An Information-Theoretic Sensor Location Model for Traffic Origin-Destination Demand Estimation Applications

Xuesong Zhou
Department of Civil and Environmental Engineering, University of Utah,
Salt Lake City, Utah, 84112, zhou@eng.utah.edu

George F. List
Department of Civil, Construction and Environmental Engineering, North Carolina State University,
Raleigh, North Carolina, 27695, gflist@ncsu.edu

Abstract
To design a transportation sensor network, the decision-maker needs to determine what sensor investments should be made, as well as when, how, where and with what technologies. This paper focuses on locating a limited set of traffic counting stations and automatic vehicle identification readers in a network so as to maximize the expected information gain for the subsequent origin-destination demand estimation problem. The proposed sensor design model explicitly takes into account several important error sources in traffic origin-destination demand estimation, such as the uncertainty in historical demand information, sensor measurement errors, as well as approximation errors associated with link proportions. Based on a mean-square measure, the paper derives analytical formulations to describe estimation variance propagation for a set of linear measurement equations. A scenario-based stochastic optimization procedure and a beam search algorithm are developed to find sub-optimal point and point-to-point sensor locations subject to budget constraints. The paper also provides a number of illustrative examples to demonstrate the effectiveness of the proposed methodology.

1. Introduction

An important mission in the deployment of intelligent transportation systems (ITS) is to build and extend sensor networks to improve transportation system observability, productivity and efficiency. Emerging advances in the fields of surveillance, telecommunications, and information science have placed transportation sensor technology at the threshold of a period of major growth. Next-generation transportation sensor networks are expected to offer more reliable and less costly channels to measure complex transportation system dynamics. For example, with streaming roadside detector counts and video camera images from various locations, traffic controllers in a transportation management center can closely monitor network-wide traffic conditions and rapidly respond to traffic disturbances due to incidents or severe weather conditions. Automatic vehicle identification (AVI) and automatic vehicle location (AVL) data, on the other hand, could be utilized by traffic planners to estimate accurate and up-to-date origin-destination (OD) traffic demand, route choice and traffic delay information. In this study, we limit our focus to the point and point-to-point sensor location problem that is intended to enhance the quality of traffic origin-destination demand estimates.

In the early stages of transportation sensor network deployment, a thorny issue is that OD demand flows could not be fully observed. Theoretically, for a least squares estimator, the measurement matrix should have full rank so that the matrix inverse exists and the unknown state can be uniquely determined from measurements (See Gelb, 1974). However, the number of traffic counting stations is often much less than the number of unknown OD pairs in a large-scale traffic network. Early studies (e.g. Van Zuylen and Willumsen, 1980) used an entropy maximization model to find the most likely OD matrix that can reproduce observed link flows. Recognizing the intrinsic under-determined nature of the OD demand estimation problem, a majority of research aims to combine sensor data with prior OD demand information in order to obtain a unique OD estimate. Based on a Bayesian statistical approach, Maher (1983) gave analytical formulations for updating the prior mean and variance estimates by assuming multivariate normal distributions for trip demand flows and link observations. Cascetta (1984) further proposed a generalized least squares (GLS) model that does not rely on any distributional assumptions on prior demand estimation errors and sensor measurement errors. To calculate asymmetric confidence intervals for the maximum entropy estimator, Bell (1985) derived expressions for the demand estimate variance covariance matrix as a function of the observation dispersion matrix. In a unified framework for static OD demand estimation, Cascetta and Nguyen (1988) systematically illustrated the linkage between the above statistical inference models. To allow flexible control of the degree of confidence in different observations and asymmetric error functions for overestimation and underestimation of observed values, List and Turnquist (1994) presented a piecewise-linear multi-objective goal programming formulation for truck demand estimation applications. Recently, many researchers have proposed various models to utilize emerging AVI sensor data to estimate OD demand, including Van der Zijpp (1997), Asakura et al. (2000), Dixon (2000), Dixon and Rilett (2002), Eisenman and List (2004), Antoniou et al. (2004) and Zhou and Mahmassani (2005), to name a few. These models essentially seek to extract valuable OD demand information from point-to-point partially observed traffic counts in the presence of low market penetration rates and/or identification errors. Along a similar line of research, Sherali et al. (2006) recently presented a quadratic zero-one optimization model for locating AVI readers to maximize the benefit factors that capture travel time variability along specified trips in a traffic network.

To design a transportation sensor network, the decision-maker needs to determine what sensor investments should be made, as well as when, how, where and with what technologies. A fundamental question arising in the evaluation of alternative plans is how to select a measure or a set of criteria that can quantify information gain from sensor measurements at various locations. A number of statistical measures have been proposed to evaluate the quality of an OD demand estimator, such as the root mean square error (RMSE) and mean absolute error (MAE), and these performance indices can be expressed as deviations in terms of OD demand or link flows. In the sensor location problem, the link flow observations are not available before installing the sensors, thus it could be more desirable to choose statistical measures related to OD demand estimation errors. However, the true OD demand matrix is also generally unknown, so existing sensor location models tend to construct indirect quality measures that do not require knowledge of the exact values of OD demand flows. Lam and Lo (1990) proposed "traffic flow volume" and "O-D coverage" criteria to determine the priority of point detector locations. Yang et al. (1991) presented a "maximum possible relative error (MPRE)" criterion to calculate the most possible deviation from an estimated demand table to the unknown true OD trip demand. In their model, both link count observations and link flow proportions are assumed to be error-free, so the possible OD demand values that reproduce

2

link observations lie in a polyhedron defined by a set of traffic measurement equations. If an OD pair is not covered by any sensor, then the resulting MPRE is infinite, which motivates an OD pair covering rule that tries to ensure that a portion of the demand flow for each OD pair is observable. Bierlaire (2002) proposed a similar "total demand scale (TDS)" measure to calculate the difference between the maximum and minimum possible total demand estimates in a polyhedron constrained by traffic measurements. Yang and Zhou (1998) further defined a maximum coverage rule in terms of geographical connectivity and OD demand population. Yim and Lam (1998) tested a number of rules in several large traffic networks. Bianco et al. (2001) presented an iterative two-stage procedure and several priority-based greedy heuristics to cover OD flows and reduce the maximum possible relative error. Based on the entropy measure for OD flows proposed by Van Zuylen and Willumsen (1980), Chung (2001) introduced OD-specific weights to take into account the information content of the prior OD estimate, and Ehlert et al. (2006) further proposed a second-best location procedure to select informative links in a traffic network with partial detector coverage. Chen et al. (2005) adopted the TDS measure to evaluate the quality of OD demand estimates and to calculate the value of possible traffic counting station locations. The MPRE concept was further extended by Yang et al. (2006) to address the screen-line traffic counting location problem. Eisenman et al. (2006) proposed a Kalman filtering-based conceptual framework to characterize the error propagation dynamics in OD demand estimation, and they developed a simulation-based approach to numerically evaluate the value of point sensors for real-time network traffic estimation and prediction applications in a large-scale network.

In the fields of electrical engineering and information science, the sensor location problem has also received increasing attention in the last decade. Various measures are used to quantify the value of sensor information, such as Shannon entropy, Rényi divergence and Kullback-Leibler divergence, depending on the underlying assumptions and application areas. For example, an early study by Hintz (1991) used a Shannon entropy-based model to locate sensors for tracking a single target moving in one dimension. Lee (1998) proposed a combinatorial optimization framework to construct an atmospheric and geological sensor network design model, which maximizes the expected conditional entropy from spatially distributed sensors by selecting design points from a design space. Recently, information-theoretic measures have been widely integrated into machine learning models (e.g. Denzler and Brown, 2002) and target detection models with mobile sensors (e.g. Zhao et al., 2002). In the above applications, the unknown system states (e.g. the position and velocity of targets) typically can be directly measured by sensors. In comparison, sensing origin-destination traffic demand flows in a transportation network is difficult in its own right, as the OD estimation problem involves complicated mapping functions and a large number of unknown variables that are spatially and temporally connected to one another.

While significant progress has been made in formulating and solving the sensor location problem for OD demand estimation, several challenging theoretical and practical issues remain to be addressed. First, most of the existing studies do not explicitly take into account various error sources in the OD estimation process. In fact, the quality of historical OD demand estimates could significantly vary depending on the date and size of the original survey conducted. If link proportions are generated from traffic assignment programs, then estimation errors in the traffic flow and route choice models could, through the traffic assignment process, propagate to the final link proportion estimate (see Cascetta, 1984 and Ashok and Ben-Akiva, 2000 for detailed discussions). Second, the optimization criteria used in the existing sensor location models typically differ from those used in OD demand estimation. Due to the inconsistency between two models, the potential of scarce sensor resources might not be fully achieved in terms of maximizing information gain for OD demand estimation. For example, a sensor location plan that maximizes OD flow coverage does not necessarily yield the least OD demand estimation error for a GLS estimator. Third, most studies focus on how to locate traffic counting stations, while the value of emerging AVI sensors in the OD estimation problem has not been systematically investigated.

By extending the traffic state learning framework proposed by Eisenman et al. (2006), this paper adopts an information-theoretic approach to examine the inherent connection between the sensor location problem and the OD demand estimation problem. Essentially, we consider these two problems as a sequential optimization process, in which the sensor design stage first determines sensor locations and sensor types, and the subsequent OD demand estimation stage infers OD trip desires when sensor measurements are available for use. Moreover, this study aims to: (1) propose a theoretically rigorous sensor location model that can recognize different uncertainty sources and minimize the overall error in the OD demand estimation stage; (2) derive efficient formulations to analytically estimate the expected information gain

from a given set of sensor locations; and (3) present a practically useful framework to assist the decision-maker in locating traffic counting stations and AVI readers in a network.

The organization of this paper is as follows. After defining the sensor location problem, we first present a variety of measurement models that utilize point and point-to-point sensor data, and then discuss several possible single-valued metrics that can quantify the information content of OD demand estimates. Based on a mean-square criterion, we present an analytical formulation to describe mean and variance propagation in OD estimation. This is followed by a mathematical programming optimization model for the sensor location problem and a beam search-based heuristic solution procedure. At the end, we give a number of examples to illustrate the proposed methodology.

2. Notation and Problem Statement

We first introduce all the sets and subscripts in the sensor location and OD demand estimation problems.

Sets
I = set of origin zones,
J = set of destination zones,
L = set of links,
L' = set of links with point observations (e.g. link counts), $L' \subseteq L$,
L'' = set of links with point-to-point observations (e.g. vehicle identification counts), $L'' \subseteq L$,
L^* = set of links with sensors, $L^* = L' \cup L'' \subseteq L$.

Size of sets
m = number of observations,
n = number of OD pairs $|I| \times |J|$,
q = number of nodes in a sensor network.

Subscripts
i, j = subscript for origin/destination zone, $i \in I, j \in J$,
l, s = subscript for link with traffic measurements,
h = subscript for sensor sequence,
k = subscript for measurement used in Kalman filter updating,
w = subscript for initial demand table.

The following notations are used to represent four important components in the proposed sensor location model, namely the available measurements, estimation variables, mapping matrices between variables and measurements, as well as estimation error terms.

Estimation variables
$d_{(i,j)}$ = demand volume with destination in zone j, originating their trips from zone i.

Measurements
c'_l = number of vehicles passing through link l,
c''_l = number of tagged vehicles passing through link l,
$c''_{(l,s)}$ = number of tagged vehicles observed on link s, traveling from link l.

Mapping matrices
$p_{(l)(i,j)}$ = link flow proportions, i.e. proportion of vehicular demand flows from origin i to destination j, contributing to the flow on link l,
$\hat{p}_{(l)(i,j)(w)}$ = estimated link flow proportions based on the wth initial OD demand table,

4

$p_{(l,s)(i,j)}$ = point-to-point flow proportions, i.e. proportion of vehicular flows from origin i to destination j, contributing to the link-to-link flow from link l to link s,

$\hat{p}_{(l,s)(i,j)(w)}$ = estimated link-to-link flow proportions based on the wth initial OD demand table,

$\hat{p}_{(h)(i,j)(w)}$ = estimated sensor-sequence flow proportions based on the wth initial OD demand table,

$\hat{a}$ = estimated market penetration rate of AVI tags, i.e. the percentage of vehicles with tags in the entire vehicle population.

