A Sensor Location Model to Optimize OD Estimation Using a Bayesian Statistical Procedure

Ning Wang\(^1\) and Pitu Mirchandani\(^2\)

\(^1\)School of Computing, Informatics and Decision Systems Engineering, Arizona State University, Tempe, AZ 85281
nwang14@asu.edu
Tel. 1-480-246-0871

\(^2\)School of Computing, Informatics and Decision Systems Engineering, Arizona State University, Tempe, AZ 85281
pitu@asu.edu
Tel. 1-520-965-2758 Fax 1-520-965-2751

Submitted for Presentation and Publication
92th TRB Annual Meetings, January 13-17, 2013
Submitted August 1, 2012

Word Count: 6000 + 4 Figure + 5 Tables = 8250 words

Abstract

For transportation planning and operations decisions, static origin-destination (OD) matrices, which specify the number of trips between each OD pair, is an essential input. One approach to estimate the OD matrices is to use data from traditional counting sensors on links with assumptions or models on how drivers choose routes on the network. The inverse problem is to locate a given number of counting sensors to obtain good OD matrix estimates. A new model for locating sensors to minimize the uncertainties in route flows estimates is presented in this paper. The model assumes a route choice set from each O to each D is known and some prior estimates of flows on the routes in this set are given. The reliabilities for each OD route prior and the link flow measurements are assumed to be given as well. Furthermore, the sensors in this problem scenario are not necessary perfect and measurement errors can occur; however the reliabilities of the sensors are assumed known. Extensive computational experiments and comparisons with some existing sensor location models indicate that the proposed model consistently gives more reliable estimates of OD flows.

1. INTRODUCTION AND PROBLEM ADDRESSED

Origin-destination (OD) matrices, sometimes referred to as OD demands, specify the number of trips made between each OD pair. From the traffic modeling viewpoint, volumes between OD pairs are generated from the node demands (upstream), and are “assigned” onto the network according to some behavioral rules, resulting in route volumes or link volumes (downstream). This is essentially the traditional four-step process to model and predict traffic loads and congestion. Because OD matrices are fundamental for various analyses for transportation
planning and operations (e.g., [1], [2]), the studies on the predicting power of producing “most likely” or “best” OD matrices has attracted significant attention in the past 30 years.

The traditional methods for obtaining an OD matrix utilize home interviews, roadside surveys or other behavioral models, such as the gravity-type trip distribution model. Because such procedures are expensive, they have seldom been researched or applied frequently. As the inverse of this traffic loading problem, OD matrices are often estimated from link flow volumes measured on some links using traffic sensors or detectors. Much research has been conducted on the relationship between the traffic counts from a subset of links and the estimated traffic demand that has generated this count data.

The quality of estimated OD matrices from link volumes depends on several factors, such as (1) the route-choice and traffic loading assumptions, (2) the quality of observed data from sensors, (3) the dependencies between link flows due to network topology and traffic loading, (4) the choice of OD estimation methods, and (5) on the link set where the sensors are located [3]. The first factor is paramount because that different traffic loading assumptions, such as equilibrium or near-equilibrium assumption, are involved in each estimation model, implicitly or explicitly. As the input information to estimate OD matrices, the quality of the observed data relates to the reliability of the counting devices, which can be improved through technologies and sampling and data collecting methods. The third factor comes from the complexity of the problem scenario. The last two factors reveal two important research questions:

- How to decide on the optimum subset of links where counting sensors are to be located;
- How to compute the best estimates of the flows of interest by using anticipated data from given candidate sensors location.

These two research questions are the two levels of the bilevel Sensor-Location Flow Estimation (SLFE) problem defined by Gentili and Mirchandani [4, 5]. The upper level is an optimization model that selects the best location set based on lower level solutions for each candidate set, while the lower level is an estimation model, also an optimization model, that selects estimates which minimize the expected estimation errors using the anticipated data from sensors. The decisions on both levels share the same objective of producing the best estimates of OD matrices which minimizes the expected estimation errors. This paper focuses on the upper level decision, which is to determine a link set on where to locate sensors to obtain high-quality OD estimations.

Recall that a traffic loading assumption is essential to the SLFE problem, in particular for the upper level sensor location problem addressed in this paper. Traffic loading or traffic assignment is a procedure to disaggregate OD trips into flows on various routes between each OD pair. Often, traffic equilibrium that assumes the observed link flows are from a flow pattern where Wardrop’s First Principle [6] holds (that is, all used routes between an origin and a destination have equal and minimum travel times, while other routes between that OD pair have no flow from that OD demand). As discussed by Wang et al. [7], a very large number of routes (some quite unusual) may result for an OD pair using the corresponding mathematical network equilibrium models for estimating OD trips. This paper will use the assumption in Wang et al. [7] that a route choice set is associated with each OD pair and includes all the routes that may be used for that OD pair. With such an assumption, the observed data need not be restricted to fit an equilibrated traffic pattern but, due to sensors errors, traveler perceptions and modeling
approximations, need only fit a traffic loading model with finite sets of OD routes. This is a more
general model since these route sets could be loaded so that equilibrium or near-equilibrium
traffic pattern results. Furthermore, by assuming knowledge of the route choice sets, the typical
OD estimation problem can be converted to a route flow estimation problem.

Measurement error is one type of uncertainty caused by the measurement devices in the
SLFE problem space. The magnitude of measurement errors can be represented by the variance
of the link flow measurements or in our prior knowledge of the sensor’s accuracy, which can be
estimated from the same type of sensors at similar locations or from other related studies.
Required by some OD estimation models, prior OD matrices include perception errors, which
can be described as the reliability for the prior OD volumes. In general, this reliability can be
based on one or more of these factors: (a) historical data (empirical uncertainty); (b) planner’s
subjective belief, the kind of uncertainties used in strategic decision making situations when
there is no empirical information, or (c) physical reasoning, (e.g., the probability we will get a
six with a toss of a fair die is 1/6). In the context of our estimation problem, we anticipate that
the uncertainties will correspond to the planners’ subjective reliabilities that are based mostly on
empirical evidences but tempered by anecdotal observations on the data they have before them.

The problem addressed in this paper can then be stated as follows: “Find a subset of links
in a network where counting sensors must be located, where OD routes, route flow priors and
their reliabilities, and the sensors’ reliabilities are given, and the objective is to minimize the
uncertainties in the a posteriori knowledge of route flows”. A new model is developed to solve
this location problem. The paper is organized as follows. Section 2 reviews some related
literature to the SLFE problem. Section 3 presents the optimization model and algorithm.
Experimental results and comparisons with other existing models are presented in Section 4.
Some concluding remarks are made in Section 5.

