Bayesian Inference for Static Traffic Network Flows with Mobile Sensor Data

Zhen Tan
Cornell University
zt78@cornell.edu

H.Oliver Gao
Cornell University
hg55@cornell.edu

Abstract

Vehicle trajectory information are becoming available from mobile sensors such as onboard devices or smart phones. Such data can provide partial information of origin-destination trips and are very helpful in solving the network flow estimation problem which can be very challenging if only link counts are used. Even with this new information, however, there is still structural bias in the maximum likelihood based approach because of uncertainties in the penetration rates. A Bayesian inference approach in which the earlier link-count-based methods are extended is proposed. We incorporate posterior simulation of route-choice probabilities and penetration rates. The results of a numerical example show that our method can infer network flow parameters effectively. Inclusion of mobile sensor data and prior beliefs based on it can yield much better inference results than when non-informative priors and only link counts are used.

1. Introduction

Network flow estimation is a key to successful transportation studies [1] and is critical for intelligent transportation systems and smart city applications [2]. Traffic intensity estimation for all directed origin-destination (OD) pairs from link counts (i.e., the number of vehicles that traverse individual links during a certain time period) has been studied extensively (e.g., [3]-[7]). However, since the number of links is typically much smaller than the number of directed OD pairs in the network, the problem is generally underdetermined (i.e., multiple or infinitely many solutions exist) [6][8]. In order to determine the most likely solution, non-Bayesian approaches include an earlier method that employs the principle of maximum entropy and minimum information [4], and more recent methods that use maximum likelihood estimation (MLE) and moments based algorithms [5]. However, these techniques give little weight to prior information, and only point values (if any) are specified in the prior, with no measure of the degree of belief in it [3].

Furthermore, methods such as MLE have iteration formulae which are very hard to compute [5]. Finally, both MLE and the method of moments [5] suffer from structural ambiguity in the case of Poisson-based likelihood, they tend to overestimate (underestimate) low (high) rates [6].

Because of these limitations, many studies have proposed using Bayesian inference methods for the network flow problem (e.g. [1][3][6][7][9]-[11]). The problem has an inherent Bayesian flavor since we make forecasts of "tomorrow's" OD flows based on "today's" OD flow estimates [6]. However, most of these studies used only historical link counts, without additional information that is relevant to the OD rates.

Naturally, data other than link flows would be useful in inferring the demand [12][13]. An important source of data could be mobile sensors such as onboard GPS devices, probe cars [2][14] or smart phones used for on-demand mobility service, all of which are becoming increasingly available today. Some OD information could also be extracted from data such as the rates of customers requesting and finishing a ride, and drivers' smart phones can now be monitored by on-mobility service providers as well [15]. Such devices can record vehicle trajectories, thus provide information on route flows and the corresponding OD addresses [16], which could be very helpful in solving an otherwise highly underdetermined link-count based OD flow estimation problem [2][8]. Thus, in this study, we use mobile sensor data along with traditional link counts to perform Bayesian inference of traffic network flows. We essentially extend the posterior simulation approach of earlier studies to the case where the Poisson rates are partially observed.

2. Literature review and our approach

The posterior distribution used in estimating fixed-routing network OD flows, which was first formulated in [3], assumes a normal prior in the traffic intensity and normal observation errors. A number of studies (e.g., [11]) have developed variants of that approach based on normality assumption. The advantage of assuming a normal distribution is that posterior

HICSS

updating can be done analytically. However, it entails intensive matrix inversion computations. To avoid this, study [11] adopted Bayesian network and iteratively infer the normal OD rates, but the convergence property of the propose approach is not clear. Also, the assumption of normality of the flow counts could break down in cases where some of the OD rates are relatively small [6]. A more realistic and commonly adopted prior for the OD rates is the Poisson distribution (e.g., [7]), but the resultant posterior cannot be easily evaluated due to difficulties in support identification [6]. Therefore, Markov Chain Monte Carlo (MCMC) methods were first proposed in [6] to infer the Poisson rates. However, the model used in [6] assumes that routing is fixed or has known probabilities for each OD pair.

The actual route-choice probabilities over different routes depend on factors such as travel times [9][17]. Hence some later studies focus on inferring route flows instead. For example, an expectation maximization (EM) based approach was proposed to do Bayesian inference [7] of route flows. The MCMC approach of [6] was extended to time-varying route flows in [9], where the route-choice probabilities are functions of the inferred historical route flow. Thus the uncertainties in the route-choice probabilities were also modeled. Similar approaches have been proposed in [1] for dynamic OD inference and in [11] for static OD inference. The study in [17] includes the key parameters of the choice probability functions in the posterior simulation. However, inferring such a long parameter vector calls for a long sequence of historical data and expensive computation [9]. Moreover, the functional form of the choice probabilities could affect the results of the inference. Thus it would be desirable to assign uncertainties to the route-choice probabilities more straightforwardly, which is a key element of the method proposed in this study. Listed in Table 1 (in chronological order) are some representative studies on Bayesian inference of link-count-based estimation of network OD flows. Several aspects of those studies are compared: the methods used for posterior computation, the assumed distributions of the (independent) OD rates, the time domain of the estimation problem, and how routing is incorporated.

Note that among all the studies listed in Table 1, only ours utilizes a partial set of actual routing data. As new data of this type become available, more informative priors can be built. However, one needs to have a good sense of the penetration rates of the mobile sensors from which the data are extracted, and those rates are nontrivial to estimate [14]. The study in [8] formulated the OD flow estimation problem using both link counts and sporadic routing data. It pointed to problems of MLE approach which are intrinsic even if the penetration rates are common across ODs. Actual penetration rates can have considerable regional variation [14], so structural bias of MLE could have negative consequences. Thus it would be desirable to have a reliable procedure for inferring the penetration rates in order to better utilize the new data.

Table 1. List of related studies

Study	Posterior comp.	Rates dist.	Time domain	Routing
[3]	analytical	Normal	static	fixed
[6]	MCMC	Poisson	static	fixed or known prob.
[7]	EM	Poisson	static	fixed.[a]
[9]	MCMC	General[b]	dynamic	formula for prob. [c2]
[11]	iterative formulas	Normal	static	formula for prob. [c1]
[1]	MCMC	Normal	dynamic	formula for prob. [c1]
This work	MCMC	Poisson[e]	static	unknown prob.[d]

a: Route rates were inferred directly. b: Poisson dist. was used in examples. c: Routing probabilities depend on the realized costs (a fixed function of the route flow) in the previous time step via a known function (c1), or an unknown function whose parameters were to be inferred (c2). d: Routing probabilities were inferred directly from a partial set of observed route rates.

Therefore, our approach extends the MCMC framework by incorporating inference on OD-specific penetration rates and route-choice probabilities. We derive the conditional posteriors of those new parameters, and we propose convenient ways to utilize partial set of actual routing data to construct informative priors. Hence our model is not only able to jointly infer more parameters as a result of including the new data, but also to do so more efficient by using priors that are more informative.

3. Data and model

We consider a network of n nodes and r directed links. Each OD pair a consists of two distinct nodes, so the number of OD pairs is $c = n(n-1)$. Let $s = (i, j)$ represent the directed link from node i to node j, and X_a be the total traffic count of OD pair a within a certain time period to our interest (which is assumed to be no less than the travel time for any OD pair). X_a's are the quantities we want to infer.

3.1. Observed and unobserved data

Let Y_s be the observed traffic count on link s over the time period for which the OD rates are to be

inferred. Note that the number of measured links is at most r, which is typically smaller than c. Thus, the network flow estimation problem is in general underdetermined if only link counts are used.

Let k_a be the number of routes can be used for OD pair a, and let the counts for those routes be $X_{a,1}$, $X_{a,2}$, ..., X_{a, k_a}, so $\Sigma_{t=1...k_a} X_{a,t} = X_a$. Let $m = \Sigma_a k_a$ be the total number of routes, which is in general larger than c and much larger than r. We collect the OD counts $\{X_a, a = 1, ..., c\}$ into an m-dimensional column vector of route counts, X, by arranging all the route flow components $X_{a,t}$ for all OD pairs in order of their OD address a, that is, $X := \{X_{a,t}, t = 1, ..., k_a; a = 1, ..., c\}$. By this construction, it is clear that the total flow for OD pair a relates to routes $t = (1 + \Sigma_{a'=1...a-1} k_{a'}), ..., \Sigma_{a'=1...a} k_{a'}$.

