Dynamic Origin–Destination Demand Estimation Using Automatic Vehicle Identification Data

Xuesong Zhou and Hani S. Mahmassani, *Member, IEEE*

***Abstract*—This paper proposes a dynamic origin–destination (OD) estimation method to extract valuable point-to-point split-fraction information from automatic vehicle identification (AVI) counts without estimating market-penetration rates and identification rates of AVI tags. A nonlinear ordinary least-squares estimation model is presented to combine AVI counts, link counts, and historical demand information into a multiobjective optimization framework. A joint estimation formulation and a one-sided linear-penalty formulation are further developed to take into account possible identification and representativeness errors, and the resulting optimization problems are solved by using an iterative bilevel estimation procedure. Based on a synthetic data set, this study shows the effectiveness of the proposed estimation models under different market-penetration rates and identification rates.**

***Index Terms*—Road vehicle identification, state estimation, traffic information systems, transportation networks.**

I. Introduction

TIME-DEPENDENT origin–destination (OD) demand information is an essential input for dynamic traffic assignment (DTA) models, used to describe and predict time-varying traffic network flow patterns, as well as to generate anticipatory and coordinated control and information supply strategies for intelligent traffic network management. In general, OD-trip-desire information can be obtained from direct-interview surveys or estimated from real-time traffic surveillance data. Populating OD-demand patterns from survey samples, however, is a resource intensive and time-consuming process, and conventional survey methods cannot provide up-to-date dynamic demand inputs required by online Advanced Traffic Management Systems (ATMS) and Advanced Traveler Information Systems (ATIS) applications. Deployment of intelligent transportation system (ITS) technologies offer more reliable and less costly channels to measure the time-varying states of transportation systems, and both real-time and archived traffic measurements provide valuable data to help capture the underlying travel decision processes.

Substantial research has been devoted to the dynamic demand-estimation problem using time-varying link counts. Early models [1]–[3] were proposed to estimate time-dependent OD flows on individual components, such as a single intersection or a freeway facility; these models seek to estimate unknown dynamic OD split fractions based on the entry and exit flow measurements, under the simplifying assumption of constant link travel time. Extending the concepts and solution methodologies of the static OD-estimation problem, Cascetta *et al.* [4] proposed a generalized least-squares (GLS) estimator for dynamic OD demand based on a simplified assignment model for a general network. Growing interest in the application of simulation-based DTA models has been accompanied by several studies on the estimation of dynamic OD-trip desires. A bilevel GLS optimization model and an iterative solution framework have been proposed by Tavana and Mahmassani [5] to estimate the dynamic OD demand and to maintain the internal consistency between the upper level demand-estimation problem and the lower level DTA problem. Tavana [6] also provided an extensive literature review of the dynamic OD-demand-estimation problem and its inherent connection to the DTA problem.

In a real traffic network, the number of independent link counts is typically less than the number of unknown time-dependent OD pairs, so dynamic OD-demand estimation purely relying on traffic link counts might lead to an underdetermined system. As a result, additional information is needed to find a unique OD-demand estimate. The use of flow counts across screen lines and cordon lines in dynamic OD-demand estimation was introduced by Chang and Wu [7] and Chang and Tao [8] to extract more information from the existing traffic surveillance and survey data. Zhou *et al.* [9] proposed a multiobjective optimization framework to combine available historical static demand information and multiday traffic link counts to estimate the variation in the traffic demand over multiple days.

Automatic vehicle identification (AVI) data represent another data source of growing importance for estimating dynamic OD-demand flows, and more generally for traffic network management. Two classes of demand-estimation problems using vehicle identification data should be distinguished: 1) the estimation of tagged vehicle demand and 2) the estimation of population demand. Several studies focus on the first class of problems. Based on the transponder tag data collected from a freeway corridor in Houston, Dixon and Rilett [10] applied the framework developed by Cascetta *et al.* [4] to calculate the link-flow proportions based on the observed travel time from AVI counts. They presented both off-line GLS models and online Kalman filtering models for estimating tagged OD demand. Antoniou *et al.* [11] introduced path-flow proportion matrices that relate OD-demand flows to subpath tag counts, and extended Ashok's framework [12] to estimate and predict tagged vehicular OD-demand flows.

Manuscript received July 1, 2005; revised September 18, 2005 and November 27, 2005. The Associate Editor for this paper was S. C. Wong.

The authors are with the Maryland Transportation Initiative and with the Department of Civil and Environmental Engineering, University of Maryland, College Park, MD 20742 USA (e-mail: xszhou@umd.edu; masmah@umd.edu).

Digital Object Identifier 10.1109/TITS.2006.869629

If only a subset of vehicles is equipped with transponder tags or only a subset of vehicles is correctly identified by the AVI readers, then the second class of problems needs to be explicitly considered in order to infer the population trip desires. Several models have been developed for the estimation of population demand using AVI counts. Recognizing low identification rates associated with license-plate-based AVI data, Van der Zijpp [13] proposed a constrained optimization formulation to jointly estimate the unknown OD-demand flows and identification rates. Along the same line, Asakura *et al.* [14] provided an off-line least-squares model to simultaneously determine the OD demand and the location-dependent identification rates, and further investigated day-to-day fluctuations in estimated OD demands. Dixon [15] proposed a three-stage procedure to estimate population OD demand from transponder-based AVI data: 1) Estimate the tagged OD-demand matrix from AVI data, 2) estimate market-penetration rates using AVI data and link counts, and then 3) scale the estimated tagged vehicle demand to the total population demand using the estimated market-penetration rates. In brief, the above models require estimating either market-penetration rates or identification rates so as to relate the AVI samples to the population demand using a multiplicative function structure. The estimation of market-penetration rates or identification rates, however, is a difficult problem in its own right, as these two types of rates are essentially time-dependent and location-dependent random variables. Moreover, the inclusion of market-penetration rates and identification rates in the demand-estimation problem could dramatically increase the number of unknown variables and impact the reliability of the final population demand estimate through the multiplicative structure.

This paper focuses on the estimation of population OD demand using partially observed AVI counts. To circumvent primary difficulties associated with estimating market-penetration rates and identification rates, this research samples population OD split fractions from point-to-point AVI counts and extracts OD-demand distribution information, instead of treating OD split fractions as unknown variables, as is the case in several early dynamic OD-estimation models [1]–[3]. A nonlinear ordinary least-squares model is first proposed to systematically combine AVI counts, link counts, and historical OD-demand information. Furthermore, two OD-demand-estimation formulations are developed to take into account the possible identification and representativeness errors. Synthetic AVI traffic counts are used to investigate the relative value of AVI counts to the OD-estimation problem under the different market-penetration rates and identification rates.

II. Problem Statement

The following notation is used to represent all the variables in the dynamic OD-demand-estimation formulation.

