Real-Time Estimation of Origin-Destination Matrices with Partial Trajectories from Electronic Toll Collection Tag Data

Jaimyoung Kwon*
Department of Statistics
California State University
Hayward, CA 94542
Tel: (510) 885-3447, Fax: (510) 885-4714
jkwon@csuhayward.edu

Pravin Varaiya
Department of Electrical Engineering and Computer Science
University of California, Berkeley CA 94720
Tel: (510) 642-5270, Fax: (510) 642-7615
varaiya@eecs.berkeley.edu

For Presentation and Publication
84th Annual Meeting
Transportation Research Board
January 2005
Washington, D.C.

March 14, 2005

Words: 4,200 (excluding Figure and Table captions)
Plus 4 Table and 4 Figures (2,000)
TOTAL: 6,200

*Corresponding Author

ABSTRACT

The origin-destination (O-D) matrix of a traffic network is usually estimated from link traffic counts combined with a sample survey. Partially observed vehicle trajectories obtained with vehicle re-identification or automatic vehicle identification (AVI) techniques such as electronic tags provide a new data source for real-time O-D matrix estimation. However, due to incomplete sampling, accurate estimation of O-D matrices from these data is not trivial. We develop a statistical model for such data and derive an unbiased estimator of the O-D matrix based on the method of moments. Further exploiting the sound statistical model, we also develop the bootstrap standard error estimate of the O-D matrix estimator. The algorithm can be computed quickly and performs very well under simulation, compared with simpler estimators. Applied to data from vehicles with electronic toll collection tags in the San Francisco Bay Area, the algorithm produces realistic time series of the hourly O-D matrix. We also study the relationship of the proposed estimator with the similar methods in the literature and discuss the extension of the methods to general, more complex networks.

Keywords: real-time O-D matrix estimation; automatic vehicle identification (AVI); electronic toll collection tags; method of moments estimation

1. INTRODUCTION

Knowledge of the origin-destination (O-D) matrix of a traffic network is useful for various planning and operations tasks. The matrix gives the number of vehicles traveling between different zones in a region. Traditionally, the O-D matrix is estimated from link traffic counts combined with a sample survey. Approaches to O-D matrix estimation using these data include network equilibrium, gravity, and distribution-assignment models (*1*).

Vehicle locations observed using vehicle re-identification or automatic vehicle identification (AVI) techniques provide a new data source for real-time O-D matrix estimation. Examples of anonymous vehicle tracking technologies based on vehicle re-identification include loop detectors with high speed scanning detector cards that generate inductive signatures (*2*), laser-based detection systems that provide vehicle lengths (*3*), and the video image processing technology that generate video-based vehicle signatures (*4*). These technologies can cover only short roadway sections and vehicles are identified indirectly from their physical features, thus data from them are of limited use for O-D matrix estimation. More extensive coverage and higher accuracy are achieved by AVI techniques, where individual vehicles are uniquely identified and can be detected more accurately. AVI may be based on such mechanisms as a transponder (or a 'tag') on the vehicle or automated license plate readers, among others. Currently, these technologies are mainly used for Electronic Toll Collection (ETC) and traffic law enforcement. Examples of ETC system deployment include the E-Z pass in the East Coast (*5*) and FasTrak in the San Francisco Bay Area (*6*). We will focus on data from tag-based AVI techniques in the current paper, though the extension of the methods to data from other AVI mechanisms or those from vehicle re-identification techniques is straightforward.

Using a tag-based AVI mechanism, one can obtain partial trip trajectories of vehicles equipped with a tag by installing tag readers or antennas at various locations in a region. The trajectories are "partial" since not all vehicles are equipped with a tag and not all vehicles with a tag are detected. Such data turned out to be a useful source for real-time travel time measurement and have been utilized with great success in such systems as (*7*) and (*8*), among others. But there has been relatively little work on utilization of such data for real-time O-D estimation, partly due to the difficulty arising from the incomplete sampling and also due to the lack of a realistic testbed.

Past works on O-D estimation from AVI data include that of Van der Zijpp (*9*), who studied a larger problem of combining AVI data with more traditional link-volume data for real-time estimation of the O-D matrix. Even though the study was very extensive, the method was not applied to real data and only simulation results were reported. Asakura et al. (*10*) developed a method that is basically identical to a subset of those studied in (*9*) and applied it to data from Han-Shin expressway obtained with the video-based license plate recognition technique. But they applied it to a very small linear subsection of the large freeway network, bypassing many difficult issues arising from the complex network topology.

This paper formulates the data in a statistically sound model and derives an unbiased OD matrix estimator based on the method of moments estimator from such data. Also, various conditions one needs to impose on the freeway network topology will be laid out explicitly in graph theory terminology, which will pave the way for future extension of the proposed methods to larger, more complex freeway networks. The bootstrap estimate of the standard error of the O-D matrix estimator is also developed, further exploiting the sound statistical model. Finally, we apply the algorithm to the data from a real, non-linear freeway network equipped with antennas that read ETC tags.