Estimation error terms
(Errors related to point measurements)
ω_l = measurement error on link l related to senor equipment and environment,

$\eta_{(l)(i,j)(w)}$ = modeling errors associated with link flow proportion $\hat{p}_{(l)(i,j)(w)}$ estimated from traffic assignment or simulation programs based on the wth initial OD demand table,

$\varepsilon'_{l,w}$ = combined error term on link l related to point measurement and modeling errors,

(Errors related to point-to-point measurements)
$\hat{\varepsilon}''_{(i,j)}$ = combined error term associated with point-to-point measurements from origin i to destination j,

$\hat{\varepsilon}''_{l,w}, \hat{\varepsilon}''_{(l,s)(w)}, \hat{\varepsilon}''_{h,w}$ = combined error term associated with point-to-point measurements on link l, link pair (l,s) and path index h, respectively, where the corresponding proportion matrices are constructed based on the wth initial OD demand table,

$\varepsilon''_{(i,j)}$ = sampling error term related to destination choice ratios for OD pair (i,j) (without involving market penetration rate estimate),

$\varepsilon''_{(l,s)(w)}$ = combined error term associated with point-to-point measurements obtained on link pair (l,s), based on the wth initial OD demand table.

Note that, if a link proportion matrix is used in the measurement equation (based on the wth initial OD demand table), then the corresponding error term should include a subscript of w, and vice versa. Finally, we summarize the above notations in the following vector and matrix form, which will be extensively used to derive the value of the information.

Vector and matrix forms in Kalman filtering framework
C = sensor measurement vector, consisting of m elements,
D = OD demand vector, consisting of n elements $d_{(i,j)}$,
$\bar{D}$ = initial OD demand vector used for generating link proportions through traffic assignment,
D^- = a priori estimate of the mean values in the demand vector, consisting of n elements,
D^+ = a posteriori estimate of the mean values in the demand vector,
$\tilde{D}$ = a posteriori demand estimate error, i.e. $\tilde{D} = D - D^+$,
P^- = a priori error covariance matrix of demand estimate, consisting of $(n \times n)$ elements,
P^+ = a posteriori error covariance matrix, i.e. conditional covariance matrix of estimation errors after including measurements,
H = sensor matrix that maps unknown demand flows D to measurements C, consisting of $(m \times n)$ elements,
K = updating gain matrix, consisting of $(n \times m)$ elements,
R = variance covariance matrix for combined errors, including measurement and modeling errors,
ε = combined error term, $\varepsilon \sim N(0, R)$.

Consider a traffic network with multiple origins $i \in I$ and destinations $j \in J$, as well as a set of nodes connected by a set of directed links. Given prior information on OD trips, the sensor location problem seeks to find a set of links $L^* = \{L', L''\}$ so that link counts c'_l are available on link $l \in L'$ and vehicle

identification data are available from AVI readers located on link $l \in L''$ for the subsequent OD demand estimation problem. The point-to-point measurements include vehicle identification counts $c''_l \ \forall \ l \in L''$ and point-to-point counts $c''_{(l,s)} \ \forall \ l,s \in L''$. The goal of the sensor location problem is to maximize information gain from the sensor set on L^*, subject to budget constraints for installation and maintenance. Note that, as AVI reader stations are usually installed on link segments in a network, the "point-to-point counts" will be equivalently referred to as "link-to-link counts" in order to maintain congruity with "link counts" from point sensors. In this study, we assume the historical OD demand information can be characterized by the *a priori* mean vector D^- and the estimation error variance matrix P^-.

3. Measurement Models

A linear measurement equation is used to relate the unknown OD demand to both point and point-to-point measurements:

$$C = HD + \varepsilon, \text{ where } \varepsilon \sim N(0, R). \tag{1}$$

Measurement vector C includes both link counts and vehicle identification counts (i.e. link-to-link counts), and the size of C depends on the set of sensor locations. Sensor matrix H provides a linear mapping between demand flows and observations, and a typical example is a link flow proportion matrix that maps OD flows to link counts. In this study, we assume sensor matrix H can be determined from traffic assignment or simulation programs, based on prior demand information. Additionally, we assume that the measurement error covariance matrix R is known. It should be remarked that, although the actual values of sensor measurements are unknown before the sensors are installed, the analyst can estimate the magnitude of measurement errors from the same type of sensors at similar locations or related studies in other areas. In short, the given conditions for the sensor location problem can be mathematically summarized as: (1) the mean D^- and covariance matrix P^- of the *a priori* demand estimate, (2) the measurement error covariance matrix R and the sensor matrix H for all possible sensor sites.

3.1. Point Measurement Models

The measurement equation for using link counts is typically expressed as:

$$c'_l = \sum_{i,j} p_{(l)(i,j)} \times d_{(i,j)} + \omega_l. \tag{2}$$

That is, the link flow count on link l is the sum of flow passing through link l from different OD pairs plus a measurement error. Since it is difficult to directly measure the true values of link flow proportions, especially in a congested traffic network, the analyst typically uses traffic assignment or simulation programs to produce an estimate of the link proportions. To adequately consider different possible base demand matrices used in traffic assignment, we use a set of initial demand matrices, and $\hat{p}_{(l)(i,j)(w)}$ represents an estimated link proportion vector generated from the wth initial OD demand table $\bar{D}_w$, that is,

$$\hat{p}_{(l)(i,j)(w)} = p_{(l)(i,j)} + \eta_{(l)(i,j)(w)}. \tag{3}$$

Substituting Eq. (3) into Eq. (2) yields

$$c'_l = \sum_{i,j} \hat{p}_{(l)(i,j)(w)} \times d_{(i,j)} + \varepsilon'_{l,w}, \tag{4}$$

where $\varepsilon'_{l,w} = \sum_{i,j} \eta_{(l)(i,j)(w)} \times d_{(i,j)} + \omega_l. \tag{5}$

The above equation indicates that the combined error term $\varepsilon'_{l,w}$ reflects the overall effect of errors from measurements and modeling, and the magnitude of the total error depends on the quality of the initial demand that generates link flow proportions through the traffic assignment program. For simplicity, the following analysis first ignores possible interactions among the error terms, and assumes the combined error is white noise. A later section will focus on how to recognize and accommodate possible estimation errors in link flow proportions generated by different initial OD demand matrices.

The link proportion formulation can be easily extended to incorporate origin/termination counts and screen-line counts. For instance, we can set up virtual links corresponding to screen lines and then construct

the related "screen-line" flow proportions, i.e. the percentage of vehicular demand flow from an origin-destination pair contributing to the flow on a screen line, to adopt a similar form as Eq. (2).

3.2. Point-to-point Measurements

In principle, the ultimate way to ensure observability of the OD demand estimation problem is to add more measurements, from point sensors, point-to-point, or semi-continuous sensors. The next task in our study is to establish measurement equations that can utilize point-to-point AVI information, integrating and extending the work by Eisenman and List (2004) and Zhou and Mahmassani (2005). We first assume (1) AVI readers can correctly identify every tagged vehicle, and (2) the tagged vehicles are a representative subset of the entire population. The first condition assumes 100% identification rates. Under the second condition, the tagged vehicles probabilistically represent the entire population. If an AVI reader and a point detector are located on the same link l, then the market penetration rate of tagged vehicles can be estimated from $\hat{a}_l = \dfrac{c''_l}{c'_l}$. The average value of $\hat{a}_l$ for all links $l \in L' \cap L''$ can be used to estimate a network-wide market penetration rate $\hat{a}$.

If AVI counts can be obtained from origin i to destination j, we have

$$c''_{(i,j)} = \hat{a} \times d_{(i,j)} + \hat{\varepsilon}''_{(i,j)}, \tag{6}$$

where $\varepsilon''_{(i,j)}$ is a combined error term that includes measurement errors for AVI tags and sampling errors associated with $\hat{a}$. Similarly, we can map OD flows to link counts, and link-to-link counts through the AVI market penetration rate, that is,

$$c''_l = \hat{a} \times \sum_{i,j} \hat{p}_{(l)(i,j)(w)} \times d_{(i,j)} + \hat{\varepsilon}''_{l,w}, \tag{7}$$

$$c''_{(l,s)} = \hat{a} \times \sum_{i,j} \hat{p}_{(l,s)(i,j)(w)} \times d_{(i,j)} + \hat{\varepsilon}''_{(l,s)(w)}. \tag{8}$$

Furthermore, an AVI sensor network can be viewed as a complete graph in which each pair of sensor nodes is connected by an edge corresponding to the sensor equation (8). The number of possible edges for a complete graph with q nodes is $q(q-1)$.

More generally, we can utilize the number of tagged vehicles passing through a sensor path sequence in an AVI sensor network to infer the OD demand flows using the following equation:

$$c''_h = \hat{a} \times \sum_{i,j} \hat{p}_{(h)(i,j)(w)} \times d_{(i,j)} + \hat{\varepsilon}''_{h,w}, \tag{9}$$

where a sensor sequence is an ordered list of a subset of the AVI sensors. For example, the subset of sensors {a, b, c} in Fig. 1 could generate the following sensor paths: *abc*, *acb*, *bac*, *bca*, *cab*, and *cba*. Accordingly, the sensor-sequence flow proportion $\hat{p}_{(h)(i,j)(w)}$ describes the proportion of vehicular demand flows from origin i to destination j contributing to the flow passing through sensor path h sequentially. Ideally, the number of possible sensor sequence arrangements for q AVI sensors is the sum of the permutations. In many cases, however, a sensor path sequence might contain unnecessarily long detours, and no vehicle would be observed traveling along such sensor sequences. For example, in the four-sensor network shown in Fig. 1, sequences such as *adbc* and *adcb* are less likely to carry any traffic flow. Thus, the above formula serves as a weak upper bound for the number of AVI measurement equations that can be included in a sensor design model. It should be also noted that, in addition to providing pair-wise counts in inferring origin-to-destination demand, the observations of sensor sequence could provide more information in calibrating the critical route choice model and link proportion matrix in the traffic assignment process, and thus could subsequently improve the overall quality of estimated traffic network flow patterns in terms of OD, path and link flows.
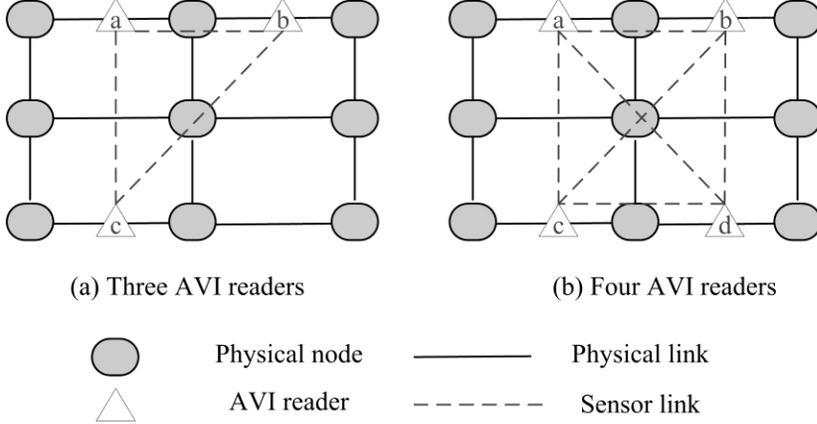
(a) Three AVI readers   (b) Four AVI readers

⬤ Physical node   ——— Physical link

△ AVI reader   – – – – – Sensor link

Figure 1 AVI sensor networks

In practice, because the rates of AVI tags essentially could vary significantly across different origin-destination pairs, estimating the true market penetration rate could be extremely difficult. The following measurement equations are intended to circumvent the need to infer the market penetration rate. If tagged vehicles are a representative subset of the entire population, then the split fractions of tagged vehicles can be used as sample estimates of the population split fractions. For example, the sampled destination choice split, that is, the percentage of tagged vehicles originating in zone i and headed to destination j, is

$$\frac{c''_{(i,j)}}{\sum_j c''_{(i,j)}} = \frac{d_{(i,j)}}{\sum_j d_{(i,j)}} + \varepsilon''_{(i,j)}. \tag{10}$$

In this case, there are $c''_{(i,j)}$ tagged vehicles destined to zone j out of $\sum_j c''_{(i,j)}$ vehicles observed

originating in zone i. Hence, the sampled destination choice split fractions essentially follow a multinomial distribution. Note that the Kalman filter is a linear filter, while Eq. (10) involves a nonlinear function. In this case, we need to apply an extended Kalman filter to handle the nonlinear measurement equations. More specifically, a first-order Taylor series expansion can be used on *a priori* estimate D^- to express the above nonlinear equation (10) in a linear fashion. Measurement equations for the link-to-link split fractions can be similarly obtained by using sampled link-to-link flow proportions:

$$\frac{c''_{(l,s)}}{c''_l} = \frac{\sum_{i,j} \hat{p}_{(l,s)(i,j)(w)} d_{(i,j)}}{\sum_{i,j} \hat{p}_{(l)(i,j)(w)} d_{(i,j)}} + \varepsilon''_{(l,s)(w)}, \tag{11}$$

where $\varepsilon''_{(l,s)(w)}$ refers to the combined error that includes the following error sources:

1) Model assumption errors related to the hypotheses regarding perfect representativeness.
2) Sensor errors (i.e. identification errors) related to link-to-link AVI count $c''_{(l,s)}$ and link count c''_l.

