Abstract—This paper proposes a two-stage optimization model to determine the origin–destination (O–D) trip matrix and the heterogeneous sensor deployment strategy in an integrated manner for a vehicular traffic network using sensor information from active (camera-based license plate recognition) and passive (vehicle detector) sensors. The first stage solves the heterogeneous sensor selection and location problem to determine the optimal sensor deployment strategy, in terms of the selection of the numbers of the two sensor types and their installation locations, to maximize the traffic information available for the O–D matrix estimation problem. The traffic information includes the observed link flow, path trajectory, and path coverage information. The second stage leverages this traffic information to determine the network O–D matrix that minimizes the error between the observed and estimated traffic flows (link, O–D, and/or path). Correspondingly, two network O–D matrix estimation models are proposed where the link-based model incorporates the flow conservation rule between O–D and link flows and uses the link-node incidence matrix, and the path-based model assumes a given link-path incidence matrix. An iterative solution procedure is designed to determine the network O–D matrix and link flow estimates. Results from numerical experiments suggest that the path-based model outperforms the link-based model in the estimation of network O–D matrices. The relative contributions of combinations of the two sensor types to the network O–D matrix estimation problem are also analyzed. They suggest that active sensors provide valuable path information to solve the O–D matrix estimation problem, but at the cost of a significantly higher unit price. The study results have key implications for heterogeneous sensor selection and location strategies.

Index Terms—Heterogeneous sensor deployment strategy, intelligent transportation systems, sensor network location problem, origin–destination matrix estimation.

I. INTRODUCTION

A

n origin–destination (O–D) trip matrix characterizes the demand pattern in a vehicular traffic network. It is a crucial component for long-term transportation planning and short-term traffic management. Hence, the accurate estimation of the O–D matrix is a well-studied problem. Traditional approaches for estimating O–D matrices are based on manual surveys, such as household interview, license plate recording, and postcard mail-back. However, they suffer from high costs and potential data sampling or recording errors. To resolve the O–D data collection issues, methods have been proposed to estimate O–D trip matrices from link traffic counts or flows [1]. They include generalized least squares [2], maximum likelihood [3], entropy maximization [4], Kalman filter [5], and Bayesian inference [6] studies. However, using link flows to infer the O–D matrix can entail multiple solutions as these problems are typically underdetermined. Two approaches are proposed in the literature to address this underdetermined problem. The first approach is to solve the problem using a bilevel model [7], [8] which estimates the O–D matrix at the upper level based on some pre-specified user route choice decision rules (for example, under user equilibrium (UE) or stochastic user equilibrium (SUE)) in the lower level model. The second approach is to determine the O–D matrix from link flows using a path flow estimator (PFE) [9], [10]. Both approaches assume that specific travelers’ path choice decision rules are known based on some traffic assignment principles. However, the path choice decisions are not easy to obtain in practice.

With the rapid development of intelligent transportation systems (ITS) and sensor technologies, various sensors have been developed for traffic monitoring and data collection. In general, traffic sensors can be classified into two types: passive and active sensors. Passive sensors, such as vehicle detectors (VDs), are used to observe point measurements (e.g., link flow, occupancy, and speed) [11]. Active sensors, such as video cameras, are used for applications such as automatic vehicle identification (AVI) and license plate recognition (LPR). Through two-way communication between roadside equipment and an onboard unit, an active sensor based system provides point-to-point measurements (e.g., travel time, vehicle trajectory, and vehicle identification) [12], [13]. Past studies have incorporated link flow, path flow and/or vehicular flow pattern information provided by active sensors for the O–D matrix estimation problem [14]–[20]. Hence, they leverage vehicle trajectory information [14], [15] and path flow information [16]–[20], in addition to link flow information, for the network O–D matrix estimation problem.

The network sensor location problem (NSLP) [21], [22] seeks to determine the minimum number of traffic sensors...
and their installation locations to completely infer the network traffic conditions, and is similar to the observability problem [23]. From a network flow observability perspective, partial link flow information collected using strategic sensor deployment can infer full link flow information [22], [24], [25], and the upper bound on the number of required sensors without path enumeration can be derived [24], [25]. It has been applied in the context of route guidance [26], travel time data collection [27], travel time estimation [28], [29], and link flow inference [22]–[24], [30]. The quantity and content of the data collected from traffic sensors can substantially affect the performance of the O–D matrix estimation models [31]–[33]. Hence, the methods used to solve the NSLP can be adapted for the O–D matrix estimation problem [21], [22], [34]. However, past studies only sequentially solved the sensor location and O–D matrix estimation problems, using traffic information from a single sensor type [21], [35]–[38] or from heterogeneous sensors [19], [39]–[41]. Castillo et al. used scanning-link information collected by plate scanning technique for link and/or route flow estimation [14], [35], [38]–[40]. In addition, Castillo et al. further defined the flow amount of information (FAO) to analyze the number of linearly independent scanning links for route flow reconstruction [39], [40]. Some of their studies solved the network O–D matrix estimation problem and the heterogeneous sensors deployment problem (HSDP) in an integrated manner either by algebraic based methods [39], [40] or a Bayesian approach [37], [38].

A strategic sensor deployment plan for different sensor types and installation locations has important implications for the performance of the network O–D matrix estimation model, and the discrepancy between the estimated and true O–D flows is crucial information to verify the appropriateness of the sensor deployment strategy. That is, solving for the sensor deployment strategy and then assuming it as given in the O–D matrix estimation problem in the sequential approach does not leverage information needs related to accurate O–D matrix estimation to guide the selection and location of deployed sensors. Hence, there is a key need to solve the heterogeneous sensor deployment problem and network O–D matrix estimation problem in an integrated manner.

