UCI-ITS-TS-WP--02-13

Temporal and Spatial Variability of Travel Time

UCI-ITS-TS-WP-02-13

Rachel R. He [1]
Henry X. Liu [2]
Alain L. Kornhauser [1]
Bin Ran [3]

[1] Department of Operations Research and Financial Engineering
Princeton University; Princeton, NJ 08544, U.S.A.
rhe@princeton.edu, alaink@princeton.edu

[2] California PATH, ATMS Center, Institute of Transportation Studies
University of California, Irvine, hliu@translab.its.uci.edu

[3] Department of Civil and Environmental Engineering
University of Wisconsin at Madison; Madison, WI 53706, U.S.A.
bran@engr.wisc.edu

August 2002

Institute of Transportation Studies
University of California, Irvine
Irvine, CA 92697-3600, U.S.A.
http://www.its.uci.edu

Temporal and Spatial Variability of Travel Time

Rachel R. He[1], Henry X. Liu[2], Alain L. Kornhauser[3], Bin Ran[4]

[1,3]Department of Operations Research and Financial Engineering
Princeton University, Princeton, New Jersey 08544
Tel: 609-258-4411
E-mail: rhe@princeton.edu, alaink@princeton.edu

[2]Institute of Transportation Studies
University of California, Irvine, CA 92697-3600
Tel: 949-824-2949
E-mail: hliu@translab.its.uci.edu

[4]Department of Civil and Environmental Engineering
University of Wisconsin, Madison, Wisconsin 53706
Tel: 608-262-0052
E-mail: bran@engr.wisc.edu

Abstract

Studying the variability of travel times is very important in helping Intelligent Transportation Systems provide drivers with useful and accurate route guidance. Much of the work done toward examining the variability in travel times has concentrated on long-term variability, such as peak hours, non-peak hours, daily and seasonal variability and concluded that travel time would follow a normal or log-normal distribution under the conditions as (1) travel times on all separate route sections are independent and (2) trip times per unit distance on all sections are identically distributed. This paper addresses the issue of short-term temporal and spatial variations and correlations of travel times. Due to the inadequacy of GPS-based probe vehicle data, the vehicle tracking data from Paramics simulation for the Orange County (CA) network is utilized in this study. The results indicated strong evidence of significant correlation between link travel times and between link travel time and link arrival time. And that conditional link travel time distributions could hardly be represented by Normal or lognormal distributions and vary by links and arrival times. The result has important implications in calibrating travel time estimation and prediction.

Key Words:

Travel time distribution, Temporal correlation, Spatial correlation

Temporal and Spatial Variability of Travel Time

Rachel R. He, Henry X. Liu, Alain L. Kornhauser and Bin Ran

1. Introduction

Route guidance is always an interest in the area of Advanced Traveler Information Systems (ATIS). A wide range of factors influences the route choices of individuals. Besides factors such as monetary cost, comfort and safety, the variability of the travel time is always regarded as an important factor. Recent empirical study by Abdel-Aty et.al. (1996) found that about 54% of the respondents in the stated preference survey indicated that travel time variability is either the most important or second most important reason for choosing their daily commuter routes. Traditional travel time functions, such as BPR function, Greenshield's model, have limits to provide accurate estimation and prediction. With the advancement of data collecting technologies, it is possible to study travel time's variability by utilizing real world data. In order to provide drivers with accurate and useful information on congested routes, anticipated travel times and alternate routes, it is necessary to ensure that two important components of ITS system are functioning successfully: travel time data collection and data analysis.

The travel time data collection part is most commonly done with the use of loop detectors and/or cameras on the roadway of interest. While many travel time data collection techniques, such as license plate matching, point detection devices, radio navigation, AVI and cellular phone, exist, one that is relatively accurate with low operating costs is the probe vehicle with a GPS-based Personal Navigation Assistant (PNA) (Boyce 1991). In a GPS-based route guidance system, the GPS receiver usually provides one to ten-second updates of latitude, longitude and time stamp as long as the antenna can "see" four of the GPS satellites, which guarantees providing accurate, continuous and automated data with low operating costs. Gathering travel times with moving vehicles produces point-to-point measures, which are more representative of highway performance than the point estimates of speed from fixed detectors, particularly for capturing stochastic times through intersections. To secure sufficient data to describe highway performance in an accurate and timely manner, a relatively large number of participating vehicles would be required. (Sen 1996, Srinivasan 1996)