2. LITERATURE REVIEW

Gentili and Mirchandani [4, 5] defined the Sensor-Location Flow Estimation (SLFE) problem as
a bilevel problem. The upper level is a location model to select the subset of links where sensors
are located; and the lower level is an estimation model to compute the best flows of interest by
using data from given candidate sensors location from upper level. Both papers identified the
connection between the sensor location and flow estimation problems.

Methods for static OD estimation from link counts can be grouped into two categories.
The first one is based on traffic modeling concepts, such as minimum information (entropy
maximizing) model and combined models for traffic planning (e.g. Van Zuylen and Willumsen
[8], Fisk [9]). The second category is based on statistical inference that the volumes and priors
are assumed to be generated by some probability distributions (or distribution free version with
parameters representing the mean and variance). An estimate of OD matrix is obtained by
estimating the parameters of the probability distributions. Statistical inference approaches
include Maximum Likelihood (Spiess [10]), generalized least squares (Cascetta [11], Bell [12]
and Yang et al. [13]) and Bayesian inference (Maher [14]). Eisenman et al. [15] and Zhou and
List [16] proposed a Kalman filtering-based framework to address OD estimation problems.
Larsson et al. [3] provided a comparative study of some existing sensor location models and classified these models based on the objective of the corresponding optimization problems to locate the sensors, which were categorized by the maximization of:

1. Link flow coverage (LFC): locate sensors so that link flow volume covered (or intercepted) is maximum (Lam and Lo [17]);
2. OD-pair coverage (ODC): locate sensors to maximize OD pairs covered (where an OD pair is “covered” if at least one route in its choice set is observed) (Yang et al. [18], Ehlert et al. [19]);
3. Route cardinality coverage (RCC): locate sensors to maximize the number of routes covered (or intercepted) (Gentili and Mirchandani [20], Castillo et al. [21]);
4. OD-demand coverage (ODDC): locate sensors to maximize total OD demand covered (Hodgson [22], Yim and Lam [23]);
5. Route flow coverage (RFC): locate sensors to maximize the total route flow intercepted (Ehlert et al. [19], Yang and Zhou [24]);

These five models rely on different input information. For example, link flow coverage model uses link flow priors, which is the summation of the average flows of routes passing such link; OD-pair coverage and route coverage models depend on the network structure that specifies which links are used by each OD pair and which links are used by each OD route, respectively; OD-demand coverage and route flow coverage need both network structure specification as above, as well as the average of prior flow volumes for all OD pairs and OD routes, respectively.

In recent years, the uncertainties in the prior and posterior flow of interests were started to be addressed in traffic sensor location problems. Fei et al. [25] applied Kalman-filtering-based framework to minimize the error of the estimated OD demand matrix in traffic in a time varying traffic demand context, and examined the sensor location problem with and without budget restrictions. Fei and Mahmassani [26] proposed a bi-objective model to address sensor location problem from the perspective of both demand coverage and demand uncertainty reduction in the context of dynamic traffic assignment. Wang et al. [7] proposed a sensor location model to maximize the variance reductions in posterior flows of interest for a Generalized Least Squares estimator. Using such a variance reduction objective for minimizing the uncertainties in estimation was first proposed by Mirchandani et al. [27] for estimating travel times by locating vehicle identification readers and was also used by Zhou and List [16] for estimating OD matrix.

Along the same lines, this paper addresses the problem to select sensors locations with the consideration of observation noise and uncertainties in priors. The objective of the problem is to minimize estimation error using Bayesian Inference technique. The major assumption about traffic loading is that OD volumes are assigned on known finite route choice sets. Equilibrium or near-equilibrium assignment may be used as apriori loading. We emphasize that the OD matrices can be uniquely determined from obtained route flow estimates in polynomial time and therefore the OD estimation problem can be considered as simply a route flow estimation problem. It should be noted that Zhou and List [16] also addressed this problem but they used Kalman Filtering methods and their prior uncertainties were focused on finite discrete scenarios for prior OD demands and the resulting equilibrium flows, each scenario having a given probability of occurring,

3. MODEL DEVELOPMENT
3.1 Notation

We will first provide the notation only so that readers familiar with this problem can quickly get to the modeling constructs. Readers unfamiliar with this topic can come back to this notation section to follow the modeling developments. Since there is significant notation, we classify notation as it relates to (a) network topology, (b) definition of route flows, (c) links flows, (d) observations or measurements, and (e) decision variables.

**Network Topology Parameters:**

- $R$: Route-choice set for a network
- $A$: Set of all the links in the network
- $A'$: Set of links where sensors are located
- $A_f$: Set of links feasible for sensors to be located
- $|R|$: Number of routes in a network
- $|A|$: Number of links in a network
- $N$: Number of sensors to locate
- $\rho_a^i \in (0,1)$: Link-route parameter. $\rho_a^i = 1$ if route $i$ uses link $a$, otherwise, $\rho_a^i = 0$

\[
H = \begin{bmatrix}
\rho_1^1 & \cdots & \rho_{|R|}^1 \\
\vdots & \ddots & \vdots \\
\rho_{|A|}^1 & \cdots & \rho_{|A|}^{|R|}
\end{bmatrix}
\]

**Route Flows:**

- $x^i$: Route flow parameter (real mean) of $i^{th}$ route
- $\mathbf{x} = (x^1, x^2, \ldots, x^{|R|})'$
- $\mu_0^i$: Mean of prior distribution of $x^i$
- $\mu_0 = (\mu_1^{(0)}, \mu_2^{(0)}, \ldots, \mu_{|R|}^{(0)})'$
- $\eta_{ij}^{(0)}$: Covariance of prior distribution between $x^i$ and $x^j$ (variance of prior distribution if $i = j$)

\[
V_0 = \begin{bmatrix}
\eta_{11}^{(0)} & \cdots & \eta_{1|R|}^{(0)} \\
\vdots & \ddots & \vdots \\
\eta_{|R|1}^{(0)} & \cdots & \eta_{|R||R|}^{(0)}
\end{bmatrix}
\]

- $u^i$: Prior’s reliability of route $i$
$u^{max}$: Bound of prior’s reliability

$\xi_i$: Real error in route flow prior of route $i$

$\mu_i$: Mean of posterior distribution of $x^i$

$\mu_1 = (\mu^{(1)}_1, \mu^{(1)}_2, ..., \mu^{(1)}_R)^\prime$

$\eta_{ij}^{(1)}$: Covariance of posterior distribution between $x^i$ and $x^j$ (variance of posterior distribution if $i=j$)

$V_1 = \begin{bmatrix} \eta_{11}^{(1)} & \cdots & \eta_{1R}^{(1)} \\ \vdots & \ddots & \vdots \\ \eta_{R1}^{(1)} & \cdots & \eta_{RR}^{(1)} \end{bmatrix}$