Apart from the link counts from fixed traffic sensors, we can also obtain a partial route counts from mobile sensors [8]. We assume that the correspondence between the vehicle trajectory data and the underlying OD pairs is known. Let $X^{trk}_{a,t}$ be the tracked flow count on route t of OD pair a, so we have the tracked count for OD pair a within the period of interest is $X^{trk}_a = \Sigma_{t=1...k_a} X_{a,t} \leq X_a$. We store all the tracked route counts in the m-dimensional column vector $X^{trk} = \{X^{trk}_{a,t}, t = 1, ..., k_a; a = 1, ..., c\}$, in the same order as in X, and let r-dimensional column vector $Y^{trk} = \{\Sigma_a Y^{trk}_{a,s}, s = 1, ..., r\}$ contain the total link counts corresponding to the observed route counts. Since $X \geq X^{trk}$ (pairwise), we can define a nonnegative column vector $X^{not} = X - X^{trk} = \{X^{not}_a, a = 1, ..., c\}$, which represents the unobserved route counts for all the OD pairs. We also define a column vector $Y^{not} = Y - Y^{trk}$, which represents the link counts that are not part of any tracked route count. We have the following key relationship:

$$Y = AX; \quad Y^{trk} = AX^{trk}, \quad (1)$$

where $A = \{A_{s,t}, s = 1, ..., r; t = 1, ..., m\}$ is the $r{\times}m$ routing matrix, with $A_{s,t} = 1$ if link s belongs to route t, and $A_{s,t} = 0$ otherwise. Equation (1) simply expresses each link count in Y (or Y^{trk}) as the sum of the counts $X_{a,t}$ (or $X^{trk}_{a,t}$) for all the routes that use link s. The difference of the two equations in (1) gives:

$$Y^{not} = AX^{not}. \quad (2)$$

Thus Y^{not} imposes a set of linear equality constraints on X^{not}. Equation (2) is utilized in the Bayesian inference model later. Note that A is singular, since the number of its rows (r) is smaller than number of its columns (m). We assume that no route contains a cycle, this indicates that X and A have finite size.

3.2. Underlying distributions

We assume that the count for OD pair a follows a Poisson distribution with mean λ_a: $X_a \sim \text{Poisson}(\lambda_a)$. We define the Poisson OD rates vector $\Lambda = (\lambda_1, ...,$

$\lambda_c)^T$. The Poisson distribution is commonly used in traffic modeling, as in [17]. It has been proposed (e.g., in [3] and [8]) that a Poisson distribution be approximated by a normal distribution, mainly because of computational concerns, but a normal approximation is rather poor when λ_a is small, and can lead to (unrealistic) negative estimates of the flow variables [7][8]. In addition, the computational effort required with use of an Maximum a posteriori (MAP) estimator or the MLE is still considerable when the variance of the approximating normal distribution depends on λ_a [7]. Hence we focus on the original Poisson model, our proposed inference framework can easily be extended to normally distributed OD counts. Also, note that dependencies between flows (e.g., gravity models) can be modeled via hierarchical models for the rates Λ and can be conveniently incorporated in the same Bayesian framework [6][18]. Therefore, in this study, we also restrict our model development based on the independent Poisson structure.

Now we introduce the other two key parameters. First, $w_a \in [0, 1]$ is the penetration rate of the mobile sensors (e.g., regular on-demand service users) for OD pair a. Let $w = (w_1, ..., w_c)^T$. Unlike in [8], we assume that w_a is OD-pair specific. Second, $p_{a,t}$ is the choice probability for route t of OD pair a, so $\Sigma_{t=1...k_a} p_{a,t} = 1$. Let $p_a = (p_{a,1}, ..., p_{a, k_a})^T$. The conditional distributions of the observed and unobserved OD counts are: $X^{trk}_a \mid X_a \sim \text{Binomial}(X_a, w_a)$ and $X^{not}_a \mid X_a \sim \text{Binomial}(X_a, u_a)$, respectively, where $u_a = 1 - w_a$. The conditional distributions of the observed and unobserved route counts are: $\{X^{trk}_{a,t}, t = 1, ..., k_a\} \mid X^{trk}_a \sim \text{Multinomial}(X^{trk}_a, p_a)$ and $\{X^{not}_{a,t}, t = 1, ..., k_a\} \mid X^{not}_a \sim \text{Multinomial}(X^{not}_a, p_a)$, respectively. This implies that the unconditional distributions are:

$$X^{trk}_a \sim \text{Poisson}(w_a \lambda_a); \ X^{not}_a \sim \text{Poisson}(u_a \lambda_a);$$
$$X^{trk}_{a,t} \sim \text{Poisson}(p_{a,t} w_a \lambda_a); \quad (3)$$
$$X^{not}_{a,t} \sim \text{Poisson}(p_{a,t} u_a \lambda_a).$$

3.3. The estimation problem and issue of MLE

The problem is to jointly infer the underlying mean OD rates Λ, the penetration rates w and the route-choice probabilities $P = (p_1^T, ..., p_c^T)^T$ from the observed link counts $Y = (Y_1, ..., Y_r)^T$ and the partial set of route counts X^{trk}. Note that we can also infer the mean route rates as in other studies (e.g., [7], [9]). Then the route-choice probabilities can be indirectly estimated by calculating the ratios of the individual route rates to the total OD rates. We instead focus on direct inference of the route-choice probabilities in order to more naturally combine those with our prior knowledge of the routing decisions.

Actually, any parameter vector $\boldsymbol{\theta} = (\boldsymbol{\Lambda}^{\mathrm{T}}, \boldsymbol{w}^{\mathrm{T}}, \boldsymbol{P}^{\mathrm{T}})^{\mathrm{T}}$ that satisfies $A\boldsymbol{\Lambda} = \boldsymbol{Y}$ and $diag(w_1\boldsymbol{p}_1^{\mathrm{T}},\ldots, w_c\boldsymbol{p}_c^{\mathrm{T}})\boldsymbol{\Lambda} = \boldsymbol{X}^{trk}$ will produce estimates of $E[\boldsymbol{Y}]$ and $E[\boldsymbol{X}^{trk}]$ that match the corresponding observed data perfectly. Therefore, since the number of links, r, is typically much smaller than the total number of routes, m, we cannot expect to obtain reliable estimates for the problem using MLE. This is mentioned in the study of the normal approximation model in [8] (and would apply to our study if we considered the product $w_a p_{a,t}$ as the "route penetration rate $q_{a,t}$" for route t of OD pair a in their context), but that study focused on the case where $q_{a,t}$ is the same for all routes. Since we use OD-specific penetration rates w_a and route-choice probabilities $\boldsymbol{p}_a$, such unreliability would be common in our problem setting, and could lead to frequent bias in estimation.

Hence, these issues have led us to focus on a situation where Bayesian approach can be fully exploited. Such an approach allows us to incorporate prior beliefs about OD intensities to guide the posterior analysis and prevent the "structure bias" of MLEs with extreme estimates.

Next we will present the Bayesian inference model. We include the prior and the inference model in both $\boldsymbol{w}$ and $\boldsymbol{P}$, which is the main contribution of this study.

4. Bayesian inference model

First note that given the route counts $\boldsymbol{X}^{not}$, we can immediately evaluate the conditional posteriors of the Poisson rates, penetration rates and route choice probabilities. Thus, we also move the latent variable $\boldsymbol{X}^{not}$ into the joint distribution.