The rest of the paper is organized as follows. Section 2 briefly describes the ETC tag reader system deployed in the Bay Area and points out the sampling problems it creates. Section 3 describes the statistical method in detail and the directed graph representation of freeway networks. The relationship between the proposed method and previous studies in (*10*) and (*9*) are also discussed. Section 4 investigates the performance of the proposed algorithm using simulated data and a real data set from San Francisco Bay Area. Section 5 discusses the possible extension of the proposed algorithm to general, more complex networks and concludes the paper.

2. BAY AREA ETC TAG READER SYSTEM

The FasTrak system operates according to the Caltrans Title 21 standard (*11*). The system has two elements, a reader and a transponder. The reader is mounted on a pole; the transponder is mounted on the vehicle windshield.

The reader transmits a RF trigger pulse to turn the transponder on. After a short time delay, the reader transmits an encoded, polling signal, which, upon detection and decoding by the transponder, provides initial information to the transponder including the type of transaction the reader wishes to conduct.

The reader then transmits an unmodulated CW (continuous wave) RF signal for the transponder to modulate with a reply data message, including the transponder ID, while backscattering to the reader. The reader then transmits an encoded acknowledge message and requests that the transponder not respond to the same polling message again for 10 seconds (*11*).

A reader system serves four transponder-reading locations, each with its own antenna (*12*). In the Bay Area deployment, all four antennas are mounted on the same sign structure, located in the freeway divider. Three of the four antennas point in one freeway direction—the major direction, while the fourth antenna points in the opposite direction—the minor direction. The area covered in the minor direction, and the number of detected vehicles, is consequently smaller.

The antennas are mounted at a height of between 16 and 24 ft, unlike at the toll booth where the antenna is close to the vehicle. As a result, only a fraction of tagged vehicles are detected.

The passage of a tagged vehicle generates a record, comprising a scrambled version of the tag (to protect privacy) and a sequence of reader locations where the vehicle is detected and the corresponding time stamp. The resulting data set is incomplete for two reasons. First, the *penetration rate*, the fraction of vehicles that carry tags, is between 15 and 40 percent, depending on the time of day. Second, the *detection rate*, the fraction of tagged vehicles that are detected by a reader is about 80% in the major direction and 40% in the minor direction. Consequently, if a tagged vehicle passes under several readers, only some of these may detect it. A proper statistical analysis must take into account the incompleteness of the data.

2. METHOD

Consider a region comprising several *directional locations* or 'nodes' on roadways (freeways or arterials), indexed $j = 1,\ldots,J$. Note that it is assumed that we have separate detection for vehicles that travel in one direction of the roadway and for those that travel in the other direction. Such data are obtained when two groups of antennas are installed right next to each other pointing in opposite directions, such as in the Bay Area ETC tag reader system. Thus we need to use caution to interpret a directional location as a "zone" in the region since trips from a single "zone" will show up in different directional locations depending on the direction of the trip. Also note that the method developed here is applicable, with little modification, to those situations in which the direction of individual trips is not known.

For two nodes j and k, write $j \rightarrow k$ if j is immediately upstream of k. This relation defines an *edge* in the freeway network viewed as a directed graph (*13*). An *O-D pair* is any pair $[j,k]$, sometimes written $[jk]$. A vehicle has the O-D pair $[jk]$ if it makes a trip that begins before j (but not before any other nodes that are considered) and ends after k (but not after any other nodes). Note that $j = k$ is also allowed, in which case we simply write $[j]$. The meaning of this is obvious from the previous interpretation: the vehicle makes a trip that begins (just) before j and ends right after j.

A *path* is a sequence of nodes $(j_1,\ldots, j_A)$ connected by edges. An O-D pair $[j, k]$ is *traversable* if a path exists starting from j and ending at k. It is *uniquely traversable* if such path is unique. A path *contains* another path if the latter is included in the former. An O-D pair *contains* another O-D pair if any path that contains the latter is contained by a path for the former. We assume the freeway network satisfies the following requirements:

1. The graph is connected, i.e., there is a path connecting every pair of nodes;
2. An O-D pair is uniquely traversable if it is traversable;
3. The graph has no *cycle*, i.e., a path with more than one edge that starts and ends at the same node.

Without these assumptions, the model formulation and methods to find the solution become very complicated. We discuss how to relax these requirements in Section 4.

Index time periods, say hours of the day, by $t = 1,\ldots,T$. Denote by $N(t)$ the total number of vehicles that make a trip during t. Index vehicles by $i = 1,\ldots, N(t)$. These vehicles are grouped into disjoint sets of vehicles corresponding to different O-D pairs. The O-D volume between nodes j and k during period t is denoted by $N_{jk}(t)$. At t the O-D matrix is simply $\{N_{jk}(t),\ j,\ k = 1,\ldots, J\}$. If a fixed time period t is considered, we drop the time index and write $N_{jk}(t)$ simply as N_{jk} or $N[jk]$.

Let ψ be *penetration probability*, i.e. the probability that an individual vehicle is equipped with a tag. Let π_j be the probability that a car with a tag traveling the link j is detected by the tag reader. We assume that for every vehicle, regardless of their O-D, the event that it is equipped with a tag and the event that it is detected are independent. Although our narrative is in terms of electronic tags, a similar formulation applies with other vehicle reidentification techniques.