I	Set of origin zones.
J	Set of destination zones.
L_{lc}	Set of links with link-count observations.
L_{vi}	Set of links with vehicle identification observations.
l	Subscript for link with traffic measurements.
i	Subscript for origin zone, $i \in I$.
j	Subscript for the destination zone, $j \in J$.
τ	Subscript for departure time intervals, $\tau = 1, 2, \ldots, T_{\mathrm{d}}$.
t	Subscript for observation time interval, i.e., sampling time interval, $t = 1, 2, \ldots, T_{\mathrm{o}}$.
tg	Superscript for tag-equipped vehicles.
id	Superscript for identified vehicles.
k	Superscript for iteration counter.
$c_{(l,t)}$	Number of vehicles on link l during observation interval t.
C	Vector of measured flows on the links, consisting of element $c_{(l,t)}$.
$c_{(l,s,t)}$	Number of vehicles observed on link s, traveling from link l during observation interval t.
$d_{(i,j,\tau)}$	Demand volume with destination in zone j, originating their trip from zone i during departure interval τ.
D	Dynamic OD-demand matrix, consisting of elements $d_{(i,j,\tau)}$.
$c_{(i,j,\tau)}$	Number of vehicles observed in destination zone j, originating their trip at zone i during departure interval τ.
$b_{(i,j,\tau)}$	Origin-to-destination split fraction, i.e., proportion of traffic departing from origin i during departure time interval τ, heading towards destination j.
$b_{(l,s,t)}$	Link-to-link split fraction, i.e., proportion of traffic passing link l during observation time interval t, heading towards link s (link s is not necessary to be a downstream link of link l).
$p_{(l,t)(i,j,\tau)}$	Link-flow proportions, i.e., proportion of vehicular demand flows from origin i to destination j, starting their trips during departure interval τ, contributing to the flow on link l during observation interval t.
$\hat{p}_{(l,t)(i,j,\tau)}$	Estimated link-flow proportions based on a DTA program.
$p_{(l,s,t)(i,j,\tau)}$	Link-to-link-flow proportions, i.e., proportion of vehicular flows from origin i to destination j, starting their trips during departure interval τ, contributing to the link-to-link flow from link l (during observation intervals t) to link s.
$\hat{p}_{(l,s,t)(i,j,\tau)}$	Estimated point-to-point-flow proportions based on a DTA program.
$\hat{P}$	Estimated flow proportion matrix that includes elements $\hat{p}_{(l,t),(i,j,\tau)}$ and $\hat{p}_{(l,s,t),(i,j,\tau)}$.
$\eta_{(l,s,t)}$	Sampling error term in estimation of link-to-link split fraction $b_{(l,s,t)}$.
$\zeta_{(l,s,t)}$	Combined error term in estimation of link-to-link split fraction $b_{(l,s,t)}$.
$\varepsilon_{(l,t)}$	Combined error term in estimation of traffic flow on link l during observation interval t.
$g_{(i,j)}$	Target demand, which is the total traffic demand during period of interest for OD pair (i,j).
G	Historical OD-demand matrix, consisting of elements $g_{(i,j)}$.

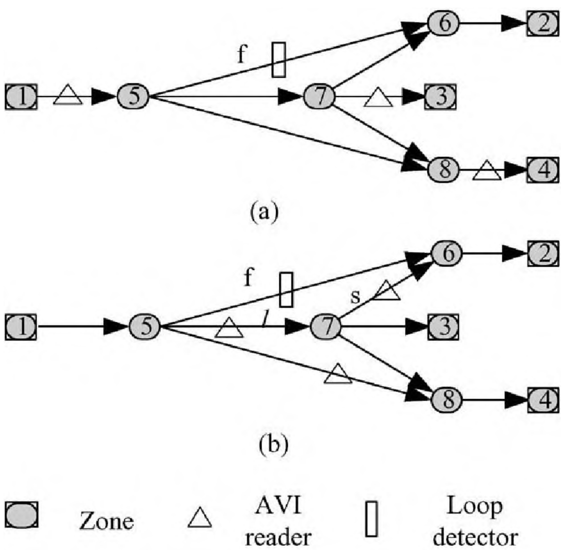

Fig. 1. Example of utilizing AVI point-to-point counts.

Consider a traffic network consisting of multiple origins $i \in I$ and destinations $j \in J$, as well as a set of nodes connected by a set of directed links. The analysis period of interest is discretized into departure time intervals $\tau = 1, 2, \ldots, T_{\rm d}$. Link counts $c_{(l,t)}$ are available on link $l \in L_{\rm lc}$ during observation interval $t = 1, 2, \ldots, T_{\rm o}$. AVI reader stations are located on link $l \in L_{\rm vi}$, and vehicle identification data include point counts $c^{\rm id}_{(l,t)} \forall\, l \in L_{\rm vi}$, $t = 1, \ldots, T_{\rm o}$ and point-to-point counts $c^{\rm id}_{(l,s,t)} \forall\, l, s \in L_{\rm vi}$, $t = 1, \ldots, T_{\rm o}$.

Note that the tg superscripts denote actual AVI tags, while the id superscripts denote those AVI tags that are actually identified by the readers. Specifically, $c^{\rm tg}_{(l,s,t)}$ denotes point-to-point AVI tags, whereas $c^{\rm id}_{(l,s,t)}$ is the number of vehicles identified by AVI detectors, respectively, traveling from link l to link s during observation interval t. The market-penetration rate of AVI tags is the percentage of vehicles with tags in the entire vehicle population. The sampling time intervals for AVI counts and traffic link counts are assumed to be the same for notation simplicity, and with no loss of generality. 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. Given the link counts, vehicle identification counts, and prior information on OD trips, the dynamic OD-demand-estimation problem seeks to find time-dependent OD-trip desires (over a time horizon of interest) so as to minimize deviations between the observed traffic flows and the assigned traffic flows (resulting from a DTA process), and deviations between estimated OD-demand flows and the historical demand matrix.

In order to circumvent the difficulties in estimating market-penetration rates of AVI tags, this study utilizes probe vehicle data to extract spatial distribution information of trip makers in a traffic network. This approach can be illustrated using a small network (shown in Fig. 1). In that example, the problem is to estimate the population OD-demand flows $d_{(1,2)}$, $d_{(1,3)}$, and $d_{(1,4)}$ from available AVI and loop detector counts. For simplicity, subscripts τ and t are dropped in this illustration. In Fig. 1(a), an AVI reader located on link (1,5) records tagged vehicle flows departing from zone 1, that is, $\sum_j d^{\rm tg}_{(1,j)} = d^{\rm tg}_{(1,2)} + d^{\rm tg}_{(1,3)} + d^{\rm tg}_{(1,4)}$. OD AVI counts $d^{\rm tg}_{(1,3)}$ and $d^{\rm tg}_{(1,4)}$ are also observed. If tagged vehicles are representative of the total population, then AVI counts can be used to estimate the OD split fractions for the population demand traveling from zone 1, leading to the following measurement equations:

$$\frac{d^{\rm tg}_{(1,3)}}{\sum\limits_j d^{\rm tg}_{(1,j)}} = \frac{d_{(1,3)}}{\sum\limits_j d_{(1,j)}} + \eta_{(1,3)} \tag{1}$$

$$\frac{d^{\rm tg}_{(1,4)}}{\sum\limits_j d^{\rm tg}_{(1,j)}} = \frac{d_{(1,4)}}{\sum\limits_j d_{(1,j)}} + \eta_{(1,4)} \tag{2}$$

where $\eta_{(1,3)}$ and $\eta_{(1,4)}$ are sampling errors.