In terms of route guidance, one is particularly interested in the route travel time rather than the individual link travel time. Since the route travel time depends on the link travel times, the error in the route travel time estimate will be the accumulation of the errors in each link travel times. However, a statistical difficulty that arises in analyzing travel time data is the lack of independence between link travel times. Study of long-term (day-to-day) travel time variations usually concluded that it would follow a normal or lognormal distribution under the conditions as (1) travel times on all separate route sections are independent and (2) trip times per unit distance on all sections are identically distributed. However, these assumptions may not be valid for the short-term estimation of route travel time. The objective of this paper is to study the temporal and spatial variability of travel time in short term and to test if travel times are independent and distributed normally.

This paper summarizes the developments of extracting link and path travel time observations from probe vehicle data for the experimental network. Statistical process is executed to build the link/path travel time distributions and to study the spatial correlation between link travel times and temporal correlation between link travel time and link arrival time. Our final goal is to utilize these probe vehicle data to generate timedependent link travel time distributions for each link as in Figure 1, as well as path travel time distributions with consideration of correlation between links.

The reminder of this paper is organized as follows. In section 2, specific method of network building and data processing are described. Section 3 examines the temporal distributions of link travel times. Then the spatial correlations between link travel times are statistically examined and path travel time derivation by different methods are compared in Section 4. Section 5 discusses the findings and concluding remarks.

2. Experimental Site

This paper describes some results found for travel time variability of private car mode. The data consist of travel times simulated by PARAMICS for 15 minutes during morning peak period. The study network is a highly congested corridor network in the city of Irvine, Orange County, California, illustrated in Figure 2. This network is coded and well calibrated in Paramics simulation by the authors for other projects. The vehicle tracking data for the Irvine Center Drive is used in this study. The locations of cars and times were recorded every second.

In this study, links are defined as the pairs between "monument" and "monument" instead of traditional link definition. "Monuments" are on half way between intersections/interchanges where the probability of the uniform flows is maximized. Because traditional link definition results in unnecessary complexity in description of link travel time, such as the stochastic signal delay, queuing time, and start-up time, we propose this "monument-to-monument" network building method not only to avoid the complexity because the link travel time in our study includes the random time spent at the intersection, but also to easily manipulate the data collected from GPS-equipped probe vehicles. We only need to find the probe data points that bracket the monument M_i, then interpolate time and assign T_i to M_i. From the set of [M_i, T_i], we can derive the set of [M_i, M_j, ΔT]. ΔT is the travel time between monuments M_i and M_j, which is one sample of stochastic link travel time in our study.

The GPS-based probe vehicle records are collected and shown in an example in Figure 3. However, the data currently available are not adequate for short-term distribution analysis yet (Pope 1999, Lin 2000), so we adopt the well developed simulation package PARAMICS in this study to simulate the traffic condition. The GPS records are increasing everyday, and it is expected that the analysis of real world travel times could be implemented soon. Nevertheless, the analysis of simulation data has given us a clear view how link travel times are temporally and spatially correlated as shown in the following sections.

3. The Conditional Distributions of Travel Times

3.1 Link Travel Time

After processing the output from PARAMICS according to our network building method, we get the observed means and standard deviations of the link travel times for 15 minutes during morning peak hours (start from 300 seconds until 1200 seconds) for the north and south bound routes shown in Table 1. The great variability of these link travel times is shown by the column of coefficient of variation. And the p-values from Shapiro-Wilk normality test for all links are less than 0.05, which means normality hypothesis for link travel time is rejected at 95% confidence level for all links aggregating for 15 minutes.

Link travel time distribution is also conditional on the link arrival time. For example, in Figure 4, Link 11 has totally different travel time distribution before and after arrival time 800 (500 seconds after the first car departed the origin). The conditional link travel time distribution before arrival time 800 is normally distributed with mean 58.56 and standard deviation 16.15 because the p-value of Shapiro-Wilk normality test equals 0.64. However the normality hypothesis for the conditional link travel time after arrival time 800 with mean 97.4 and standard deviation 21.62 is severely challenged because the p-value of normality test equals 0.0037.

It is also observed that different link exhibits different travel time patterns. For example, in Figure 5, Link 12 has totally different travel time distribution from Link 11 although their lengths are similar (Link 11 is 421.49 meters long while Link 12 is 397.98 meters long) and traffic volumes are similar (202 vehicles are observed on Link 11 while 201 are observed on Link 12). And both the normal hypotheses for the conditional link travel times before and after arrival time 800 are rejected for p-value <0.0001.

