Transportation Research Part C 71 (2016) 184–197

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Data fusion algorithm for macroscopic fundamental diagram estimation

Lukas Ambühl [*], Monica Menendez

Traffic Engineering Group, Institute for Transport Planning and Systems, ETH Zurich, Switzerland

ARTICLE INFO

Article history:
Received 22 March 2016
Received in revised form 24 July 2016
Accepted 27 July 2016

Keywords:
MFD estimation
Simulation
Loop detector data (LDD)
Floating car data (FCD)
Fusion
Probe penetration estimation

ABSTRACT

A promising framework that describes traffic conditions in urban networks is the macroscopic fundamental diagram (MFD), relating average flow and average density in a relatively homogeneous urban network. It has been shown that the MFD can be used, for example, for traffic access control. However, an implementation requires an accurate estimation of the MFD with the available data sources.

Most scientific literature has considered the estimation of MFDs based on either loop detector data (LDD) or floating car data (FCD). In this paper, however, we propose a methodology for estimating the MFD based on both data sources simultaneously. To that end, we have defined a fusion algorithm that separates the urban network into two subnetworks, one with loop detectors and one without. The LDD and the FCD are then fused taking into account the accuracy and network coverage of each data type. Simulations of an abstract grid network and the network of the city of Zurich show that the fusion algorithm always reduces the estimation error significantly with respect to an estimation where only one data source is used. This holds true, even when we account for the fact that the probe penetration rate of FCD needs to be estimated with loop detectors, hence it might also include some errors depending on the number of loop detectors, especially when probe vehicles are not homogeneously distributed within the network.

1. Introduction and background

Recurring traffic congestion is one of the major problems in urban traffic (Buitelaar et al., 2007). As an example, in Zurich 85% of urban congestion is recurring - not an uncommon value for a European city (Keller and Wuethrich, 2012; Buitelaar et al., 2007). Even without considering congestion externalities, such as noise and green house gas emissions, it is obvious that strategies relieving urban networks from congestion are in need.

A promising framework for urban traffic management and control is the macroscopic fundamental diagram (MFD). The MFD relates average flow and average density of an urban network (Daganzo, 2007; Daganzo and Geroliminis, 2008). Its properties and applications have been studied extensively with simulations (e.g. Ji et al., 2010; Mazloumian et al., 2010; Courbon and Leclercq, 2011; Daganzo et al., 2011; Knoop et al., 2012; Gayah and Dixit, 2013; Aboudolas and Geroliminis, 2013) and also with real traffic data (e.g. Geroliminis and Daganzo, 2008; Buisson and Ladier, 2009; Cassidy et al., 2011; Geroliminis and Sun, 2011; Leclercq et al., 2014; Tsubota et al., 2015). For real data, different types of sensors can be used to estimate the aforementioned parameters. The most commonly used data sources in this case are loop detectors and

---

* Corresponding author.
E-mail address: lukas.ambuehl@ivt.baug.ethz.ch (L. Ambühl).

http://dx.doi.org/10.1016/j.trc.2016.07.013
0968-090X/

floating cars. Loop detector data (LDD) is obtained with static sensors, which evaluate the traffic condition at pre-defined locations. Floating car data (FCD) is collected from vehicles driving in the network (Leclercq et al., 2014). Both can be collected in real-time. In the following, a more detailed review of recent studies regarding the MFD estimation using LDD or FCD is given. A first part will concentrate on empirical studies, whereas a second part will concentrate on results based on simulation/theory.

Shortly after the theoretical introduction of the MFD, FCD was used to prove its existence. A first emperical study approximated the MFD in a 10 km^2 network within the city of Yokohama to create the first MFD from real data (Geroliminis and Daganzo, 2008). The authors used LDD and GPS data of taxis. They used LDD to estimate the taxi penetration rate. Using LDD from the City of Toulouse, it was found that the location of the loop detectors within the links is of essence to a well-defined MFD (Buisson and Ladier, 2009). Their analysis showed that the distance between the loop detector and the traffic signal has a strong impact on the shape of the resulting MFD. If the loop detector is placed close to a traffic signal, the periodic queue upstream of the signal leads to an overestimation of the density in the network. Recently, a Chinese study investigated the MFD of Changsha, China with real data from loop detectors and GPS data from taxis (Beibei et al., 2016). They used traffic counts from LDD and traffic densities from FCD. Furthermore, they found that the spatial distribution of the taxi proportion varies significantly.