The loop detector on link (5,6) captures partial OD flow $d_{(1,2)}$, so demand $d_{(1,2)}$ can be related to link count $c_{(f)}$ using link-flow proportion $p_{(f),(1,2)}$, where f denotes link (5,6), and $\varepsilon_{(f)}$ is the combined error in estimation of traffic flow on link l.

$$c_{(f)} = p_{(f),(1,2)} d_{(1,2)} + \varepsilon_{(f)} \tag{3}$$

Combining the above three measurement equations, one can create a system of nonlinear equations to estimate the unknown population OD-demand flows $d_{(1,2)}$, $d_{(1,3)}$, and $d_{(1,4)}$, without estimating the market-penetration rate of vehicle tags. Fig. 1(b) shows a more general case, where direct OD tagged vehicle counts are unavailable. Denote links (5,7) and (7,6) as links l and s, respectively. Link-to-link counts $c_{(l,s)}$ cover partial OD flows $d_{(1,2)}$, $d_{(1,3)}$, $d_{(1,4)}$, while link counts $c_{(l)}$ record partial OD flows $d_{(1,2)}$. The resulting split fraction for link pair (l, s), which represents the percentage of vehicular flow traveling on link l contributing to the flow on link s, is

$$\begin{aligned} \frac{c^{\rm id}_{(l,s)}}{c^{\rm id}_{(l)}} &= \frac{\sum\limits_{i,j} p_{(l,s)(i,j)} d_{(i,j)}}{\sum\limits_{i,j} p_{(l)(i,j)} d_{(i,j)}} + \eta_{(l,s)} \\ &= \frac{p_{(l,s)(1,2)} d_{(1,2)}}{p_{(l)(1,2)} d_{(1,2)} + p_{(l)(1,3)} d_{(1,3)} + p_{(l)(1,4)} d_{(1,4)}} + \eta_{(l,s)}. \end{aligned} \tag{4}$$

In this case, the information on OD-demand distributions can be partially revealed from the link-to-link split fractions, i.e., the ratio of the link-to-link AVI counts to the link AVI counts. Because link proportions are determined by the route choice behavior and traffic-flow propagation mechanism (both captured in the network traffic assignment process), the above link-to-link split formulation introduces additional difficulty and uncertainty in demand estimation, compared to the previous formulation that only uses OD split factions. In this study, a simulation-based DTA program, namely DYNASMART-P [16], is used to estimate the link flow and link-to-link-flow proportions.

III. Nonlinear Least-Squares Formulation

This section formulates a dynamic OD-estimation model based on the following two assumptions.

1) AVI readers can correctly identify every tagged vehicle.
2) Tagged vehicles are a representative subset of the entire population.

The first condition assumes 100% identification rates and indicates $c^{\mathrm{id}}_{(l,s,t)} = c^{\mathrm{tg}}_{(l,s,t)}$ and $c^{\mathrm{id}}_{(l,t)} = c^{\mathrm{tg}}_{(l,t)}$. Under the second condition, the tagged vehicles probabilistically represent the entire population, and the split fractions of tagged vehicles can be used as sample estimates for the population split fractions. These assumptions will be relaxed in the next section. To construct a rigorous statistical inference model, the following discussion examines the properties of a random AVI sample for estimating split fractions. Recognizing that there are $c^{\mathrm{id}}_{(l,s,t)}$ tagged vehicles choosing link s out of $c^{\mathrm{id}}_{(l,t)}$ vehicles observed on link l at time t, link-to-link identified vehicle count $c^{\mathrm{id}}_{(l,s,t)}$ essentially follows a binomial distribution with the sample size of $c^{\mathrm{id}}_{(l,t)}$ and the success probability of $c_{(l,s,t)}/c_{(l,t)}$, i.e., $c^{\mathrm{id}}_{(l,s,t)} \sim \mathrm{Binomial}\left[c^{\mathrm{id}}_{(l,t)}, (c_{(l,s,t)}/c_{(l,t)})\right]$. Since

$$\frac{c_{(l,s,t)}}{c_{(l,t)}} = \frac{\sum\limits_{i,j,\tau} p_{(l,s,t)(i,j,\tau)} d_{(i,j,\tau)}}{\sum\limits_{i,j,\tau} p_{(l,t)(i,j,\tau)} d_{(i,j,\tau)}} \tag{5}$$

the sample proportion $c^{\mathrm{id}}_{(l,s,t)}/c^{\mathrm{id}}_{(l,t)}$ is an unbiased estimator of split fractions for the population proportion. That is

$$\begin{aligned}\frac{c^{\mathrm{id}}_{(l,s,t)}}{c^{\mathrm{id}}_{(l,t)}} &= \frac{c_{(l,s,t)}}{c_{(l,t)}} + \eta_{(l,s,t)} \\ &= \frac{\sum\limits_{i,j,\tau} p_{(l,s,t)(i,j,\tau)} d_{(i,j,\tau)}}{\sum\limits_{i,j,\tau} p_{(l,t)(i,j,\tau)} d_{(i,j,\tau)}} + \eta_{(l,s,t)}\end{aligned} \tag{6}$$

and the mean and variance of the sampling error $\eta_{(l,s,t)}$ are $\mathrm{E}(\eta_{(l,s,t)}) = 0$ and

$$\mathrm{Var}\left(\eta_{(l,s,t)}\right) = \frac{b_{(l,s,t)} \cdot \left(1 - b_{(l,s,t)}\right)}{c^{\mathrm{id}}_{(l,t)}} \tag{7}$$

where $b_{(l,s,t)} = c_{(l,s,t)}/c_{(l,t)}$. Equation (7) indicates that the variance of the sampling error $\eta_{(l,s,t)}$ decreases as the size of the sample $c^{\mathrm{id}}_{(l,t)}$ increases. If link volume $c_{(l,t)}$ is observed from a loop detector for link $l \in L_{\mathrm{vi}}$, a linear measurement equation can be obtained as

$$\frac{c^{\mathrm{id}}_{(l,s,t)}}{c^{\mathrm{id}}_{(l,t)}} = \frac{\sum\limits_{i,j,\tau} p_{(l,s,t)(i,j,\tau)} d_{(i,j,\tau)}}{c_{(l,t)}} + \eta_{(l,s,t)}. \tag{8}$$

Alternatively, if AVI readers cover the entry links of origin i and the exit links of destination j, then AVI OD counts are directly used to infer destination distributions without involving link flow and link-to-link-flow proportions

$$\frac{c^{\mathrm{id}}_{(i,j,\tau)}}{c^{\mathrm{id}}_{(i,\tau)}} = \frac{d_{(i,j,\tau)}}{\sum\limits_{j} d_{(i,j,\tau)}} + \eta_{(i,j,\tau)}. \tag{9}$$

The covariance of sampling errors can be analyzed as follows. First, denote links s and s' as two distinct links reachable from link l. If links s and s' are two independent choice alternatives for vehicles traveling on link l at time t, then link-to-link AVI counts $c^{\mathrm{id}}_{(l,s,t)}$ and $c^{\mathrm{id}}_{(l,s',t)}$ follow a multinomial distribution, leading to the covariance of sampling errors as

$$\mathrm{Cov}\left(\eta_{(l,s,t)}, \eta_{(l,s',t)}\right) = \frac{b_{(l,s,t)} \cdot b_{(l,s',t)}}{c^{\mathrm{id}}_{(l,t)}}. \tag{10}$$

If certain vehicles in link-to-link flows $c_{(l,s,t)}$ also appear in flows $c_{(l,s',t)}$, links s and s' cannot be viewed as independent choice alternatives for vehicles traveling on link l at time t. In this case, link counts $c_{(l,s,t)}$ and $c_{(l,s',t)}$ can be partitioned into three mutually exclusive categories $\omega \in \Omega$.

$\omega = 1$: choose only s; $\omega = 2$: choose only s'; $\omega = 3$: choose both s and s'.

