ARTICLE IN PRESS

Transportation Research Part C xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Transportation Research Part C

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

Optimal number and location of Bluetooth sensors considering stochastic travel time prediction

Hyoshin Park *, Ali Haghani

Department of Civil and Environmental Engineering, University of Maryland, College Park, MD 20742, United States

ARTICLE INFO

Article history:
Received 14 October 2014
Received in revised form 4 March 2015
Accepted 15 March 2015
Available online xxxx

Keywords:
Bluetooth technology
Sensor deployment
Dynamic relocation
Stochastic programming
Travel time prediction
Clustering

ABSTRACT

Determining the optimal number and location of sensors is essential to effectively manage traffic on highways. Optimal solutions dealing with dynamic traffic patterns and relocation of sensors have received little attention. In this study, existing *fixed* sensors are used to estimate travel time prediction errors at candidate locations where we deploy *portable* sensors. Potential sampling error of each candidate location is also counted in selecting optimal locations. A two-stage stochastic formulation considers uncertainty of traffic conditions based on scenarios generated by principal component analysis and clustering analysis to uncover the underlying spatial correlations and temporal patterns. The first stage decision, determining the optimal number of sensors, is made before the deployment. The second stage, evaluating the expected travel time prediction errors, specifies sensor arrangements in each scenario. A dynamic model has predefined rearrangement stages. At each stage, sensor locations are modified as the pattern of travel time error changes over time, considering sensor acquisition and relocation expenses. The deterministic and stochastic solutions serve as a lower bound and an upper bound for the dynamic solution. Higher relocation expense leads to more sensors being used, while higher sensor costs leads to fewer sensors being used with more frequent relocations.

1. Introduction

1.1. Motivation

Determining the optimal location of sensors on highways is essential for reliable travel time estimation and prediction as a fundamental input for the Intelligent Transportation Systems (ITS). The quality of traffic data collected by sensors has a direct influence on the reliability of network performance. ITS technologies attempt to relieve congestion by assisting drivers with this information (Stathopoulos and Karlaftis, 2003).

Two types of sensors exist to detect traffic on the road. First, *point* sensors (e.g., inductive loop detectors) provide traffic speeds at their locations. These fixed sensors in pavement or roadside on highway segments represent instantaneous travel time. On the other hand, *point-to-point* sensors (e.g., automatic vehicle identification (AVI) readers) provide direct measurement of experienced travel time between successive devices.

Essential issues affecting the quality of traffic information are the differences between the instantaneous and experienced travel time during congestion. This could be caused by technical characteristic or limitation of sensors. The point sensors are

* Corresponding author. Tel.: +1 3015310142; fax: +1 3014052585.
E-mail addresses: hspark@umd.edu (H. Park), haghani@umd.edu (A. Haghani).

http://dx.doi.org/10.1016/j.trc.2015.03.023
0968-090X/

Please cite this article in press as: Park, H., Haghani, A. Optimal number and location of Bluetooth sensors considering stochastic travel time prediction. Transport. Res. Part C (2015), http://dx.doi.org/10.1016/j.trc.2015.03.023

prone to various errors caused by malfunctioning. The point-to-point sensors only provide average speed over the entire segment. It takes time for an actual trip to be realized and for the travel time to become available. An approximate relationship between travel time from point sensors and point-to-point sensors can be studied (Zhang and Rice, 2003).

Although the sensing technology has been improved, an optimal sensor placement is still a challenge. Road networks represent dynamic evolution and correlation of traffic states. Temporal-spatial diversity of the travel demand originates from different departing and routing decisions that commuters make across different times of a day, or different days of a week. The difficulties of the traffic reliability improvement arise from inflexibility of stationary sensors that cannot respond to varying uncertainties over time. A sensor with high-precision technologies is expensive, and thus it is uneconomical to install additional fixed sensors without a strategic plan. While the random variables are still unknown, the planning decision can be made based on scenarios obtained by probability of expected values. This motivates us to formulate the sensor location problems using stochastic programming.