3) Sampling errors for the split fractions $\dfrac{c''_{(l,s)}}{c''_l}$.

4) Estimation errors related to link flow and link-to-link flow proportions from the traffic assignment program based on the wth initial demand table, which can be further caused by inconsistency in various assumptions of the route choice behavior, traffic flow propagation, as well as input data errors related to traffic control and information strategies.

In general, in order to avoid link proportion estimation errors, it is advantageous to locate AVI sensors to cover the entry/exit links of traffic analysis zones so that we can directly measure the origin-destination flows. Moreover, one can combine AVI origin-to-destination counts for OD pair (i,j) and traffic link counts on link l to directly estimate the link flow proportions:

$$\hat{p}_{(l),(i,j)} = \frac{c''_{(l)(i,j)}}{c''_{(i,j)}} \,. \tag{12}$$

4. Measures of Information

One of the fundamental questions in both OD demand estimation and sensor location problems is which criteria should be selected to drive the underlying optimization processes. Essentially, the OD demand estimation problem is to find a new estimate D^+ that can combine and utilize information from prior estimates and sensor measurements. By definition, the posterior error covariance matrix is

$$P^+ = E\{[\tilde{D} - E(\tilde{D})][\tilde{D} - E(\tilde{D})]^T\} \,. \tag{13}$$

If the estimator is unbiased (i.e. $E(\tilde{D}) = 0$), then the above equation reduces to

$$P^+ = E(\tilde{D}\tilde{D}^T) = E\left[(D - D^+)(D - D^+)^T\right] \,. \tag{14}$$

In the following, we examine two commonly used estimation criteria, namely, the mean-square error and entropy. The classic Kalman filter aims to minimize the mean-square error, that is, the Euclidean norm square of $\tilde{D}$:

$$E\left\|\tilde{D}\right\|^2 = E(\tilde{D}^T\tilde{D}) = E\left[(D - D^+)^T(D - D^+)\right], \tag{15}$$

and equals the trace of the variance and covariance matrix:

$$tr(P^+) = tr(E[\tilde{D}\tilde{D}^T]) \,. \tag{16}$$

Entropy is another commonly used measure of information. For a discrete variable, Shannon's original entropy is defined as the number of ways in which the solution could have arisen. This definition has been used in previous OD demand estimation models that assume error-free measurements and link flow proportions, e.g. that of Van Zuylen and Willumsen (1980). For a continuously distributed random vector D, on the other hand, the entropy is measured by $-E(\ln f(D))$, where f is the joint density function for D. If D follows a normal distribution, then its entropy is quantified as

$$\beta + \frac{1}{2}\ln(det(P^+)) \,, \tag{17}$$

where β is a constant that depends on the size of D, that is, the number of unknown OD pairs in the context of OD demand estimation. The entropy measure is proportional to the log of the determinant of the covariance matrix. By ignoring the constant β and the monotonic logarithm function, we can simplify the entropy-based information measure for the posterior demand estimate as $det(P^+)$. Geometrically, the determinant of a variance covariance matrix can be interpreted as a measure of the volume of a hyperellipsoid for unknown demand variables centered at D^+, that is, the axis directions of the ellipsoid are given by the eigenvectors of P^+, and the lengths of the axes are proportional to the square root of the eigenvalues. The solid ellipsoid of D values satisfying

$$(D - D^+)^T(P^+)^{-1}(D - D^+) \le \chi_n^2(\alpha) \tag{18}$$

has a probability of $1 - \alpha$, where $\chi_n^2(\alpha)$ is the upper $(100 \times \alpha)^{th}$ percentile of a Chi-square distribution with n (i.e. number of OD pairs) degree of freedom. The detailed description of (18) can be found in Bryson and Ho (1975).

$$\begin{array}{c} \begin{array}{cc} d_{1,2} & d_{1,3} \end{array} \\ \begin{array}{c} d_{1,2} \\ d_{1,3} \end{array} \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} \end{array} \qquad \begin{bmatrix} 4 & 1 \\ 1 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$$

$$\lambda_1 = 4, \lambda_2 = 1 \qquad\qquad \lambda_1 = 4.3, \lambda_2 = 0.7 \qquad\qquad \lambda_1 = 1, \lambda_4 = 4$$

$$Tr = 5 \quad Det = 4 \qquad\qquad Tr = 5 \quad Det = 3 \qquad\qquad Tr = 5 \quad Det = 4$$

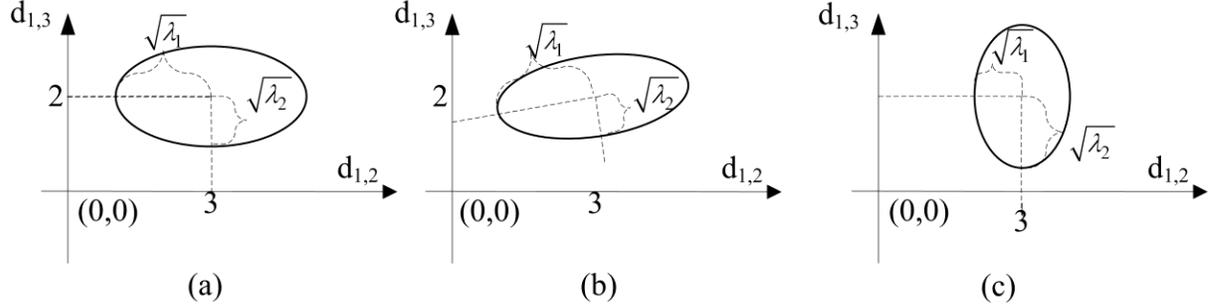(a)   (b)   (c)

Figure 2 Likelihood ellipses for a demand vector of two OD pairs

In comparison, the trace of a covariance matrix, which is used in the mean-square criterion, corresponds to the circumference of the rectangular region that encloses the ellipsoid. Both trace and determinant measures, in fact, use single numerical values to describe the amount of variations in random variables. Moreover, $tr(P^+)$ and $det(P^+)$ are, respectively, the sum and product of the eigenvalues associated with the covariance matrix P^+. Johnson and Wichern (2002) offer detailed discussions on the strengths and weaknesses of the determinant and trace measures as descriptive summaries of random variable variations.

As an illustration, Fig. 2 shows likelihood ellipses for a demand flow vector of two OD pairs with the same mean vector [3, 2] for the OD pair $d_{1,2}$ and $d_{1,3}$, and covariance matrices $\begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 4 & 1 \\ 1 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$, respectively. These three matrices correspond to the same trace value of 5 but different determinants of 4 and 3. Clearly, the trace function only utilizes the variance information, while the determinant function captures the correlation among the random variables. It should be remarked that, as single-valued measures, both the trace and determinant functions are unable to detect and distinguish different correlation structures. For example, Figs. 2-(a) and 2-(c) have the same trace and determinant values.

In this study, the trace of the demand estimation error covariance matrix (i.e. mean-square error) is selected as the information measure for both OD estimation and sensor location problems for the following reasons. First, as shown below, the commonly used GLS OD demand estimator actually optimizes the mean-square error function. Thus, by applying the same mean-square criterion for both the OD estimation and sensor location problems, we can build an internally consistent optimization framework. Second, with closed form formulations for updating the *a posteriori* mean and variance matrix of OD demand estimates, the mean-square criterion is more numerically tractable than the entropy criterion, although the latter form appears to provide better information to describe the covariance matrix.

5. Least Mean-Square OD Demand Estimator

This section discusses how to characterize the propagation of the mean and error covariance matrix in a mean-square estimator for a given sensor location set. The OD demand estimation problem of interest can be stated as follows: given link counts, vehicle identification counts, and prior information on OD trips, we want to find OD trip desires (over a time horizon of interest) that minimize deviations between observed traffic flows and assigned traffic flows (resulting from a traffic assignment process), and deviations

between estimated OD demand flows and the historical demand matrix. In the context of dynamic OD demand estimation, the analyst needs to identify the number of trips for each OD pair in a spatial dimension, and for each departure time interval in a temporal dimension. Without loss of generality, we limit our focus to the problem of static OD demand estimation.

Given *a priori* statistics D^- and P^-, sensor location L^*, as well as measurement vector C, if the linear model (1) adequately describes the mapping between OD demand and observations, and the combined errors ε belong to white noise processes that are uncorrelated with initial demand flows, we can derive an optimal demand estimator with respect to the mean-square criterion $tr(P^+)$ as below. A more comprehensive assessment of the linear mean-square estimator and Kalman filter can be found in Gelb et al. (1974), Lewis (1986) and Wiki (2007).

First, we assume the estimator takes a linear updating form of

$$D^+ = D^- + K(C - \hat{C}) = D^- + K(C - HD^-).$$
(19)

Essentially, the term $C - HD^-$ is the error of the prior estimate, which is also known as the innovation residual or measurement residual. The above equation can be viewed as the update phase in Kalman filtering, in which measurement information from the current stage is used to correct/refine the *a priori* estimate D^- to obtain a new estimate D^+.

Substituting Eq. (19) into $P^+ = E\left[(D - D^+)(D - D^+)^T\right]$ gives

$$P^+ = Cov[D - D^- - K(C - HD^-)].$$
(20)

Substituting C from the linear sensor equation (1) into the above equation, we have

$$P^+ = Cov[D - D^- - K(HD + \varepsilon - HD^-)] = Cov[(I - KH)(D - D^-) - K\varepsilon].$$
(21)

Assuming the measurement error term $\varepsilon \sim N(0, R)$ is uncorrelated with the other terms leads to

$$P^+ = Cov[(I - KH)(D - D^-)] - Cov[K\varepsilon]$$
$$= (I - KH)Cov(D - D^-)(I - KH)^T - KCov(\varepsilon)K^T.$$
(22)
$$= (I - KH)P^-(I - KH)^T - KRK^T$$

where R is the covariance matrix of the measurement errors. The above formula describes the propagation of error covariances for any given updating matrix K. To minimize the trace of the posterior variance-covariance matrix, we need to find an optimal matrix K that satisfies

$$\frac{\partial}{\partial K} tr(P^+) = -2(I - KH)P^- H^T + 2KR = 0.$$
(23)

Solving the above matrix derivative equation for K, the optimal weighting matrix is

$$K = P^- H^T (HP^- H^T + R)^{-1}.$$
(24)

Substituting the above optimal gain matrix back into Eq. (22) leads to

$$P^+ = (I - KH)P^- = P^- - KHP^-.$$
(25)

It is worth noting that many equivalent formulations exist for the above equation, for example,

$$P^+ = \left((P^-)^{-1} + H^T R^{-1} H\right)^{-1}.$$
(26)

To show the equivalency of Eq. (25) and Eq.(26), one can apply the Matrix Inversion Lemma (MIL) as follows:

$$\left((P^-)^{-1} + H^T R^{-1} H\right)^{-1}$$

$$\overset{MIL}{=} P^- - P^- H^T (HP^- H^T + R)^{-1} HP^-.$$

$$= P^- - KHP^-$$

Recall that, the matrix inversion lemma is $(A + XBX^T)^{-1} = A^{-1} - A^{-1}X(X^T A^{-1}X + B^{-1})^{-1}X^T A^{-1}$, where $A = (P^-)^{-1}$, $X = H^T$, and $B = R^{-1}$ in our case.

Eq. (26) is commonly used to illustrate how the information is accumulated and updated in the Kalman filter. This simple additive form updates the prior belief state $(P^-)^{-1}$ by a linear combination of observation information. In the case of a complete lack of historical demand information, we can set $(P^-)^{-1} = 0$. To obtain P^+, Eq. (26) requires the inversion of two $n \times n$ matrices, while Eqs. (24) and (25) only require the inversion of one $m \times m$ matrix. If the number of observations is less than the number of OD pairs, that is,

$m < n$, then the error covariance updating formula based on Eqs. (24) and (25) is more numerically efficient than Eq. (26).