A study explored the possibilities of using mobile probe data with a microscopic simulation of downtown Orlando (Gayah and Dixit, 2013). The authors randomly selected a fraction of vehicles in the network to serve as mobile probes. Using average speeds of the mobile probes, the average flow and density were estimated based on an a priori known MFD. It was found that the higher the mobile probe penetration rate and the longer the observation period, the more accurate the MFD estimations were. Nagle and Gayah (2015) estimated the penetration rate of mobile probes using an approach of Herrera et al. (2010). For their abstract grid network simulation, they placed loop detectors on almost all links. The probe penetration rate was estimated by combining the trajectory information from mobile probes and the total vehicle counts from loop detectors. Then, MFDs were estimated at different penetration levels, and their accuracy was evaluated with the use of root mean squared errors.

Using an abstract network, the empirical findings of Buisson and Ladier (2009) were confirmed by further investigating different cases of spatial distribution of loop detectors within links (Courbon and Leclercq, 2011). The authors showed that averages from loop detectors may overestimate up to around 100% of the real density value, depending on flow. They concluded that if loop detectors are well spread within the links (i.e., the average location of a loop detector is regularly distributed within the links for all links in the network), they can cover then different traffic states. This, in turn, means that the average density estimate, as needed for an MFD, is representative. In a recent study, the effects of heterogeneously distributed GPS-equipped mobile probe vehicles on the estimation of the probe penetration rate were investigated (Du et al., 2016). The authors used simulations of an abstract grid network with loop detectors and proposed a new probe penetration estimation method. In a first step, k-means clustering is applied to the measured probe penetration rates for different origin-destinations (OD). In a second step, the harmonic mean for these OD penetration rates is calculated in order to increase the accuracy of the probe penetration rate estimation for heterogeneously distributed mobile probes. This in turn, increases the accuracy of the estimated MFD. Another study successfully used incomplete LDD information of a simulated network of Chania, Greece to inform a gating control (Keyvan-Ekbatani et al., 2013). The authors used information from virtual loop detectors installed in the middle of the links. Ortigosa et al. (2014) studied the influence of the location and number of measurement points for an MFD perimeter control. Unlike discussed by Buisson and Ladier (2009), the placement of the loop detector within the link was not in question. In fact, the loop detectors were placed optimally within a link - queues upstream of a traffic signal did not distort the density measurements. They employed a microscopic simulation of downtown Zurich, created an MFD with only a subset of all the links (partial MFD), and calculated the accuracy thereof compared to a MFD obtained with full information.

Most research so far focused on the estimation of the MFD with either FCD *or* LDD. Leclercq et al. (2014), however, presented a cross-comparison of link data and probe data. They included a correction method for the densities of links improperly measured by loop detectors. With the aid of traffic flow theory they then showed significant improvements in the density estimation on an abstract urban corridor. Moreover, they showed that estimating the MFD is significantly improved when using LDD for flow values and FCD for speed values compared to an estimation with LDD only. This combined estimation method is also the basis for the equivalent estimation method which was applied in the empirical study in China mentioned earlier (Beibei et al., 2016) - where LDD was used for flow measures and FCD for density measures. Nagle and Gayah (2015) compared the use of LDD and FCD. Using a microscopic simulation of an abstract grid with one-way streets and a simple perimeter control, they estimated the MFD based on the methods proposed by Geroliminis and Daganzo (2008) and Gayah and Dixit (2013), respectively. Their analysis revealed that mobile probes and loop detectors can both adequately be used to inform a perimeter control scheme.

As shown with the literature reviewed above, most of the existing research has concentrated on the estimation of the MFD with a single data source. Additionally, and contrary to reality, some of the existing literature needs detailed knowledge, for instance, a priori knowledge of the FCD network coverage. Moreover, the accuracy achieved with these methods is somewhat unsatisfactory. Therefore, since both types of data sources usually co-exist, we show in this paper that fusing them improves the overall estimation of the MFD.

The remaining sections of this paper are organized as follows: First, we introduce the basic methodology for estimating an MFD based on FCD and LDD, separately. Then, we propose a fusion algorithm which combines both data sources for different