The problem of solving the heterogeneous sensors deployment problem and network O–D matrix estimation problem is referred to as the HSDP-OD problem in this paper. The corresponding problem using traffic information from a single sensor type is labeled the NSLP-OD problem. Past studies to solve the NSLP-OD or HSDP-OD problem usually use two steps in a sequential approach. The first step determines a strategy for locating sensors in the network, and the second step solves the O–D matrix estimation problem based on the sensor deployment from the first step [14], [19], [21], [35], [36], [38]–[41]. Hence, the sequential approach is a one-shot procedure as the O–D matrix estimation results are not fed back to the NSLP/HSDP stage. Castillo et al. [37] propose a Wardrop-minimum variance (WMV) method assignment problem and Bayesian network (BN) approach for the link and O–D flow estimation problem. The NSLP-OD problem is solved by sequentially selecting one link as the equipped link to update the O–D matrix estimate until the budget constraint is violated. However, some specific input data or model assumptions are typically required in a sequential approach for the NSLP-OD or HSDP-OD problem. For example, a link-path incidence matrix is necessary for the developed sensor location and O–D matrix estimation models [14], [21], [35]–[38], [40], [41]; this may entail solving a user equilibrium traffic assignment problem to identify the paths [37]. However, assumptions on the availability of some additional input data may not be realistic from a data availability perspective. For instance, the sensor location models require prior O–D/path flow information [35]–[37], [40], [41], or link flow proportions [37], [41] to determine the sensor locations. The O–D matrix estimation models assume the prior O–D/path information [14], [21], [35]–[38], [41], or path assignment probability to be known [41], or that link flow proportions can be estimated [21], [37].

Unlike previous studies for the NSLP-OD problem [14], [21], [35]–[41], this study specifically investigates the HSDP-OD problem. And, unlike the sequential approach used previously for the HSDP-OD problem [19], [41], this study addresses the HSDP-OD problem in an integrated manner using a two-stage optimization model where the error on the O–D matrix estimate in the second-stage model is fed back to modify the sensor deployment strategy in the first-stage model until some pre-specified error thresholds are met. In the first stage, the HSDP model determines the optimal numbers of active (camera-based license plate recognition) and passive (vehicle detector) sensors and their installation locations to maximize the traffic information available for the O–D trip matrix estimation. This traffic information consists of the observed link flows, path trajectories and path coverage information. In the second stage, the O–D matrix estimation model leverages this traffic information to determine the network O–D matrix that minimizes the error between the observed and estimated traffic flows (link, O–D and/or path). Two network O–D matrix estimation models are proposed. The link-based model determines the O–D matrix based on the link-node incidence matrix under flow conservation rules without requiring strong assumptions related to prior knowledge/data. The path-based model assumes a given link-path incidence matrix and leverages active sensor information. A specially designed feedback mechanism is proposed to update parameters (weights of the objective function terms) in the heterogeneous sensors deployment model of the first stage based on the performance of the O–D matrix estimates in the second stage. Through this feedback mechanism, the proposed two-stage model captures the interactions between the heterogeneous sensor deployment strategy and the network O–D matrix estimation problem. This establishes a bridge between the HSDP and the O–D matrix estimation problem, and represents an integrated approach to solve the HSDP-OD problem. The results from an empirical study using the Sanmin network in Taiwan indicate that the proposed integrated optimization model can provide network O–D matrix estimates as well as the numbers and locations of the two sensor types consistent with the corresponding objectives at the two stages.

The remainder of this paper is organized as follows. Section II characterizes the heterogeneous sensor based traffic information and presents the integrated model for the HSDP-OD problem, including the formulations of the heterogeneous sensors deployment and O–D matrix estimation models.
Section III describes the solution procedure with the feedback mechanism to solve the integrated model. Section IV discusses the results of numerical experiments based on a small test network and the Sanmin network. Finally, concluding remarks are presented in Section V.

II. Model Formulation

This paper proposes an integrated two-stage optimization model for the HSDP-OD problem. The first stage is the heterogeneous traffic sensors deployment model which seeks to determine the sensor deployment strategy that optimizes the traffic information available to the second-stage O–D matrix estimation problem. The weights of the objective function terms in this model are functions of the errors between the observed and estimated link and path/O–D flow data determined in the second-stage model. The second stage is the O–D matrix estimation model that leverages the traffic information obtained in the first-stage model to determine the O–D matrix that minimizes the errors between the observed and estimated traffic flow data. The traffic information from the first stage and the traffic flow data errors from the second stage integrate this two-stage optimization model. The two-stage model for the HSDP-OD problem is described hereafter.

A. Heterogeneous Sensor Based Traffic Information

The passive sensors used in this study are VDs and the active sensors are LPRs. The heterogeneous sensors deployment model determines the optimal heterogeneous sensor deployment strategy, in terms of the selection of the numbers of LPRs and VDs and their installation locations, to maximize the available traffic information for the O–D matrix estimation problem. The observed traffic information using the VDs and LPRs can be categorized into three types: (i) link flow information, (ii) path trajectory information, and (iii) path coverage information.

Link Flow Information: Link flow information can be collected on links equipped with either VR or LPR sensors. Hence, the number of VDs and LPRs deployed can be regarded as the number of pieces of observed link flow information, and are formulated using 0-1 integer decision variables as follows.

\[ x_{ij} = \begin{cases} 1, & \text{if link } ij \text{ is equipped with a VD} \\ 0, & \text{otherwise} \end{cases}, \quad \forall i,j \in \mathcal{N} \quad (1) \]

\[ y_{ij} = \begin{cases} 1, & \text{if link } ij \text{ is equipped with an LPR} \\ 0, & \text{otherwise} \end{cases} \]

\[ \mathcal{A}: \text{link set}; \]
\[ \mathcal{N}: \text{node set}. \]

Path Trajectory Information: For a given link-path incidence matrix, each element in this matrix indicates if a specific path passes through a link. Hence, such an element in a link-path incidence matrix can be expressed in Eq. (2).

\[ \delta_{ij}^p = \begin{cases} 1, & \text{if path } p \text{ contains link } ij \\ 0, & \text{otherwise} \end{cases} \quad (2) \]

The path trajectory information of a given O–D pair can be collected by LPR sensors installed at various links along this specific path, and is formulated by the LPR decision variable and the delta function as follows.

\[ \text{Path trajectory information: } \sum_{p \in \mathcal{P}} \sum_{ij \in \mathcal{A}} y_{ij} \cdot \delta_{ij}^p \quad (3) \]

where \( \mathcal{P} \) is the path set.

In addition to its capability to collect link flow information, an LPR sensor is able to actively track a vehicle’s identification. Thereby, installing LPR sensors at some strategic links would provide full or partial path flow information of a given O–D pair. As a result, the flow on a given O–D pair can be (partially) observed by summing up the corresponding (partially) collected path flows.