For an O-D pair $[j,k]$, we define

$$M[jk] = (\text{Number of cars detected at } j \text{ and } k \text{ but not before } j \text{ nor after } k). \tag{1}$$

Note that terms like 'before (after) a zone' are well defined thanks to the requirements above on the freeway network. We can rewrite $M[jk]$ as the sum

$$\begin{aligned} M[jk] = \textstyle\sum\{\sum\{1(&\text{ Car } i \text{ has O-D path } [l,m] \\ &\text{and vehicle } i \text{ is equipped with a tag} \\ &\text{and vehicle } i \text{ is detected at } j \text{ and } k \text{ but not before } j \text{ nor after } k) \\ &: \text{over all vehicles } i\} \\ &: \text{over all O-D pairs } [l,m] \text{ that contain } [j,k]\}. \end{aligned} \tag{2}$$

The notation '1(·)' means the indicator function. Although the accounting identity (2) appears complicated, it allows the computation of various moments in explicit form. The expectation is given by

$$E(M[jk]) = \textstyle\sum\{N_{lm}\, \psi\, p_{l(jk)m} : \text{over all O-D pairs } [l,m] \text{ that contain } [j,k]\}. \tag{3}$$

Here

$$p_{l(jk)m} = (1-\pi_l)\ldots(1-\pi_{j-1})\, \pi_j \pi_k\, (1-\pi_{k+1})\ldots(1-\pi_m) \tag{4}$$

is the probability that a car with path $[l,m]$ is detected at link j and k but not before j nor after k. The detection probability is assumed to be given. Similarly, the elements of the variance-covariance matrix of the elements of M are

$$\begin{aligned} \text{Cov}(M[j,k], M[j',k']) = \textstyle\sum\{&N_{lm}\, \psi^2\, [1([jk]=[j'k'])\, p_{l(jk)m} - p_{l(jk)m}\, p_{l(j'k')m}] \\ &: \text{over all O-D pairs } [l,m] \text{ that contain both } [j,k] \text{ and } [j',k']\}. \end{aligned} \tag{5}$$

Define the vectors

$$\boldsymbol{M} = (M[jk] \text{ for all O-D pairs } [jk]), \tag{6}$$
$$\boldsymbol{N} = (N[j,k] \text{ for all O-D pairs } [jk]). \tag{7}$$

Note that the elements in both vectors are arranged in the same order as the index $[jk]$. In other words,

$$\boldsymbol{M}_{\theta[jk]} = M[jk] \text{ and } \boldsymbol{N}_{\theta[jk]} = N[jk], \tag{8}$$

in which $\theta[jk]$ is the order corresponding to the O-D $[jk]$.

For large number of vehicles N_{jk}, by the Central Limit Theorem, $\boldsymbol{M}$ approximately has a multivariate normal distribution,

$$\boldsymbol{M} \sim \mathrm{N}(\boldsymbol{AN}, \Sigma(\boldsymbol{N})), \tag{9}$$

in which the matrix $\boldsymbol{A}$ and the function $\Sigma(\cdot)$ are specified by (3) and (5) respectively. More precisely,

$$\boldsymbol{A}_{[jk],[lm]} = p_{l(jk)m} 1([l, m] \text{ contains } [j, k]), \tag{10}$$

and $\Sigma(\boldsymbol{N})$ is the matrix whose $\theta[jk]$th row and $\theta[j'k']$th column element is $Cov(M[j, k], M[j', k'])$. Since

$$\mathrm{E}(\boldsymbol{M}) = \boldsymbol{AN}, \tag{9}$$

the method of moment estimator (*14*) of the unknown $\boldsymbol{N}$ is given by

$$\boldsymbol{\eta} = \boldsymbol{A}^{-1}\boldsymbol{M}. \tag{10}$$

Note that the same formulas as (4) and (12) were independently proposed by (*9*) (their equation (24)) and (*10*) (their equation (2)), although neither authors made it clear that it is the method of moment estimator. The multinomial model is approximated by a multivariate normal distribution in the Kalman filter context in (*9*) and the issue of estimation of the standard error of such an estimator is not considered in detail. The other work of (*10*) lacked rigorous statistical formulation.

The approximate error distribution of the estimator $\boldsymbol{\eta}$ can be obtained from (9)-(12) as

$$\boldsymbol{\eta} \sim \mathrm{N}(\boldsymbol{N}, \boldsymbol{A}^{-1}\Sigma(\boldsymbol{N})\,(\boldsymbol{A}^{-1})^{\mathrm{T}}).$$

Computing $\Sigma(\boldsymbol{N})$ is complicated, and so we use the bootstrap procedure (*15*) to estimate the error of the estimator. Given an estimate $\boldsymbol{\eta}$, one simulates bootstrap samples $\boldsymbol{M}^{*(b)}$, b=1,…,B assuming $\boldsymbol{\eta}$ to be the true parameter. For each of those bootstrap samples, one estimates $\boldsymbol{\eta}^{*(b)}$ again and uses the bias and standard error of the bootstrap sample as an estimate of the bias and standard error of the bootstrap sample as an estimate of the bias and standard error of the estimator itself.

4. ANALYSIS

We apply the algorithm to a simulated data set as well as to a real ETC tag data from the San Francisco Bay Area.