4.1. Joint distribution

By our assumption of the independent Poisson model, together with equation (2), the independence of $\boldsymbol{Y}^{not}$ and $\boldsymbol{X}^{trk}$, and the independence of $\boldsymbol{X}^{not}$ and $\boldsymbol{X}^{trk}$, we have the following joint distribution:

$$
\begin{aligned}
p(\boldsymbol{X}^{trk}, \boldsymbol{X}^{not}, \boldsymbol{Y}, \boldsymbol{\theta}) &= p(\boldsymbol{X}^{trk}, \boldsymbol{Y}^{not}, \boldsymbol{X}^{not}, \boldsymbol{\theta}) \\
&= p(\boldsymbol{\theta})p(\boldsymbol{X}^{trk}|\boldsymbol{\theta})p(\boldsymbol{X}^{not}|\boldsymbol{\theta})p(\boldsymbol{Y}^{not}|\boldsymbol{X}^{not}, \boldsymbol{\theta}) \\
&= p(\boldsymbol{\theta})p(\boldsymbol{X}^{trk}|\boldsymbol{\theta})p(\boldsymbol{X}^{not}|\boldsymbol{\theta})\mathbf{1}_{\{\boldsymbol{Y}^{not}=A\boldsymbol{X}^{not}\}},
\end{aligned} \quad (4)
$$

$$
p(\boldsymbol{X}^{trk}|\boldsymbol{\theta}) = \prod_{a=1}^{c}\prod_{t=1}^{k_a} \frac{(p_{a,t}w_a\lambda_a)^{X_{a,t}^{trk}} e^{-p_{a,t}w_a\lambda_a}}{X_{a,t}^{trk}!},
$$

$$
p(\boldsymbol{X}^{not}|\boldsymbol{\theta}) = \prod_{a=1}^{c}\prod_{t=1}^{k_a} \frac{(p_{a,t}u_a\lambda_a)^{X_{a,t}^{not}} e^{-p_{a,t}u_a\lambda_a}}{X_{a,t}^{not}!},
$$

where the indicator function $\mathbf{1}_{\{f\}}$ takes the value 1 when f is true, and 0 otherwise. Hence if the prior for each OD count, λ_a, the prior for each penetration rate, w_a,

and the prior on the route-choice probabilities, $\boldsymbol{p}_a$, as well as the above indicator function, are all given, then the joint distribution can be characterized. Each of the prior models can be specified. For convenience in posterior evaluation, we would like to use independent conjugate priors for $\boldsymbol{\Lambda}$, $\boldsymbol{w}$ and $\boldsymbol{P}$ across OD pairs, which will be discussed in detail in Section 4.3.

4.2. Conditional posteriors

Under the assumption of the independent Poisson model, we can easily obtain the full conditional distributions for $\boldsymbol{\Lambda}$, $\boldsymbol{w}$ and $\boldsymbol{P}$, since they will all have conjugate priors. Some additional work is needed to obtain the conditional distribution for $\boldsymbol{X}^{not}$, which doesn't have an analytical form.

By (4), the full conditional distribution for the Poisson rates is:

$$
p(\boldsymbol{\Lambda}|\boldsymbol{X}^{trk}, \boldsymbol{X}^{not}, \boldsymbol{Y}, \boldsymbol{P}, \boldsymbol{w}) = p(\boldsymbol{\Lambda}|\boldsymbol{X}^{trk}, \boldsymbol{X}^{not})
$$

$$
= \prod_{a=1}^{c} p(\lambda_a|X_a^{trk} + X_a^{not}), \quad (5)
$$

$$
p(\lambda_a|X_a^{trk} + X_a^{not}) \propto p(\lambda_a)p(X_a^{trk} + X_a^{not}|\lambda_a)
$$

$$
\propto p(\lambda_a)\lambda_a^{X_a^{trk}+X_a^{not}} e^{-\lambda_a}.
$$

Thus, conditional on $\boldsymbol{X}^{not}$ and $\boldsymbol{X}^{trk}$, we get X_a by summing over the entries of $\boldsymbol{X}^{not}$ and $\boldsymbol{X}^{trk}$ that correspond to OD pair a, and then we can easily simulate $\boldsymbol{\Lambda}$ as a set of independent draws from the implied univariate posteriors. If the prior $p(\lambda_a)$ is a gamma distribution or a mixture of gammas, these draws are made from the corresponding gamma (or mixture of gamma) posteriors.

Similarly, from (4) we can get the full conditional distribution for the penetration rates:

$$
p(\boldsymbol{w}|\boldsymbol{X}^{trk}, \boldsymbol{X}^{not}, \boldsymbol{Y}, \boldsymbol{P}, \boldsymbol{\Lambda}) = p(\boldsymbol{w}|\boldsymbol{X}^{trk}, \boldsymbol{X}^{not}, \boldsymbol{P}, \boldsymbol{\Lambda})
$$

$$
= \prod_{a=1}^{c} p(w_a|X_a^{trk}, X_a^{not}, \lambda_a, \boldsymbol{p}_a), \quad (6)
$$

$$
p(w_a|X_a^{trk}, X_a^{not}, \lambda_a, \boldsymbol{p}_a) \propto p(w_a)p(X_a^{trk}, X_a^{not}|w_a, \lambda_a, \boldsymbol{p}_a)
$$

$$
\propto p(w_a)\prod_{t=1}^{k_a} w_a^{X_{a,t}^{trk}}(1-w_a)^{X_{a,t}^{k}}
$$

$$
= p(w_a)w_a^{X_a^{trk}}(1-w_a)^{X_a^{k}}.
$$

Thus, conditional on $\boldsymbol{X}^{not}$ and $\boldsymbol{X}^{trk}$, we can easily simulate $\boldsymbol{w}$ as a set of independent draws from the implied univariate posteriors. If the prior $p(w_a)$ follows beta distribution, these draws are also made from the updated beta posteriors.

The vector $\boldsymbol{Y}^{not}$ in (2) implies a set of r linear equality constraints on $\boldsymbol{X}^{not}$. As a result, the posterior of $\boldsymbol{X}^{not}$ can be reduced to a set of r linear equations that deliver precise values of r elements of $\boldsymbol{X}^{not}$ given specified values of the remaining $m - r$ elements. The marginal posteriors for these $m - r$ elements can be directly evaluated and used as the centerpiece for the

MCMC simulation. The algebraic structure of general network, which is derived in [6], is summarized below, and is critical to the ensuing inferential development. First we reorder the columns of the routing matrix A as $A = (A_1\ A_2)$, where A_1 is a non-singular $r \times r$ matrix. Then we have

$$X_1 = A_1^{-1}(Y^{not} - A_2 X_2), \qquad (7)$$

where the elements of X^{not} are reordered as $(X_1^T\ X_2^T)^T$ in accord with the partition $A = (A_1\ A_2)$. Equation (7) is a simple consequence of equation (2) and the invertibility of A_1. We reorder the entries of the vector P (whose t^{th} element is P_t) to match the reordered X^{not}. We use the reordered X^{not} and P in the sequel.

Based on this property, for any θ, and fixed Y^{not}, the conditionals $[X^{not}|\theta, Y^{not}]$ are concentrated in the sub-space of dimension m-r defined by the partition (7) of the routing matrix A. After column reordering, the posterior has the form $[X_1 | X_2, \theta, Y^{not}][X_2 | \theta, Y^{not}]$, where $[X_1 | X_2, \theta, Y^{not}]$ is degenerate at $X_2 = (X^{not}_{r+1}, \ldots, X^{not}_m)^T$ and $X_1 = A_1^{-1}(Y^{not} - A_2 X_2)$. Hence the full conditional for X_2 (over the support defined by $X^{not}_t \geq 0$ for all $t = 1, \ldots, m$) is

$$p(X_2 | \theta, Y^{not}) \propto \prod_{t=1}^{m} \frac{(P_t w_{a(t)} \lambda_{a(t)})^{X^{not}_t}}{X^{not}_t!}, \qquad (8)$$

where $a(t)$ denotes the index of the OD pair that corresponds to X^{not}_t (the tth element of the reordered X^{not}). Equation (8) means that the full conditional for X_2 is simply the product of the independent Poisson priors for all the X^{not}_t, which are constrained by (2) and rewritten in the form in (7). Using (8), we have that the full conditional for each $X_{2,t}$ ($t = 1, \ldots, m - r$) (over the support defined by $X_{2,t} \geq 0$, $t = 1, \ldots, m - r$):

$$p(X_{2,t} | X_{2,-t}, \theta, Y^{not}) \propto \frac{(P_t \lambda_a)^{X^{not}_t}}{X^{not}_t!} \prod_{l=1}^{r} \frac{(P_l \lambda_{a(l)})^{X_{1,l}}}{X_{1,l}!}, \qquad (9)$$

where $X_{2,-t}$ denotes the elements in X_2 other than $X_{2,t}$. Note that $X_{2,t}$ corresponds to X^{not}_{t+r} of the reordered vector X^{not}. Identifying the support of (9) requires study of the linear constraints $X_1 \geq 0$ for $X_1 = A_1^{-1}(Y - A_2 X_2)$, as discussed in [6], which amounts to $X^{not}_t \geq 0$ for $t = 1, \ldots, r$). Note that because of incorporation of mobile sensor data, the underlying support of X^{not}_t ($t = 1, \ldots, r$) will be smaller than the corresponding total route flow, which makes the inference problem computationally more efficient.