The error covariance updating equations (25) and (26) clearly show the linkage between the *a priori* uncertainty and the *a posteriori* uncertainty. *KH* in Eq. (25) measures the degree of uncertainty reduction due to inclusion of new measurements, while $H^T R^{-1} H$ in Eq. (26) corresponds to the value of additional information from sensors. For the sensor location problem, the most important and useful property of the above linear mean square error updating formula is that the posterior covariance matrix P^+ is independent of the specific value of the measurements *C*, although the conditional mean estimate D^+ is determined by the detailed values of sensor data *C*. As a result, we can calculate P^+ and the expected information gain *before* installing any sensors and obtaining measurements from them.

Let H_k be the kth row vector of matrix *H*, corresponding to the kth measurement. If measurement errors are uncorrelated, then $R = \text{diag}\{r_k\}$ and it is easy to show that

$$(P^+)^{-1} = (P^-)^{-1} + H^T R^{-1} H = (P^-)^{-1} + \sum_k (\frac{1}{r_k} H_k^T H_k) .\qquad(27)$$

Similar to Eq. (26), the above error covariance updating equation is commonly used in the information form of the Kalman filter, in which the information matrix $(P^+)^{-1}$ is recursively updated by the incoming sensor information $\frac{1}{r_k} H_k^T H_k$ from measurement *k*. Recall that the product of $\frac{1}{r_k} H_k^T H_k$ is an $n \times n$ matrix, where *n* is the number of OD pairs. For OD pairs that have flows passing through the selected sensor links, the corresponding link proportion is positive. As a result, different sensor locations could lead to positive values at the different cells in the variance covariance matrix of posterior OD demand estimates, where these cells correspond to the OD pairs traversing the sensor links. In a later section, Fig. 3 illustrates how different single sensors affect the OD estimation uncertainty; Fig. 6 shows how multiple sensors could improve the overall estimation quality. Under the measurement error independence assumption, one can avoid matrix inversion in calculating the gain matrix K_k for the kth measurement by using the sequential updating formula

$$K_k = \frac{P^- H_k^T}{H_k P^- H_k^T + r_k^{-1}} .\qquad(28)$$

Because the sensor location stage needs to incorporate as much sensor information as possible, the importance of a sensor at a certain location depends on the *value of the information/knowledge* that it can provide for the OD estimation stage. In light of Eq. (27), several key factors affect the possible information gain as shown in the following.

Sensor error: A large combined error variance *R* yields a small increase in R^{-1} and accordingly a small reduction in OD demand estimation uncertainty in terms of $(P^+)^{-1}$. On the other hand, sensors with less noise allow us to find an OD demand estimate with greater accuracy.

Sensor coverage: If an OD pair is measured by at least one sensor, then the diagonal element associated with that OD pair in matrix $H^T R^{-1} H$ should be positive. For instance, in the case of point sensors, the diagonal element for OD pair (*i,j*) is $\sum_{l \in L'} [p_{(l)(i,j)}]^2$. If a set of sensors covers all the OD pairs in a network, then we have positive values for all the diagonal elements. If $H^T R^{-1} H$ has full rank, then its inverse exists and we can obtain a unique estimate even without prior information (i.e. $(P^-)^{-1} = 0$). It

should be noted that positive diagonal elements in matrix $H^T R^{-1} H$ do not imply that the matrix has full rank (e.g. $H^T R^{-1} H = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ in a two OD pair case).

Marginal information gain: To obtain a meaningful marginal information gain with respect to the prior OD demand estimate, one should not only focus on finding a sensor matrix with adequate information, but also seek to ensure that the information content of the sensor matrix $H^T R^{-1} H$ reduces the existing demand uncertainty. Consider an OD pair (*i,j*) with high uncertainty in the historical OD demand; that is, the diagonal element for OD pair (*i,j*) in $(P^-)^{-1}$ is small. In this case, even a minor diagonal value for OD pair (*i,j*) in $H^T R^{-1} H$ could generate a considerable uncertainty reduction in the final variance matrix P^+. In contrast, if the prior variance of OD pair (*i,j*) is already very small, a large amount of information from $H^T R^{-1} H$ does not necessarily produce a significant marginal quality improvement.

It should be noted that the variance covariance updating formulation used in the proposed framework is based on two critical assumptions: (1) a linear measurement model for utilizing point and point-to-point sensor data, and (2) unbiased OD demand estimators. It should be remarked that the traffic assignment process that maps OD demands to link and path flows is a highly nonlinear process. This means a bi-level optimization framework is needed, especially for congested networks. Interested readers are referred to Yang (1995), Tavana and Mahmassani (2001) and Tavana (2001) for detailed assessments on bi-level OD demand estimation formulations. Moreover, the estimated link proportions from traffic assignment programs could be affected by numerous error sources, and the mean of the resulting combined errors in the measurement model can be non-zero, leading to a biased demand estimator. In these cases, it is difficult to use a closed-form formula to characterize the estimation error dynamics, while a simulation-based approach in conjunction with a bi-level OD demand estimator offers a valid and feasible alternative to sample the error propagation process and estimate possible information gain, as discussed in Eisenmen et al. (2006). In addition, the simulation approach can explicitly consider negativity constraints, while the proposed variance updating formulation does not impose the non-negativity constraints on the OD demand flow. To take advantage of these two frameworks, the analyst can first use the proposed simplified model to select a set of promising locations, and then use simulation to evaluate the information gains in detail.

6. Scenario-based Sensor Location Model

Similar to many classical Kalman filtering applications (e.g. target tracking), the preceding analysis assumes a deterministic mapping matrix *H*. However, the link proportions matrix used in the OD demand estimation process above is generated from an initial demand estimate $\bar{D}$ through a traffic assignment model. Different seeds of the OD demand table could lead to different network assignment results and link proportion matrices. In other words, the uncertainty associated with the prior demand estimate implies that the mapping matrix *H* in the sensor location problem under consideration is non-deterministic.

Moreover, the sensor location problem should design the sensor location scenario based on a predicted future demand. In medium and long-term urban transportation planning applications, since the future demand forecast depends on a number of uncertain and dynamic factors such as population growth, land use and the other socioeconomic attributes, it is quite difficult to provide a single unbiased and up-to-date demand estimate. If only one inaccurate initial OD demand matrix is used to estimate the link proportion matrix and then determine the final location scenario, the error in the demand input could propagate throughout the estimation process and result in a biased location decision.

If traffic observations are available after the sensors are installed, one can apply a bi-level OD demand estimation procedure to iteratively use the measurements to adjust and update the OD demand estimate so as to approximate a more accurate link proportion matrix. Although the uncertainty associated with the link proportion matrix could be reduced through the aforementioned updating process, a satisfactory sensor location solution should explicitly recognize the estimation errors in both the initial OD demand matrix and the resulting link proportion matrix estimates. To this end, this study uses a set of initial demand instances to generate multiple samples of the link proportion matrix and seeks to find a solution that maximizes the mean value of the sensor information or minimizes the mean uncertainty of the final OD demand estimate under different scenarios. This approach can be viewed as an adaptation of the scenario-based stochastic

13

optimization method, which constructs representative scenarios to characterize uncertain parameters, assigns a probability to each scenario, and then solves a deterministic equivalent to optimize the expected objective function. It should be remarked that, if the mean values of the *a priori* OD demand estimate are used to construct a single instance of the link proportion matrix, then this scheme leads to an "Expected-Value (EV)" solution from the perspective of stochastic optimization. The discussions of the value of the stochastic solution over the EV solution and their properties are beyond the scope of this paper. Interested readers are referred to the book by Birge and Louveaux (1997) for a more general and detailed discussion.

In this study, we first select a set of initial OD demand matrices $\bar{D}_w$ based on the distribution of the historical OD demand estimate characterized by mean D^- and variance P^-, and then perform traffic assignment using $\bar{D}_w$ to generate a set of link proportion samples, e.g. $\hat{p}_{(l)(i,j)(w)}$. By selecting the trace of the covariance matrix as a proxy that captures the information contained in the estimated demand, we can further formulate the sensor location problem as the following:

$$\mathbf{Max} \ z = E_w\left\{tr[(P_w^+)^{-1}]\right\} = E_w\left\{tr[(P^-)^{-1} + H'_w{}^T R^{-1} H'_w + H''_w{}^T R''^{-1} H''_w]\right\} \tag{29}$$

Subject to
Budget constraint:
$$\beta' \sum x'_l + \beta'' \sum x''_l \leq \beta \tag{30}$$
where

H'_w = mapping matrix between OD demand and point measurements, based on the link flow proportion matrix generated from the wth initial OD demand table.

H''_w = mapping matrix between OD demand and point-to-point measurements, based on the link flow proportion matrix generated from the wth initial OD demand table.

x'_l = 1 if a link count sensor (point sensor) is installed on link l, 0 otherwise.

x''_l = 1 if an AVI sensor (point-to-point sensor) is installed on link l, 0 otherwise.

β', β'' = installation and maintenance costs for point sensors and point-to-point sensors.

β = total available budget for building or extending the sensor network.

Essentially, the goal of the above sensor location model is to add sensor information from spatially distributed measurements to minimize the expected uncertainty associated with the OD demand estimate for different initial demand seeds. Note that the AVI link-to-link counts are available if both links have AVI readers installed. Matrix H''_w for AVI counts can be constructed from Eq. (6) – (11), depending on the underlying market penetration rate and the locations of AVI readers.

Obviously, the complexity of solving the proposed problem is determined by the evaluation of the objective function, which can be decomposed into three major steps: (1) calculating the *a posteriori* variance matrix P^+, (2) calculating the inverse of the covariance matrix $(P^+)^{-1}$, and (3) calculating the trace of the inverse. The first step involves two matrix multiplications: $H^T R^{-1}$ and $(H^T R^{-1})H$, H is an ($m \times n$) matrix, R is an ($m \times m$) matrix. The first step has a worst-case complexity of $O(n^2 m)$, and using the Gaussian elimination method to calculate the inverse of matrix P_w^+ leads to an $O(n^3)$ operation. In this study, we use

$$tr(P_w^+)^{-1} = \text{sum of} \ \frac{1}{\text{eigenvalue}(P^+)} \ , \tag{31}$$

where a FORTRAN subroutine EVLRG from the IMSL library (Rice, 1983) is used to compute the eigenvalues in the above equation. When $n < m$, the bottleneck in this process is in obtaining the eigenvalues for the variance matrix in step 3.

For a large-scale network application, two strategies are available to reduce the size of the problem and thus the computational time. First, one can focus on OD pairs with significant volumes. Note that many OD pairs in the OD demand table have zero values. Second, one can aggregate original OD demand zones into a set of super zones within a manageable size, and this strategy is especially suitable for a subarea analysis where many OD zones outside the study area can be consolidated together.

Similar to solving a discrete network design problem, a branch-and-bound search procedure can be used to solve the integer programming problem. In the search tree, each node represents a decision to install a

loop detector or an AVI sensor on a link. As finding the optimal sensor location solution in a large-scale network is computationally prohibitive, this study presents a beam search heuristic algorithm to reduce the size of the search tree. Based on a breadth-first node selection mechanism, a beam search algorithm branches from the nodes level by level. At each level, it keeps only θ promising nodes, and prunes the other nodes permanently in order to limit the total number of nodes to be examined. θ is typically referred to as the beam width, and the total computational time of the beam search algorithm is proportional to the selected beam width.

The notation used in the beam search solution procedure is given as follows:
$\bar{L}'$, $\bar{L}''$ = sets of candidate links for point sensor installation and point-to-point sensor installation, respectively,

ANL= active node list that contains nodes ready for branching,

u = parent node index used in the beam search process,

v', v'' = child node index with a new point sensor or a point-to-point sensor, respectively,

θ = number of promising nodes to be kept at each level,

Each search node u includes the following attributes:
$L'(u)$, $L''(u)$ = sets of sensor links with point observations and point-to-point observations, respectively, at node u,

$z(u)$ = value of the objective function z shown in Eq. (29) at node u,

$t(u)$ = search level index at node u.

Algorithm 1. Beam search

Step 0: (Preprocessing)
Construct a set of initial OD demand tables, generate related P^-_w H'_w and H''_w for the wth initial OD demand table.
Determine the candidate point and point-to-point sensor sets $\bar{L}'$ and $\bar{L}''$.

Step 1: (Initialization)
$ANL=\varnothing$. Create the root node u with $L'(u) = L''(u) = \varnothing$, $t(u) = 0$. Insert the root node into the ANL.

Step 2: (Stopping criterion)
Terminate and output the best feasible solution under one of the following conditions:
i)   if all of the active nodes in ANL have been visited,
ii)   the number of active nodes in memory is exceeded.

Step 3: (Node generation and evaluation)
For each node u at level t in ANL, remove it from ANL, and generate child nodes as follows.