Simulation

Consider a region with three directional locations j = 1, 2, 3. The freeway network is defined by edges 1→2 and 2→3. The six O-D pairs are listed in Table 1, together with the paths containing them. Even for this simple network, all O-D pairs except [1, 3] are contained in multiple paths.

Assume the detection probability is 0.5 for all three readers, i.e. $\pi_1=\pi_2=\pi_3=0.5$ and N_{jk}=1,000 for all O-D [j,k]. Also assume all vehicles are equipped with a tag, so the penetration rate ψ=1. We simulate $M[j,k]$ as the sum of binomial random variables with distribution bin(N_{lm}, ψp_{lm}), summed over all O-D pair $[l,m]$ containing the path $[j,k]$.

For comparison, we also compute the naïve estimator

$$\eta'_{jk} = (\text{ Number of vehicles that are observed at both } j \text{ and } k) / \pi_j \pi_k \,. \quad (11)$$

The naive estimator gives a rough idea about the number of vehicles that travel between two locations, but it is obviously incorrect for estimation of ***N***.

Figure 1 and Table 2 show the distribution of the estimates from the two methods estimated with 200 simulation runs. Moment estimators are very accurate, exhibiting less than 1% bias for all O-D pairs. The relative errors are less than 10%. As expected, the naive estimator performs very poorly by comparison. Except for O-D pair [1, 3], it has huge biases. On the other hand, their standard errors are of similar magnitude as those of the moment estimators. To summarize, a naive approach cannot be used for analysis of O-D patterns, while the moment estimator provides very accurate estimates for all O-D pairs.

Bay Area ETC Data

The left plot in Figure 2 shows the topology of the freeway network and reader system locations, which can be viewed as zone labels. The data were collected over 24-hours of Thursday, 24 July, 2003. The current reader coverage is incomplete. It only includes the northern part of East Bay and a small part of San Francisco. Thus, for example, ‘zone 2’ includes not only the area between zones 3 and 9 but also the huge area that lies beyond 580 West. The reader coverage is being extended.

The graph representation in the right plot of Figure 2 shows all the nodes and edges. Note there are two directional locations for each reader system. In the graph, if we allow trips between nodes 7 (east and west) and 5 (north and south), denoted by arrows in dashed lines, there are cycles (7E→5N→3E and 7W→3W→5S), so we do not allow these trips. For illustration, we consider only eastbound/southbound traffic involving nodes 1E, 3E, 2E, 9E, 5S, and 7E (gray nodes). The graph is then completely described by five edges: 1→5, 1→3, 3→7, 3→2, and 2→9. There are J = 6 nodes and 15 traversable O-Ds. For each O-D (path), the set of containing O-Ds (paths) is shown in Table 3. Note that the path-containing relationship is much more complicated than the network considered for simulation.

The penetration rate ψ is taken to be 15% and we use this value in the application. The detection rate parameters π_j are estimated using the total volume data obtained from the loop detectors installed at the location close to each antenna. We use the PeMS system (*16*) to obtain the loop data. From these data, we find π_j=0.8 for all the locations.

We first apply the algorithm to estimate the O-D matrix for the whole 24-hour period. Table 4 shows the estimates and the bootstrap estimates of its bias and variance. Most biases are controlled under 1% of the estimates and the relative errors under 5%, except for O-Ds [1, 1] and [3, 2]. Figure 3 shows the bootstrap distribution of the estimates.

We then estimate hourly O-D matrices by splitting samples into hours according to the vehicle's last observed timestamp and applying the algorithm for each hourly sample. Figure 4 shows the result.

The trend clearly reveals the changing pattern of vehicle travel demand for each O-D pair over the day. For example, the O-D [1, 3] (San Francisco to East Bay) and [5] (East Bay to San Jose) have higher demand in the morning while the O-D [2] (East Bay to Richmond) and [9] (Northeast Bay to The Carquinas bridge) have higher demand in the evening.

Computation Times

The computation required for the proposed algorithm only involves solving (12) for a matrix $\boldsymbol{A}$ of a moderate size, which can be done almost instantaneously in modern computers. Computing the bootstrap standard error estimate simply requires repeating the estimation process multiple (usually hundreds) times for re-sampled data and can also be done quickly, usually within seconds. Such fast computation implies that, using the proposed algorithm, an O-D matrix estimate as well as its bootstrap standard error can be made available in real-time for traffic operation and management purposes.

5. DISCUSSIONS AND CONCLUSION

We developed a method of moments algorithm for estimating real-time OD matrices from partially observed vehicle trajectories data, together with a bootstrap procedure to estimate the standard error of the estimator.

We restricted our attention to freeways with relatively simple topology, in which the corresponding graph is connected, all traversable O-Ds are uniquely traversable, and there is no cycle. These requirements are too strict for many real world freeway networks and can be relaxed with varying difficulty. For a disconnected graph, one only needs to decompose it into connected components and apply the algorithm to each component. For non-uniquely traversable O-Ds, the accounting equation (2) needs to be appropriately modified. If the route choice probabilities are known, such modification is relatively straightforward, but if they are unknown, they become extra parameters in the model and the extension is not trivial anymore. When there are cycles, one promising approach would be to remove non-likely paths from consideration to reduce the dimension of $\boldsymbol{M}$ and $\boldsymbol{N}$ and simplify $\boldsymbol{A}$. Overall, these are all interesting challenges and we are currently working on extending the current algorithm to cope with such complex networks.