**Link Flows:**

$v_a$: Real link flow on link $a$

$\mathbf{v} = (v_1, v_2, ... v_{|A|})^\prime$

$\bar{v}_a$: Potential link flow observation if a sensor is located on link $a$

$\mathbf{\bar{v}} = (\bar{v}_1, \bar{v}_2, ... \bar{v}_{|A|})^\prime$

$\epsilon_a$: Measurement error on link $a$

$\tau_a^2$: Variance of link flow measurement on link $a$

$\Sigma = \begin{bmatrix} \tau_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \tau_{|A|}^2 \end{bmatrix}$

$t_a$: Sensor’s reliability when located on link $a$

$t^{max}$: Bound of sensors’ reliability

$\bar{v}_a$: Mean of Link flow prior means on link $a$

**Observations by sensors:**

$y_n$: Link flow observed by sensor $n$

$\mathbf{Y} = (y_1, y_2, ..., y_N)^\prime$

$g_n$: An instance of link flow observed by sensor $n$

$\mathbf{g} = (g_1, g_2, ..., g_N)^\prime$

$\varphi_n^2$: Variance of observation by sensor $n$
**Decision Variables for Location Models:**

- \( y_a \in (0,1): y_a = 1 \) if locating sensor on link \( a \), otherwise \( y_a = 0 \) (used for existing location models)

- \( z_{na} \in (0,1): z_{na} = 1 \) if locating \( n^{th} \) sensor on link \( a \), otherwise \( z_{na} = 0 \).

\[
\mathbf{z} = \begin{bmatrix}
    z_{11} & \cdots & z_{1|A|} \\
    \vdots & \ddots & \vdots \\
    z_{N1} & \cdots & z_{N|A|}
\end{bmatrix}
\]

\[
\mathbf{z}_n = [z_{n1}, ..., z_{n|A|}]
\]

Among these parameters, \( \rho_a^l \) are obtained from definitions of the given route choice sets. \( \mathbf{x} \) is a column vector for the means or real route flow with elements \( x^i \) for \( i = 1, \ldots, |R| \). \( \mathbf{x} \) has a prior distribution and we will discuss later. \( \hat{\mathbf{v}}_a \) represents the flow observed on link \( a \) in the link set \( A \), and it is the elements of \( \hat{\mathbf{v}} \), which is a \(|A|\)-dimension column vector. The relationship between \( \hat{\mathbf{v}}_a \) and \( \mathbf{x} \) is

\[
\hat{\mathbf{v}}_a = \sum_{k \in R} \rho_a^k x^k + \varepsilon_a \quad \forall a \in A \tag{1}
\]

or

\[
\hat{\mathbf{v}} = \mathbf{Hx} + \mathbf{\varepsilon} \tag{2}
\]

where \( \varepsilon_a \) is the random measurement error from sensor if one is located on link \( a \). Equation (1) indicates that the observation of link flow is equal to the sum of all route flows that use the link, plus a random measurement error from the sensor. This relationship always exists for entire link set \( A \), however, we can measure only those in set \( A' \) where sensors are located. If link \( a \) is not covered by a sensor, equation (1) will not be active in the analysis (for estimating route flows).

The decision variables in this problem are binary and defined as \( z_{na} \) (\( n = 1, \ldots, N \), \( a = 1, \ldots, |A| \)), where \( n \) is the subscript for sensors and \( a \) is the subscript for links. \( N \) is the number of sensors’ (a given limit) and \( |A| \) is the total number of links. \( z_{na} = 1 \) means the \( n^{th} \) sensor is located on link \( a \). Note that each sensor can only be located on only one link; and for each specific link \( a \), we only allow at most one sensor. These restrictions can be expressed as

\[
\sum_{a=1}^{|A|} z_{na} = 1, \quad n = 1, \ldots, N \tag{3}
\]

\[
\sum_{n=1}^N z_{na} \leq 1, \quad \forall a \in A \tag{4}
\]

We can define a matrix \( \mathbf{z} \) for the decision variables, using \( z_{na} \) as elements. The dimension of \( \mathbf{z} \) is \( N \) by \(|A|\), and the number of binary variables for each problem is \( N^*|A| \).

Since \( \hat{\mathbf{v}} \) is a column vector with dimension \(|A|\), the rows in (2) that can be accessed by the estimation model include only the rows (i.e., links) where a sensor is located, in particular when \( z_{na} = 1 \). By the definition of \( \mathbf{z} \), the matrix multiplication \( \mathbf{z}\hat{\mathbf{v}} \) selects the links with sensors.

Take an example of a four-link network with link flow volume \( \hat{\mathbf{v}} = (1,20,30,4) \), \( \mathbf{z} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \) indicates that the first sensor is located on link 2 (in row 1, “1” appears in...
column 2 of \( z \) and the second sensor is located on link 4 (in row 2, “1” appears in column 4 of \( z \)). Hence \( \hat{z} = (20,4)' \) simply selects the volumes for link 2 and 4.

Therefore, we can modify equation (2) by equation (5)

\[
Y = \hat{z} = z(H \mathbf{x} + \varepsilon) = zH \mathbf{x} + z\varepsilon \tag{5}
\]

which eliminates the links without sensors and captures actual observations collected by candidate sensors \( z \).

Recall that route flow \( \mathbf{x} \) has a prior information. In this paper, we assume the prior distributions are multivariate normal (MVN) which provides accurate approximation for the true distribution of traffic volumes (Maher [14]). Define prior distribution \( \mathbf{x} \sim \text{MVN}(\mu_0, \mathbf{V}_0) \), where \( \mu_0 \) and \( \mathbf{V}_0 \) are mean and variance of prior distribution. We assume the random measurement error follows a normal distribution with mean zero and variance \( \tau_a^2 \) and observations on different links are independent, that is \( \varepsilon \sim \text{MVN}(0, \Sigma) \), where \( \Sigma \) is a diagonal “dispersion” matrix for link flow observations. Note that the parameter \( \Sigma \) can be estimated from the same type of sensor at similar locations or related studies in other areas. Because of equation (2), \( \hat{\mathbf{v}} | \mathbf{x} \sim \text{MVN}(H \mathbf{x}, \Sigma) \) is the conditional probability of link flows \( \hat{\mathbf{v}} \) given route flow \( \mathbf{x} \). This conditional probability can be used for the likelihood function of the potential link flow observations \( \hat{\mathbf{v}} \). Again, because of the nature of SLFE problem, which is a sequential decision-making procedure for first locating sensors and then estimating flow of interest, the conditional probability that is used for computing the posterior distribution, is only available for the links with sensors at the estimation stage. Hence, given sensors’ location, the accessible measurement errors considering the sensors’ location is \( z\varepsilon \sim \text{MVN}(0, z\Sigma z') \) (Mardia [2826]), and from equation (5) link flow observations \( Y \) can be expressed as \( Y \sim \text{MVN}(zH \mathbf{x}, z\Sigma z') \).