3.2 Path Travel Time

In this example, there are total 25 observations for northbound path travel time while there is no observation for southbound path travel time. Northbound path travel time during the observed 15 minutes is normally distributed with mean 697.84 and standard deviation 84.62 since the p-value of Shapiro-Wilk Normality test is 0.6445, as shown in Figure 6.

From Figure 6, it is also observed that the path travel time is conditional on the departure time. Path travel time patterns are different before and after departure time 450 (150 seconds after the first car departed the origin). Both the normal hypotheses for the conditional path travel times before and after departure time 450 can't not be rejected with p-value 0.1846 and 0.5468, respectively. The difference is the path travel time before 450 is normally distributed with mean 714.44 and standard deviation 43.36, while the conditional path travel time after departure time 450 is normally distributed with mean 688.5 and standard deviation 100.97.

Until now, we have separately derived conditional distributions from adequate observations for link and path travel times. How can we combine link travel times to get valid path travel time distribution? Is it appropriate to assume the independence of link travel times? In next section, spatial correlation between links is discussed.

4. Spatial Correlations in Travel Times

It is observed that it is not appropriate to assume that link travel times are independent. For example, in Figure 7, the correlation of travel times on link 10 and 11 is very strong with p-value of F-statistic equals 1.794×10^{-6}.

An analysis of the correlation coefficients between travel times on adjacent links found some evidence of significant correlation between link travel times. The analysis was based on computing and testing correlation coefficients for each pair of links, using the data points available for each pair. For example, for link 2 and 3, sample mean and covariance for both links are calculated as:

$$\overline{M} = \binom{\overline{X}}{\overline{Y}} = \binom{19.86}{17.84},$$

$$\frac{\sum_{i=1}^{n}(M_i - \overline{M})(M_i - \overline{M})'}{n-1} = \begin{pmatrix} 23.28 & -2.19 \\ -2.19 & 67.77 \end{pmatrix}.$$

The correlation calculations for northbound and southbound paths are summarized in Table 2 and 3, respectively. For the northbound route, 5 (link 3 and 4, 4 and 5, 5 and 6, 6 and 7, 10 and 11) of the 12 correlation coefficients were significant at the 5% level. For the southbound route, 6 (link 20 and 21, 21 and 22, 22 and 23, 25 and 26, 27 and 28, 29 and 30) of the 13 correlation coefficients were significant at the 5% level as shown in the column of F-statistic.

Next we compare the estimation of path travel time accumulated from link travel times with and without consideration of link correlations. Without consideration of link correlations and link travel time's dependency on link arrival time, the mean and variance of path travel time are just the summation of links' means and variances listed in Table 1, which are 511.12 and 12690.63 respectively. It is clear to see that they are far away from the path travel time distribution directly observed from the probe vehicle data, which are 697.84 and 7160.54. If we assume link and path travel times follow Normal distribution, we can derive path travel time from sample mean and covariance in Table 2, which gave us a path travel time normally distributed with mean 519.55 and variance 11524.06, which is a little bit better than the result under independence assumption, but still far away from observations.

To include link correlations and link travel time's dependency on link arrival time into our study, we found the best way to derive path travel time distribution is utilizing empirical conditional link travel time distributions, such as those in Figure 4 and 5, because no commonly available distributions with simple parameters can fit those observations well. And the result is shown in Figure 8 with mean 643.18 and variance 43501.

The comparison results are concluded in Table 4. It is very clear that path travel time estimation from empirical link travel time distributions with consideration of correlation is much closer to the distribution directly observed from the probe vehicle data, in terms of mean, than both the path travel time distributions derived from link travel times without consideration of correlation and under normal assumption. This gives more evidence of rejecting independence assumption and normal assumption. However because of the small sample size of path travel time (25 observations), the observed path travel time variance is much smaller than the path travel time distribution derived from empirical link travel times (more than 100 observations on each link) with consideration of correlations. It is expected with larger sample size that the variance of path travel time will increase.

5. Discussions

Examining and understanding the temporal and spatial variability of real world travel time is very important for calibrating travel time estimation and prediction. This paper addresses the issue of short-term temporal and spatial variation and correlation in travel times. Due to the inadequacy of GPS-based probe vehicle data, the output from Paramics simulation for the Irvine network is utilized. In this study we have the following findings.