Based on (4), we can also get the full conditional distribution for P. We have the conditional likelihood for each vector p_a given the computed route counts $X^{trk}_{a,t}$, $X^{not}_{a,t}$ as well as Poisson rate λ_a and the penetration rate w_a. On this basis, the likelihood function factorizes into a set of c components of the form in (4), and (since w_a, λ_a are independent of p_a) the full conditional of each p_a is:

$$p(P | X^{trk}, X^{not}, Y, w, \Lambda) = p(P | X^{trk}, X^{not}, w, \Lambda)$$

$$= \prod_{a=1}^{c} p(p_a | X^{trk}_a, X^{not}_a, w_a, \lambda_a), \qquad (10)$$

$$p(p_a | X^{trk}_a, X^{not}_a, w_a, \lambda_a) \propto p(p_a) p(X^{trk}_a, X^{not}_a | p_a, w_a, \lambda_a)$$

$$\propto p(p_a) \lambda_a^{X^{trk}_a + X^{not}_a} e^{-\lambda_a} \prod_{t=1}^{k_a} \frac{p_{a,t}^{X^{trk}_{a,t} + X^{not}_{a,t}}}{X^{trk}_{a,t}! X^{not}_{a,t}!}$$

$$\propto p(p_a) \prod_{t=1}^{k_a} p_{a,t}^{X^{trk}_{a,t} + X^{not}_{a,t}},$$

where the second last line follows from $\Sigma_{t=1\ldots k_a} X^{not}_{a,t} = X^{not}_a$, $\Sigma_{t=1\ldots k_a} X^{trk}_{a,t} = X^{trk}_a$ and $\Sigma_{t=1\ldots k_a} p_{a,t} = 1$. Thus, it is easy to see that when the prior $p(p_a)$ is a Dirichlet distribution, so is the conditional posterior of p_a [19]. Specifically, suppose the prior $p(p_a)$ is a Dirichlet distribution with concentration parameters $\gamma_{a,1}$, $\gamma_{a,2}$, …, γ_{a,k_a}. Then the posterior of p_a conditional on $X_a = X^{trk}_a + X^{not}_a = \Sigma_{t=1\ldots k_a} X_{a,t}$ is another Dirichlet distribution with new concentration parameters $\gamma_{a,1} + X_{a,1}$, $\gamma_{a,2} + X_{a,2}$, …, $\gamma_{a,k_a} + X_{a,k_a}$. Hence in this case the sampling of p_a from its posterior distributions is also straightforward.

4.3 Prior information

We could simply assign a uniform prior or a highly diffuse gamma prior for Λ; however, that could lead to estimation biases due to the structural ambiguity of the Poisson-based likelihood without additional constraints [6]. Therefore, updating of the prior estimates based on the historical traffic data and experience is helpful to constrain the problem and overcome the identification difficulties [3] [6].

In the absence of historical survey data that are reliable, we propose to use the penetration information to form a prior for Λ. Since the penetration itself is to be inferred, we assume a beta prior for $w_a \in [0, 1]$, $w_a^0 \sim \text{Beta}(x_a^0, y_a^0)$ with mean $z_a^0 = x_a^0/(x_a^0 + y_a^0)$. Then since we already have a partial set of observations $X^{trk}_a = \Sigma_t X^{trk}_{a,t}$, it is natural to specify the prior mean of λ_a as $d_a^0 = \max(1, X_a^{trk}) / z_a^0$. Therefore, we use the gamma distribution with rate $\alpha_a^0 = b > 0$ and shape $\beta_a^0 = bd_a^0$ for the prior of λ_a. The smaller the value of b, the more diffuse the prior.

As a starting point, we may need to "guess" the parameters (x_a^0, y_a^0) of the beta prior for the penetration rate w_a. They may be estimated from historical on-demand survey data [14] or user registration data. If such data are not available, a natural way to do this is to look at the total link counts versus the tracked link counts. If we assume that the penetration rates are similar across different ODs, we can use a common prior for all the ODs ($a = 1, \ldots, c$) with parameters

$$\begin{cases} x_a^0 = \sum_{s=1}^{r} Y_s^{trk} = \mathbf{1}_r^T \boldsymbol{A} \boldsymbol{X}^{trk}; \\ y_a^0 = \sum_{s=1}^{r} Y_s^{not} = \sum_{s=1}^{r}(Y - Y_s^{trk}) = \mathbf{1}_r^T(\boldsymbol{Y} - \boldsymbol{A}\boldsymbol{X}^{trk}), \end{cases} \quad (11)$$

where $\mathbf{1}_r$ is an r-dimensional column vector of 1's. This implies that the prior mean for w_a, $z_a^0 = x_a^0/(x_a^0 + y_a^0)$, is just the ratio of the total link counts that are tracked to the total observed link count, which is consistent with the recommendation in [8] and [14]. In addition, since $\text{Var}(w_a^0) = x_a^0 y_a^0 (x_a^0 + y_a^0)^{-2}(x_a^0 + y_a^0 + 1)^{-1}$, such specification leads to a smaller variance (i.e., our prior beliefs on w_a are more certain) when the total link count is larger.

In practice, it is also possible that the penetration rates vary from one geographical area to another [14]. If we assume there are significant spatial differences among the penetration rates, we can use a separate beta prior for each OD pair a in the following way:

$$x_a^0 = \sum_{s:A_{sa}=1} Y_s^{trk}; \ y_a^0 = \sum_{s:A_{sa}=1} Y_s^{not} = \sum_{s:A_{sa}=1}(Y_s - Y_s^{trk}). \quad (12)$$

This is simply the proportion of tracked counts among the total counts on the links that are used by OD pair a. However, if different paths have common links, the number of tracked counts on a specific link would be the sum of the tracked counts from all the OD pairs that uses that link [14]. Hence this is just a rough prior estimate that can be used when other data are hard to get. Also the larger the total counts on links in $\{s: A_{sa} = 1\}$, the more confident we are of this prior estimate.