Path Coverage Information: When different paths share the same path trajectories, it is difficult to identify the observed path flows of these paths. Based on the path trajectory information of a given link-path incidence matrix, a distinction function [35, 42] is introduced to distinguish between paths, and expressed in Eq. (4).

\[ d (p^0, p^1, \delta_{ij}^p) = \begin{cases} 1, & \text{if } \delta_{ij}^{p^0} \neq \delta_{ij}^{p^1} \\ 0, & \text{otherwise} \end{cases} \quad (4) \]

The distinction function identifies the difference between two paths by comparing their trajectories. Specifically, if \( d (p^0, p^1, \delta_{ij}^p) \) is 1, it means that link \( ij \) is equipped with an LPR sensor, and this LPR-equipped link is able to distinguish between paths \( p^0 \) and \( p^1 \) based on a given link-path incidence matrix.

Similar to the definition of the path trajectory information, the path coverage information is formulated by the LPR decision variable and the distinction function as follows.

\[ \text{Path coverage information: } \sum_{ij \in \mathcal{A}} y_{ij} \cdot d (p^0, p^1, \delta_{ij}^p), \forall p^0, p^1 \in \mathcal{P} \quad (5) \]

When the value of Eq. (5) is 0, it indicates that a current LPR sensor deployment configuration cannot distinguish the difference between paths \( p^0 \) and \( p^1 \). When the value of Eq. (5) is 1 or greater than 1, it means that at least one LPR sensor can distinguish the difference between paths \( p^0 \) and \( p^1 \). In addition, in order to maximize the capacity of the path differentiation capability of an LPR sensor, a path coverage variable, \( m_{p^0, p^1} \) is defined as follows.

\[ \sum_{ij \in \mathcal{A}} y_{ij} \cdot d (p^0, p^1, \delta_{ij}^p) \geq m_{p^0, p^1}, \forall p^0, p^1 \in \mathcal{P} \quad (6) \]
where, 
\[
m_{p^{0}, p^{1}} = \begin{cases} 
1, & \text{if paths } p^{0} \text{ and } p^{1} \text{ can be differentiated by LPR sensors} \\
0, & \text{otherwise.}
\end{cases}
\]

Eq. (6) depicts that the path coverage variable is the lower bound for the number of pieces of path coverage information, where the collected path coverage information for the entire network by deploying an LPR sensor at link \( ij \) is as large as possible.

### B. Heterogeneous Sensors Deployment Model

The first-stage heterogeneous sensors deployment model is formulated as an integer program to determine the numbers of LPRs and VDs and their installation locations to maximize the available traffic information subject to constraints on the available budget, network topology, and set covering rules. To quantify the relative contributions of the three types of traffic information to the estimation of the O–D matrix in the second stage, weights are introduced for the corresponding terms in the first-stage objective function. Thereby, the heterogeneous sensor deployment strategy in the first stage, through the three types of traffic information observed using these sensors, is linked to the second-stage objective of determining the O–D matrix with the minimum amount of error. The first stage model formulation is as follows.

Max  \[
\begin{align*}
\sum_{ij \in A} \left( \sum_{p \in P} w_{\alpha} \cdot \zeta(\hat{i}_{ij}) \cdot (x_{ij} + y_{ij}) \right) \\
+ \sum_{ij \in A} \left( \sum_{p \in P} w_{\beta} \cdot \zeta(\hat{t}^{rs}) \cdot y_{ij} \cdot \delta_{ij}^{p} \right) \\
+ \sum_{ij \in A} \left( \sum_{p \in P} w_{\gamma} \cdot \zeta(\hat{t}^{rs}, \hat{r}^{s1}) \cdot m_{p^{0}, p^{1}} \right)
\end{align*}
\]

s.t.  \[
\sum_{ij \in A} \left( x_{ij} + y_{ij} \right) \cdot \delta_{ij}^{p} \geq 1, \ \forall p \in P
\]
\[
x_{ij} + y_{ij} \leq 1, \ \forall i, j \in A
\]
\[
C_{D} : \sum_{ij \in A} x_{ij} + C_{L} : \sum_{ij \in A} y_{ij} \leq C
\]

where,  
\[
C_{D} \quad \text{the cost of a VD (passive sensor);}
\]
\[
C_{L} \quad \text{the cost of an LPR (active sensor);}
\]
\[
C \quad \text{total budget;}
\]
\[
\alpha_{ij} \quad \text{initial weight of link flow information for link } ij;
\]
\[
\beta_{ij}^{p} \quad \text{initial weight of the } p^{th} \text{ path trajectory information for link } ij;
\]
\[
\gamma_{p^{0}, p^{1}}^{\hat{t}^{rs}, \hat{r}^{s1}} \quad \text{initial weight of the path coverage information for the } p^{0} \text{ path of } t^{rs}, \text{ and the } p^{1} \text{ path of } t^{rs};
\]
\[
w_{\alpha} \quad \text{weight for link flow information;}
\]
\[
w_{\beta} \quad \text{weight for path trajectory information;}
\]
\[
w_{\gamma} \quad \text{weight for path coverage information;}
\]
\[
\zeta(\hat{i}_{ij}) \quad \text{error function based on link flow estimate at link } ij, \text{ where } \hat{i}_{ij} \text{ is the estimated flow on link } ij;
\]
\[
\zeta(\hat{t}^{rs}) \quad \text{error function based on O–D flow estimate for O–D pair } rs, \text{ where } \hat{t}^{rs} \text{ is the estimated flow for O–D pair } rs;
\]
\[
\zeta(\hat{t}^{rs}, \hat{r}^{s1}) \quad \text{error function based on O–D flow estimates for O–D pairs } r_{0}s_{0} \text{ and } r_{1}s_{1}.
\]

