits extremely robust performance in terms of sensor failure, showcasing both the lowest mean error increase and standard deviation (as summarized in Table 1).

Table 1: Simple statistics of relative observability errors

	Average	Standard Deviation
Max. Indep. Route set	**0.015**	**0.003**
Castillo's Alg. 1	0.027	0.005
K-Shortest Path	0.022	0.006
K-Indep. Shortest Path	0.026	0.004

The results obtained through this simple exploration further reinforce the methodological conclusions drawn in our previous work: topologically, route set generation techniques have a very strong influence in how robust link flow inference problems are wrt. total amount of sensors and, indeed, sensor failure. Future research includes extending the results reported in this work by (i) investigating how the % of failing sensor influences the overall robustness of the approaches (ii) validating these results on multiple networks and, finally, (iii) evaluating how flow estimation techniques might be influenced by these results.

References

Ahmed, A., Watling, D., Ngoduy, D., 2014. Significance of Sensor Location in Real-time Traffic State Estimation. Procedia Eng., Fourth International Symposium on Infrastructure Engineering in Developing Countries, IEDC 2013 77, 114–122. doi:10.1016/j.proeng.2014.07.012

Castillo, E., Calviño, A., Lo, H.K., Menéndez, J.M., Grande, Z., 2014. Non-planar hole-generated networks and link flow observability based on link counters. Transp. Res. Part B Methodol. 68, 239–261. doi:10.1016/j.trb.2014.06.015

Castillo, E., Conejo, A.J., Menéndez, J.M., Jimenez, P., 2008a. The observability problem in traffic network models. Comput.-Aided Civ. Infrastruct. Eng. 23, 208–222.

Castillo, E., Menéndez, J.M., Jiménez, P., 2008b. Trip matrix and path flow reconstruction and estimation based on plate scanning and link observations. Transp. Res. Part B Methodol. 42, 455–481. doi:10.1016/j.trb.2007.09.004

Li, X., Ouyang, Y., 2011. Reliable sensor deployment for network traffic surveillance. Transp. Res. Part B Methodol. 45, 218–231. doi:10.1016/j.trb.2010.04.005

Rinaldi, M., Corman, F., Viti, F., 2015. Assessing the Effect of Route Information on Network Observability Applied to Sensor Location Problems. Transp. Res. Procedia, 18th Euro Working Group on Transportation, EWGT 2015, 14-16 July 2015, Delft, The Netherlands 10, 3–12. doi:10.1016/j.trpro.2015.09.050

Rinaldi, M., Viti, F., 2017. Exact and approximate route set generation for resilient partial observability in sensor location problems. Transp. Res. Part B Methodol. Under Review.

Viti, F., Rinaldi, M., Corman, F., Tampère, C.M.J., 2014. Assessing partial observability in network sensor location problems. Transp. Res. Part B Methodol. 70, 65–89. doi:10.1016/j.trb.2014.08.002

Viti, F., Verbeke, W., Tampère, C., 2008. Sensor Locations for Reliable Travel Time Prediction and Dynamic Management of Traffic Networks. Transp. Res. Rec. J. Transp. Res. Board 2049, 103–110. doi:10.3141/2049-12

Yen, J.Y., 1971. Finding the K Shortest Loopless Paths in a Network. Manag. Sci. 17, 712–716. doi:10.1287/mnsc.17.11.712

Zhou, X., List, G.F., 2010. An Information-Theoretic Sensor Location Model for Traffic Origin-Destination Demand Estimation Applications. Transp. Sci. 44, 254–273. doi:10.1287/trsc.1100.0319