The proposed algorithm performed well. The simulation study shows that the proposed moment estimator and bootstrap standard errors are accurate under ideal settings. When applied to data from vehicles equipped with ETC transponders and 'readers' installed at various locations in San Francisco Bay Area, it produces credible hourly OD matrices very quickly.

The estimation method proposed here is similar to the existing algorithm previously developed by (*9*) and (*10*). Differences and contributions of the current paper, when compared to the previous works, are the following:

1. More rigorous statistical formulation,
2. Bootstrap standard error estimation,
3. Applicability to freeway networks that are more complicated than linear segments,
4. Explicit specification of the conditions required on the freeway network in graph theory terminology for the algorithm to work, which paves the way for future extension to larger, more complex freeway networks, and
5. Application to a real-world situation using relatively novel data type from ETC tags.

Combined with the wider deployment of tag readers and greater penetration of tags in the vehicle population, the proposed methods would enable district-wide O-D real-time reporting in timely and accurate manner.

5. ACKNOWLEDGEMENT

This study is part of the PeMS project, which is supported by grants from Caltrans to the California PATH Program. We are very grateful to John Wolf of Caltrans for his support. We thank the Metropolitan Transportation Commission and Chao Chen for help in securing the ETC data.

REFERENCES

[1] T. Abrahamsson. Estimation of Origin-Destination Matrices using Traffic Counts-A Literature Survey. Technical Report IR-98021, International Institute for Applied Systems Analysis, 1998. http://netec.mcc.ac.uk/WoPEc/data/Papers/wopiasawpir98021.html, Accessed 7/19/2004.

[2] C. Sun, S.G. Ritchie, W. Tsai, and R. Jayakrishnan. Use of vehicle signature analysis and lexicographic optimization for vehicle reidentification on freeways. *Transportation Research, Part C*, 7:167185, 1999.

[3] Larson, J.E., Van Katwyk, K., Liu, C., Cheng, H.H., Shaw, B., and Palen, J. *A real-time laser-based prototype detection system for measurement of delineations of moving vehicles*. UCB-ITS-PWP-98-20, PATH, University of California, Berkeley, CA, 1998.

[4] C.A. MacCarley. *Video-based vehicle signature analysis and tracking system phase 2: algorithm development and preliminary testing.* UCB-ITS-PWP-98-10, PATH, University of California, Berkeley, CA, 1998.

[5] E-z pass website. http://www.ezpass.com/ , Accessed 11/1/2004.

[6] FasTrak website. http://www.511.org/fastrak/ , Accessed 7/19/2004.

[7] 511.org traffic information website. http://traffic.511.org/ , Accessed 11/1/2004.

[8] Houston real-time traffic map website. http://traffic.houstontranstar.org/ , Accessed 11/1/2004.

[9] N.J. Van der Zijpp. Dynamic od-matrix estimation from traffic counts and automated vehicle identification data. In *Transportation Research Record* 1607, pages 8794, Washington DC, 1997. TRB, National Research Council.
[10] Y. Asakura, E. Hato, and M Kashiwadani. Od matrices estimation model using avi data and its application to the han-shin expressway network. *Transportation*, 27(4):419-438, 2000.
[11] California Department of Transportation. http://www.dot.ca.gov/hq/traffops/elecsys/title21/docs/t21updat.htm , Accessed 7/19/2004.
[12] Sirit Technologies. Automatic Vehicle Identification. Identity Flex Installation Manual.
[13] S. Even. *Graph Algorithms*. Computer Science Press, 1979.
[14] P. Bickel and K.A. Doksum. *Mathematical Statistics: Basis Ideas and Selected Topics*. Prentice-Hall, 2000.
[15] B. Efron and R. J. Tibshirani. *An Introduction to the Bootstrap*. Chapman & Hall/CRC, 1994.
[16] PeMS website. http://pems.eecs.berkeley.edu , Accessed 2/1/2005.

LIST OF TABLES

LIST OF FIGURES

TABLE 1 O-D Path and Paths Containing the O-D Path for the Simulated Linear Freeway

O-D path	***Paths containing the O-D path***
(1)	(1), (1,2), (1,2,3)
(1,2)	(1,2), (1,2,3)
(1,2,3)	(1,2,3)
(2)	(1,2), (1,2,3), (2), (2,3)
(2,3)	(1,2,3), (2,3)
(3)	(1,2,3), (2,3), (3)

TABLE 2 Bias and Standard Error of the Moment and Naive Estimators from Simulation with 200 Runs when True *N*=1,000

	Bias		***SE***	
O-D pair	***Moment Estimator***	***Naive***	***Moment Estimator***	***Naive***
1,1	-1.0	2007.6	56.2	51.8
1,2	2.8	1005.2	76.8	76.5
1,3	-1.6	-3.3	58.6	55.5
2,2	-1.0	3001.7	79.5	59.1
2,3	2.5	997.8	73.9	72.6
3,3	-4.1	2003.0	61.0	51.1