3.2 Route Flow Estimation Model with Link Flow Observations

As the lower level of SLFE problem, the procedure of estimating flow variables from observations will be conducted once the sensors are located, which is analogous to obtaining a posteriori knowledge based on observation evidence and prior knowledge. If the flows of interest are route flow volumes, the corresponding problem is to reproduce the estimated link flows that are close to the observed link flows from the counting sensors

Originated from the Bayesian statistical approach by Maher [14], the Bayesian estimator we will use in this paper is developed as follows:

With given locations \( z \), if \( \mathbf{x} \sim \text{MVN}(\mu_0, \mathbf{V}_0) \) and \( Y = zH \mathbf{x} + z\varepsilon \), where \( \varepsilon \sim \text{MVN}(0, \Sigma) \), so that \( Y | \mathbf{x} \sim \text{MVN}(zH \mathbf{x}, z\Sigma z') \), the posterior distribution of \( \mathbf{x} \) given observations \( Y = \mathbf{g} \) is also MVN.

The probability densities of \( \mathbf{x} \) and \( Y | \mathbf{x} \) are:
By Bayesian theorem, the posterior as:

\[ f(x|Y = g) \propto p(x)f(Y = g|x) \sim MVN(\mu_1, V_1) \]

Therefore, \( f(x|Y = g) \)

\[
\propto \exp \left[ -\frac{1}{2} (g - zHx)'(z\Sigma z')^{-1}(g - zHx) \right] \exp \left[ -\frac{1}{2} (x - \mu_0)'V_0^{-1}(x - \mu_0) \right]
\]

\[
= \exp \left[ -\frac{1}{2} (g'(z\Sigma z')^{-1}g - x'H'z'(z\Sigma z')^{-1}x) + x'V_0^{-1}x - \mu_0'V_0^{-1}x - x'V_0^{-1}\mu_0 + \mu_0'V_0^{-1}\mu_0 \right]
\]

\[
= \exp \left[ -\frac{1}{2} (x - \mu_1)'V_1^{-1}(x - \mu_1) \right]
\]

Equating the corresponding terms gives that

\[ V_1^{-1} = V_0^{-1} + H'z'(z\Sigma z')^{-1}zH \] (8)

\[ V_1^{-1}\mu_1 = V_0^{-1}\mu_0 + H'z'(z\Sigma z')^{-1}g \] (9)

Further matrix operation from (9) and (10) gives that

\[ \mu_1 = \mu_0 + V_0H'z'(z\Sigma z' + zHV_0H'z')^{-1}(g - zH\mu_0) \] (10)

\[ V_1 = V_0 - V_0H'z'(z\Sigma z' + zHV_0H'z')^{-1}zHV_0 \] (11)

Therefore, for the prior distribution of route flow \( x \sim MVN(\mu_0, V_0) \) and the posterior distribution \( x|Y = g \sim MVN(\mu_1, V_1) \), equations (10) and (11) update for the mean vector and dispersion matrix for any given sensors' location \( z \).

The expected change in posterior variance is

\[ E[V_1 - V_0] = -V_0H'z'(z\Sigma z' + zHV_0H'z')^{-1}zHV_0 \] (12)

From (12), the posterior dispersion matrix varies according to the different sensors' location matrix \( z \). The trace of (12) is the change in total variance using above Bayesian' approach to update route flows information.

### 3.3 Route Flow Estimation Model with Independent Link Flow Observations

When using equations (10) and (11) as the updating equations for mean and dispersion matrix of route flow, the calculations for the inverse matrix are complicated. Maher indicated that because
of the sequential nature of Bayes’ Theorem, one can iteratively conduct the Bayesian procedure using independent individual observations, such that the posterior from one observation becomes the prior for the next. The sequence of selecting link from which to select measurements does not have effect the posterior distribution in the final step.

Suppose that a single observation of the $n$th sensor is defined by $\mathbf{z}_n$, which is a $|A|$-dimension row vector—the $n^{th}$ row from $\mathbf{z}$ in particular. If $z_{na} = 1$, the $n^{th}$ sensor is located on link $a$ and then the dispersion matrix $\mathbf{\Sigma}$ contains a single element $\mathbf{z}_n\mathbf{\Sigma}\mathbf{z}_n'$ which can be defined as variance brought in by sensor $n$, simplified as $\varphi_n^2 = \sum_{a=1}^{|A|} z_{na}^2 \tau_a^2$, because for each $n$, there is one and only one $z_{na}$ to be non zero. Matrix $\mathbf{H}$ becomes a row vector with only $a^{th}$ row from original $\mathbf{H}$ left. The reduced matrix $\mathbf{H}$ can be described by

$$\mathbf{z}_n\mathbf{H} = \left[ \sum_{a=1}^{|A|} z_{na} \rho_a^1, \sum_{a=1}^{|A|} z_{na} \rho_a^2, \ldots, \sum_{a=1}^{|A|} z_{na} \rho_a^{|R|} \right] = [h_1^n, h_2^n, \ldots, h_1^n].$$

Because the inverse of a scalar $\mathbf{z}_n\mathbf{\Sigma}\mathbf{z}_n' + \mathbf{z}_n\mathbf{H}\mathbf{V}_0\mathbf{H}'\mathbf{z}_n'$ is the reciprocal, the updating equations become:

$$\mathbf{\mu}_1 = \mathbf{\mu}_0 + \frac{\mathbf{v}_0\mathbf{H}'\mathbf{z}_n'}{\mathbf{z}_n\mathbf{\Sigma}\mathbf{z}_n' + \mathbf{z}_n\mathbf{H}\mathbf{V}_0\mathbf{H}'\mathbf{z}_n'} (g_n - \mathbf{z}_n\mathbf{H}\mathbf{\mu}_0) \quad (13)$$

$$\mathbf{V}_1 = \mathbf{V}_0 - \frac{\mathbf{v}_0\mathbf{H}'\mathbf{z}_n'\mathbf{z}_n\mathbf{H}\mathbf{V}_0'}{\mathbf{z}_n\mathbf{\Sigma}\mathbf{z}_n' + \mathbf{z}_n\mathbf{H}\mathbf{V}_0\mathbf{H}'\mathbf{z}_n'} \quad (14)$$