Scan through $\bar{L}'$; if a link l in $\bar{L}'$ is not in $L'(u)$, then generate a new node v', where $L'(v') = L'(u) \cup l$, $L''(v') = L''(u)$ and $t(v') = t(u)+1$.

Scan through $\bar{L}''$; if a link l in $\bar{L}''$ is not in $L''(u)$, then generate a new node v'', where $L''(v'') = L''(u) \cup l$, $L'(v'') = L'(u)$ and $t(v'') = t(u)+1$.

For each newly generated node v, calculate the objective function $z(v)$ as the mean value of the information based on different initial OD demand tables. If the budget constraint is satisfied for a newly generated ode, add it into the ANL.

Step 4: (Node filtering)

Select θ best nodes from the ANL in the search tree, and go back to Step 2.

To estimate the overall computational complexity of the proposed beam search algorithm, we can calculate the total number of nodes to be examined in the search tree. If only point sensors are considered, the budget constraint allows q point sensors to be installed. Each level in the search tree corresponds to a decision to install one sensor; the final search tree should include a total of $q+1$ levels, while level 0 is the dummy root node. Essentially, the total computational time of the beam search algorithm is determined by the number of sensors to be installed, the beam width θ as well as the size of the candidate sensor links.

7. Numerical Examples

In this study, we first use a series of examples in a 6-node network to demonstrate the proposed methodology, and then consider more realistic settings based on a simplified network in Irvine, CA and a large-scale network in Research Triangle Park, NC. In Fig. 3, OD pair 1 travels from node 1 to node 2 and OD pair 2 travels from node 1 to node 3. The first OD pair has two routes. 70% of the flow travels along path {1, 4, 5, 2} while the remaining 30% travels along path {1, 4, 6, 5, 2}. Both OD pairs have a flow volume of 20 units. Let us assume $P^- = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$, meaning that OD pair 1 has a larger *a priori* variance than OD pair 2. We first consider three alternative one-sensor location scenarios that do not involve estimation of the route choice percentages: scenario (a) covers OD pair 1, scenario (b) covers OD pair 2, and scenario (c) covers both OD pairs. The measurement matrices for the above three cases are, respectively, [1,0], [0,1] and [1,1], where the row in matrix H refers to a measurement and the column refers to an OD pair.

base case

$P^- = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$

$tr = 5 \quad det = 4$

(a) one sensor for OD pair (1->2)

$R = 1$

$H = [1,0] \quad H^T R^{-1} H = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

$P^+ = \begin{bmatrix} 0.8 & 0 \\ 0 & 1 \end{bmatrix}$

$tr = 1.8 \quad det = 0.8$

(b) one sensor for OD pair (1->3)

$R = 1$

$H = [0,1] \quad H^T R^{-1} H = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

$P^+ = \begin{bmatrix} 4 & 0 \\ 0 & 0.5 \end{bmatrix}$

$tr = 4.5 \quad det = 2$

(c) one sensor for both OD pairs

$R = 2^2$

$H = [1,1] \quad H^T R^{-1} H = \begin{bmatrix} 0.25 & 0.25 \\ 0.25 & 0.25 \end{bmatrix}$

$P^+ = \begin{bmatrix} 2.22 & -0.44 \\ -0.44 & 0.89 \end{bmatrix}$

$tr = 3.11 \quad det = 1.78$
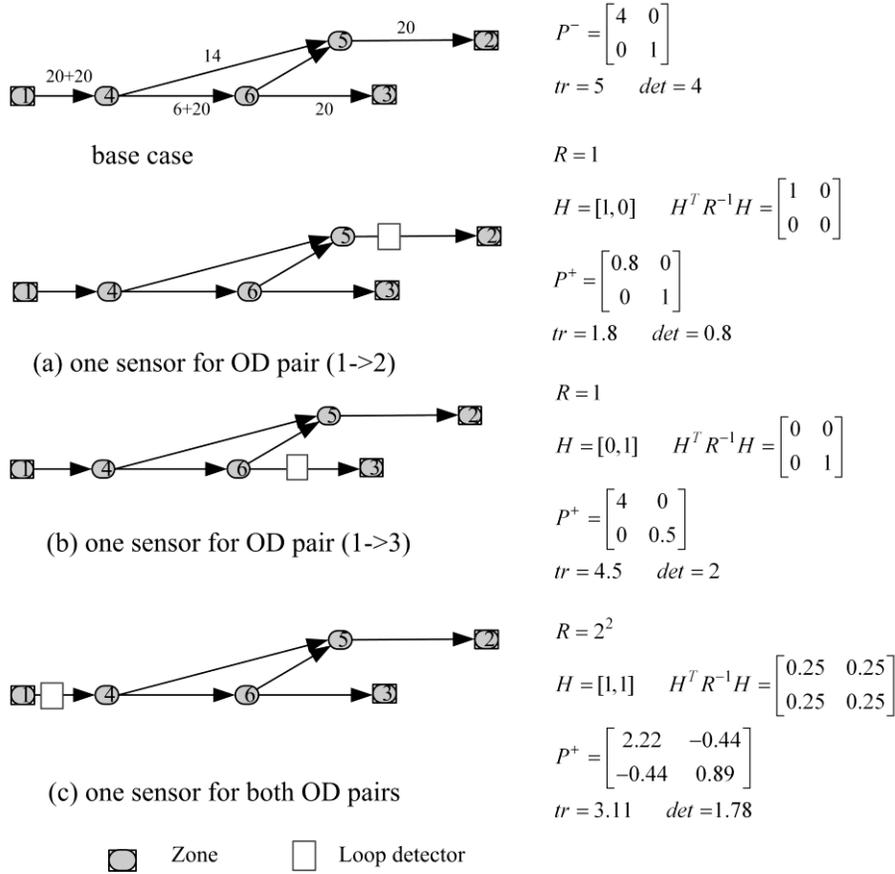
Zone   Loop detector

Figure 3 Examples of single point sensor location scenarios

For simplicity, we assume the standard deviation of the measurement error for a point sensor is 5% of the corresponding true flow volume. Fig. 3 shows the resulting sensor information matrices $H^T R^{-1} H$, posterior variance matrices, traces and determinates. It is interesting to note that the sensor information matrix for scenario (c) is $\begin{bmatrix} 0.25 & 0.25 \\ 0.25 & 0.25 \end{bmatrix}$, which contains considerable correlation between the estimate errors of the two OD pairs. The trace values for the three scenarios are 1.8, 4.5 and 3.11, which quantify different magnitudes of uncertainty reduction compared to the original $tr(P^-)$ of 5. Specifically, scenario (a) allocates the scarce sensor resource to the critical OD pair with the largest variance, producing a better total uncertainty reduction than scenarios (b) and (c). In contrast, although scenario (c) fully covers both OD flows, it does not produce the maximum information gain in terms of both trace and determinant criteria. Actually, even if we assume $R=1$ in scenario (c), which is identical to scenarios (a) and (b), the resulting $P^+ = \begin{bmatrix} 1.33 & -0.67 \\ -0.67 & 0.83 \end{bmatrix}$ and the trace and determinant are 2.16 and 0.66, respectively. These two values are still larger than the measures in scenario (a). The examples above reveal that the sensor location problem is more complicated than a simple network coverage problem, and that it is important to recognize and capture the uncertainties in the prior OD demand estimates and the available measurements.

$$R = (0.7)^2 I$$
$$H = [1,0] \qquad H^T R^{-1} H = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$
$$P^+ = \begin{bmatrix} 0.8 & 0 \\ 0 & 1 \end{bmatrix}$$
(d) error-free link proportion of 0.7 $\qquad tr = 1.8 \qquad det = 0.8$

$$R = (0.7 + 0.3)^2 I$$
$$H = [1,0] \qquad H^T R^{-1} H = \begin{bmatrix} 0.49 & 0 \\ 0 & 0 \end{bmatrix}$$
(e) STD of link proportion estimation errors $\qquad P^+ = \begin{bmatrix} 1.35 & 0 \\ 0 & 1 \end{bmatrix}$
(from traffic assignment) = 0.3 $\qquad tr = 2.35 \qquad det = 1.35$

$$R = (0.7 + 0.1)^2 I$$
$$H = [1,0] \qquad H^T R^{-1} H = \begin{bmatrix} 0.765 & 0 \\ 0 & 0 \end{bmatrix}$$
(f) STD of link proportion estimation errors $\qquad P^+ = \begin{bmatrix} 0.98 & 0 \\ 0 & 1 \end{bmatrix}$
(from AVI data) = 0.1 $\qquad tr = 1.98 \qquad det = 0.98$
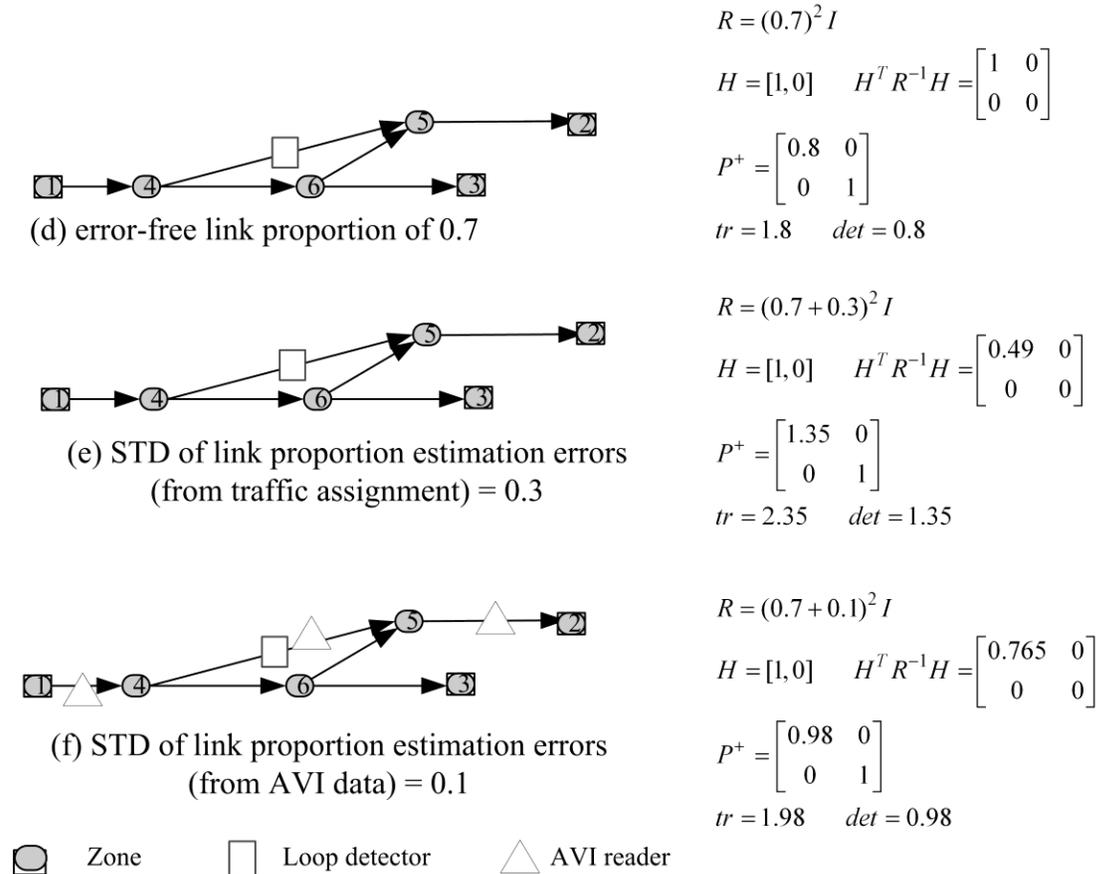
Zone   Loop detector   AVI reader
Figure 4 Examples of single point sensor location scenarios with route choice