As for the prior for the routing probabilities $\boldsymbol{p}_a$ (since they are taken to be random variables in this study), the assumptions will depend on the context. We may choose a uniform prior or an informative Dirichlet prior based on the experience or historical data. As noted earlier, the information obtained from the mobile sensor provides vehicle trajectory data, which can give us a partial set of observations of route flows $X^{trk}_{a,t}$, and corresponding partial set of OD flows X^{trk}_a. Thus the fractions of the individual routes being used for OD pair a from these partial observations, $f_{a,1}, \ldots, f_{a,k_a}$, (where $f_t = X^{trk}_{a,t}/X^{trk}_a$) can be calculated if $X^{trk}_a > 0$; otherwise, we assume that $f_{a,1} =, \ldots, = f_{a,k_a} = 1/k_a$. Therefore a natural choice is to use a Dirichlet prior distribution for $\boldsymbol{p}_a$ with $(f_1, \ldots, f_{k_a})^T$ as its mean. We assume that the larger the number of observed trips with known trajectories, the more certain our belief that they are the "true" fractions, this belief is encoded in the prior $\boldsymbol{p}_a^0$. If $\boldsymbol{p}_a^0 \sim \text{Dirichlet}(\boldsymbol{\gamma}_a^0)$, then the mean and variance of its component $p_{a,t}^0$ are:

$$\text{E}(p_{a,t}^0) = \frac{\gamma_{a,t}^0}{\gamma_{a,0}^0}; \quad \text{Var}(p_{a,t}^0) = \frac{\gamma_{a,t}^0(\gamma_{a,0}^0 - \gamma_{a,t}^0)}{(\gamma_{a,0}^0)^2(\gamma_{a,0}^0 + 1)},$$

where $\gamma_{a,0}^0 = \Sigma_{t=1\ldots k_a} \gamma_{a,t}^0$. It follows that if we use $\boldsymbol{\gamma}_a^0 = (\max(X^{trk}_{a,1},1),\ldots, \max(X^{trk}_{a,k_a},1))^T$ as the parameter

vector in $\boldsymbol{p}_a^0 \sim \text{Dirichlet}(\gamma_a^0)$, we can achieve the desired properties that are mentioned earlier, that is, $\text{E}(p_{a}^0{}_{,t}) = f_t$, and the larger the observation $X^{trk}_{a,t}$ and X_a^{trk}, the more certain we are of the value of $p_a^0{}_{,t}$. An example will be presented in Section 6. Note that if $X^{trk}_a = 0$, we have a uniform prior on $\boldsymbol{p}_{a,t}$, for example, γ_a^0 is a k_a-dimensional vector of 1's.

5. MCMC implementation

We inherit the MCMC framework introduced in [6] to infer the OD rates, which combines the Metropolis-Hastings (M-H) step within an overall Gibbs sampling framework. However, we extend the method by including simulation of unknown route-choice probabilities and penetration rates, both of which are OD-specific. This is done following their full conditionals. In the posterior simulation, we attach informative priors to each of the OD rates and route-choice probabilities using the mobile sensor data.

The compatibility of the independent Poisson models for the priors for the route counts implies that the construction of the conditional posteriors for route counts are structurally similar with those for the fixed routing case. Hence, the MCMC analysis for simulation of the route counts under fixed routing proposed in [6] still apply and will produce the full set of unobserved route counts $\boldsymbol{X}^{not} = \{X^{not}_{a,t}, t = 1, \ldots ,k_a; a = 1, \ldots, c\}$. In particular, the results on convergence of the MCMC also hold in our problem settings.

5.1 Overall Gibbs sampling

Given the observed data $\boldsymbol{Y}$ and $\boldsymbol{X}^{trk}$, we can first compute $\boldsymbol{Y}^{not} = \boldsymbol{Y} - \boldsymbol{Y}^{trk} = \boldsymbol{Y} - \boldsymbol{A}\boldsymbol{X}^{trk}$. Then given the initial prior distribution parameters $x_a^0, y_a^0, \alpha_a^0, \beta_a^0, \gamma_a^0$, we can use the procedure in Algorithm 1 to simulate the posterior distribution $[\boldsymbol{X}^{not}, \boldsymbol{\theta} \mid \boldsymbol{Y}^{not}, \boldsymbol{X}^{trk}]$. Note that K is the total number of iterations, which is predefined.

5.2 M–H Step for unobserved route flows

As discussed in [6], the sampling step for $X_{2,t}^k$ (line 10 in Algorithm 1) is not easy except for certain small networks. Larger networks involve a large support in (9), which can lead to an excessive computational burden if only link counts are considered [7]. However, this computational cost can be reduced significantly thanks to the availability of a partial set of link flow observations $\boldsymbol{X}^{trk}$. Therefore, we can embed an M–H step into the Gibbs sampling framework to infer only part of $\boldsymbol{X}^{not}$ (namely, $\boldsymbol{X}_2$). Let $p_t(\cdot)$ be the unnormalized conditional posterior in (9), and let $q_t(X_{2,t})$ be the fixed

proposal distribution for each single element. A candidate value $X_{2,t}^*$ is drawn from $q_t(\cdot)$ and accepted with probability:

$$\eta = \min\left[1, \frac{p_t(X_{2,t}^*)q_t(X_{2,t})}{p_t(X_{2,t})q_t(X_{2,t}^*)}\right]. \qquad (13)$$

Note that the unnormalized density $p_t(\cdot)$ is evaluated only at candidate draws, hence for a given proposal distribution $q_t(\cdot)$, it is unnecessary to identify the actual support of $p_t(\cdot)$, or to evaluate it completely across the support, which is different from direct Gibbs sampling [6]. When all the OD rates are large, we can use a normal approximation such as the one in [7], and the computational cost can be reduced in this step.

| Algorithm 1 MCMC posterior simulation |
|---|
| 1: $k \leftarrow 0$ |
| 2: **for** $a = 1, ..., c$ **do** |
| 3:   Sample $w_a^k \sim \text{Beta}(x_a^k, y_a^k)$ |
| 4:   Sample $\lambda_a^k \sim \text{Gamma}(\alpha_a^k, \beta_a^k)$ |
| 5:   Sample $\boldsymbol{p}_a^k$ from Dirichlet(γ_a^k) |
| 6: $\boldsymbol{w}^k \leftarrow \{w_a^k\}, \boldsymbol{\Lambda}^k \leftarrow \{\lambda_a^k\}, \boldsymbol{P}^k \leftarrow \{\boldsymbol{p}_a^k\}$ |
| 7: $\boldsymbol{\theta}^k \leftarrow (\boldsymbol{\Lambda}^{k\text{T}}, \boldsymbol{w}^{k\text{T}}, \boldsymbol{P}^{k\text{T}})^\text{T}$ |
| 8: **for** $t = r + 1, ..., m$ **do** |
| 9: $\boldsymbol{X}_{2,-t} \leftarrow (X_{2,r+1}^k, ..., X_{2,t-1}^k, X_{2,t+1}^{k-1} ..., X_{2,m}^{k-1})^\text{T}$ |
| 10:   Sample $X_{2,t}^k \sim p(\cdot \mid \boldsymbol{X}_{2,-t}, \boldsymbol{\theta}^k, \boldsymbol{Y}^{not})$ by (9) |
| 11: $\boldsymbol{X}_1^k \leftarrow \boldsymbol{A}_1^{-1}(\boldsymbol{Y}^{not} - \boldsymbol{A}_2\boldsymbol{X}_2^k)$ by (7) |
| 12: $\boldsymbol{X}^{not\,k} \leftarrow (\boldsymbol{X}_1^{k\text{ T}} \boldsymbol{X}_2^{k\text{ T}})^\text{T}$ |
| 13: **for** $a = 1, ..., c$ **do** |
| 14: $x_a^{k+1} \leftarrow x_a^k + X^{trk}{}_a, y_a^{k+1} \leftarrow y_a^k + X^{not}{}_a,$ |
| 15: $\alpha_a^{k+1} \leftarrow \alpha_a^k + \sum_t X^{trk}{}_{a,t}^k + X^{not}{}_a,^k + 1, \beta_a^{k+1} \leftarrow \beta_a^k + 1$ |
| 16:   **for** $t = 1, ..., k_a$ **do** |
| 17: $\gamma_{a,t}^{k+1} \leftarrow \gamma_{a,t}^k + X^{trk}{}_{a,t} + X^{not}{}_a,^k$ |
| 18: **if** $k < K$ **then** |
| 19: $k \leftarrow k + 1,$ **go to** line 2 |
| 18: **else** |
| 19:   Stop |

Here we use the Poisson distribution as the proposal distribution. In the first step, one of the components of the vector $\boldsymbol{X}_2$ is chosen, and its value range is determined to be either zero or a large set of values. This is performed by calculating $\min_t\{Y^{not}{}_t - \sum_{t' \neq t} A_{t,t'}X^{not}{}_{t'}\}$, which is what remains after subtracting from the observed $Y^{not}{}_t$ the values of the remaining components of $\boldsymbol{X}_2$ and taking the minimum among those differences. If the support is larger than the single point zero, Poisson($p_{a(t),t}u_{a(t)}\lambda_{a(t)}$) is used as the proposal distribution, and the ratio of the two likelihood functions of the data is computed, with the numerator a function of the new candidate $X_{2,t}^*$ and the denominator a function of the old value $X_{2,t}$, keeping everything else fixed. That is, from (9), we obtain the acceptance probability for candidate $X_{2,t}^*$:

$$\eta = \min\left[1, \frac{\prod_{t=1}^{r}(p_{a(t),t}\lambda_{a(t)})^{X_{1,t}^*} / X_{1,t}^*!}{\prod_{t=1}^{r}(p_{a(t),t}\lambda_{a(t)})^{X_{1,t}} / X_{1,t}!}\right], \qquad (14)$$

where $X_{1,t}^*$ is an element from the newly computed $\boldsymbol{X}_1$ as a function of $\boldsymbol{X}_{2,-t}$ and $X_{2,t}^*$ through (7). We should ensure that all the elements of the newly computed vector $\boldsymbol{X}_1$ are nonnegative. If any element of $\boldsymbol{X}_1$ is negative, we simply reject this candidate $X_{2,t}^*$ at that sampling. This is more straightforward and saves considerable computation cost compared to the "iterative, trial and error search process" used in [6] where a uniform proposal is used.