The elements $\mu_1^{(i)}$ and $\eta_1^{(i)} (i=1,\ldots,|R|; j=1,\ldots,|R|)$ of $\mathbf{\mu}_1$ and $\mathbf{V}_1$ are given by

$$\mu_1^{(i)} = \mu_0^{(i)} + \frac{\sum_{j=1}^{|R|} h_{nj}^i \eta_0^{(j)}}{\sum_{j=1}^{|R|} h_{nj}^i \left( \sum_{j=1}^{|R|} h_{nj}^i \eta_0^{(j)} \right)} (g_n - \sum_{r=1}^{|R|} h_{nr}^i \mu_0^{(r)}) \quad (15)$$

$$\eta_1^{(i)j} = \eta_0^{(i)j} - \left( \frac{\sum_{j=1}^{|R|} h_{nj}^i \eta_0^{(j)}}{\sum_{j=1}^{|R|} h_{nj}^i \left( \sum_{j=1}^{|R|} h_{nj}^i \eta_0^{(j)} \right)} \right) \frac{\left( \sum_{j=1}^{|R|} h_{nj}^i \eta_0^{(j)} \right)}{\sum_{j=1}^{|R|} h_{nj}^i \left( \sum_{j=1}^{|R|} h_{nj}^i \eta_0^{(j)} \right)} \quad (16)$$

Denoting $\sum_{j=1}^{|R|} h_{nj}^i \eta_0^{(j)}$ by $S_n^i$ ($i = 1, \ldots, |R|, n = 1, \ldots, N$) and $\sum_{j=1}^{|R|} h_{nj}^i S_n^j$ by $T_n$. $\sum_{r=1}^{|R|} h_{nr}^i \mu_0^{(r)}$ is how we calculated link flow prior mean from prior means of route flows passing that link and we denote it as $\bar{\nu}_n$. Then equations (13) and (14) become

$$\mu_1^{(i)j} = \mu_0^{(i)j} + \frac{s_n^i}{\varphi_n^2 + \tau_n^2} (g_n - \bar{\nu}_n) \quad (17)$$

$$\eta_1^{(i)j} = \eta_0^{(i)j} - \frac{s_n^i \eta_0^{(j)}}{\varphi_n^2 + \tau_n^2} \quad (18)$$

$s_n^i$ describes the variation on flow of route $i$ influenced by the $n$th sensor in the current iteration. $T_n$ is the variation on all routes influenced by sensor $n$’s observation. $\varphi_n^2$ is the variation of measurement sensor $n$ itself. The ratio $\frac{s_n^i}{\varphi_n^2 + \tau_n^2}$ in (17) is actually the weight of link flow partitioned into specific route $i$ passing it. If sensor’s location is decided ($n$ is fixed), $\varphi_n^2 + \tau_n$ will be fixed. The route $i$ with large variation $S_n^i$ will be given large weight for sharing the information from observation $g_n$. The rationale behind is that if the prior route information is
not reliable (i.e., has a large variance), it requires more information from observations in order to
produce a reliable posterior and vice versa. If the observation is not reliable \( (\varphi_n^2 \text{ is large}) \), the
corresponding ratio will become lower and the observation will be considered with less weight in
computation of the distributions of the corresponding routes.

3.4 Locating Sensors to Maximize Variance Reduction

Because the trace of equation (12) measures the total variance reduction of posterior from prior,
using Bayesian procedure the strategy for sensor location will be to locate sensors on a subset of
links which has largest reduction in the variance of the posterior distributions. The objective
function of the location model is to maximize the trace of equation (12) by selecting different
decision variable \( z \). This new sensor location model to maximize variance reduction using
Bayesian estimator can be formulated as

\[
[MVR] \quad \max \; \text{tr}(V_0 H' z' (z \Sigma z' + z H V_0 H' z')^{-1} z H V_0) \tag{19}
\]

\[
\text{s.t. } \sum_{a=1}^{[A]} z_{na} = 1, \quad n = 1, ..., N, \sum_{a=1}^{[A]} z_{na} = 1, \quad n = 1, ..., N \tag{3}
\]

\[
\sum_{a=1}^{N} z_{na} \leq 1, \quad \forall a \in A, \sum_{a=1}^{N} z_{na} \leq 1, \quad \forall a \in A \tag{4}
\]

\[
z_{na} \in \{0, 1\}, \quad n = 1, ..., N, a \in A \tag{20}
\]

We can simplify the formulation for independent observations. Suppose \( N \) sensors will be
located in the network. Instead of solving [MVR] directly, we can use an iterative method that
locate one link with capability to reduce most posterior variance at a time. After selecting one
link, the route flow mean and variance are updated using (13) and (14). The posterior becomes
the prior and process repeats for another unselected link. We define \( A_f \) as the set of links which
have not been selected yet and \( \hat{A}_f \) is the set of selected links (and cannot be selected again). The
process is repeated until all \( N \) links are determined. For each sensor \( n \), its location is decided by
the model, formulated as:

\[
[MVRS] \quad \max \sum_{i=1}^{[R]} \frac{(s^i_n)^2}{\varphi_n^2 + T_n} \tag{21}
\]

\[
\text{s.t. } h^i_n = \sum_{a=1}^{[A]} z_{na} \rho^i_a, \quad \forall i = 1, ..., [R] \tag{22}
\]

\[
\varphi_n^2 = \sum_{a=1}^{[A]} z_{na} \tau_a^2 \tag{23}
\]

\[
S^i_n = \sum_{j=1}^{[R]} h^j_n \eta^j_{ij}(0), \quad \forall i = 1, ..., [R] \tag{24}
\]

\[
T_n = \sum_{i=1}^{[R]} h^i_n S^i_n \tag{25}
\]

\[
\sum_{a=1}^{[A]} z_{na} = 1 \tag{26}
\]

\[
z_{na} = 0, \quad \forall a \in \hat{A}_f \tag{27}
\]
The objective function (21) in formulation MVRS maximizes total variance reduction \( \sum_{i=1}^{\mid R \mid} (\eta_{ii}^{(0)} - \eta_{ii}^{(1)}) \). Constraints (22)-(25) are introducing notations to simplify objective function (21). Constraint (26) restricts sensor \( n \) will be located on only one link. Constraint (27) forces the decision of sensor’s location be made only from the feasible links’ set \( A_f \). Constraint (20) is the binary constraint for decision variables.

3.5 Algorithm

Locating only one sensor in the network for problem MVRS can be solved by a greedy algorithm: for each candidate link, calculate the objective value \( \sum_{i=1}^{\mid R \mid} \frac{(s_{i,n})^2}{\sigma_{n,i}^2 + r_n} \) and select the one with largest objective value. The solution of MVRS can be found by \( O(|A|) \) time.