TABLE 3 O-D Path and Paths Containing the O-D Path for the Bay Area Network

O-D path	***Paths containing the O-D path***
(1)	(1), (1,3), (1,3,2), (1,3,2,9), (1,3,7), (1,5)
(1,3)	(1,3), (1,3,2), (1,3,2,9), (1,3,7)
(1,3,2)	(1,3,2), (1,3,2,9)
(1,3,2,9)	(1,3,2,9)
(1,3,7)	(1,3,7)
(1,5)	(1,5)
(2)	(1,3,2), (1,3,2,9), (2), (2,9), (3,2), (3,2,9)
(2,9)	(1,3,2,9), (2,9), (3,2,9)
(3)	(1,3), (1,3,2), (1,3,2,9), (1,3,7), (3), (3,2), (3,2,9), (3,7)
(3,2)	(1,3,2), (1,3,2,9), (3,2), (3,2,9)
(3,2,9)	(1,3,2,9), (3,2,9)
(3,7)	(1,3,7), (3,7)
(5)	(1,5), (5)
(7)	(1,3,7), (3,7), (7)
(9)	(1,3,2,9), (2,9), (3,2,9), (9)

TABLE 4 Estimate of 24-hour O-D Matrix and Standard Errors from Bay Area ETC Data for 24 July, 2003

O-D path	*Moment Estimator (vehicles/day)*	*Bootstrap Bias*	*Bootstrap Relative Bias (Percent)*	*Bootstrap S.E.*	*Bootstrap Relative Error (Percent)*
(1)	1256	10.5	0.8	207.3	16.5
(1,3)	25205	-21.6	-0.1	374.3	1.5
(1,3,2)	4610	-1.1	0	201.6	4.4
(1,3,2,9)	9825	16.6	0.2	215.1	2.2
(1,3,7)	4494	-3.9	-0.1	150.7	3.4
(1,5)	6000	-11.4	-0.2	183.7	3.1
(2)	29679	-5.2	0	361.8	1.2
(2,9)	19006	52.7	0.3	338	1.8
(3)	22218	-21.4	-0.1	342.5	1.5
(3,2)	2990	12.8	0.4	166.2	5.6
(3,2,9)	3848	9.3	0.2	179.6	4.7
(3,7)	3632	2.4	0.1	155.9	4.3
(5)	17575	1.6	0	274.9	1.6
(7)	3564	8.7	0.2	142.9	4
(9)	24621	-41.9	-0.2	351.8	1.4