Fig. 4 considers a more general case, in which the route choice percentage (i.e. link proportion) associated with link (4, 5) needs to be estimated from either traffic assignment programs or AVI data. Assuming there is no error in the flow proportion estimate, scenario (d) constructs a lower bound for the uncertainty reduction attainable by locating a point sensor on link (4, 5). It should be noted that the above

benchmark scenario gives exactly the same posterior variance and covariance matrix as scenario (a) in Fig. 3. This indicates that if link flow proportions are error-free and the standard deviation of the measurement error is proportional to the link volume, then locating a sensor on a low-volume link can still extract significant OD demand information. In scenario (e), we assume that the standard deviation of measurement errors in the link proportion matrix (due to errors from traffic assignment) is 0.3. Clearly, a considerable amount of information loss is observed by comparing scenarios (d) and (e), as the corresponding trace value increases from 1.8 to 2.35. We assume that, by utilizing AVI counts to calibrate the link proportion, scenario (f) significantly reduces the standard deviation of the measurement error to 0.1. As a result, the overall OD demand estimation accuracy is considerably improved, while the associated trace value is only greater than the lower bound in scenario (d) by 10%.
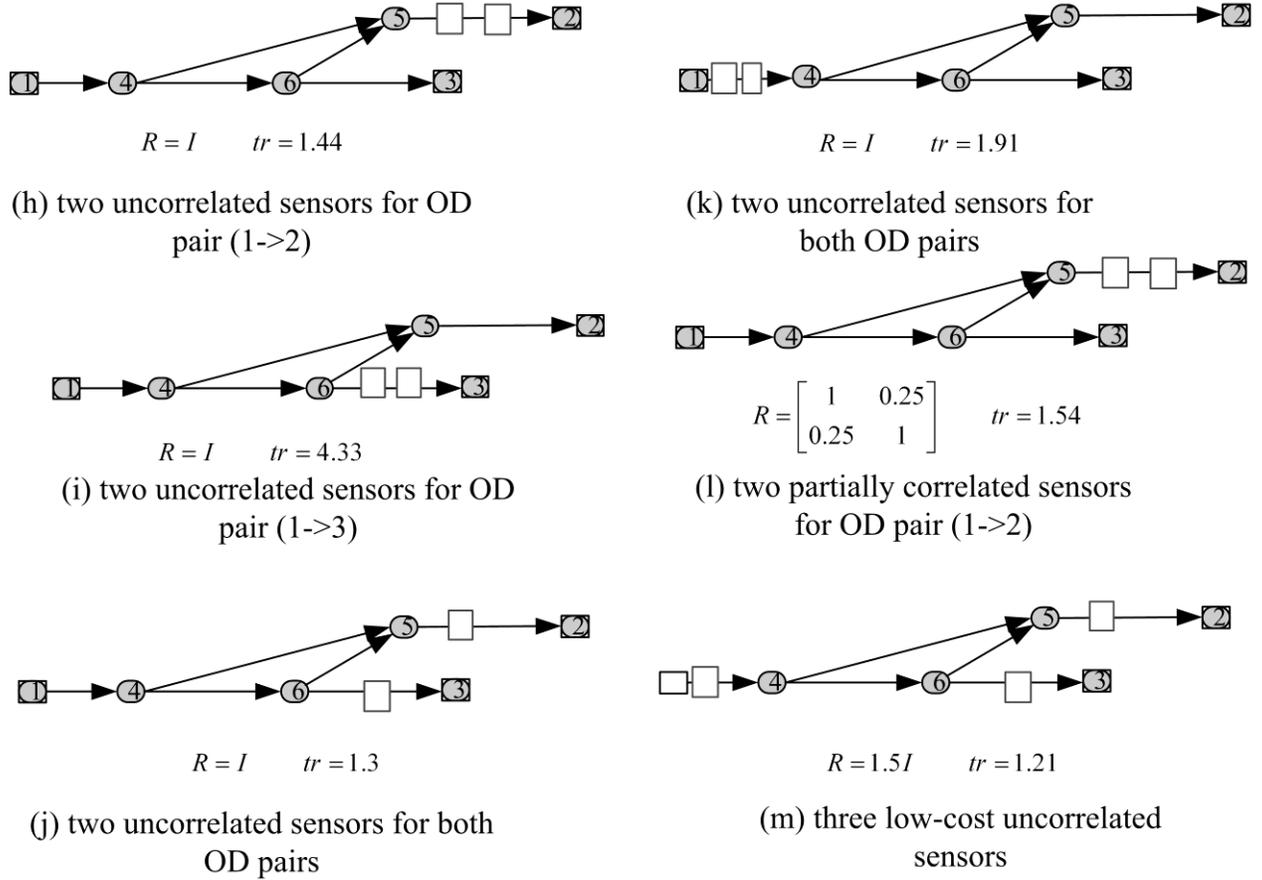

$R = I \qquad tr = 1.44$

(h) two uncorrelated sensors for OD pair (1->2)

$R = I \qquad tr = 1.91$

(k) two uncorrelated sensors for both OD pairs

$R = I \qquad tr = 4.33$

(i) two uncorrelated sensors for OD pair (1->3)

$R = \begin{bmatrix} 1 & 0.25 \\ 0.25 & 1 \end{bmatrix} \qquad tr = 1.54$

(l) two partially correlated sensors for OD pair (1->2)

$R = I \qquad tr = 1.3$

(j) two uncorrelated sensors for both OD pairs

$R = 1.5I \qquad tr = 1.21$

(m) three low-cost uncorrelated sensors

Figure 5 Examples of multiple point sensor location scenarios

Scenarios (h) to (l) in Fig. 5 examine a wide range of scenarios with two sensors. Scenarios (h) and (i) still focus on individual OD pairs, while scenarios (j) and (k) cover both OD pairs. Scenario (l) assumes that the measurement errors between the two sensors are correlated to each other. Overall, scenario (j) is able to systematically balance the information needs of the different OD pairs. Interestingly, while both scenarios (j) and (k) cover two OD pairs, scenario (j) dominates scenario (k) in the sense that the former provides a smaller value for each element in the variance and covariance matrix. Comparing scenarios (h) and (l), as expected, correlated measurement errors in the latter case lead to some information loss. Scenario (m) considers three inexpensive roadside sensors with relatively large levels of measurement noise. If the total cost of these three sensors is lower than the cost of the two expensive and accurate sensors in scenario (j), then scenario (m) is actually a more cost-efficient and reliable alternative. Again, these examples show the advantage of the proposed methodology in systematically evaluating the trade-off between the accuracy of individual sensors and network-wide reliability.
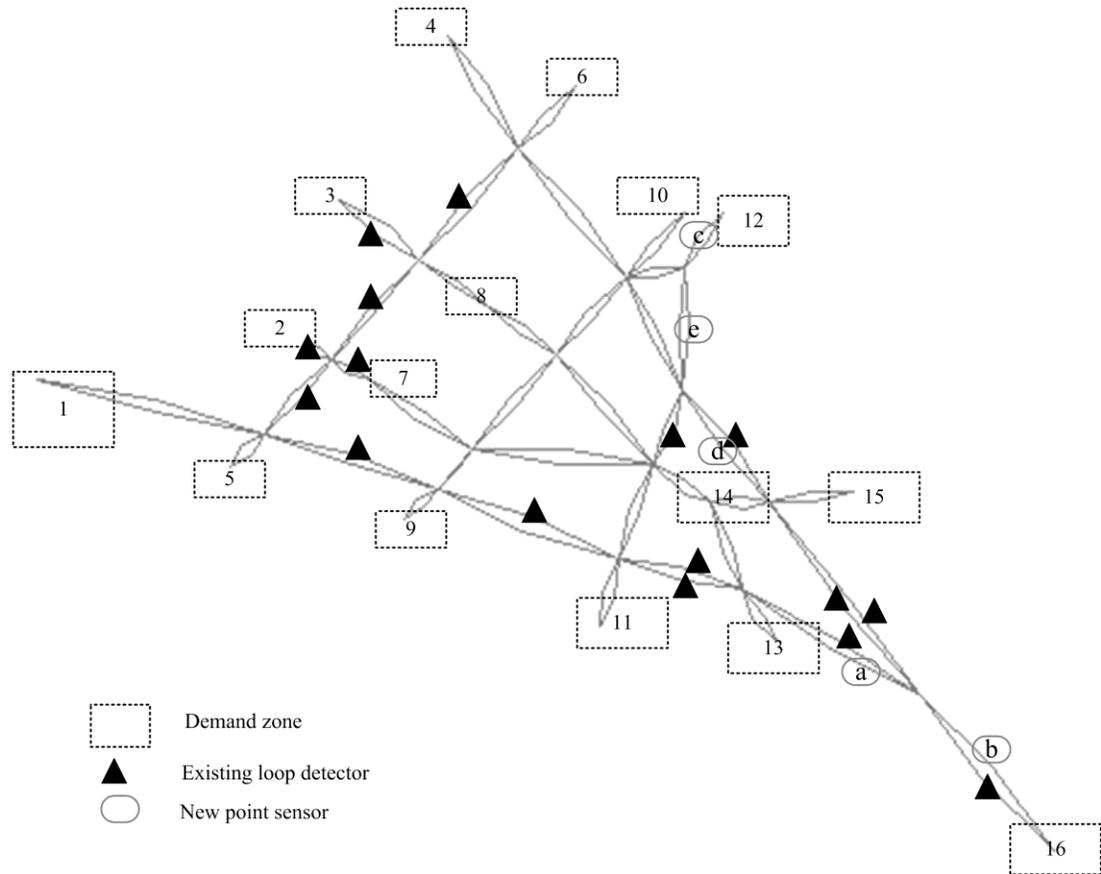
18

(n) one AVI reader

$$R = 0.1^2$$

$$H = [0.1, 0.1] \quad H^T R^{-1} H = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$P^+ = \begin{bmatrix} 1.33 & -0.67 \\ -0.67 & 0.83 \end{bmatrix}$$

$$tr = 2.16 \quad det = 0.66$$

(o) two AVI readers

$$R = \begin{bmatrix} 0.1^2 & \\ & 0.05^2 \end{bmatrix}$$

$$H = \begin{bmatrix} 0.1 & 0.1 \\ 0.1 & 0 \end{bmatrix} \quad H^T R^{-1} H = \begin{bmatrix} 5 & 1 \\ 1 & 1 \end{bmatrix}$$

$$P^+ = \begin{bmatrix} 0.21 & -0.10 \\ 0.73 & 0.55 \end{bmatrix}$$

$$tr = 0.76 \quad det = 0.11$$

(p) three AVI readers

Zone △ AVI reader

$$R = \begin{bmatrix} 0.1^2 & & \\ & 0.05^2 & \\ & & 0.05^2 \end{bmatrix}$$

$$H = \begin{bmatrix} 0.1 & 0.1 \\ 0.1 & 0 \\ 0 & 0.1 \end{bmatrix} \quad H^T R^{-1} H = \begin{bmatrix} 5 & 1 \\ 1 & 5 \end{bmatrix}$$

$$P^+ = \begin{bmatrix} 0.20 & -0.03 \\ -0.03 & 0.17 \end{bmatrix}$$

$$tr = 0.37 \quad det = 0.03$$

Figure 6 Examples of AVI sensor location scenarios

Scenarios (n-p) in Fig. 6 briefly demonstrate the information gain by installing AVI sensors. These cases assume that the market penetration rate of tagged vehicles is 10%, and the standard deviation of the combined error terms is 2.5% of the tagged vehicular flow volume. By progressively adding AVI sensors in the network, we are able to significantly reduce the uncertainty of the OD demand estimates.

Figure 7 Irvine simplified network

A simplified Irvine, CA test bed network is shown in Fig. 7. It includes 16 OD zones, 31 nodes and 80 directed links (32 freeway links and 48 arterial links). The time of interest is the morning peak period (6:30 am − 8:30 am). A historical static traffic OD demand table is used to construct an *a priori* mean estimate. The Irvine test bed network has a triangular structure, while the OD demand flows among zones 1, 4 and 16 account for about 36% of total OD trips. As the estimation variance and covariance matrix is unavailable in the planning data, we assume that the standard deviation of the *a priori* demand variance is 30% of the corresponding demand volume in the historical demand table. Actual traffic link counts were measured on 10 freeway links and 6 arterial links in May 2001, but no real-world AVI traffic measurements were recorded. Essentially, the Irvine network data set exemplifies a common situation in many applications, where an outdated traffic OD demand matrix is available, and only a set of OD pairs is covered by detectors installed in the network. We assume the standard deviation of link flow estimation errors is 5% of the assigned link volume.