The resulting covariance can be expressed in terms of variance and covariance between disjoint sets.

$$\begin{aligned}\mathrm{Cov}&\left(\eta_{(l,s,t)}, \eta_{(l,s',t)}\right) \\ &= \sum_{\omega\in\Omega}\sum_{\omega'\in\Omega} \mathrm{Cov}\left(\eta_{(l,\omega,t)}, \eta_{(l,\omega',t)}\right) \\ &= \frac{1}{c^{\mathrm{id}}_{(l,t)}} \sum_{\omega\in\Omega}\sum_{\omega'\in\Omega} b_{(l,\omega,t)} b_{(l,\omega',t)}\end{aligned} \tag{11}$$

where $b_{(l,\omega,t)}$ is the proportion of link flow from link l at time t heading towards category $\omega \in \Omega$.

A complete measurement equation of link-to-link split fractions can be obtained by substituting the estimates of flow proportion matrices from a DTA problem into (6).

$$\frac{c^{\mathrm{id}}_{(l,s,t)}}{c^{\mathrm{id}}_{(l,t)}} = \frac{\sum\limits_{i,j,\tau} \hat{p}_{(l,s,t),(i,j,\tau)} d_{(i,j,\tau)}}{\sum\limits_{i,j,\tau} \hat{p}_{(l,t)(i,j,\tau)} d_{(i,j,\tau)}} + \zeta_{(l,s,t)} \tag{12}$$

where $\zeta_{(l,s,t)}$ refers to the combined error in estimation of link-to-link split fraction $b_{(l,s,t)}$. The combined error term $\zeta_{(l,s,t)}$ includes the following error sources:

1) model assumption errors related to the hypotheses on perfect representativeness and 100% identification rates;
2) sensor errors (i.e., identification errors) related to link-to-link AVI count $c^{\mathrm{id}}_{(l,s,t)}$ and link count $c^{\mathrm{id}}_{(l,t)}$;
3) sampling errors $\eta_{(l,s,t)}$;
4) aggregation errors related to time-varying OD-demand flows;
5) estimation errors related to link flow and link-to-link-flow proportions from the DTA program, which can be further caused by inconsistency in DTA assumptions on the route choice behavior, traffic-flow propagation, as well as input data errors related to traffic control and information strategies.

Because the split fractions only carry information on OD-demand distributions, it is necessary to combine other information sources that describe OD population demand volumes in order to estimate a complete OD matrix. Obviously, the observed traffic volume on link l during time interval t can be related to the OD-demand flows using the link-flow proportions, corresponding to a measurement equation based on link counts

$$c_{(l,t)} = \sum_{i,j,\tau} \hat{p}_{(l,t)(i,j,\tau)} \cdot d_{(i,j,\tau)} + \varepsilon_{(l,t)}. \quad (13)$$

If a static OD-demand matrix is available from existing survey data or other planning applications, the formulation proposed by Zhou *et al.* [9] can be adopted here to consider the deviation between the static demand and the sum of dynamic demand over the study period as

$$g_{(i,j)} = \sum_{\tau} d_{(i,j,\tau)} + \xi_{(i,j)}. \quad (14)$$

In addition, OD-demand flows should satisfy nonnegativity constraints

$$d_{(i,j,\tau)} \geq 0 \qquad \forall i, j, \tau. \quad (15)$$

If values of split fractions are considerably small and the number of observations in the AVI point sample is large, the binomial probabilities can be approximated by a Poisson distribution. If the AVI point sample size is sufficiently large with a moderate value of the split fraction, error terms $\zeta_{(l,s,t)}$ can be assumed to follow a normal distribution with zero mean according to the central limit theorem. The bilevel dynamic OD-estimation framework developed by Tavana and Mahmassani [5] and Zhou *et al.* [9] can be adopted here to minimize the combined deviations with respect to link counts, historical static demand, and AVI split fractions, subject to the (definitional) DTA constraint and nonnegativity constraints for demand variables.

$$\min Z = \left[Z_1(D,C) + Z_2(D,G) + Z_3(D,C^{\text{id}})\right] \quad (16)$$

$$\text{s.t.} \quad \hat{P} = \text{assignment } [D] \text{ from DTA} \quad (17)$$

$$d_{(i,j,\tau)} \geq 0 \qquad \forall\, i, j, \tau \quad (18)$$

where

$$Z_1(D,C) = w_1 \sum_{l \in L_{\text{lc}}, t} \left[c_{(l,t)} - \sum_{i,j,\tau} \hat{p}_{(l,t)(i,j,\tau)} \cdot d_{(i,j,\tau)}\right]^2 \quad (19)$$

$$Z_2(D,G) = w_2 \sum_{i,j} \left[g_{(i,j)} - \sum_{\tau} d_{(i,j,\tau)}\right]^2 \quad (20)$$

$$Z_3(D,C^{\text{id}}) = w_3 \sum_{l \in L_{\text{vi}}, t} \sum_{s \in L_{\text{vi}}} \left[\frac{c^{\text{id}}_{(l,s,t)}}{c^{\text{id}}_{(l,t)}} - \frac{\sum_{i,j,\tau} \hat{p}_{(l,s,t)(i,j,\tau)} d_{(i,j,\tau)}}{\sum_{i,j,\tau} \hat{p}_{(l,t)(i,j,\tau)} d_{(i,j,\tau)}}\right]^2 \quad (21)$$

and w_1, w_2, and w_3 are positive weights associated with, respectively, the deviations with respect to link counts, historical static demand, and observed split fractions.

IV. Representativeness Biases and Identification Errors

The two idealized assumptions in the above analysis, that is, 100% identification rates and perfect representativeness of AVI samples, could be difficult to satisfy in many applications. First, recognition rates vary significantly among different AVI technologies. Active tags, especially used for toll collection purposes, can provide a satisfactory > 99% identification rate, but license plates and passive tags typically have relatively low identification rates ranging from 50% to 80%. On the other hand, the AVI sample data might not be a perfectly representative image of the underlying population. For instance, tag users and nontag users might belong to different socioeconomic groups with heterogeneous preferences in terms of the value of time, especially when tag users could experience less congestion on dedicated lanes in a toll plaza by paying one-time charges or monthly fees for transponder tags. If an electronic toll collection system exists on commuters' daily routes, network users are more likely to purchase and use AVI transponders to avoid congestion, as opposed to network users who rarely use toll roads. The representativeness of AVI data, essentially, should be verified on a case-by-case basis. This section presents two formulations intended to recognize departures from the above assumptions. The first formulation presents a general and flexible framework for incorporating such deviations when additional information is available, while the second formulation considers a situation with limited information.

A. *Joint Estimation of OD-Demand Volume and Fixed-Effect Parameters*

Under low identification rates, observed split fractions could be considerably smaller than true split fractions of tagged vehicles. The effect of representativeness biases of AVI samples, on the other hand, could cause observed split fractions to either overstate or understate the population split fractions. When these two assumptions for the ordinary least-squares formulation are not attainable, there is a great need to establish a flexible estimation framework that can accommodate possible departure from the idealized conditions. A natural approach is to establish a joint estimation model as the following:

$$b^{\text{id}}_{(l,s,t)} = \hat{b}_{(l,s,t)} \alpha_{(l,s)} + \zeta_{(l,s,t)} \quad (22)$$

where

$$b^{\text{id}}_{(l,s,t)} = \frac{c^{\text{id}}_{(l,s,t)}}{c^{\text{id}}_{(l,t)}}$$

$$\hat{b}_{(l,s,t)} = \frac{\sum_{i,j,\tau} \hat{p}_{(l,s,t)(i,j,\tau)} d_{(i,j,\tau)}}{\sum_{i,j,\tau} \hat{p}_{(l,t)(i,j,\tau)} d_{(i,j,\tau)}}$$

and a fixed-effect parameter $\alpha_{(l,s)}$ is introduced to take into account the systematic impact due to representativeness biases and low identification rates in estimating link-to-link split fractions. The validity of the two idealized assumptions can