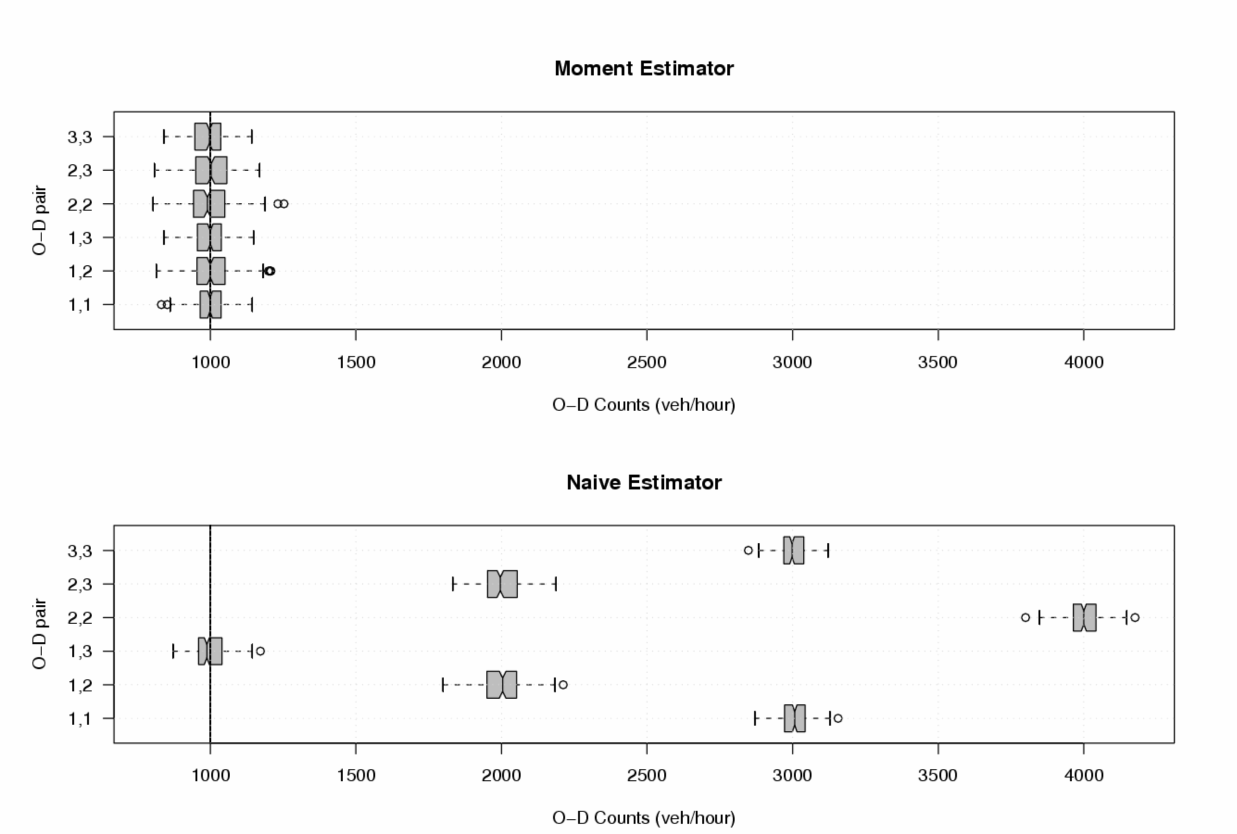

FIGURE 1 Simulation results for method of moment estimator (top) and naive estimator (bottom). For each O-D pair, the distribution of the estimator is shown as a boxplot. The vertical line is the true value N_{jk}=1,000.

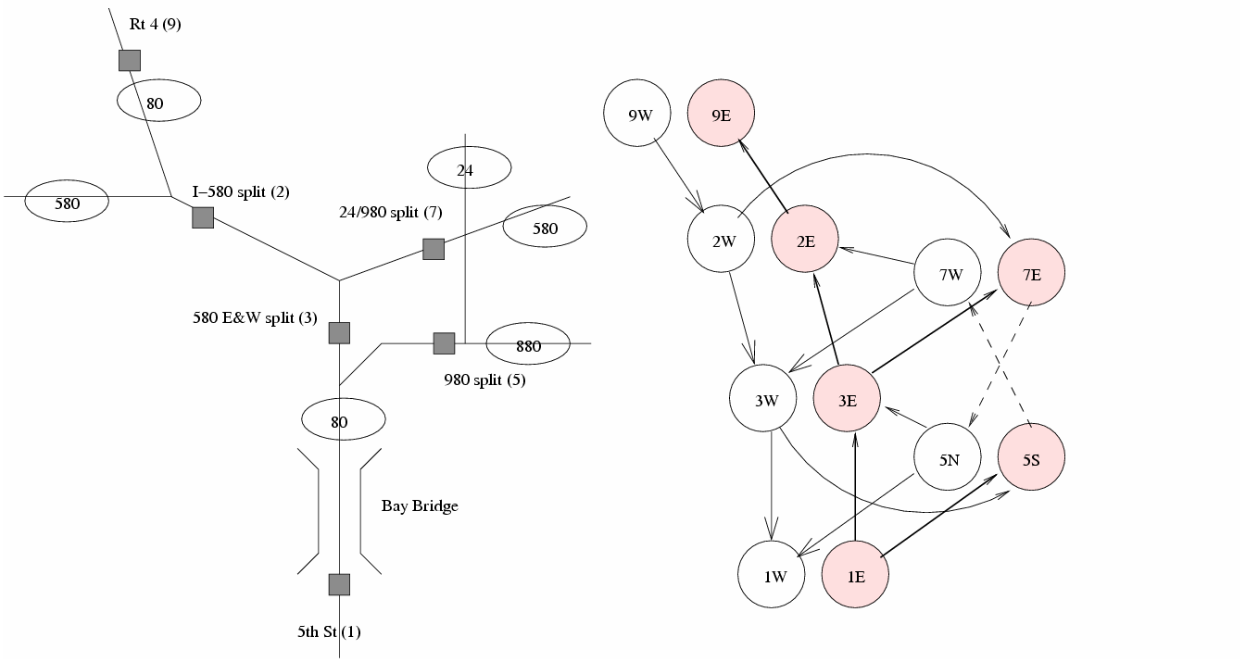

FIGURE 2 Map of the Study Site in San Francisco Bay Area with approximate location of the tag readers (left) and graph representation of the freeway.