Eq. (7) is the objective function whose goal is to maximize the number of pieces of observed traffic information on link flow, path trajectory, and path flow coverage. The decision variables are \( x_{ij}, y_{ij}, \) and \( m_{p^{0}, p^{1}} \). It is assumed here that maximizing the number of pieces of these three types of information collected using the VDs and LPRs will maximize the traffic information available to the second stage problem. The three terms in the objective function correspond to the observed link flow information, path trajectory information, and path coverage, respectively. The corresponding weights \( (w_{\alpha}, w_{\beta}, w_{\gamma}) \) reflect the relative importance of each type of information for the O–D matrix estimation problem. It is important to note here that each of the weights depends on a pre-determined initial weight and errors on link flow and O–D matrix estimates in the second-stage model. In the solution procedure described in Section III, these weights are iteratively updated when the O–D matrix estimation results are obtained in the second stage. Eq. (8) introduces the path coverage variable, \( m_{p^{0}, p^{1}} \) and states that if a path can be distinguished by at least one LPR sensor based on the distinction function from a given link-path incidence matrix, this path is covered. The number of the path coverage variables is dependent on the size of the path set. This number can be large, and some analytical approaches have been proposed to reduce this size [35]. Eq. (9) illustrates the set covering rule, which indicates that each path should be observed by at least one VD or LPR sensor [35], [42]. Eq. (10) reflects that a link is at most equipped with one sensor, which can be either active or passive. Eq. (11) is the budget constraint that incorporates the unit costs of both sensor types. Generally, the \( C_{L} / C_{D} \) ratio is approximately 100 since the system infrastructure of an active-type sensor system requires higher initial and maintenance costs.

### C. Network O–D Matrix Estimation Models Using Heterogeneous Sensor Based Traffic Information

When a link is equipped with a VD or LPR sensor \( (x_{ij} \text{ or } y_{ij}) \), the corresponding link flow information, \( \pi_{ij} (x_{ij}, y_{ij}) \) can be collected and/or observed. If an LPR sensor deployment configuration is determined, the collection of \( y_{ij} \)'s and \( m_{p^{0}, p^{1}} \)'s, which are respectively denoted as \( Y \) and \( M \), can identify the path flows based on the mapping of the observed path trajectory information and path coverage variable in a given link-path incidence matrix. The observed path flow information is defined as \( T_{p}^{\pi}(Y, M) \).
The heterogeneous sensor deployment model developed in the first stage of the integrated model can provide subsets of link and/or path flows. The network O–D matrix estimation model developed in the second stage leverages this traffic information on the observed link/path flows and path trajectory/coverage information. Accordingly, two versions of the network O–D matrix estimation model are developed. One is the link-based model which abides by the link flow conservation rule based on a link-node incidence matrix. The other is the path-based model which estimates the path flows in light of a given link-path incidence matrix. Both models are formulated as nonlinear least squares (NLS) programs that minimize the errors on the estimated link/path flows and/or O–D matrix. The models are discussed hereafter.

**Link-Based O–D Matrix Estimation Model:** The link-based model is developed based on a non-proportional traffic assignment principle, and is formulated as a nonlinear program with linear constraints as follows.

\[
\begin{align*}
\text{Min} \quad & \sum_{(i,j) \in A} \left[ \frac{\hat{v}_{ij} - \sum_{ij \in A} \hat{v}_{ij}}{\sum_{ij \in A} \hat{v}_{ij}} - \frac{\tau_{ij}(x_{ij}, y_{ij})}{\sum_{ij \in A} \tau_{ij}(x_{ij}, y_{ij})} \right]^2 \\
+ & \sum_{r} \sum_{s} \left[ \frac{\sum_{r} \sum_{s} \sum_{p} f_{ps}^r}{\sum_{r} \sum_{s} \sum_{p} \hat{f}_{ps}^r(Y, M)} - \frac{\sum_{r} \sum_{s} \sum_{p} \hat{T}_{ps}^r(Y, M)}{\sum_{r} \sum_{s} \sum_{p} T_{ps}^r(Y, M)} \right]^2 \\
\text{s.t.} \quad & \hat{V} \cdot L = 0 \\
& \sum_{i} \hat{v}_{ir} + \sum_{s \in S} \hat{v}_{is} = \sum_{j} \hat{v}_{rj}, \forall r \in R \\
& \sum_{i} \hat{v}_{is} = \sum_{r \in R} \hat{r}_{r}^s + \sum_{j} \hat{v}_{sj}, \forall s \in S \\
& \hat{r}_{ps}^s \geq 0, \forall r \in R, s \in S, p \in P \\
& \hat{v}_{ij} \geq 0, \forall i, j \in A
\end{align*}
\]

where,
\[
\begin{align*}
\hat{v}_{ij} & \quad \text{estimated flow on link } ij; \\
\hat{V} & \quad \text{a vector of estimated link flows;}
\tau_{ij}(x_{ij}, y_{ij}) & \quad \text{observed link } ij \text{ flow based on the deployment of VD or LPR sensors;}
\hat{T}_{ps}^r(Y, M) & \quad \text{observed } p \text{th path flow between O–D pair } rs \text{ based on the deployment of LPR sensors;}
\hat{r}_{ps}^s & \quad \text{estimated O–D flow between path } rs; \\
L & \quad \text{link–node incidence matrix;}
R & \quad \text{a set of origin nodes;}
S & \quad \text{a set of destination nodes;}
Y & \quad \text{the collection of } y_{ij};
M & \quad \text{the collection of } m_{ij}.
\end{align*}
\]

Eq. (12) is the objective function which minimizes the errors between the (partially) observed and estimated link and O–D flows. The flow estimation errors are normalized to circumvent the effect of difference in the orders (in terms of the absolute values) of the link and O–D flow estimates. In Eq. (12), two categories of the traffic information are collected. One is the observed link flow, \(\tau_{ij}(x_{ij}, y_{ij})\) which is provided by the VD and LPR sensors. The other is the observed path flow, \(\hat{T}_{ps}^r(Y, M)\) which is determined by the LPR sensor deployment configuration and path coverage variable conditions (i.e., Eq. (8)). The observed O–D flows are the summation of the partially or fully observed path flows from LPR sensors. Because sensors cannot be deployed on all links of a network due to the budget constraint, the observed traffic information may include some (partial) link and/or path flows for a given O–D pair. Eq. (13) is the link flow conservation constraint based on a link-node incidence matrix while origin and destination nodes are excluded from the flow conservation rule. When origin or destination nodes are intermediate nodes, the flow conservation rule for the origin and destination nodes is described by Eqs. (14) and (15), respectively. Eq. (14) states that the flow departing from an origin node is equal to sum of the flow originating at the origin and the pass-through flow(s). Similarly, for the destination nodes, Eq. (15) states that the incoming link flows to a destination node are composed of the destination flows and pass-through flow(s). Eq. (16) is the inequality constraint indicating that, for a particular O–D pair, the estimated O–D flow should be greater than or equal to the O–D flow partially or fully observed by LPR sensors. Eqs. (17) and (18) are the non-negativity constraints on the estimated O–D and link flows, respectively.