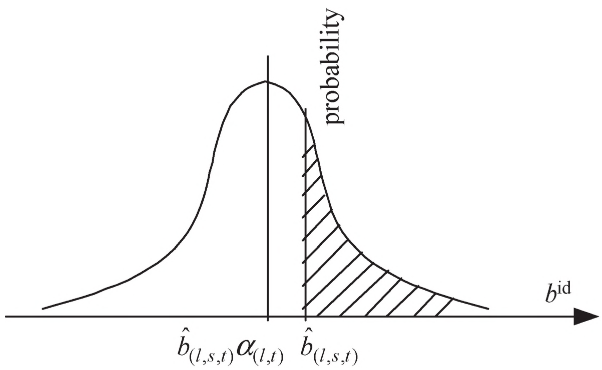

Fig. 2. Conceptual view of distribution of observed split fractions.

be measured using the following statistical procedure. A full estimation model incorporates a fixed-effect parameter for each possible link pair (l, s) explicitly, leading to a new objective function that jointly estimates the unknown OD-demand and fixed-effect parameters

$$Z_3(D, C^{\mathrm{id}}) = w_3 \sum_{l \in L_{\mathrm{vi}}, t} \sum_{s \in L_{\mathrm{vi}}} \left[b^{\mathrm{id}}_{(l,s,t)} - \hat{b}_{(l,s,t)} \alpha_{(l,s)} \right]^2. \quad (23)$$

The null hypothesis (H_0) states that the systematic deviation is zero, corresponding to a reduced model with fixed-effect parameters of 1. The alternative hypothesis (H_1) states that the systematic deviation is nonzero, corresponding to the full model. A standard F-statistic test can be applied in this context.

$$\boldsymbol{H}_0 : \alpha_{(l,s)} = 1 \quad (24)$$

$$\boldsymbol{H}_1 : \alpha_{(l,s)} \neq 1 \quad (25)$$

The probability density function of observed split fractions for link pair (l, s) could be plotted in Fig. 2, in which $\hat{b}_{(l,s,t)} \alpha_{(l,s)}$ is the center of the distribution. Alternatively, we can consider a two-sided penalty function to minimize the positive and negative deviations from observed AVI point-to-point split fractions.

$$Z_3(D, C^{\mathrm{id}}) = \sum_{l \in L_{\mathrm{vi}}, t} \sum_{s \in L_{\mathrm{vi}}} \left(w_3^+ \left[b^{\mathrm{id}}_{(l,s,t)} - \hat{b}_{(l,s,t)} \alpha_{(l,s)} \right]^+ + w_3^- \left[b^{\mathrm{id}}_{(l,s,t)} - \hat{b}_{(l,s,t)} \alpha_{(l,s)} \right]^- \right) \quad (26)$$

where w_3^+ and w_3^- are penalty terms for positive and negative deviations, respectively. This objective function can be viewed as an adaptation of a goal programming approach, which assigns priority factors for overestimation and underestimation from the specified goals. The two-sided weights can be in turn interpreted as the relative confidence and preference of a decision maker on the possible sign of systematic deviations due to representativeness and identification errors.

The inclusion of the fixed-effect parameters considerably complicates the model structures, and the existence of numerous error sources, such as assignment modeling errors and temporal fluctuation of demand flows, might lead to inconclusive estimates for $\alpha_{(l,s)}$ with large variance. Additionally, if only total AVI counts for the entire planning horizon are available (e.g., in the context of static OD estimation), then fixed-effect parameters, which are designed to deal with time-series data, would not be appropriate for inclusion in the estimation model.

B. *One-Sided Linear-Penalty Models Using Imprecise Knowledge*

In many instances, the traffic planner knows the likely range of fixed-effect parameters but with little information about the exact values of representativeness and identification errors. For instance, license plate sample data have perfect representativeness but low identification rates, meaning that $\alpha_{(l,s)} \leq 1$ for sure. In an AVI system where passive tags are offered to the public at no charge, both representativeness errors and identification errors can coexist but the impact of low identification rates is more likely to be dominating, also implying that $E(\alpha_{(l,s)}) \leq 1$. In these two cases, it is desirable to design a population OD-demand estimator that can utilize the above imprecise knowledge from the planner. A simple one-sided penalty formulation, that is

$$Z_3(D, C^{\mathrm{id}}) = \sum_{l \in L_{\mathrm{vi}}, t} \sum_{s \in L_{\mathrm{vi}}} \left(w_3^+ \left[b^{\mathrm{id}}_{(l,s,t)} - \hat{b}_{(l,s,t)} \right]^+ \right) \quad (27)$$

is proposed to accomplish this task. Instead of penalizing all positive and negative deviations with respect to uncertain $\hat{b}_{(l,s,t)} \alpha_{(l,s)}$, this estimator only penalizes positive residuals between observed split fractions and estimated split fractions (shown in the shaded region in Fig. 2). In other words, estimated split fractions are penalized only if they are less than the corresponding observed split fractions.

It is easy to see that negative deviations between observed split fractions and estimated split fractions can be caused by either demand estimation errors or low identification rates of vehicle identification tags. Compared to the two-sided formulation (26), the simplified one-sided formulation (27) utilizes less information from AVI counts. However, by omitting negative deviations and using $\hat{b}_{(l,s,t)}$ instead of $\hat{b}_{(l,s,t)} \alpha_{(l,s)}$ in the deviation function, this new objective function is able to eliminate the need for the planner to exactly estimate unknown representativeness errors and identification rates (i.e., fixed-effect parameters $\alpha_{(l,s)}$). It should be noticed that, as fixed-effect parameters $\alpha_{(l,s)}$ become smaller, the difference between $\hat{b}_{(l,s,t)} \alpha_{(l,s)}$ and $\hat{b}_{(l,s,t)}$ becomes larger, and consequently residuals of $[b^{\mathrm{id}}_{(l,s,t)} - \hat{b}_{(l,s,t)}]$ are less likely to be positive. That is to say, the larger representativeness errors or/and identification errors, the less information could be extracted from AVI counts using this one-sided formulation.

As part of multiobjective demand estimation, the value of w_3^+ still needs to be jointly determined by considering the relative confidence placed on link counts and historical demand information. For fixed values of w_1 and w_2, a larger w_3^+ means that the decision maker has more confidence in the quality of AVI data, and vice versa. Compared to the least-squares form, another advantage of the linear-penalty function form is that it is much "smoother" with respect to outlying observations, in the sense that large deviations cannot substantially impair the estimation performance.

V. Bilevel Estimation Algorithms and Identification Conditions

The proposed bilevel programming problem can be solved by an iterative solution algorithm, extended from the solution framework in [5], [6], and [9].

Step 1) (Initialization) $k = 1$. Start from an initial guess of the traffic demand matrix, obtain flow proposition matrix $\hat{P}^1$ from the DTA simulator.

Step 2) (Optimization) Substitute flow proportion matrix $\hat{P}^k$ to solve the upper level estimation problem.

Step 3) (Simulation) Use estimated demand $\hat{D}^k$ to run the DTA simulator so as to generate new flow proportions $\hat{P}^{k+1}$.

Step 4) (Evaluation) Calculate the deviation between simulated link flows and observed link counts, the deviation between estimated demand $\hat{D}^k$ and target demand G, as well as the deviation between estimated link-to-link split fractions and observed link-to-link split fractions.

Step 5) (Convergence test) If the convergence criterion is satisfied (estimated demand is stable or no significant improvement in the overall sum of deviations), stop; otherwise $k = k + 1$ and go to Step 2).