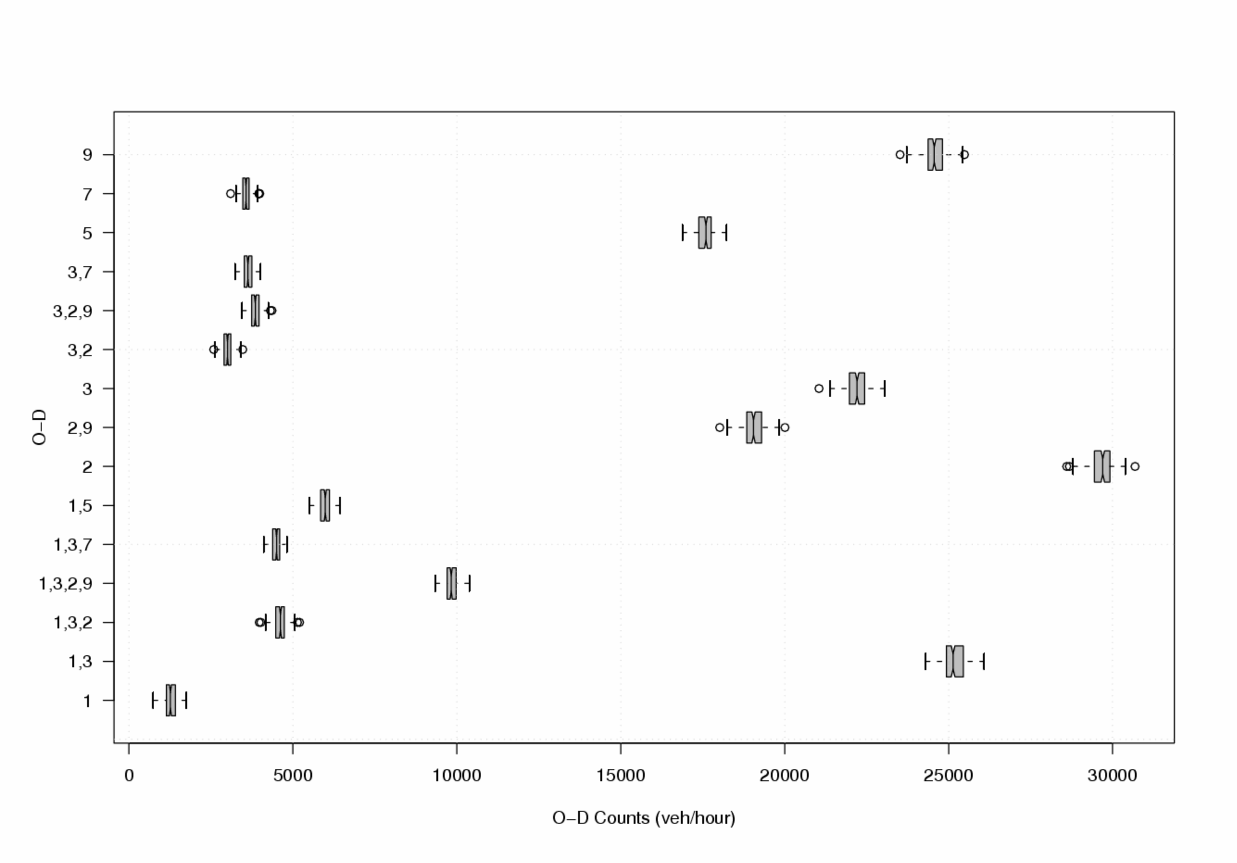

FIGURE 3 Moment estimator and bootstrap distribution of the estimator for Bay Area data.

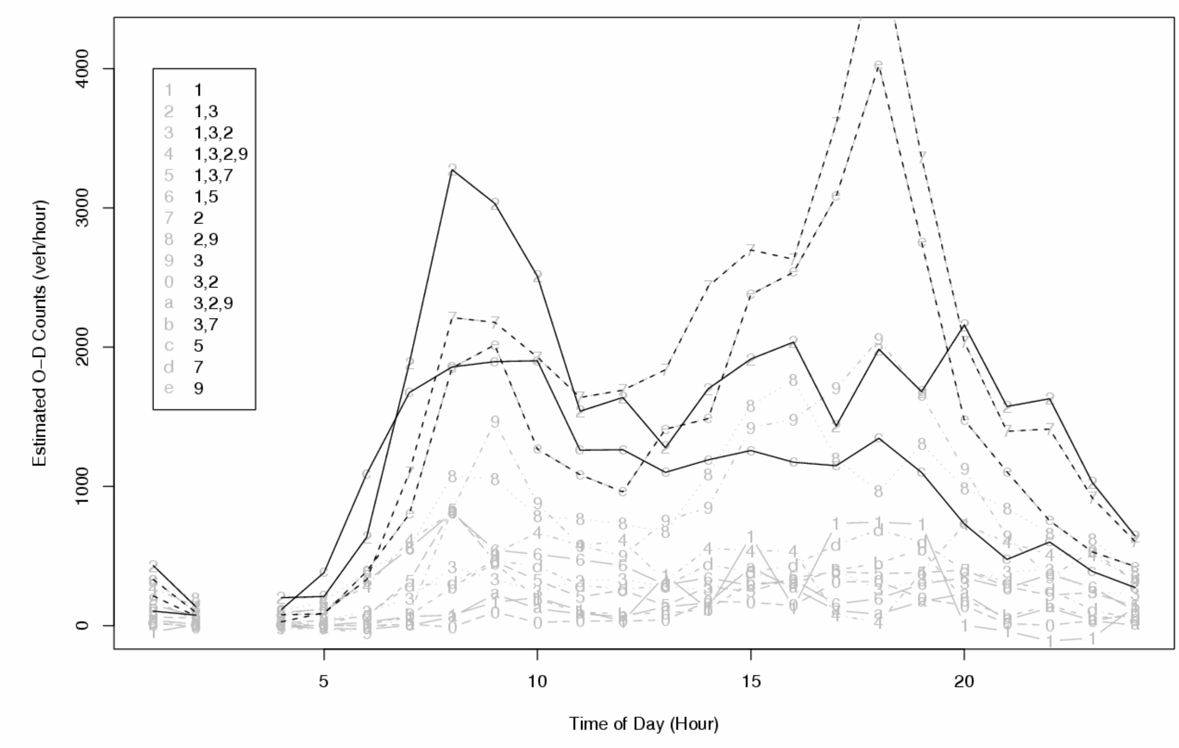

FIGURE 4 Estimate of hourly O-D volumes for Bay Area data. Two black solid lines correspond to O-D pairs (1,3) and (5) that have higher morning traffic and black dotted lines correspond to (2) and (9) that have higher evening traffic.