Therefore, the entire algorithm for solving [MVR] is

BAYESIAN LOCATION ALGORITHM

Input: \( \mu_0, V_0, H, N \)
Output: \( z = (z_1, z_2, ..., z_N)' \), \( \mu_1, V_1 \)

(1) Set \( n := 1 \), where \( n \) is the sensors’ index \( (1, 2, ..., N) \).
    Set \( A_f = A \).
    Set \( A' = \emptyset \).

(2) For each link \( a \in A_f \), calculate \( \sum_{i=1}^{\mid R \mid} \frac{(s_{i,n})^2}{\sigma_{n,i}^2 + r_n} \).
(3) Select the link \( a \in A_f \) with largest value calculated in step 2 and \( z_n \) is a row vector with only non-zero element in \( a \)th column. Remove link \( a \) from \( A_f \) and add link \( a \) into \( A' \).
(4) Using observation on link \( a \),
    update posterior using equations (14) and (15) for \( \mu_1, V_1 \).
(5) Increment \( n \) by 1
(6) If \( n = N + 1 \) then stop.
    Else set \( \mu_0 = \mu_1, V_0 = V_1 \) and go to Step 2.

The running time for this algorithm is polynomial because the entire algorithm will run for \( N \) loops and within each loop consisting of preprocessing (step 2) and search (step 3) can be conducted in polynomial time.
The optimality of above algorithm can be proved by the sequential nature of Bayes’ Theorem. If we have an observation from single sensor \( z_i \), the posterior of route flow \( X \) can be derived as:

\[
P(X|z_1) \propto P(X)l(X|z_1)
\]

where \( P(X) \) is the route flow prior, and \( l(z_1|X) \) is the likelihood for \( X \) on observation \( z_1 \). Next, the posterior for \( X \) can be updated using a second observation from sensor \( z_2 \).

\[
P(X|z_1, z_2) \propto P(X)l(X|z_1, z_2) \propto P(X)l(X|z_1)l(X|z_2) \propto P(X|z_1)l(X|z_2)
\]

Above relations hold when \( z_1 \) and \( z_2 \) are independent. It is now confirmed that updating directing from \( P(X) \) to \( P(X|z_1, z_2) \) by a single application of Bayes’ theorem with observation \( z_1 \) and \( z_2 \) is equivalent to a two-stage process that sequentially updates posterior by individual sensor. Therefore, the posterior with all \( N \) observations (solution for model MVR) can be obtained by a series of application of Bayesian approach with individual sensors (solution for model MVRS).

4. EXPERIMENTAL RESULTS

4.1 Experiment Setup

We applied the proposed MVR model and algorithm to two problem scenarios and the performances were compared with three existing models: Flow Coverage (LFC), Route Flow Coverage (RFC) from Larsson [3] and LOC model from Wang et al. [7]. The reason to select LFC and RFC in comparison is that both models are widely used in practice and rely on network structures and the magnitude of prior information; LOC model relies on the variance or degree of believe of prior information (no measurement errors considered). All the experiments were conducted in Matlab.

Figure 1 shows the scheme of the experimental procedure. First (Steps 1-2), problem scenarios are generated by defining the network (supply), loading the assumed traffic flows (demand) and obtaining the “actual” mean of network traffic \( x^i \) and \( v^i \). Assumed route flow prior distribution (mean and variance), sensors’ reliabilities and potential link flow observations are generated by linking the reliabilities to the actual values (Step 3). For each route, the prior’s reliability was first randomly generated by \( u_i \in (0, u_{\text{max}}) \); then the prior’s variance was calculated by \( \eta_{ii}^{(0)} = x^iu_i^2 \); the prior’s mean was generated from a normal random generator \( \mu_{ii}^{(0)} \sim N(x^i, \eta_{ii}^{(0)}) \); all covariance were assumed to be zero in the initial priors. Similar logic was used to generate sensors’ reliabilities and potential observations. For each link, the sensor’s reliability \( t^a \) was randomly generated within range \([0, t_{\text{max}}]\) and variance of link flow measurement was calculated by \( \tau_{ii}^2 = v^at^a \); the potential observation on such link was generated from a normal random generator \( \hat{v}_a \sim N(v^a, \tau_{aa}^2) \). In Step 4, MVR and other location models are solved using Matlab to obtain location decisions \( Z \). Then, with these decisions, posterior mean and variance route flows are obtained using the Bayesian estimation model in section 3.2 or 3.3 (Step 5). To evaluate each location model (Step 6), the posterior means are compared with actual route flows in terms of the bias. Total variance (trace of \( V_k \)) are compared as well. The
evaluation criterion is the sum-squared error (SSE) of the posterior distribution, which is defined
by the summation of variance and squared bias:

\[ SSE = BIAS^2 + Variance = \sum_{i \in R} \left( \mu_i^{(1)} - x^i \right)^2 + \sum_{i \in R} \eta_i^{(1)} \]  

where \( x^i \) is the true mean value of the route flow \( i \). Because SSE measures the combination of
bias and variance, if the magnitude of variance is significantly lower than the squared bias, the
value of SSE will be dominated by the effect of bias. Because our proposed MVR already
ensures the maximum variance reduction, which is the minimum of variance of posterior, the
criterion SSE will be sensitive to show the effect of bias provided by the comparative locational
models.

Finally, note that for each network scenario, we can have different problem instances by
locating different number of sensors between 1 and (say) \( N \), the number of possible sensor sites.
Furthermore, we can have priors with different qualities, in particular prior variance and bias. In
addition, sensors can have various reliabilities. Hence, each problem instance in our
computational experiments is defined by (1) network supply, (2) traffic demand, (3) reliability of
priors on traffic flows to be estimated, (4) reliability of sensors, and (5) number of sensors to be
located. Solving the location model for each problem instance, for the four models being
compared, gives us a score for each model, as defined by the SSE.

Two sets of network scenarios were used in the computational experiments.

1. A grid network with 16 OD pairs, 43 routes and 48 links (grid network 1, left side of
Figure 2).
2. A grid network with 16 OD pairs, 204 routes and 112 links (grid network 2, right
side of Figure 2).


Figure 1: Procedure for the computational experiments
4.2 Numerical Results

Four network scenarios were tested with two initial prior distributions. We simulated two sets of link flow observations for each grid network. When generating the experiments, $u^\text{max}$ is set to be 0.15 and $t^\text{max}$ is 0.1. This parameter setting reflects that we assume counting devices are more accurate than our prior knowledge in practice.