1.2. Literature review

The main perspectives of existing studies are traffic network flow and travel time prediction/estimation.

The first approach is explained by flow observability, or flow estimation (Zhou and List, 2010; Gentili and Mirchandani, 2012). Locating sensors in the traffic origin–destination (OD) demand estimation problem was proposed by Fei and Mahmassani (2011). Yang and Zhou (1998) selected traffic counting locations on the basis of complete OD pair separation. Liu and Danczyk (2009) placed detectors to minimize the traffic measurement errors. Procedures involving matrix algebra were developed based on link-path incidence (Castillo et al., 2008; Hu et al., 2009) or link-node incidence matrices (Ng, 2012; Hu and Liou, 2014). A stochastic model was used by Fei et al. (2013) to find near-optimal sensor location solutions of OD flow estimation with non-recurring traffic events.

Several studies considered travel time and OD estimation. Chen et al. (2004) used a genetic algorithm to minimize the number of AVI readers and maximize the number of individual readings. Sherali et al. (2006) proposed a nonlinear mixed integer optimization model to determine the optimal placement of AVI readers. Bartin et al. (2007) used a weighted summation of the speed variations. Mirchandani et al. (2009) proposed greedy heuristics to maximize the total vehicle-miles monitored, and to minimize the variance of predicted travel times. Asudegi and Haghani (2013) formulated a multi-objective optimization problem. Xing et al. (2013) selected a path travel time uncertainty criterion to construct a joint sensor location and travel time estimation/prediction framework. From different perspective, reliable location models were developed to allow facility failure (Daskin, 1983) and probabilistic sensor failures (Li and Ouyang, 2011).

In previous studies, aggregated errors have been used for sensor location problems without indication of stochastic features. Diverse sources of random errors cause uncertainties in the system. The difference between the reported speed data and ground truth observations (Bluetooth sensors) vary significantly over time windows. Data from different detectors are highly correlated among themselves and related to prevailing traffic conditions which exhibit short-term fluctuations (Stathopoulos and Karlaftis, 2003).

In contrast to existing models that used *fixed* sensors, our model benefits from *portable* sensors. Since 2008, the University of Maryland has been conducting ground truth testing of the traffic data across several road segments in East Coast. Bluetooth technology is proved to be a cost-efficient speed data collection approach (Haghani et al., 2010). Ease of shipping, handling and installation make relocation an option. While AVI readers have issues with privacy concerns, the machine access control (MAC) address of a cell phone, camera, or other electronic devices is not linked to a specific person through any type of central database. The MAC address, which is unique for each Bluetooth device, and the time of the detection are logged when the device is detected at a Bluetooth sensor (Young, 2008). From the logged times for each MAC identifier at the two Bluetooth sensor stations, the travel time for that specific MAC identifier on that road segment is calculated. These sensors are node-based, and two of them are required to find the travel time on a road segment. Currently, locations of sensor installation are chosen with a high-likelihood of observing congestion (Haghani et al., 2013).

1.3. Proposed approach

Dynamic deployment of portable sensors cover links with high uncertainty of traffic state and provide more reliable travel time data. We address fundamental challenges caused by temporal and spatial variability of travel time errors. The distribution of link travel time errors in a particular day is not a linear summation of the link travel time errors in all time periods of a day. In addition, drivers' inconsistent speed causes dependencies between consecutive links.

As shown in Fig. 1, two types of travel time errors from two types of sensors are used to decide optimal location of sensors. The benefit function is indicated as travel time errors that are weighted by travel demand for each segment. We assume that travel demand follows lognormal distribution. However, sum of travel time errors for different time interval with random variables may not follow a normal distribution, but may exhibit asymmetries. We make no assumption regarding either joint or marginal distributions due to unknown distribution of errors. Since travel time errors are closely related with traffic conditions, we assume that travel time errors in some segments are correlated. Clustering method is adapted to uncover the underlying temporal patterns and spatial correlations. Extracted principal components are used to classify freeway segments into several groups and partition time of day into sequential time intervals.