The multiobjective optimization techniques presented by Zhou *et al.* [9] can be applied here to determine the weights in the upper level objective function. Standard nonlinear optimization algorithms, such as the projected gradient algorithm, can be applied to solve the proposed nonlinear estimation problem. To further construct a computationally feasible algorithm for the upper level estimation problem, one can linearize the nonlinear function of split fraction $b_{(l,s,t)}$ based on a first-order Taylor series approximation around previous estimate of $\hat{D}^{k-1}$ at iteration $k - 1$.

The proposed dynamic OD-demand-estimation problem has $|I| \times |J| \times T_{\text{d}}$ unknown demand variables. Loop detectors can provide at most $|L_{\text{lc}}| \times T_{\text{o}}$ independent link volume measurements, and prior information on static OD demand can be viewed as $|I| \times |J|$ observations on the unknown dynamic OD-demand matrix. Additionally, AVI data provide at most $|L_{\text{vi}}| \times |L_{\text{vi}}| \times T_{\text{o}}$ link-to-link tagged vehicle counts, which dramatically alleviate the underspecification problem of OD estimation. To identify a unique solution for the dynamic OD-estimation problem, the number of independent observations should be greater than the number of unknown demand variables, leading to a necessary condition for uniquely estimating a dynamic OD-demand matrix as

$$|L_{\text{lc}}| \times T_{\text{o}} + |I| \times |J| + |L_{\text{vi}}| \times |L_{\text{vi}}| \times T_{\text{o}} \geq |I| \times |J| \times T_{\text{d}}. \tag{28}$$

It should be noted that AVI data only provide OD-demand distribution information, so OD-demand volume information from loop counts and historical OD tables must be added to identify a unique solution. Several factors, moreover, could significantly decrease the number of independent observations in an actual AVI data set. First, many OD flows might not be captured by any AVI readers in a general network with partial coverage of AVI detectors. In addition, time-dependent point-to-point AVI measures from two adjacent readers can be highly correlated, and consequently produce limited information on OD-demand distributions.

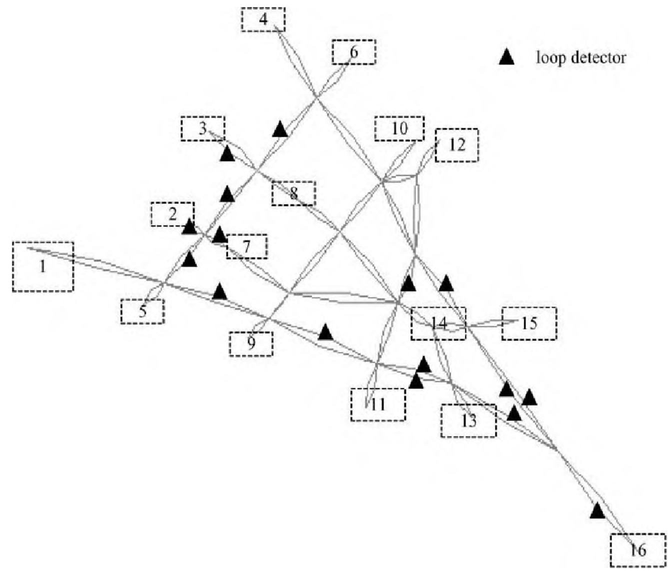

Fig. 3. Irvine simplified network.

VI. Numerical Experiments

A. Experiment Design

This section is intended to evaluate the performance of the proposed dynamic OD-demand-estimation models under different levels of market-penetration rates, identification rates, and AVI detector coverage. The experiments are conducted based on a simplified Irvine test bed network, as shown in Fig. 3, which 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 A.M.–8:30 A.M.).

Actual traffic link counts are measured on 16 links; at 30-s interval on ten freeway links, and at 5-min interval on six arterial links, but no real-world AVI traffic measurements are currently available in this data set. In order to capture a realistic OD-demand pattern for the underlying city network, this study first uses actual link counts and a historical static demand table to estimate the OD traffic demand matrix, and then uses the estimated matrix as the "true" OD demand in the following experiments. The "true" OD demand is loaded onto the network using a DTA simulation program (i.e., DYNASMART-P) to generate both link counts and point-to-point counts as the "ground-truth" observations in the synthetic data set. The DTA simulator is also used to provide link-flow proportions and link-to-link-flow proportions to the OD-estimation program. Note that, to ensure the internal consistency between link-flow measurements and point-to-point flow measurements, this study uses simulated link counts as estimation input, instead of the actual link-flow observations from the field data. In addition, stochastic disturbances that follow an independent normal distribution with zero mean are added into simulated link counts so as to emulate the effect of measurement errors, and the standard deviation of random errors is set to 10% of the corresponding simulated link volume. AVI readers are assumed to cover all the entry/exit links of each OD-demand zone, indicating that OD AVI counts are available for each OD pair. In addition, both departure time interval in the dynamic OD-demand matrix and the AVI observation time interval are set to 5 min. Essentially, the lower level DTA problem can be viewed as a nonlinear constraint, and the proposed bilevel OD-demand framework could lead to multiple locally optimal solutions. As a result, the final estimation quality is sensitive to initial OD-demand values. If the initial guess is very close to the assumed actual OD flows, the estimation process is more likely to recover the true demand pattern with less iteration. On the other hand, if the initial demand values dramatically deviate from the ground truth demand flows, the iterative solution algorithm might converge to other locally optimal solutions. For each experiment in this study, the initial demand is assumed to be 50% of the assumed actual values. The OD pair from zone 16 to zone 4 has the largest demand volume with an average of 560 vehicles for every 5-min time interval, and demand flows for most of OD pairs range from 0 to 200 vehicles (per 5-min interval).

To quantify the accuracy of estimation results, the root mean squared error (RMSE) is used as a performance measure

$$\mathrm{RMSE} = \sqrt{\frac{\sum\limits_{i,j,\tau} \left(\hat{d}_{(i,j,\tau)} - d_{(i,j,\tau)}\right)^2}{T_{\mathrm{d}} \times |I| \times |J|}} \tag{29}$$

where $d_{(i,j,\tau)}$ = "true" demand volume for OD pair (i, j) during departure interval τ, and $\hat{d}_{(i,j,\tau)}$ = estimated demand volume during departure interval τ. Each reported value in the following experiments represents the mean of RMSE from five random replications.

B. Effect of Market-Penetration Levels of AVI Tags

To test the proposed nonlinear ordinary least-squares model without fixed-effect parameters, as shown in (21), identification rates in the first set of experiments are assumed to be 100%, and the following two scenarios are evaluated. 1) All OD pairs have exactly the same market-penetration rate. 2) The market-penetration rates are assumed to follow an independent uniform distribution with a range [0.75 $\bar{\beta}$, 1.25 $\bar{\beta}$] among different OD pairs at different departure times, where $\bar{\beta}$ is the mean market-penetration rate for all OD pairs.