**Path-Based O–D Matrix Estimation Model:** The O–D matrix estimation model can alternatively be formulated as a path-based model under a given link-path incidence matrix by using path flows as the decision variables. It is developed by using a path flow estimator, and is formulated as a nonlinear program with linear constraints as follows.

\[
\begin{align*}
\text{Min} \quad & \sum_{(i,j) \in A} \left[ \frac{\hat{v}_{ij} - \sum_{ij \in A} \hat{v}_{ij}}{\sum_{ij \in A} \hat{v}_{ij}} - \frac{\tau_{ij}(x_{ij}, y_{ij})}{\sum_{ij \in A} \tau_{ij}(x_{ij}, y_{ij})} \right]^2 \\
+ & \sum_{r} \sum_{s} \left[ \frac{\sum_{r} \sum_{s} \sum_{p} f_{ps}^r}{\sum_{r} \sum_{s} \sum_{p} \hat{f}_{ps}^r(Y, M)} - \frac{\sum_{r} \sum_{s} \sum_{p} \hat{T}_{ps}^r(Y, M)}{\sum_{r} \sum_{s} \sum_{p} T_{ps}^r(Y, M)} \right]^2 \\
\text{s.t.} \quad & \hat{F} \cdot \delta = 0 \\
& \hat{f}_{ps}^r \geq \hat{T}_{ps}^r, \forall r \in R, s \in S, p \in P \\
& \hat{f}_{ps}^r \geq 0, \forall r \in R, s \in S, p \in P \\
& \hat{u}_{ij} \geq 0, \forall (i, j) \in A
\end{align*}
\]

where,
\[
\begin{align*}
\hat{F} & \quad \text{a vector of estimated path flows;}
\delta & \quad \text{link-path incidence matrix;}
\hat{f}_{ps}^r & \quad \text{estimated path flow for the } p \text{th path of O–D pair } rs.
\end{align*}
\]
Eq. (19) is similar to Eq. (12) and includes two minimization terms: (i) the normalized errors between observed and estimated link flows, and (ii) the normalized errors between observed and estimated path flows. Eq. (20) is the flow conservation constraint under a given link-path incidence matrix. Eq. (21) is the inequality constraint indicating that, for a particular O–D pair, the estimated path flow should be greater than or equal to the path flow partially or fully observed by LPR sensors. Eqs. (22) and (23) are the non-negativity constraints for both path and link flows.

Note that the link-based O–D matrix estimation model estimates the network O–D flows using the simplified flow relationship between origin/destination nodes and adjacent links (see Eqs. (14) and (15)), and infers link flows based on a link-node incidence matrix (see Eq. (13)). For small networks, composed of a few nodes and links, a link-based O–D matrix estimation model could provide satisfactory solutions for both the link and O–D flow estimates. If a network is large or has a complex structure, the link-based model may not adequately capture the flow relationship between origin/destination nodes and intermediate links. Then, a path-based approach, such as the PFE method, which assumes a known link-path incidence matrix, could be used to solve the network O–D matrix estimation problem.

III. SOLUTION PROCEDURE

An iterative solution procedure with a feedback mechanism is proposed to solve the two-stage optimization model for the HSDP-OD problem. The sensor deployment strategy is determined in the first stage. A set of link and O–D flows is computed in the second stage based on the multiple pieces of traffic information provided by the two types of traffic sensors. Then, a feedback mechanism is used to update the weights in Eq. (7) of the heterogeneous sensors deployment model based on the performance of the O–D matrix estimation model in the second stage.

When the optimal strategy for the deployment of the two sensor types is determined in the first stage, these sensor deployments provide observations on the link flows \( \pi_{ij}(x_{ij}, y_{ij}) \) and path flows \( \mathbf{f}_p(Y, M) \). Based on these flow observations, the flows on the unsequenced links and the O–D flow are computed. Then, the errors associated with link and O–D flow estimates are used to update the weights in Eq. (7) by using functions that link the errors and the weights, as illustrated earlier in this section. This study adopts the mean absolute percent error (MAPE) [43] to evaluate the link flow and O–D matrix estimation errors as follows:

\[
\varepsilon(\hat{v}_{ij}) = \text{MAPE}(\hat{v}_{ij}) = \frac{|\pi_{ij} - \hat{v}_{ij}|}{\pi_{ij}} \times 100\%
\]

\[
\varepsilon(\hat{f}^s) = \text{MAPE}(\hat{f}^s) = \frac{|\mathbf{f}^s - \hat{f}^s|}{\mathbf{f}^s} \times 100\%
\]

where,

\( \hat{v}_{ij} \) estimated flow on link \( ij \);
\( \pi_{ij} \) observed flow on link \( ij \);
\( \hat{f}^s \) estimated O–D flow between pair \( rs \);
\( \mathbf{f}^s \) observed O–D flow between pair \( rs \).

The error functions in terms of the MAPEs are used in this study in two aspects: (i) they represent the criteria to evaluate the accuracy of the link flow and O–D matrix estimates, and (ii) they are inputs to update the weights in the first-stage model.

The details of the iterative solution procedure are summarized hereafter:

**Step 0:** Obtain the input data, including network topology, unit costs of VD and LPR sensors, total budget, and initial weights for the three types of traffic information. Initialize the errors \( (\varepsilon(\hat{v}_{ij}))_0 = 0, \varepsilon(\hat{f}^s))_0 = 0 \).

**Step 1:** Determine the sensor deployment strategy using the heterogeneous sensor deployment model (Eqs. (7)–(11)) given the input data of Step 0. The decision variables in Eqs. (7)–(11) are 0–1 variables, which provide the numbers of VDs and LPRs selected and the corresponding deployment locations. Since Eqs. (7)–(11) represent an integer program, an integer programming solver (Gurobi Optimizer 5.6, Gurobi Optimization, Inc.) is used to solve the problem. The procedure is compiled in MATLAB R2013a (The MathWorks, Inc.).