First, link and path travel time distributions are conditional on arrival times. During certain arrival time range, link travel time may follow a normal distribution, however, it will change to other distribution afterward. Path travel time in this study shows consistency to be normal distribution with different mean and variance during different departure time range.

Second, link travel times are different from link to link even if those links are similar in length and traffic volume. Assumption of identically distributed travel time per unit distance in long-term travel time analysis is not valid in short-term analysis.

Third, in most cases, the conditional link travel time distributions could hardly be represented by Normal distributions. And from observation, they could hardly be represented by any commonly available distributions with simple parameters.

Fourth, link travel times are not spatially independent. Plus the effect of dependency on arrival time, the correlation between links does have effect not only on the estimation of path travel time mean but also on the variance estimation. It is indispensable to include temporal and spatial variability of travel times when developing advanced route guidance system with consideration of the stochastic attributes of transportation network and driver's risk taking behavior.

Because of the complexity of real world travel time, the best way to describe conditional link travel time distribution and correlations between links is utilizing empirical distributions by drawing histogram from sample data as we have done in calculating path travel time in Section 4. The shortcoming of using empirical distributions is data storage and manipulation may be costly, however, the empirical conditional link travel time distributions can be directly incorporated into advanced route guidance system, yielding benefits to the drivers.

References

[1] Abdel-Aty, Kitamura, M. R. and Jovanis, P. (1996) *Investigation Effect of Travel Time Variability on Route Choice using Repeated Measurement Stated Preference Data.* Transportation Research Record 1493, 39-45.

[2] Boyce, D., Hicks, J. and Sen, A. (1991) *In-vehicle Navigation System Requirements for Monitoring Link Travel Time in a Dynamic Route Guidance Systems*, ADVANCE working paper.

[3] Lin, I. (2000) *Analysis of the Sampling Mechanism for Providing Travel Time Information*, Thesis, Princeton University.

[4] Pope, K. (1999) *Determining Link Travel Times: the Trip-Speed Assignment Technique*, Thesis, Princeton University.

[5] Sen, A. (1996) *Frequency of Probe Reports*, ADVANCE evaluation report.

[6] Sen, A., et al. (1996) *Travel Time Prediction and Performance of Probe and Detector Data*, ADVANCE evaluation report.

[7] Srinivasan, K. and Jovanis, P. (1996) *Determination of Number of Probe Vehicles Required for Reliable Travel Time Measurement in Urban Network*, Transportation Research Record 1537, 15-22.