Fig. 4 shows the change of the solution quality in response to increasing market-penetration rates. When the market-penetration rate is zero, OD demand is estimated only using dynamic link-count data, corresponding to a do-nothing case. At a market-penetration rate of 1%, the additional AVI information does not generate significant error reductions compared to the do-nothing case under both penetration scenarios, which can be explained by the fact that the number of tagged vehicles observed during each observation time interval at such a low market-penetration level is still too small to provide reliable samples for point-to-point split fractions. When average market-penetration rates are higher than 2.5%, both constant and random market-penetration schemes can produce considerable error reductions. It is noticeable that, when the market-penetration rate reaches 5%, the proposed model produces nearly 20% error reductions even under random tag penetration. The quality improvement in terms of RMSE further rises to around 40% at a 10% market-penetration rate, and marginal error reductions become relatively smaller beyond this market-penetration level. As expected, the random market-penetration scenario leads to a less perfectly representative sample data set, so the resulting estimation errors are relatively larger than those of the constant market-penetration scenario. However, the experiment results clearly demonstrate that, even under the random market-penetration scheme, the proposed OD-estimation model can still effectively utilize AVI information to enhance the observability of the dynamic OD-estimation problem, as long as the average market-penetration rates are sufficiently high enough to provide reliable samples for point-to-point split fractions.

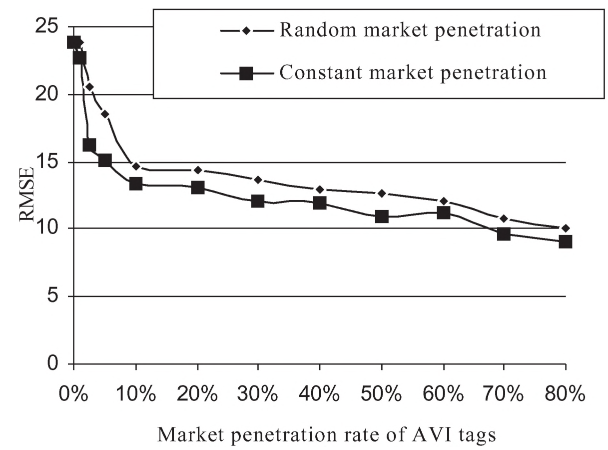

Fig. 4. Dynamic OD estimation as a function of different market-penetration rates.

C. Effect of Identification Rates

In the next set of experiments, the market-penetration rates are assumed to follow an independent uniform distribution with a range [7.5%, 12.5%] and the identification rates of AVI readers are assumed to follow a random uniform distribution with mean γ and a range $[\gamma - 0.1, \gamma + 0.1]$.

Table I summarizes the estimation errors and the corresponding percentage improvements compared with the do-nothing case (without AVI information) for the joint estimation model with fixed-effect parameters, as shown in (23), and the one-sided positive penalty formulation, as shown in (27). The experimental results show that the joint estimation model only produces a marginal performance improvement by using AVI information. The difficulty in applying this model can be attributed to two factors due to the inclusion of the fixed-effect parameters: the increase in the number of unknown variables and the resulting high nonlinearity in the multiplicative model structure. In contrast, even under a medium level of

TABLE I
PERFORMANCE OF OD-ESTIMATION MODELS IN THE PRESENCE OF IDENTIFICATION ERRORS

Identification Rates		0.8-1.0	0.7-0.9	0.6-0.8	0.5-0.7	0.4-0.6
Joint Estimation Model	RMSE	21.5	22.0	22.8	23.73	23.65
	% Improvement	9.4%	7.6%	4.2%	0.2%	0.5%
One-Sided Penalty Model	RMSE	18.12	19.17	20.84	22.22	23.02
	% Improvement	23.7%	19.4%	12.4%	6.5%	3.2%

identification rates [0.7, 0.9], the estimator with the one-sided linear-penalty form is still able to reduce error by nearly 20%, revealing that this parsimonious structure is quite robust to the imperfect observations. As identification rates decrease, the observed values of split fractions are considerably smaller than the corresponding true split fractions of tagged vehicles, and the estimated OD demand is less likely to be restricted by the one-sided penalty function. As a result, the relative value of AVI data tends to be insignificant and estimation errors from the one-sided model become larger. The one-sided linear-penalty formulation, in general, presents a tractable and intuitive approach for incorporating partially observed point-to-point sensor data with considerable identification errors.

D. AVI Detector Coverage

The above experiments use a complete AVI coverage scheme, which requires extensive detector installation and maintenance efforts for a general traffic network. Clearly, it is more desirable to maintain the estimation quality while minimizing the total number of AVI detectors in the network. A set of experiments is conducted below to reveal possible relations between the estimation quality and the percentage of covered OD-demand flows by the AVI readers. Sixteen detector-location schemes are randomly generated with identification rates of 100% and uniformly distributed market-penetration rates between [7.5%, 12.5%].

As indicated in Fig. 5, the estimation errors decrease with the increasing AVI coverage of OD-demand flows. The strong correlation between these two attributes suggests a basis for optimizing AVI detector locations for the estimation of population OD-demand flows, that is, it is advantageous to first cover the OD pairs with large demand flows.

The following experiments represent a preliminary attempt to optimize the AVI location for OD-demand-estimation purposes. Recognizing that the Irvine test bed network in the study has an obvious triangular structure, where the OD-demand flows among zones 1, 4, and 16 account for 36.4% of the total OD-trip desires, we first locate AVI detectors on the entry/exit links for these critical OD zones on the boundary. Next, zone 12 is equipped with AVI detectors to maximize the coverage on the remaining untracked OD flows in the network. In the same way, zones 13, 15, 5, 2, and 11 are sequentially added into the coverage plan, and the corresponding OD-demand estimation errors at each step are reported in Table II. The estimation error reduction as a function of the number of zones covered is further plotted in Fig. 6. Compared to the do-nothing case, nearly 30% error reduction is obtained by locating the AVI detectors to cover the six major zones, which capture 61.4% of the total OD flows in the study network. For the remaining OD zones carrying a lesser amount of significant amount of trip flows, only small marginal estimation error reductions could be obtained. According to this preliminary study, locating AVI detectors on major OD-demand zones with large traffic attraction/production can capture the essential OD distribution pattern in the network, and consequently improve the quality of OD estimates.

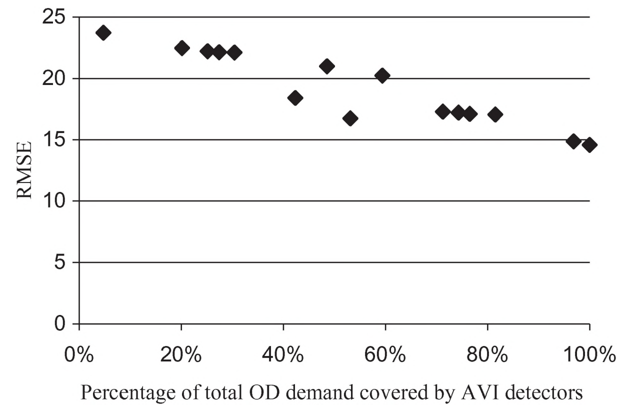

Fig. 5. Relationships between estimation performance and AVI detector coverage.

TABLE II
ESTIMATION PERFORMANCE UNDER DIFFERENT AVI LOCATION SCHEMES

# of Zones Covered	3	4	5	6	7	16
Zones Covered	1,4,16	+12	+13	+15	+5	All
% Demand Coverage	36.4%	47.3%	52.5%	61.4%	73.1%	100%
RMSE	20.90	19.12	18.41	16.81	16.65	14.78
% Improvement	12.1%	19.6%	22.6%	29.3%	30.0%	37.9%

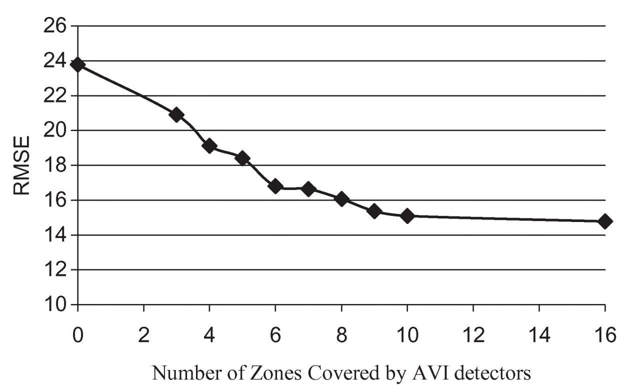

Fig. 6. Estimation errors as a function of the number of zones covered by AVI detectors.