**Step 2:** Obtain the flow information available from the first-stage decision variables, such as link flows and path flows.

**Step 3:** Incorporate the flow information obtained in Step 2 into the O–D matrix estimation model used, where Eqs. (12)–(18) represent the link-based model, and Eqs. (19)–(23) represent the path-based model. Determine the solution as the estimated link and O–D matrix that minimizes the difference between the (partially) observed and estimated link and O–D flows. For the path-based model, the O–D matrix estimate is determined by minimizing the error between the (partially) observed and estimated link and path flows based on the given link-path incidence matrix. Both O–D matrix estimation models are formulated as NLS problems, and solved by a nonlinear programming solver (Gurobi Optimizer 5.6, Gurobi Optimization, Inc.). The procedure is compiled in MATLAB R2013a (The MathWorks, Inc.).

**Step 4:** Calculate the MAPEs of the estimated link and O–D flows (Eqs. (24), (25)).

**Step 5:** Set the termination criteria that the average MAPEs of the link and O–D flow estimates are less than pre-specified thresholds. If the criteria are met for the link and O–D flow estimates, proceed to Step 7; otherwise, proceed to Step 6.

**Step 6:** Update the weights in Eq. (7) (the observed link flow, path trajectory and path coverage information) using Eqs. (26)–(28) based on the MAPE calculations in Step 4, and feed them back to Step 1.

\[
w_\alpha[\alpha_{ij}, \varepsilon(\hat{v}_{ij})]_{k+1} = \alpha_{ij} + \varepsilon(\hat{v}_{ij})_k
\]

\[
w_\beta[\beta_{ij}^p, \varepsilon(\hat{f}^s)]_{k+1} = \beta_{ij}^p + \varepsilon(\hat{f}^s)_k
\]

\[
w_\gamma[\gamma_{ij}^{r_0}, \varepsilon(\hat{r}_1^{r_1})]_{k+1} = \gamma_{ij}^{r_0} + \varepsilon(\hat{r}_1^{r_1})_k
\]

**Step 7:** Output the final results for the integrated two-stage model for the HSDP-OD problem in terms of estimated link and O–D flows.
Note that the updated weights at each iteration guide the search direction for the candidate links to be equipped with different types of sensors in a network-wide sense. These weight updating functions (i.e., Eqs. (26)–(28)) factor the relative importance of different sources of link flow and path trajectory/coverage information in an increasing order of the magnitude. The general principle for selecting the initial weights is to guarantee $\alpha_{ij} < \beta_{ij}^p < \gamma_{t_0}^{r_0} r_{t_1}^{s_1} \gamma_{t_1}^{r_1} r_{t_2}^{s_2}$.

IV. NUMERICAL ANALYSIS

To evaluate the proposed HSDP-OD model, different initial values are specified for the weights of the three different information types in Eq. (7) based on the amount of information they are expected to provide for the second stage. As the path coverage information provides inputs related to both path flows and O–D flows, it is assigned the highest weight. The path trajectory information provides data on specific vehicular movements which can be more insightful for O–D flows than the link flow information. Hence, it is assigned the second highest weight. Based on this, for the experiments in this section, the initial value of $\alpha_{ij}$ is 2, $\beta_{ij}^p$ is 10, and $\gamma_{t_0}^{r_0} r_{t_1}^{s_1}$ is 20 in the heterogeneous sensors deployment model. The HSDP-OD model is evaluated using two networks under a budget constraint: a hypothetical fishbone network [22] and a real network (Sanmin, Taiwan).

A. Fishbone Network

The fishbone network contains 10 nodes, 18 unidirectional links and 4 O–D pairs (see Fig. 1). To evaluate the HSDP-OD model in the fishbone network, we assume that the unit costs of LPR and VD are 80 and 20, respectively, and the total budget is 470. In solving the HSDP-OD model, the termination criteria are set as follows: the average MAPE of the O–D flow estimate is less than 5% and the average MAPE of the link flow estimate is less than 15%.

As previously stated, for a small network a link-based O–D matrix estimation model may suffice to provide satisfactory flow estimation results. Hence, the link-based O–D matrix estimation model is used in the HSDP-OD problem for the fishbone network. Fig. 2 shows the MAPEs for the link and O–D flow estimates for several iterations.

Fig. 2 indicates that the accuracy of the O–D flow estimates does not depend only on the accuracy of the link flow estimates in the link-based model. A comparison of the link and O–D flow estimates of the fourth and fifth iterations reveals that the accuracy of the link flow estimates in the fourth iteration was higher than that in the fifth iteration, but the accuracy of the O–D flow estimates was higher in the fifth iteration. However, there are several cases where higher accuracy in the link flow estimates also leads to higher accuracy in the O–D flow estimates as shown in the first, third, and sixth iterations. Fig. 2 illustrates that after the fourth iteration, there is a cyclic pattern every seven iterations. The fourth to ninth iterations exhibited the first cycle, and the tenth to fifteenth iterations follow the same trend, and so on. This is a reflection of the iterative solution procedure with the feedback mechanism as the MAPEs in the second stage lead to changes in the values of the weights in the first-stage model, which leads to modifications in the sensor deployment. The solution can be terminated at a specific iteration in which the error termination criteria are met. In Fig. 2, the optimal results are found in the sixth iteration, where the average MAPE for the link flow estimates is 13.5%, and the average MAPE for the O–D flow estimates is 2.3%. The corresponding sensor deployment configuration is shown in Fig. 3. The results indicate that under a budget constraint eight sensors (including five LPR sensors and three VDs) are required and the sensor deployment rate is 45%. Due to the budget constraint in the HSDP-OD model, link flow information for the unequipped links is inferred based on the observed link flows using the flow conservation equations (see Eqs. (13)–(15)), and hence the accuracy gap in the link flow estimates.