Table 1 OD Pairs with the 10 Highest Variances under the Existing Sensor Location Scheme

Origin	Destination	Estimated hourly volume based on link counts	STD of posterior variance for existing sensors	% reduction in variance due to exiting sensors	# of existing sensors covered
1	16	4000	364.65	69.61	2
16	4	6820	347.45	83.02	2
12	4	1152	345.60	0	0
4	16	2480	286.84	61.45	2
16	12	832	247.79	0.73	2
15	4	880	218.57	17.21	1
12	16	680	197.36	3.25	2
16	1	4800	195.26	86.44	4
4	15	604	181.20	0	0
4	12	444	133.20	0	0

Table 2 Optimal sensor location plan and estimation quality improvement

Sensor type	Point sensors					AVI sensors		
Sensor location plan	a	a, b	a,b,c	a,b,c,d	a, b,c,d,e	zone 1, 16	zone 1,4,16	zone 1,4,12,16
Total uncertainty reduction (%)	5.87	12.96	18.43	22.15	23.88	10.95	19.67	26.17

For OD pairs with the 10 largest variances in the posterior OD demand estimate (with existing sensors), Table 1 shows several important statistics, namely, the estimated hourly volume, the standard deviation of the *a posteriori* variance and the percentage reduction with respect to the standard deviation of the *a priori* variance. The sensor coverage for each OD pair is measured in terms of the number of point detectors that intercept at least 10% of the corresponding OD demand flow. Obviously, although large-volume OD pairs, for example (1,16), (16,4), and (16,1), have already been covered by a certain number of detectors, they are still associated with large estimation errors due to the magnitude of the original uncertainty. Overall, in this test data set, 7 of the 10 OD pairs with large variance have been already covered by sensors. On the other hand, for those OD pairs with minor volumes, which are not covered by any sensors, the associated variances could be considerably small. In this case, the question is, do we locate additional sensors to cover the unobserved minor OD pairs or still focus on the covered OD pairs with large uncertainties? In fact, as illustrated previously in the small network case, maximizing the sensor network coverage does not necessarily lead to the largest improvement in the overall OD demand estimation quality.

Suppose several point sensors and AVI sensors are added in addition to the existing point sensors. Table 2 lists optimal sensor location plans and the corresponding percentages of uncertainty reduction for different numbers of sensors. The uncertainty is measured by the square root of the total trace value for the whole OD matrix. The experiments indicate that in a two-sensor plan, a significant improvement is obtained by locating sensors at points *a* and *b* to intersect the OD pairs (1,16) and (16,4), which have been covered by two sensors but which still have significant uncertainty. In the optimal three-point-sensor plan, an additional sensor is placed on link *c* to detect the OD flows generated by zone 12, especially the uncovered OD pair (12, 4) with large volume. The optimal 5-point-sensor plan essentially covers the OD pairs associated with the 5 largest variances.

In the next set of experiments, we locate AVI detectors on the entry/exit links of OD zones in order to use the AVI OD estimation equation (6). The market penetration rate of tagged vehicles is assumed to be 5%, and the standard deviation of the combined error terms is 5% of the tagged vehicular flow volume. As shown in Table 2, the optimal AVI reader location plans seek to cover the zones with large traffic attraction/ production rates. Particularly, by locating AVI detectors in OD zones 1, 4, 12 and 16, we are able to cover 12 OD pairs in a virtual sensor network, which includes 8 of the 10 major OD pairs with large variance. In addition, the results also show that the marginal reduction in uncertainty afforded by the AVI

detectors is considerably greater than that provided by the point sensors.

The remaining experiments aim to further exploit the performance of the proposed algorithm in a large-scale network. Extracted from the Triangle Regional Travel Demand Model, NC, this study network includes 758 OD demand zones (i.e. 758×758 OD pairs), 3460 nodes and 8562 links and around 256 thousand passengers traveling in this area within a 2-hour morning peak horizon. As shown in Fig. 8, this network is served by two major interstate freeways (I-85 and I-40) and several secondary highways.
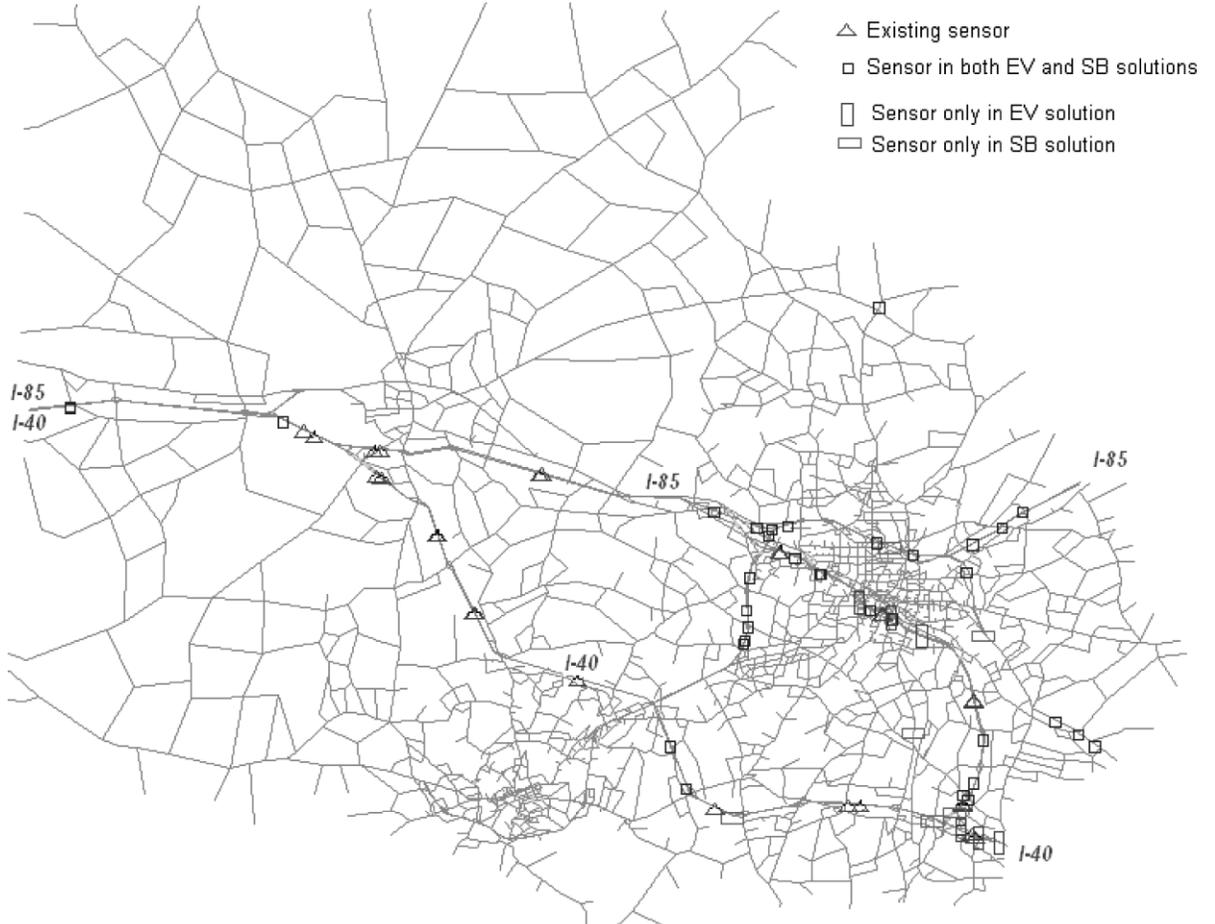
Figure 8 Research Triangle Park sub-network (NC) and sensor locations in expected-value (EV) and scenario-based (SB) solutions

As showed by Cascetta and Nguyen (1988), if the number of possible destinations is large and the sampling rate is sufficiently low, the number of sampled trips $y_{(i,j)}$ can be modeled as a Poisson probability distribution with parameter $\gamma \times d_{(i,j)}$, where γ is the survey sampling rate. In this study, we assume the prior OD demand estimate is calculated from the survey samples (i.e. $\hat{d}_{(i,j)} = \frac{y_{(i,j)}}{\gamma}$) with a sample rate of 10%, so the corresponding estimate variance is $Var(\hat{d}_{(i,j)}) = \frac{Var(y_{(i,j)})}{\gamma^2} = \frac{\gamma \times d_{(i,j)}}{\gamma^2} = \frac{d_{(i,j)}}{\gamma}$.

For simplicity, the covariance between OD pairs in the following experiments is assumed to be zero, and interested readers are refereed to Cascetta and Nguyen (1988) for a more elaborate treatment on the variance covariance matrix of the prior demand estimate under different sampling strategies (e.g. origin-based simple random sampling). All the experiments are performed on a computer system equipped with an

Intel Pentium M 2.13 GHz CPU and 2GB RAM.

To reduce the computational complexity of this large-scale sensor location problem, a beam width of 10 is used in the beam search algorithm and we only select a subset of OD pairs in a decreasing order of OD volume to construct the OD demand covariance matrix used in solution evaluation. The selected OD pairs are defined as critical OD pairs in the following analysis. We first examine the total demand coverage and average computational time for evaluating a single node in the search tree as a function of the number of critical OD pairs.

As shown in Fig. 9, focusing on 1000 and 3500 OD pairs with largest OD volumes could cover about 20% and 40%, respectively, of the total OD demand in the study area. On the other hand, the average computational time significantly increases as more OD pairs are considered in the variance-covariance matrix. Specifically, if 1000 critical OD pairs are selected, evaluating each node takes about 0.1 seconds. There are 1739 links in the study network with at least 2 lanes, and if we consider all of them as candidate sensor links, then determining a new point sensor location in the proposed beam search algorithm requires evaluation of approximately (beam width)×(number of candidate links) nodes, i.e. 10×1739×0.1 sec or 29 min.
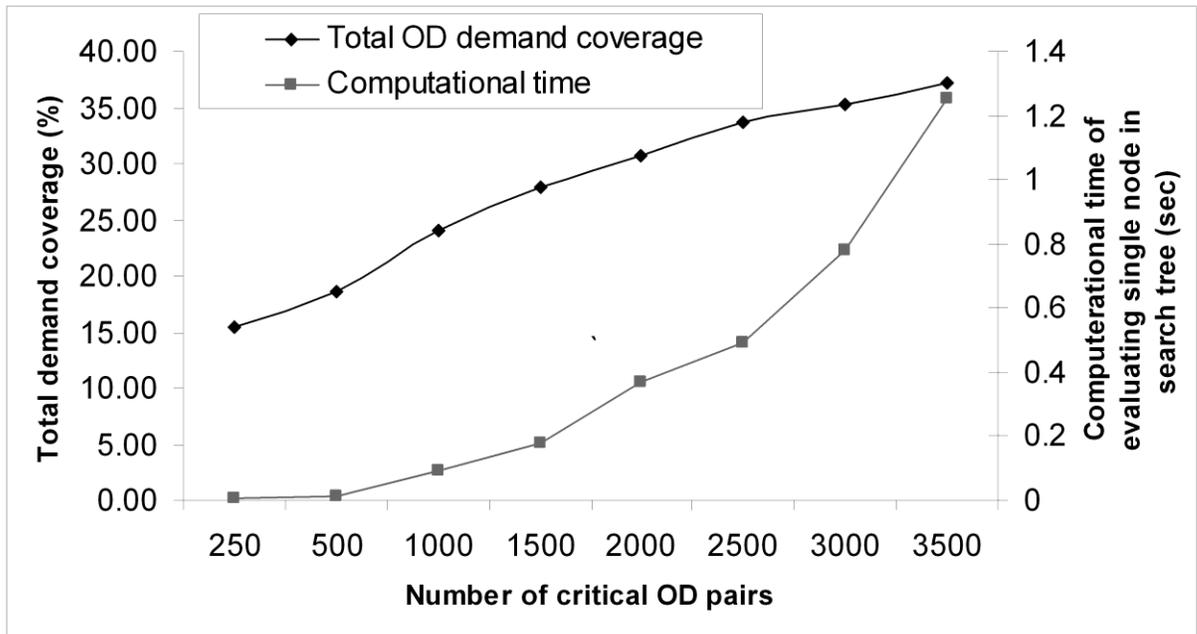

Figure 9 Demand coverage and computational time as a function of number of critical OD pairs