6. A numerical example

We used a classic simple network such as the one in [5][6] for a preliminary case study. The network has 4 nodes and 7 directed links, as shown in Figure 1. There are 12 OD pairs and 17 routes. Our task was to do Bayesian inference of the 12 OD rates and the $17 - 12 = 5$ independent route-choice probabilities.
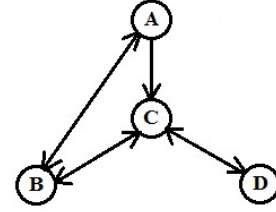
Figure 1. Example network.

We took the actual values of $\boldsymbol{\Lambda}^* = \{\lambda_a^*\}$ to be $5 \times (1, 2, 3, 4, 5, 6, 1, 2, 3, 4, 5, 6)^\text{T}$, and the route-choice probabilities $\boldsymbol{p}^* = (p_{11}^*, p_{21}^*, p_{31}^*, p_{51}^*, p_{61}^*)^\text{T}$ to be $(0.2, 0.43, 0.5, 0.57, 0.8)$. In addition, we assumed that the extra OD information from the mobile sensor data contains the partial set of OD trips $(X^{trk}_1, ..., X^{trk}_{12})^\text{T} = (5, 5, 2, 4, 5, 5, 8, 9, 5, 6, 10, 12)^\text{T}$ and the corresponding partial route counts $\boldsymbol{X}^{trk} = (1, 4, 2, 3, 1, 1, 4, 3, 2, 4, 1, 1, 3, 5, 6, 10, 12)^\text{T}$, and that the real penetration rates of mobile sensors for the 12 OD pairs are $\boldsymbol{w}^* = (0.5, 0.4, 0.15, 0.2, 0.15, 0.15, 0.2, 0.35, 0.2, 0.2, 0.3, 0.3)^\text{T}$. The procedure used for the posterior simulation is depicted in Algorithm 2.

The outputs obtained prior to execution of line 10 in Algorithm 2 are $(X^*_1, ..., X^*_{12})^\text{T} = (5, 7, 12, 16, 28, 30, 5, 8, 18, 24, 27, 26)^\text{T}$, $\boldsymbol{X}^* = (1, 4, 3, 4, 6, 6, 16, 16, 12, 24, 6, 5, 8, 18, 24, 27, 26)^\text{T}$, $\boldsymbol{X}^{not} = (0, 0, 1, 1, 5, 5, 12, 13, 10, 10, 20, 5, 4, 5, 13, 18, 17, 14)^\text{T}$, $\boldsymbol{Y} = (11, 81, 49, 32, 68, 60, 77)^\text{T}$, $\boldsymbol{Y}^{trk} = (5, 17, 13, 8, 24, 12, 28)^\text{T}$, and $\boldsymbol{Y}^{not} = (6, 64, 36, 24, 60, 48, 49)^\text{T}$. The column reordering step (line 10 in Algorithm 2) can be done by

QR decomposition of matrix A. We used the "qr()" function in MATLAB to do this automatically, and we obtained the re-permuted column indices for A as 6, 15, 10, 2, 11, 12, 9, 8, 7, 3, 5, 1, 13, 14, 4, 16, 17; the first 7 column indices correspond to A_1 (and X_1), and the other 10 correspond to A_2 (and X_2).

Algorithm 2 Procedure of our numerical experiment

1: **for** $a = 1, \ldots, c$ **do**
2:     Simulate true OD rates $X_a^* \sim \text{Poisson}(\lambda_a^*)$
3: **for** $a = 1, \ldots, c$ **do**
4:     **for** $t = 1, \ldots, k_a$ **do**
5:         Compute the true route counts: $X_{a,t}^* \leftarrow p_{a,t}^* \, X_a^*$
6:         $X_{a,t}^* \leftarrow \max\{\text{round}(X_{a,t}^*), X_{a,t}^{trk}\}$ //so that $X_{a,t}^{not} \geq 0$
7: Compute the unobserved route counts: $X^{not} \leftarrow X^* - X^{trk}$
8: Compute the observed link counts: $Y \leftarrow AX^*$
9: $Y^{trk} \leftarrow AX^{trk}$, $Y^{not} \leftarrow Y - Y^{trk}$
10: Reorder the columns of the routing matrix: $A \leftarrow (A_1, A_2)$
11: Reorder the vector accordingly: $X^{not} \leftarrow (X_1^T, X_2^T)^T$
12: Choose K
13: **Run** Algorithm 1

We re-initialized and re-ran the MCMC simulation from several possible starting points to validate that convergence was rapidly achieved. Across all M–H variants, the rate of acceptance is between 25 and 40%. We also looked at the trace plots and basic MCMC diagnostics to assure convergence. The auto correlations were monitored and thinning of 20 was chosen. We summarized the analysis by running a total of $K = 50,000$ iterations, including burn-in of 2000, and estimated the posterior based on 2400 samples. We compared two scenarios: (1) no mobile sensor data considered and non-informative priors used; (2) mobile sensor data considered and informative priors (based on those data) used.

6.1 Results with no mobile sensor data

In this case, inference of w was not needed. We assumed uniform (non-informative) priors for Λ and p. Figures 2(a), 2(b), and 2(c) show the marginal posterior distributions of the 2400 posterior samples of Λ, X and p, respectively. For comparison, the dashed curves in Figure 2(a) show the posteriors of the OD Poisson rates $\{\lambda_a\}$ conditional on the actual OD counts $\{X_a^*\}$, the underlying actual values $\{X_a^*\}$ are indicated with triangles on the lower horizontal axis in Figure 2(b); and the actual route-choice probabilities are indicated with dashed vertical lines in Figure 2(c). In Figure 2(a), we see that the simulated posteriors for the Poisson rates are much more diffuse than those for the actual OD counts, and that most of the modes are not close to the corresponding modes of the conditional distributions that used the actual $\{X_a^*\}$. The estimated results for most of the OD counts are satisfactory, only X_4 and X_{12} are significantly overestimated, while X_{10}

and X_{11} are considerably underestimated. The full marginal uncertainties associated with the OD flows are reflected in the histograms in Figure 2(b). Figure 2(c) also shows considerable uncertainty associated with the route-choice probabilities, although the posterior modes of p_1 and p_4 are quite close to their actual values, and their posterior densities are rather high for a wide region near the modes.