The performance of the optimal sensor deployment configuration (five LPR sensors and three VDs; the result of the sixth iteration in Fig. 2) is benchmarked against three scenarios with only one sensor type considered (the homogeneous sensor case using VDs). In the “partial VD only” scenario (the replacement scenario), the five LPR sensors of the optimal solution are replaced by five VDs, resulting in eight VDs. This scenario would illustrate the marginal benefit of using LPR sensors to replace the VDs in the optimal solution. In the alternative “partial VD only” scenario (the fixed scenario), we also execute the two-stage optimization model with eight VDs. In the
“full VD only” scenario, the number of VDs allowable by the budget constraint is used. Since only eighteen links exist, this results in the usage of eighteen VDs, implying perfect observation of link flow information in this hypothetical example. This scenario ensures consistency with the budget constraint. The results of the four scenarios are shown in Table I. The average MAPEs for the link and O–D flow estimates of the partial VD only scenario (replacement) are 24.1% and 119.2%, respectively. The average MAPEs for the link and O–D flow estimates of the partial VD only scenario (fixed) are 14.8% and 96.4%, respectively. The average MAPEs for the link and O–D flow estimates of the full VD only scenario are 0 and 16.7%, respectively. For the partial VD only scenarios (both the replacement and fixed scenarios), full observability of link flows cannot be achieved due to an insufficient number of VDs, since VDs only provide link flow information. The partial path information provided by the five LPR sensors can be used to identify partial path flows. The partial path flow information reduces the average MAPE for the O–D flows from 16.7% to 2.3%. This result illustrates that the additional budgetary investment to replace five VDs by five LPRs results in a significant enhancement to the accuracy of the O–D matrix estimate. Hence, there can be performance enhancement benefits by using active sensor types that provide valuable data beyond the link flow information generated by VDs. The benefit of deploying LPR sensors is that some crucial traffic information can be obtained for the O–D matrix estimation model, such as path trajectories/flows, which cannot be observed by VDs.

### B. The Sanmin Network

The real network considered is located in the Sanmin district in Kaohsiung City, Taiwan, and is referred to as the Sanmin network in this study [44]. The Sanmin network consists of 72 nodes, 202 unidirectional links and 156 O–D pairs (see Fig. 4). The Sanmin network is located in the central district of Kaohsiung City in Southern Taiwan.

The true O–D flow data was obtained from a home survey of the Kaohsiung metropolitan area [45]. For model evaluation purposes, we obtained only the traffic data relevant to the Sanmin district. This dataset is used to evaluate the performance of the HSDP-OD model. Based on the cost databases from the RITA (U.S. Department of Transportation), the unit costs for an LPR and a VD were set at $70,000 and $600, respectively, and the total budget was assumed to be $3,000,000. The termination criteria for the solution procedure are: the average MAPE of the link flow estimates is less than 15% and the average MAPE of the O–D flow estimates is less than 20%. As the Sanmin network is relatively large, the termination criterion is less tight for the O–D flow estimates compared to that of the fishbone network. Both the link-based model and the path-based O–D matrix estimation models are analyzed, and the termination criteria for both models are the same (15%–20%). Fig. 5 shows the results of the link and O–D flow estimates across different iterations under the link-based model. Fig. 6 shows the results for the path-based model.

As in the fishbone network, Fig. 5 illustrates that the accuracy of O–D flow estimates using the link-based O–D flow estimation model depends not just on the accuracy of the link flow estimates (see the seventh and eighth iterations). As the flow conservation rule between O–D and link flows (Eqs. (13)–(15)) is the only key characteristic of the link-based O–D matrix estimation model, it exhibits limited capability to identify the possible O–D flow estimates in a large bidirectional

<table>
<thead>
<tr>
<th>Data Source</th>
<th>O-D Pair</th>
<th>1-9</th>
<th>1-10</th>
<th>2-9</th>
<th>2-10</th>
<th>Avg. MAPE for O-D Flow Estimates</th>
<th>Avg. MAPE for Link Flow Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>True</td>
<td>98</td>
<td>223</td>
<td>227</td>
<td>387</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 VDs (Rep)</td>
<td>ESTIMATED</td>
<td>322</td>
<td>0</td>
<td>14</td>
<td>610</td>
<td>277.6</td>
<td>100</td>
</tr>
<tr>
<td>MAPE (%)</td>
<td>94.1</td>
<td>57.6</td>
<td>119.8</td>
<td>24.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 VDs (Fix)</td>
<td>ESTIMATED</td>
<td>321</td>
<td>0</td>
<td>152</td>
<td>472</td>
<td>227.6</td>
<td>100</td>
</tr>
<tr>
<td>MAPE (%)</td>
<td>35.9</td>
<td>22.0</td>
<td>96.4</td>
<td>14.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 VDs</td>
<td>ESTIMATED</td>
<td>129</td>
<td>192</td>
<td>206</td>
<td>418</td>
<td>31.6</td>
<td>13.9</td>
</tr>
<tr>
<td>MAPE (%)</td>
<td>13.1</td>
<td>8.0</td>
<td>16.7</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 VDs &amp; 5 LPRs</td>
<td>ESTIMATED</td>
<td>98</td>
<td>215</td>
<td>229</td>
<td>395</td>
<td>3.6</td>
<td>3.4</td>
</tr>
<tr>
<td>MAPE (%)</td>
<td>0</td>
<td>2.1</td>
<td>2.3</td>
<td>13.5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 4. The Sanmin network.

Fig. 5. MAPEs of the link-based model.
Fig. 6. MAPEs of the path-based model.

network. This is illustrated by the MAPEs in most iterations of Fig. 5 being above 20%. Despite this deficiency, the link-based O–D estimation model terminates at the fifteenth iteration as the termination criterion is met. The average MAPE of the link flow estimates is 12.7%, and the average MAPE of the O–D matrix estimates is 18.5%.

The results for the path-based O–D matrix estimation model are shown in Fig. 6. The performance in terms of the estimation accuracy of both the link and O–D flows is satisfactory. The similarity in the trends of the two MAPE plots in Fig. 6 indicates that the accuracy of the O–D flow estimates depends on that of the link flow estimates as the spatial relationship between links and paths is pre-determined by the given link-path incidence matrix. The path-based model terminates at the eighth iteration; the average MAPE of the link flow estimates is 0.0% and that of the O–D flow estimates is 1.2%. The sensor deployment configuration corresponding to the link-based model is shown in Fig. 7. It indicates that 142 sensors (42 LPR sensors and 100 VDs) are required, and the sensor deployment rate is 70%. The sensor deployment configuration corresponding to the path-based model is shown in Fig. 8, and again 142 sensors (42 LPR sensors and 100 VDs) are required. However, the sensor location deployment pattern for the path-based model is different from that of the link-based model.