Table 1 shows the SSE results for the four comparative location models. The table gives the averaged SSE for every five instances locating five numbers of sensor locations. For example, the first row in Table 1 represents the average SSE scores for $N=1$ to $N=5$. The best scores are in bold in Table 1. In general, our proposed MVR outperforms other models in most cases, especially when few sensors are in use. The row marked as “Total” in Table 1 sums all of the SSE scores and emphasizes on the advantage of MVR compared to other models for grid network 1. Observe, however, that LOC model sometimes has the best score, and often is the second best when MVR is the best.
Table 1: SSE Results for Grid Network 1

<table>
<thead>
<tr>
<th></th>
<th>obs_1</th>
<th></th>
<th>obs_2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MVR</td>
<td>LOC</td>
<td>LFC</td>
<td>RFC</td>
</tr>
<tr>
<td>prior_1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-5</td>
<td>481.04</td>
<td>486.86</td>
<td>821.60</td>
<td>832.85</td>
</tr>
<tr>
<td>6-10</td>
<td>455.34</td>
<td>462.25</td>
<td>808.11</td>
<td>566.57</td>
</tr>
<tr>
<td>11-15</td>
<td>441.49</td>
<td>441.76</td>
<td>474.10</td>
<td>502.76</td>
</tr>
<tr>
<td>16-20</td>
<td>437.53</td>
<td>437.47</td>
<td>472.21</td>
<td>491.71</td>
</tr>
<tr>
<td>21-25</td>
<td>439.34</td>
<td>440.04</td>
<td>469.49</td>
<td>478.66</td>
</tr>
<tr>
<td>26-30</td>
<td>440.92</td>
<td>441.22</td>
<td>452.77</td>
<td>455.85</td>
</tr>
<tr>
<td>31-34</td>
<td>441.26</td>
<td>441.88</td>
<td>445.33</td>
<td>447.31</td>
</tr>
<tr>
<td>Total</td>
<td>3136.91</td>
<td>3151.49</td>
<td>3943.63</td>
<td>3775.71</td>
</tr>
<tr>
<td>prior_2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-5</td>
<td>114.36</td>
<td>156.33</td>
<td>140.54</td>
<td>122.59</td>
</tr>
<tr>
<td>6-10</td>
<td>136.22</td>
<td>150.23</td>
<td>162.73</td>
<td>147.97</td>
</tr>
<tr>
<td>11-15</td>
<td>125.14</td>
<td>150.63</td>
<td>139.41</td>
<td>151.94</td>
</tr>
<tr>
<td>16-20</td>
<td>129.06</td>
<td>145.52</td>
<td>133.88</td>
<td>152.74</td>
</tr>
<tr>
<td>21-25</td>
<td>140.50</td>
<td>140.12</td>
<td>137.63</td>
<td>155.63</td>
</tr>
<tr>
<td>26-30</td>
<td>137.87</td>
<td>135.17</td>
<td>136.93</td>
<td>146.41</td>
</tr>
<tr>
<td>31-34</td>
<td>134.66</td>
<td>134.12</td>
<td>135.57</td>
<td>143.75</td>
</tr>
<tr>
<td>Total</td>
<td>917.80</td>
<td>1012.13</td>
<td>986.69</td>
<td>1021.03</td>
</tr>
</tbody>
</table>

We have similar comparisons in Table 2 among the SSE scores for the four location models applied to the grid network 2 where there are many more routes for each OD pair. The MVR model performs the best in most cases and on the overall performance in the rows marked “Total”.
Table 2: SSE Results for Grid Network 2

<table>
<thead>
<tr>
<th>Prior</th>
<th>obs_1</th>
<th>obs_2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MVR</td>
<td>LOC</td>
</tr>
<tr>
<td>1-10</td>
<td>31.470</td>
<td>33.111</td>
</tr>
<tr>
<td>11-20</td>
<td>21.893</td>
<td>22.277</td>
</tr>
<tr>
<td>21-30</td>
<td>19.821</td>
<td>19.816</td>
</tr>
<tr>
<td>31-40</td>
<td>17.797</td>
<td>18.327</td>
</tr>
<tr>
<td>Total</td>
<td><strong>143.456</strong></td>
<td>146.632</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Prior</th>
<th>obs_1</th>
<th>obs_2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MVR</td>
<td>LOC</td>
</tr>
<tr>
<td>Total</td>
<td>73.503</td>
<td><strong>72.153</strong></td>
</tr>
</tbody>
</table>

Because MVR minimize the posterior variance in each experiment instance, Figures 3 and 4 plot the variances in route flow posteriors when locating different number of sensors using MVR. Because all curves are convex in Figures 3 and 4 and start converging after half of the links covered, it indicates that most variance reductions are performed by the first few sensors when using MVR and the slope (speed of change in variance reduction) monotonically decrease. Within Figure 3 or 4, we can obtain that the quality of priors and observations affect the posterior variances of route flows.
In practice, the decision-makers are more interested in locating sensors only on few links due to budget considerations. Table 3 and 4 tabulate the problem instances when at most half the links can be detectorized, that is, average scores for all instances from $N=1$ to $N=1/2|A|$ and from $N=1$ to $N=1/4|A|$. In each cell of the table, the number on the top represents the value of SSE, while the rest two values are variance and bias related to the instance. Table 3 shows that the average value of SSE from MVR and LOC are always lower than LFC and RFC. Between MVR and LOC, the average SSEs are close because when more noise sensors used, more bias could exist in the estimates. Table 4 focuses on model performance when locating fewer sensors. MVR outperforms among all the instances in terms of SSE, variance and bias.
Table 3: Average SSEs for the Four Location Models for the Problem Instances for N=1 to N=1/2|A|