VII. Concluding Remarks

Growing use of AVI technologies provides valuable point-to-point flow observations for estimating dynamic OD-trip desires. This paper has presented a novel OD-demand-estimation approach to effectively exploit OD-demand distribution information from AVI counts. A nonlinear ordinary least-squares model combines AVI counts with other available information sources into a multiobjective optimization framework. A joint estimation formulation with fixed-effect parameters and a one-sided linear-penalty formulation is further developed to deal with the possible identification and representativeness errors. The resulting models are solved using an iterative bilevel estimation framework. Based on a synthetic data set using the simplified Irvine test bed network, this study evaluates the performance of new estimation models and provides the following key findings.

1) Sufficient market penetration is required to obtain reliable information from AVI counts.
2) In the presence of identification errors, a parsimonious structure accounting for imprecise information can provide more robust estimates than a complex joint estimation model.
3) It is advantageous to locate AVI detectors on major OD-demand zones with large traffic attraction/production so as to capture the essential OD distribution pattern in the network.

This research has investigated possible benefits of AVI data for off-line OD-estimation applications through experimental control using synthetic data. The real-world AVI data, of course, are expected to provide valuable insight on the actual performance of alternative estimation models. Online DTA applications also call for further development of efficient and effective real-time demand estimation and prediction models and algorithms using point-to-point AVI data.

References

[1] M. Cremer and H. Keller, "Dynamic identification of flows from traffic counts at complex intersections," in *Proc. 8th Int. Symp. Transportation and Traffic Theory*, V. F. Hurdle *et al.*, Eds., Toronto, ON, Canada, 1981, pp. 121–142.

[2] ——, "A systems dynamics approach to the estimation of entry and exit O-D flows," in *Proc. 9th Int. Symp. Transportation and Traffic Theory,* J. Volmuller and R. Hamerslag, Eds., Delft, The Netherlands, 1984, pp. 431–450.

[3] ——, "A new class of dynamic methods for identification of origin-destination flows," *Transp. Res.*, vol. 21B, no. 2, pp. 117–132, 1987.

[4] E. Cascetta, D. Inaudi, and G. Marquis, "Dynamic estimators of origin-destination matrices using traffic counts," *Transp. Sci.*, vol. 27, no. 4, pp. 363–373, 1993.

[5] H. Tavana and H. S. Mahmassani, *Estimation of Dynamic Origin-Destination Flows From Sensor Data Using Bi-Level Optimization Method*, 2001, Washington, DC: National Research Council. Transportation Research Board CD-ROM Paper Preprints (TRB Paper No. 01-3241).

[6] H. Tavana, "Internally-consistent estimation of dynamic network origin-destination flows from intelligent transportation systems data using bi-level optimization," Ph.D. dissertation, Dept. Civil Eng., Univ. Texas, Austin, 2001.

[7] G. L. Chang and J. Wu, "Recursive estimation of time-varying origin-destination flows from traffic counts in freeway corridors," *Transp. Res.*, vol. 28B, no. 2, pp. 141–160, 1994.

[8] G. L. Chang and X. Tao, "Estimation of dynamic O-D distributions for urban networks," in *Proc. 13th Int. Symp. Transportation and Traffic Theory*, J. B.Lesort, Ed., Lyon, France, 1996, pp. 1–20.

[9] X. Zhou, X. Qin, and H. S. Mahmassani, *Dynamic Origin-Destination Demand Estimation Using Multi-Day Link Traffic Counts for Planning Applications*, Transportation Research Record, No. 1831, pp. 30–38, 2003.

[10] M. P. Dixon and L. R. Rilett, "Real-time OD estimation using automatic vehicle identification and traffic count data," *J. Comput.-Aided Civil Infrastruct. Eng.*, vol. 17, no. 1, pp. 7–21, Jan. 2002.

[11] C. Antoniou, M. Ben-Akiva, and H. Koutsopoulos, *Incorporating Automated Vehicle Identification Data Into Origin-Destination Estimation*, 2004, Washington, DC: National Research Council. Transportation Research Board CD-ROM Paper Preprints (TRB Paper No. 04-3341).

[12] K. Ashok, "Estimation and prediction of time-dependent origin-destination flows," Ph.D. dissertation, Dept. Civil Environmental Eng., MIT, Cambridge, 1996.

[13] N. J. Van der Zijpp, *Dynamic OD-Matrix Estimation From Traffic Counts and Automated Vehicle Identification Data*, Transportation Research Record, No. 1607, pp. 87–94, 1997.

[14] Y. Asakura, E. Hato, and M. Kashiwadani, "OD matrices estimation model using AVI data and its application to the Han-Shin expressway network," *Transportation*, vol. 27, no. 4, pp. 419–438, 2000.

[15] M. P. Dixon, "Incorporation of automatic vehicle identification data into synthetic OD estimation process," Ph.D. dissertation, Dept. Civil Eng., Texas A&M Univ., College Station, 2000.

[16] H. S. Mahmassani, "Dynamic network traffic assignment and simulation methodology for advanced system management applications," *Netw. Spatial Econ.*, vol. 1, no. 3–4, pp. 267–292, Sep. 2001.

Xuesong Zhou received the B.S. degree in railroad management from Northern Jiaotong University, Beijing, China, in 1995, and the Ph.D. degree in civil engineering from the University of Maryland, College Park, in 2004.

He is currently a Research Assistant Professor in the Department of Civil and Environmental Engineering at the University of Maryland. He has authored or coauthored several papers on traffic-flow theory, traffic origin–destination (OD) demand estimation and prediction, as well as railroad-train scheduling. His current research interests include intelligent transportation systems (ITS), modeling, optimization and simulation of transportation systems, and real-time dynamic traffic estimation and prediction.

Dr. Zhou is a member of INFORMS.

Hani S. Mahmassani (M'00) received the B.S. degree in civil engineering from University of Houston, Houston, TX, in 1976, the M.S. degree in transportation engineering from Purdue University, West Lafayette, IN, in 1978, and the Ph.D. degree in transportation systems from the Massachusetts Institute of Technology, Cambridge, in 1982.

He is the first holder of the Charles Irish Sr. Chair in Civil and Environmental Engineering at the University of Maryland, College Park, where he is also the Founding Director of the Maryland Transportation Initiative. He joined the University of Maryland in August 2002, after 20 years in the faculty at the University of Texas at Austin, where he was most recently the A. Abou-Ayyash Centennial Professor in Transportation Engineering, Professor of Management Science and Information Systems, and Director of the Advanced Institute for Transportation Infrastructure Engineering and Management. He specializes in multimodal transportation systems analysis, planning and operations, dynamic network modeling and optimization, dynamics of user behavior and telematics, telecommunication–transportation interactions, and real-time operation of logistics and distribution systems.

Dr. Mahmassani is a past President of the Transportation Science Section of the Institute for Operations Research and the Management Sciences, and a past President of the International Association for Travel Behavior Research. He chairs several technical and professional committees in the U.S. and internationally and is the Editor-in-Chief of *Transportation Science* and Associate Editor of *Transportation Research C (Emerging Technologies)*, as well as the IEEE Transactions on Intelligent Transportation Systems.