6.2 Results with mobile sensor data

In this case, we used informative priors for Poisson rates Λ, the penetration rates w and the route-choice probabilities p. Specifically, we used the partial set of observations of the route counts to build our initial belief about the underlying parameters. In particular, we chose to estimate the OD-specific prior means of the penetration rates as $z^0 = (0.33, 0.30, 0.25, 0.23, 0.23, 0.22, 0.25, 0.30, 0.28, 0.28, 0.33, 0.36)^T$, and the routes-usage fractions as $f = (f_{11}, f_{21}, f_{31}, f_{51}, f_{61})^T = (0.2, 0.4, 0.5, 0.6, 0.8)^T$, which served as the prior means for the route-choice probabilities p. It can be seen that most of the elements of z^0 are reasonably close to the actual values of the elements of w, but the variation across OD pairs was smaller than w^* because of our use of equation (12). Nevertheless, f turned out to be quite close to the actual route-choice probabilities, which demonstrates the value of including the extra information in the first place: it helps us construct good informative priors. For the prior gamma distribution for λ_a, we chose the shape parameter bd_a^0 and rate parameter b for some $b > 0$, where $d_a^0 = \max(1, X_a^{trk}) / z_a^0$. For the prior Dirichlet distribution for p_a, we chose its tth concentration parameter as $\alpha_{a,t}^0 = gX_{a,t}^{trk}$. This led to the prior means of the probabilities in p to be the route-choice fractions in f estimated from the mobile sensor data. The smaller the values of b and g, the more diffuse the priors. In parallel with the first scenario, Figures 3(a), 3(b), and 3(c) show the simulated posteriors of Λ, X and p respectively, with $b, g = 0.5$.

As is clear from Figure 3, the posterior modes are almost all very close to the actual values, and the posterior distributions are more concentrated around the modes, resulting in much better inference compared to the non-informative case. If smaller values of b and g are used, the posterior distributions will be flatter, and a larger discrepancy between the posterior mode of X_a and the actual value X_a^* or between the posterior mode of λ_a and the conditional posterior of λ_a under the actual X_a^*, will be detected. However, the results are still considerably better than the first scenario according to our analysis in which b and g were both set to 0.1 and both set to 0.05.
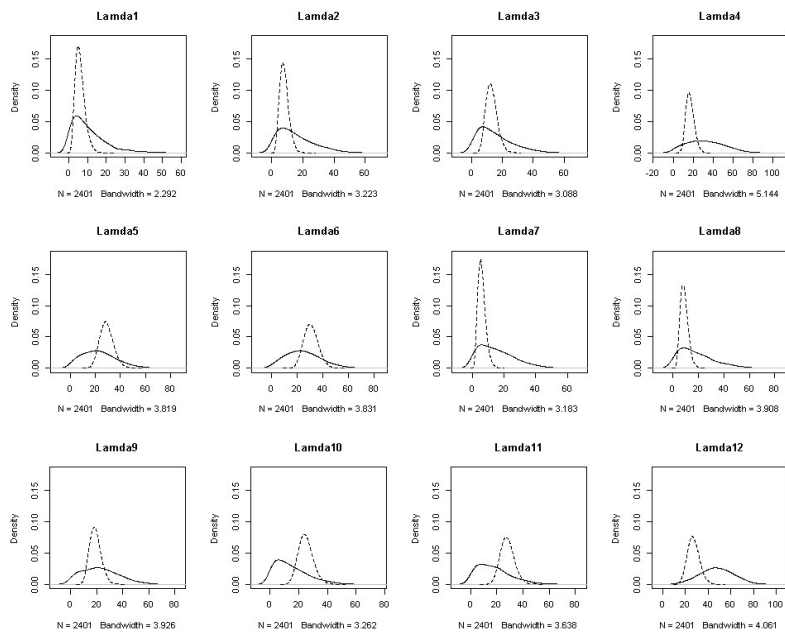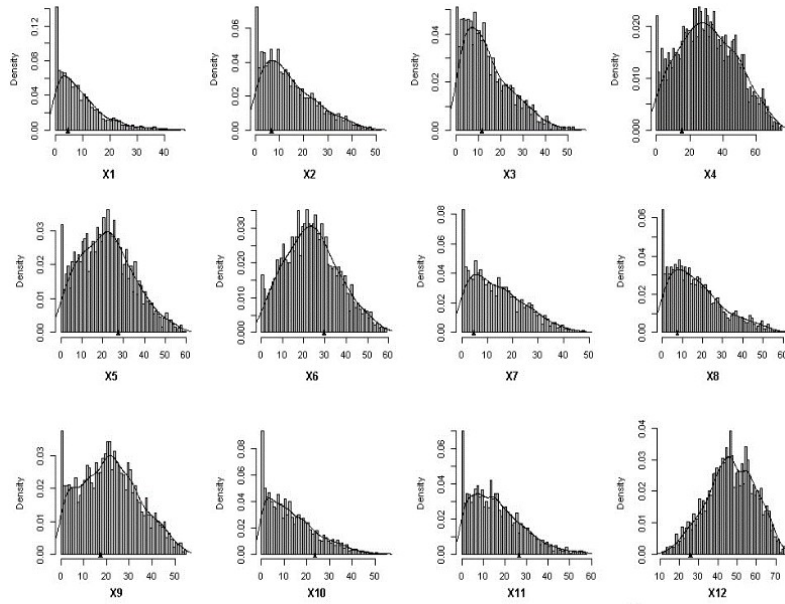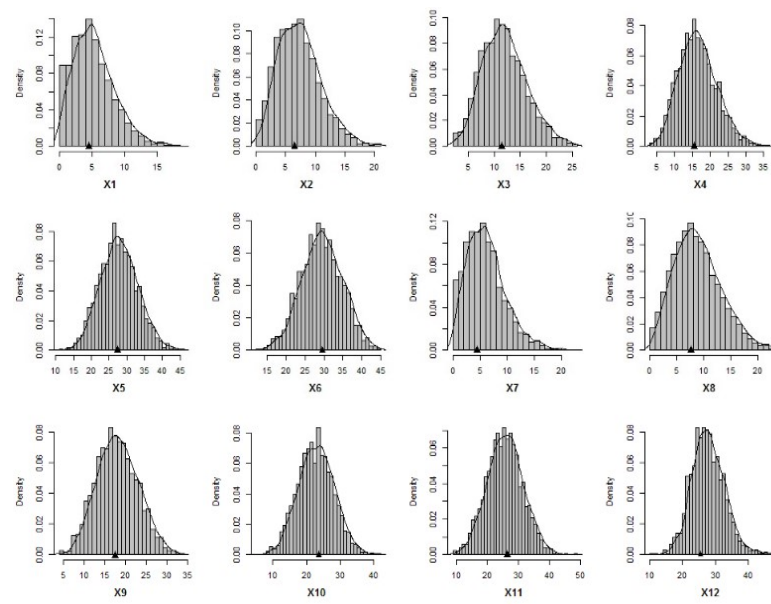
(a). Posteriors of Poisson rates λ_a. Dashed lines are the conditional posteriors were the actual X_a are known.
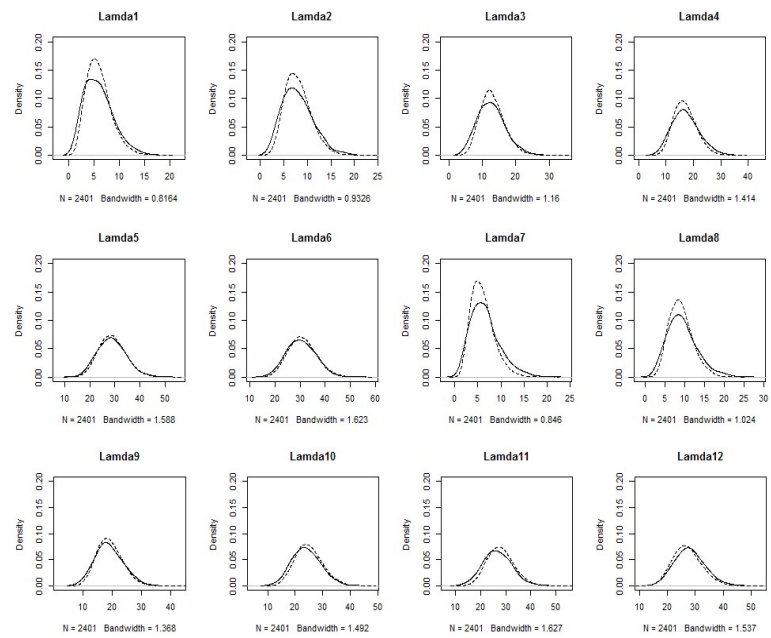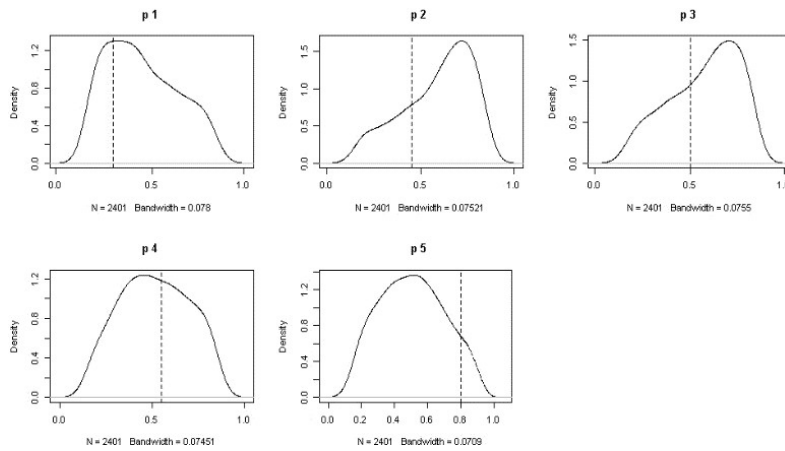
(a). Posteriors of rates λ_a. Dashed lines are the conditional posteriors were the actual X_a are known.