<table>
<thead>
<tr>
<th>(1st 50%)</th>
<th>MVR</th>
<th>LOC</th>
<th>MFC</th>
<th>RFC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SSE</td>
<td>SSE</td>
<td>SSE</td>
<td>SSE</td>
</tr>
<tr>
<td></td>
<td>Var</td>
<td>Bias</td>
<td>Var</td>
<td>Bias</td>
</tr>
<tr>
<td>prior_1,obs_1</td>
<td>456.73</td>
<td>460.47</td>
<td>674.34</td>
<td>618.45</td>
</tr>
<tr>
<td></td>
<td>7.00</td>
<td>21.21</td>
<td>7.03</td>
<td>21.29</td>
</tr>
<tr>
<td>prior_1,obs_2</td>
<td>268.31</td>
<td>264.66</td>
<td>581.99</td>
<td>482.60</td>
</tr>
<tr>
<td>prior_2,obs_1</td>
<td>125.29</td>
<td>152.03</td>
<td>145.76</td>
<td>142.06</td>
</tr>
<tr>
<td></td>
<td>7.27</td>
<td>10.86</td>
<td>7.48</td>
<td>12.02</td>
</tr>
<tr>
<td>prior_2,obs_2</td>
<td>109.06</td>
<td>113.49</td>
<td>130.58</td>
<td>126.16</td>
</tr>
<tr>
<td></td>
<td>7.04</td>
<td>10.10</td>
<td>7.15</td>
<td>10.31</td>
</tr>
<tr>
<td>prior_1,obs_1</td>
<td>23.52</td>
<td>24.16</td>
<td>36.05</td>
<td>43.31</td>
</tr>
<tr>
<td></td>
<td>2.01</td>
<td>4.64</td>
<td>2.03</td>
<td>4.70</td>
</tr>
<tr>
<td>prior_1,obs_2</td>
<td>24.99</td>
<td>24.63</td>
<td>39.95</td>
<td>41.78</td>
</tr>
<tr>
<td></td>
<td>2.04</td>
<td>4.79</td>
<td>2.06</td>
<td>4.75</td>
</tr>
<tr>
<td>prior_2,obs_1</td>
<td>11.83</td>
<td>11.46</td>
<td>16.16</td>
<td>14.09</td>
</tr>
<tr>
<td></td>
<td>1.90</td>
<td>3.15</td>
<td>1.91</td>
<td>3.09</td>
</tr>
<tr>
<td>prior_2,obs_2</td>
<td>9.58</td>
<td>10.02</td>
<td>21.13</td>
<td>15.92</td>
</tr>
<tr>
<td></td>
<td>1.90</td>
<td>2.77</td>
<td>1.91</td>
<td>2.85</td>
</tr>
</tbody>
</table>
Table 4: Average SSEs for the Four Location Models for the Problem Instances for N=1 to N=1/4

<table>
<thead>
<tr>
<th>Network 1</th>
<th>MVR</th>
<th>LOC</th>
<th>MFC</th>
<th>RFC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SSE</td>
<td>SSE</td>
<td>SSE</td>
<td>SSE</td>
</tr>
<tr>
<td></td>
<td>Var</td>
<td>Bias</td>
<td>Var</td>
<td>Bias</td>
</tr>
<tr>
<td>prior_1,obs_1</td>
<td>469.48</td>
<td>475.78</td>
<td>816.39</td>
<td>721.42</td>
</tr>
<tr>
<td>prior_1,obs_2</td>
<td>280.05</td>
<td>280.14</td>
<td>816.26</td>
<td>648.46</td>
</tr>
<tr>
<td>prior_2,obs_1</td>
<td>125.46</td>
<td>153.29</td>
<td>150.35</td>
<td>133.21</td>
</tr>
<tr>
<td></td>
<td>8.30</td>
<td>10.82</td>
<td>8.51</td>
<td>12.03</td>
</tr>
<tr>
<td>prior_2,obs_2</td>
<td>112.52</td>
<td>123.60</td>
<td>136.18</td>
<td>125.04</td>
</tr>
<tr>
<td></td>
<td>8.17</td>
<td>10.22</td>
<td>8.29</td>
<td>10.74</td>
</tr>
<tr>
<td>prior_1,obs_1</td>
<td>27.61</td>
<td>28.80</td>
<td>43.85</td>
<td>43.51</td>
</tr>
<tr>
<td></td>
<td>2.23</td>
<td>25.38</td>
<td>2.24</td>
<td>26.56</td>
</tr>
<tr>
<td>prior_1,obs_2</td>
<td>29.92</td>
<td>30.54</td>
<td>45.82</td>
<td>40.55</td>
</tr>
<tr>
<td></td>
<td>2.26</td>
<td>27.66</td>
<td>2.27</td>
<td>28.27</td>
</tr>
<tr>
<td>prior_2,obs_1</td>
<td>11.51</td>
<td>12.14</td>
<td>24.24</td>
<td>14.45</td>
</tr>
<tr>
<td></td>
<td>2.10</td>
<td>9.41</td>
<td>2.11</td>
<td>10.02</td>
</tr>
<tr>
<td>prior_2,obs_2</td>
<td>13.39</td>
<td>12.80</td>
<td>19.67</td>
<td>14.74</td>
</tr>
<tr>
<td></td>
<td>2.10</td>
<td>11.28</td>
<td>2.11</td>
<td>10.69</td>
</tr>
</tbody>
</table>

Because of the random effects in the experiment framework, average score is only one way to compare the performance. Further comparisons were conducted using statistical paired-$t$ tests to compare SSE scores with those of other models. Here one is interested in testing if the mean of estimation scores using MVR is significantly lower than that of each of the other models. The hypothesis for the paired-$t$ test is:

$$H_0: \mu_{MVR} \geq \mu_{one \ of \ other \ models}$$

$$H_1: \mu_{MVR} < \mu_{one \ of \ other \ models}$$

If the P-value for a paired-$t$ test is lower than 0.05, it means we can reject $H_0$ at 95% confidence level and have a statistical indication that MVR is better than the compared model. In addition, if the 95% confidence interval does not include zero, it is a statistical proof that both means are different. We conducted three similar paired-$t$ tests to compare SSE scores of MVR versus the scores of LOC, LFC and RFC separately; the P-values and 95% confidence intervals by Minitab are shown in Table 5. The table shows that MVR significantly outperforms all the other models because the P-values are essentially zero and none of the confidence intervals include the value of zero.
Table 1: P-values for Paired-t Tests Comparisons of SSE using Four Location Models

<table>
<thead>
<tr>
<th>Alternative hypothesis: $\mu_{MVR} &lt; \mu$</th>
<th>p-value</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_{LOC}$</td>
<td>0.000</td>
<td>(-2.18307, -0.83229)</td>
</tr>
<tr>
<td>$\mu_{LFC}$</td>
<td>0.000</td>
<td>Upper bound is -21.9645</td>
</tr>
<tr>
<td>$\mu_{RFC}$</td>
<td>0.000</td>
<td>Upper bound is -21.1212</td>
</tr>
</tbody>
</table>

5. CONCLUSIONS

In this paper, for SLFE problem, uncertainties in priors and observations are considered in the decision stage for sensor location while a Bayesian statistical procedure is applied at the estimation stage to consider both, the prior information and observations. A new sensor location model was developed and evaluated. In this model, prior knowledge on the volumes on routes is modeled by distributions. A sub model for selecting a single link at a time was developed to simplify the solution procedure. The proposed algorithm iteratively solves the sub model which makes the original problem solvable in polynomial time.

The model directly maximizes the reduction in the posterior distributions of the route flows. The amount variance reduction is related to the measurement noise in the sensors and the uncertainties in the prior distributions. In the experiments described in the paper four location models were compared. Our earlier model, LOC [7] performed well but in most cases our new model MVR performed better than LOC, and statistically significantly so with a p-value less than or equal to 0.05. The evaluation showed that MVR also performed significantly better than the other two models LFC and RFC.

REFERENCES