Fig. 7. Optimal sensor deployment configuration for the link-based model.

Based on the optimal sensor deployment configuration, the link-based and path-based O–D matrix estimation models are compared using partial and full VDs only deployments (142 and 202 VDs, respectively) and both VDs and LPRs. The partial VD only deployment scenario is assumed to follow the optimal sensor location deployment strategies under the link-based model (i.e., the fifteenth iteration in Fig. 5) and the path-based model (i.e., the eighth iteration in Fig. 6). The results are shown in Table II.

Table II illustrates that under the link-based O–D matrix estimation model, using both VDs and LPRs reduces the average MAPE for O–D flow estimates from 63.8% (partial VD only: replacement), 98.0% (partial VD only: fixed) or 74.6% (full VD only) to 18.5%. Similarly, under the path-based O–D matrix estimation model, using both VDs and LPRs reduces the average MAPE of O–D flow estimates from 31.8% (partial VD
only: replacement), 46.6% (partial VD only: fixed), or 31.8% (full VD only) to 1.2%. The underlying reasoning is similar to that discussed for the fishbone network case related to the valuable data that LPRs can additionally observe compared to VDs. Further, the path-based O–D matrix estimation model which explicitly captures path-related information by using the additional path flow information provided by the LPR sensors outperforms the link-based O–D matrix estimation model.

C. Discussion

The link-based O–D matrix estimation model expresses the relationship between O–D and link flows by using the flow conservation rules (Eqs. (13)–(15)). The link-based model has two issues. First, it cannot adequately identify possible O–D flows in a large bidirectional network. Second, the accuracy of the O–D flow estimates does not depend only on the accuracy of link flow estimates. The flow conservation rules cannot completely describe the spatial relationships between O–D and link flows. Two approaches can be used to address these issues. The first approach is to incorporate multiple traffic data sources that can provide information not only on link flows, but also on other aspects such as path trajectories/flows (note the reduced MAPEs for the link-based model in Table II when both VDs and LPRs are used). The second approach leverages the knowledge of a known link-path incidence matrix as part of a path-based O–D matrix estimation model, which can capture the causal relationships among the O–D, path and link flows (note the performance of the path-based models in Table II using VDs only or VDs and LPRs).

V. CONCLUDING COMMENTS

This study proposes a two-stage optimization model for the HSDP-OD problem. The heterogeneous sensors deployment model in the first stage incorporates three sources of observed traffic information (link flow, path trajectory and path coverage) into an integer program. The O–D matrix estimation model in the second stage is constructed as link-based and path-based NLS programs. An iterative solution procedure with a feedback mechanism is proposed to generate the optimal solutions for the first- and second-stage models. Compared to the past studies in this domain that make various assumptions as part of the modeling process, the only assumption made here is for the path-based O–D matrix estimation model, which assumes a given link-path incidence matrix in both stages of the proposed HSDP-OD model. As stated earlier, the incidence matrix can be obtained using a traffic assignment procedure.

Results from the numerical experiments suggest that the usage of LPR sensors in addition to the VDs can significantly enhance the accuracy of the O–D matrix estimation. That is, there is value to considering heterogeneous sensors that provide observations of traffic data beyond just link flows, such as path trajectories, and partial path or O–D flows, etc. Also, integrating the determination of the sensor deployment strategy with that of the accuracy of estimating the O–D matrix enables a more holistic perspective to addressing both problems by leveraging their interactions.

In this study, we solved the heterogeneous sensors deployment and network O–D matrix estimation problems in a two-stage iterative model. The feedback criteria between the first and second stages are based on the MAPEs of link and O–D flow estimates. Two potential future research directions are as follows. First, the feedback criteria used in the two-stage model can be based on other representations of the errors. Second, because the partially measured link or path flows by different sensors are associated with various degrees of measurement errors, and the assignment mappings between a set of unknown O–D flows and path/link flows are random variables determined by travelers’ route choice decisions, the HSDP-OD problem can be modeled as a state estimation problem under traffic uncertainties.

ACKNOWLEDGMENT

The authors are solely responsible for the contents of this paper.

REFERENCES


Shou-Ren Hu received the B.S. degree in transportation and communication management science from National Cheng Kung University, Tainan, Taiwan, in 1989; the M.S. degree in civil engineering from National Taiwan University, Taipei, Taiwan, in 1991; and the Ph.D. degree in civil engineering from Purdue University, West Lafayette, IN, USA, in 1996. He is currently an Associate Professor with the Department of Transportation and Communication Management Science, National Cheng Kung University. His research interests include intelligent transportation systems, transportation network modeling, highway safety, and risk analysis at railway level crossings. Dr. Hu is a Registered Traffic Engineer in Taiwan.

Srinivas Peeta received the B.S. degree in civil engineering from Indian Institute of Technology Madras, Chennai, India, in 1988; the M.S. degree in civil engineering from California Institute of Technology, Pasadena, CA, USA, in 1989; and the Ph.D. degree in civil engineering from The University of Texas at Austin, TX, USA, in 1994. He is currently a Professor of civil engineering at Purdue University, West Lafayette, IN, USA, and the Director of NEXTRANS Center, the USDOT Region 5 University Transportation Center, USA. His research interests include intelligent transportation systems, operations research, control theory, and computational intelligence techniques. Prof. Peeta serves on the Editorial Advisory Boards of the journals Transportation Research Part B, Intelligent Transportation Systems Journal, and Transportmetrica B. He serves as Area Editor of Transport Telematics for Networks and Spatial Economics. He has previously served as Chair of the Transportation Network Modeling Committee of the Transportation Research Board of the National Academies. He is also a member of the International Federation of Automatic Control Technical Committee on Transportation Systems.

Han-Tsung Liou received the B.S. degree in traffic and transportation engineering and management from Feng Chia University, Taichung, Taiwan, in 2006, and the M.M. degree in transportation and communication management science and the Ph.D. degree in transportation and communication management science from National Cheng Kung University, Tainan, Taiwan, in 2008 and 2015, respectively. He is currently a research associate at National Cheng Kung University. His research interests include intelligent transportation systems, travel demand analysis, and network optimization.