This preliminary study uses the existing 28 counting stations in this study area as the base case, and considers the following two solution strategies in a point-sensor only deployment plan: (1) the Expected-Value (EV) solution that uses the base demand table obtained from the Triangle Regional Travel Demand model, (2) the Scenario-Based (SB) stochastic solution that evaluates three scenarios with three different demand levels, namely 80%, 100% and 120%, applied to the base demand table. It should be remarked that an ideal stochastic optimization procedure requires a large number of scenarios to representatively describe the probability sample space. Given limited computational resources, this research only selects 3 scenarios to illustrate the importance of the stochastic optimization methodology.
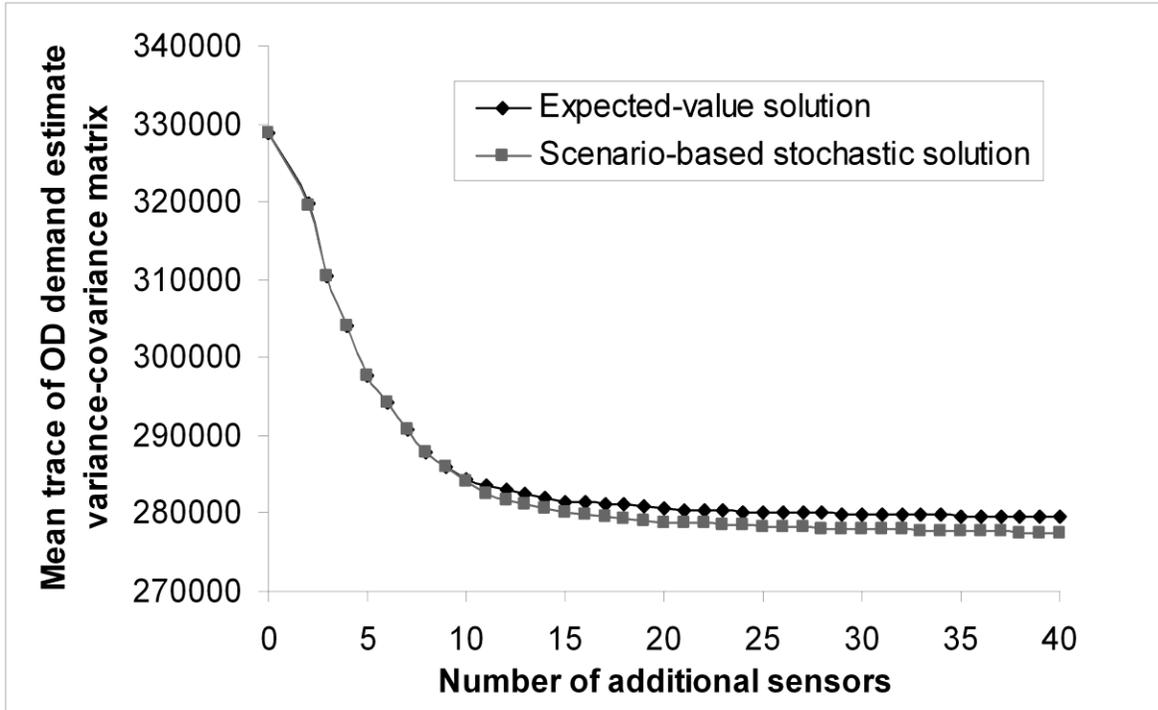
Figure 10. Uncertainty reductions as a function of the number of additional sensors

By comparing the expected-value solution and scenario-based solution, interestingly, we find the two solutions (in terms of sensor locations) are exactly the same for the first 10 additional sensors, and these sensors are located at freeway/highway links to cover major OD pairs. Both solutions lead to dramatic total uncertainty reductions in terms of the mean trace of the OD demand estimate variance-covariance matrix (the mean is computed based on the three scenarios with different demand levels). It should be noted that after 10 new sensors are added (i.e. 28+10=38 sensors in total), the marginal value of additional sensors drops significantly because most of the high volume OD pairs have been covered and extra sensors can only further intercept minor OD pairs with relatively low volumes.

EV and SB solutions begin to differ slightly from each other when more than 10 new sensors are considered. As shown in Fig. 8, the EV solution tends to place sensors on major freeway corridors which carry most of the critical OD flows in the base OD demand matrix. In contrast, the SB solution starts to locate a number of sensors on secondary highway and major arterial corridors, as the high demand scenario diverts considerable traffic from the major freeway system to alternative routes in order to reach user equilibrium in the network. As shown in Fig. 10, in terms of the mean trace reduction, the SB solution consistently outperforms the EV solution by around 4% after installing 20 additional sensors. More importantly, the EV solution with 40 additional sensors is still worse than the scenario-based solution with 20 additional sensors, which highlights the need to systematically consider the impact of uncertain demand estimates in traffic sensor network design.

8. Conclusions

With a particular emphasis on the OD demand estimation problem, this paper proposed an information-theoretic sensor location model that aims to maximize the expected information gain from a set of point and point-to-point sensors in a network, subject to budget constraints. Based on a linear measurement model, the new model explicitly takes into account several important sources of error in the OD demand estimation process, such as the uncertainty associated with prior demand estimates, measurement errors, and the estimation errors for link flow proportions (if they are generated from traffic assignment programs). After thoroughly examining several possible measures of information gain, this paper selected a mean-square minimization criterion to construct a joint sensor location and OD demand

estimation framework. In particular, the first sensor design stage determines sensor locations and sensor types, and the subsequent OD demand estimation stage infers OD demand flows from available sensor measurements. This unified framework ensures that these two stages are internally consistent in the sense that they share the same objective function and the same formulation for updating the estimate mean and covariance matrix. This study proposes a scenario-based stochastic optimization procedure to explicitly recognize the estimation error associated with the initial OD demand tables that are used to generate link proportions. A heuristic beam-search search algorithm is used to solve the combinatorial sensor selection problem. A number of illustrative examples and case studies based on real-world traffic networks are used to demonstrate the effectiveness of the proposed methodology.

Our on-going research includes several major directions. First, this study only focuses on the sensor design problem for OD demand estimation applications, and a natural extension is to assist sensor design decisions for other network-wide traffic state estimation domains, such as measuring and forecasting point-to-point travel times, path flows and route delays. Second, the proposed sensor location model is specifically designed for the minimum variance criterion, and our future work should consider other crucial factors for real-world sensor network design, such as minimizing maximum identification errors. Furthermore, the offline model developed in this study could be extended to a real-time traffic state estimation and prediction framework with agile sensors, by incorporating demand process models that can evolve over time. Because a large-scale network application increases the computational complexity quite steeply, it is undoubtedly critical to develop efficient and effective approximation and heuristic schemes along this research direction. The numerical experimental results in this study also demonstrate several computational challenges in applying the proposed information-based sensor location strategy in large-scale real-world networks, and these challenges call for more in-depth research for developing efficient and effective heuristic methods.

References

Antoniou, C., M. Ben-Akiva, H. N. Koutsopoulos. 2004. Incorporating automated vehicle identification data into origin-destination estimation. *Journal of the Transportation Research Board*, 1882, 37-44.

Asakura, Y., E. Hato, M. Kashiwadani, 2000. OD matrices estimation model using AVI data and its application to the Han-Shin expressway network. *Transportation*, 27(4), 419-438.

Ashok, K., M. Ben-Akiva. 2000. Alternative approaches for real-time estimation and prediction of time-dependent origin-destination flows. *Transportation Science*, 34(1), 21-36.

Bell, M.G.H. 1985. Variances and covariances for origin-destination flows when estimated by log-linear models. *Transportation Research Part B*, 19, 497-507.

Bianco, L., G. Confessore, P. Reverberi. 2001. A network based model for traffic sensor location with implications on O/D matrix estimates. *Transportation Science*, 35(1), 50-60.

Bierlaire, M. 2002. The total demand scale: A new measure of quality for static and dynamic origin-destination trip tables. *Transportation Research Part B*, 36, 755-851.

Birge, J.R., F. Louveaux. 1997. *Introduction to Stochastic Programming*. Springer Verlag, London.

Bryson, A.E., Y.-C. Ho.1975. *Applied Optimal Control*, Hemisphere, New York.

Cascetta, E. 1984. Estimation of trip matrices from traffic counts and survey data: A generalized least squares estimator. *Transportation Research Part B*, 22, 289-99.

Cascetta, E., S. Nguyen. 1988. A unified framework for estimating or updating origin/destination matrices from traffic counts. *Transportation Research Part B*, 22, 437-455.

Chen, A., P. Chootinan, W. Recker. 2005. Examining the quality of synthetic O-D trip table estimated by path flow estimator. *Journal of Transportation Engineering*, 131, 506-513.

Chung, I.-H. 2001. An optimum sampling framework for estimating trip matrices from day-to-day traffic counts. Ph.D. Thesis, University of Leeds.

Denzler, J., C. M. Brown. 2002. Information theoretic sensor data selection for active object recognition and state estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(2), 145-157.

Dixon, M.P. 2000. Incorporation of automatic vehicle identification data into synthetic OD estimation process, Ph.D. Dissertation, Texas A&M University, College Station, TX.

Dixon, M. P., L. R. Rilett. 2002. Real-time OD estimation using automatic vehicle identification and traffic count data. *Journal of Computer-Aided Civil and Infrastructure Engineering*, 17(1), 7-21.

Ehlert, A., M.G.H. Bell, S. Grosso. 2006. The optimisation of traffic count locations in road networks *Transportation Research Part B* 40, 460-479.

Eisenman, S.M., G.F. List. 2004. Using probe data to estimate OD matrices. *Seventh International IEEE Conference on Intelligent Transportation Systems* (on CD), Washington, DC. USA.

Eisenman, S.M., X. Fei, X. Zhou, H.S. Mahmassani. 2006. Number and Location of Sensors for Real-Time Network Traffic Estimation and Prediction: A Sensitivity Analysis. *Journal of Transportation Research Board*, 1964, pp. 260-269.

Gelb, A. 1974. *Applied Optimal Estimation*. MIT Press, Cambridge, MA.

Hintz, K.J., E.S. McVey. 1991. Multi-process constrained estimation, *IEEE Transactions on Systems, Man, and Cybernetics*, 21 (1), 434-442.

Johnson, R., D. Wichern. 2002. *Applied Multivariate Statistical Analysis* (5th ed.). Prentice Hall. New York.

Lam, W. H. K., H. P. Lo. 1990. Accuracy of O-D estimates from traffic counting stations. *Traffic Engineering and Control*, 31, 358-367.

Lee, J. 1998. Constrained maximum-entropy sampling. *Operations Research*, 46, 655-664.

Lewis, R. 1986. *Optimal Estimation with an Introduction to Stochastic Control Theory*, John Wiley & Sons.

List, G.F., M. Turnquist. 1994. Estimating truck travel patterns in urban areas. *Transportation Research Record*, 1430, 1-9.

Maher, M.J. 1983. Inferences on trip matrices from observations on link volumes: A Bayesian statistical approach. *Transportation Research Part B*, 17, 435-447.

Rice, J. R. 1983. Numerical Methods, Software, and Analysis: IMSL Reference Edition, McGraw-Hill, New York.

Sherali, H.D., J. Desai, H. Rakha. 2006. A discrete optimization approach for locating Automatic Vehicle Identification readers for the provision of roadway travel times. *Transportation Research Part B*, 40, 857-871.

Tavana, H. 2001. Internally-consistent estimation of dynamic network origin-destination flows from intelligent transportation systems data using bi-level optimization. Ph.D. Dissertation, The University of Texas at Austin.

Tavana, H., H. S. Mahmassani. 2001. Estimation of dynamic origin-destination flows from sensor data using bi-level optimization method. Transportation Research Board CD-ROM Paper Preprints (TRB Paper No. 01-3241), National Research Council, Washington, D.C.

Van der Zijpp, N.J. 1997. Dynamic OD-matrix estimation from traffic counts and automated vehicle identification data. *Transportation Research Record*, 1607, 87-94.

Van Zuylen, H.J., L.G. Willumsen. 1980. The most likely trip matrix estimated from traffic counts. *Transportation Research Part B*, 14, 281-293.

Wiki. 2007. Kalman filter. Anonymous. http://en.wikipedia.org/wiki/Kalman_filter (accessed on June 16, 2007).

Yang, H., Y. Iida, T. Sasaki. 1991. An analysis of the reliability of an origin-destination trip matrix estimated from traffic counts. *Transportation Research Part B*, 25, 351-363

Yang, H. 1995. Heuristic algorithms for the bi-level origin-destination matrix estimation algorithm. *Transportation Research Part B*, 29, 231-242.

Yang, H., J. Zhou. 1998. Optimal traffic counting locations for origin destination matrix estimation. *Transportation Research Part B*, 32. 109-126.

Yang, H., C. Yang, L. Gan. 2006. Models and algorithms for the screen line-based traffic-counting location problems. *Computers and Operations Research*, 33(3), 836-858.

Yim, K. N., W. H. K. Lam.1998. Evaluation of count location selection methods for estimation of OD matrices. *Journal of Transportation Engineering*, 124(4), 376-383.

Zhao, F., J. Shin, J. Reich. 2002. Information-driven dynamic sensor collaboration. *IEEE Signal Processing Magazine*, 19, 61-72.

Zhou, X., H. S. Mahmassani. 2005. Dynamic OD demand estimation using automatic vehicle identification data. *IEEE Transactions on Intelligent Transportation Systems*, 7(1), 105- 114.