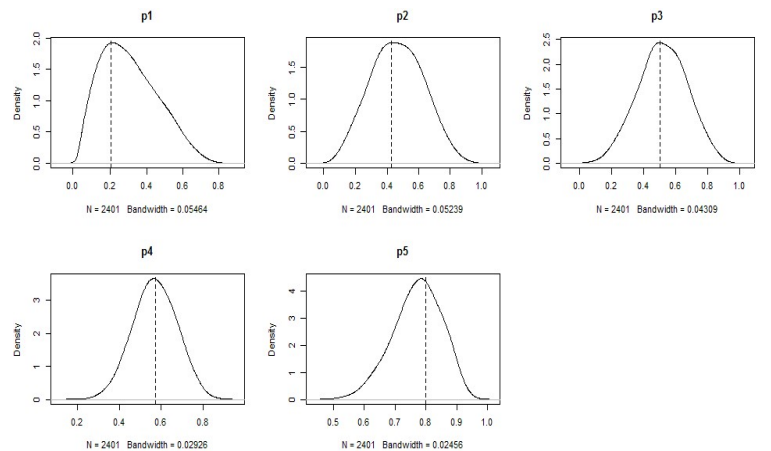(b). Posteriors of partial OD counts X_a. The actual values are indicated by the triangles on the horizontal axes.

(b). Posteriors of partial OD counts X_a. The actual values are indicated by the triangles on the horizontal axes.

(c). Posteriors of $\boldsymbol{p} = (p_{11}, p_{21}, p_{31}, p_{51}, p_{61})^\mathsf{T}$. The dashed lines correspond to the actual values.

Figure 2. Results (no mobile sensor data)

(c). Posteriors of $\boldsymbol{p} = (p_{11}, p_{21}, p_{31}, p_{51}, p_{61})^\mathsf{T}$. The dashed lines correspond to the actual values.

Figure 3. Results (with mobile sensor data)

7. Conclusion

We presented a Bayesian inference model for network flows based on both the link count and a partial set of OD rates and route counts from mobile sensor data. We incorporate inference models for penetration rates and route choice probabilities in addition to OD rates, which is a new contribution to the whole framework. We propose practical procedures for building prior beliefs about the underlying parameters based on mobile sensor data and their market penetrations. According to our numerical study, under the informative priors constructed on the basis of mobile sensor data, the posterior modes of the OD rates and route choice probabilities are almost all very close to their true values. The inference results are more accurate with smaller posterior uncertainties compared to the case where no partial set of routing data and hence only a simple prior is used. We are now experimenting on larger networks using realistic data sets. Computational efficiency is being examined empirically for the scalability of the proposed method.

8. References

[1] A.R. Pitombeira-Neto, C. Loureiro, and L.E. Carvalho, "Bayesian Inference on Dynamic Linear Models of Day-to-Day OD Flows in Transportation Networks", arXiv preprint arXiv:1608.06682, 2016.

[2] P. Cao, T. Miwa, T. Yamamoto, et al. "Bilevel Generalized Least Squares Estimation of Dynamic Origin-Destination Matrix for Urban Network with Probe Vehicle Data", Jour. of the Trans. Res. Board 2333, 8, Aug. 2013, pp. 66–73.

[3] M.J. Maher, "Inference on Trip Matrices from Observations on Link Volumes: a Bayesian Statistical Approach", Trans. Res. Part B, 17(6), Dec. 1983, pp. 435–447.

[4] H.J. Van Zuylen and L.G. Willumsen, "The Most Likely Trip Matrix Estimated from Traffic Counts", Trans. Res. Part B, 14(3), Sep. 1980, pp. 281–293.

[5] Y. Vardi, "Network tomography: Estimating Source-Destination Traffic Intensities from Link Data", J. Amer. Stat. Asso., 91(433), Mar. 1996, pp. 365–377.

[6] C. Tebaldi and M. West, "Bayesian Inference on Network Traffic using Link Count Data", J. Amer. Stat. Asso., 93(442), Jan. 1998, pp. 557–573.

[7] B. Li, "Bayesian Inference for Origin-Destination Matrices of Transport Networks using the EM Algorithm", Technometrics, 47(4), Nov. 2005, pp. 399-408.

[8] K. Parry, and M. L. Hazelton, "Estimation of Origin–Destination Matrices from Link Counts and Sporadic Routing Data", Trans. Res. Part B, 46(1), Jan. 2012, pp. 175–188.

[9] K. Parry and M.L. Hazelton. "Bayesian Inference for Day-to-Day Dynamic Traffic Models", Trans. Res. Part B, 50, Apr. 2013, pp. 104–115.

[10] M. L. Hazelton, "Statistical Inference for Time Varying Origin–Destination Matrices", Trans. Res. B, 42(6), Jul. 2008, pp. 542–552.

[11] L. Cheng, S. Zhu, Z. Chu, et al., "A Bayesian Network Model for Origin-Destination Matrices Estimation Using Prior and Some Observed Link Flows", Dis. Dyn. in Nat. and Soc., 2014, Apr. 2014, pp. 1–9.

[12] M. S. Iqbal, C. F. Choudhury, P. Wang, et al., "Development of Origin-Destination Matrices using Mobile Phone Call Data", Trans. Res. Part C, 40, Mar. 2014, pp. 63–74.

[13] E. Castillo, J. M. Menéndez, and P. Jiménez, "Trip Matrix and Path Flow Reconstruction and Estimation based on Plate Scanning and Link Observations", Trans. Res. Part B, 42(5), Jun. 2008, pp. 455–481.

[14] M. Van Aerde, B. Hellinga, L. Yu, et al., "Vehicle Probes as Real-Time ATMS Sources of Dynamic O–D and Travel Time Data", Large Urban Systems—Proceedings of the ATMS Conference, St. Petersburg, Fla., 1993, pp. 207–230.

[15] L. Eadicicco, "Uber Is Tracking Drivers' Phones to Watch for Speeding", Jun. 2016, http://time.com/4387031/uber-driver-app-tracking.

[16] J. Cui, F. Liu, J. Hu, et al., "Identifying Mismatch between Urban Travel Demand and Transport Network Services using GPS Data: A Case Study in the Fast Growing Chinese City of Harbin", Neurocomputing, 181(C), Mar. 2016, pp. 4–18.

[17] Y. Sheffi, Urban Transportation Network: Equilibrium Analysis with Mathematical Programming Method. Prentice-Hall, Upper Saddle River, NJ, 1985.

[18] L. Carvalho, "A Bayesian Statistical Approach for inference on static Origin-Destination Matrices in Transportation Studies", Technometrics, 56(2), May 2014, pp 225–237.

[19] A. Gelman, J. Carlin, H. Stern, et al., Bayesian Data Analysis (3rd edition). Chapman & Hall, London, UK, 2013.