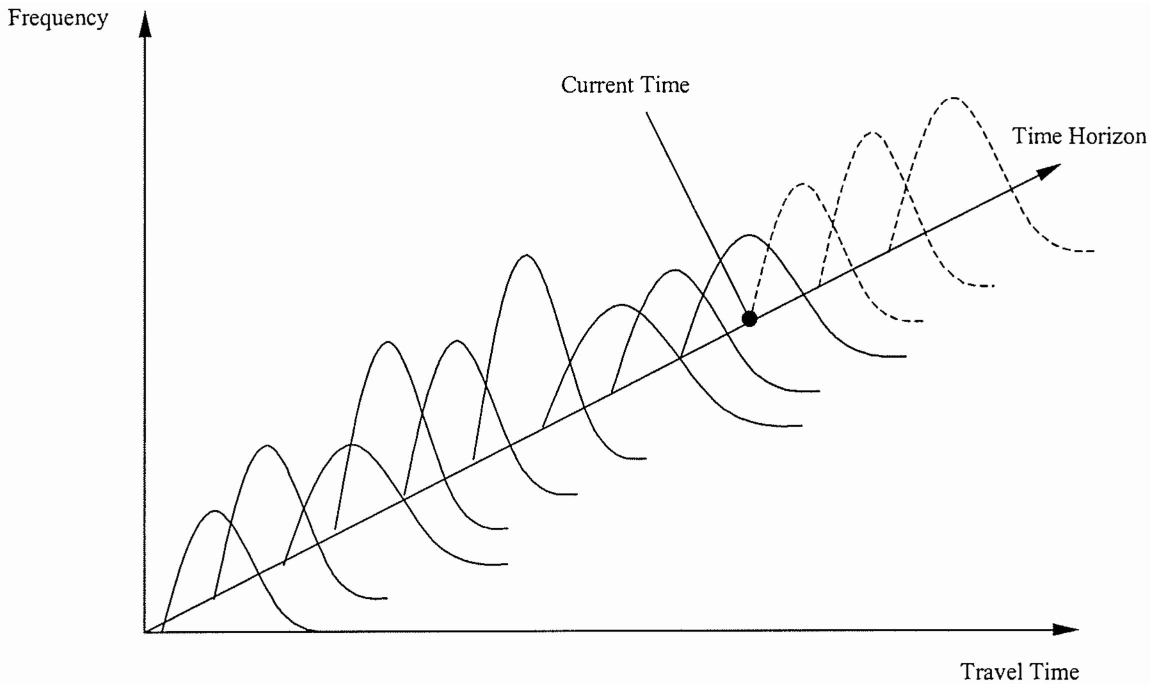

Figure 1. Dynamic and Stochastic Attributes of A Link

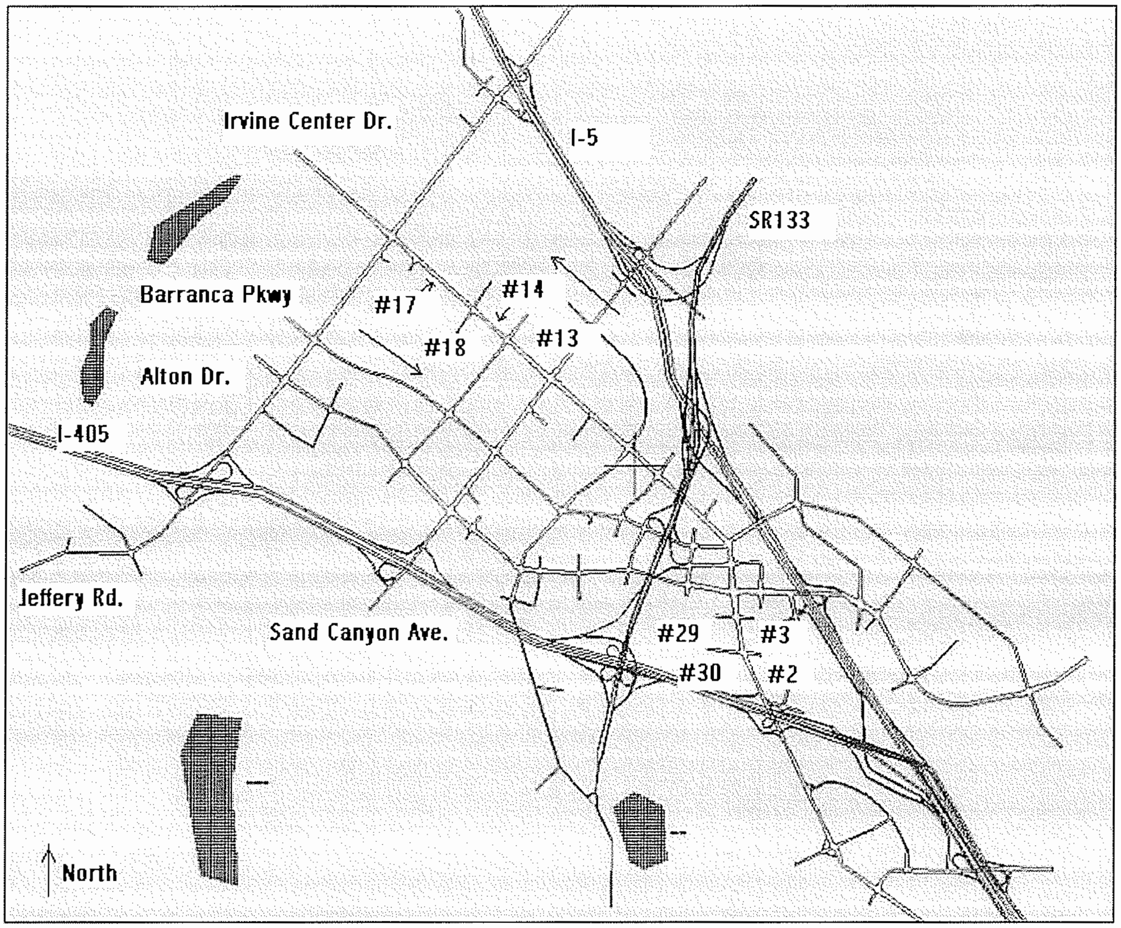

Figure 2. Network in the City of Irvine, California

(Coded in Paramics simulation. The vehicle tracking data for the Irvine Center Drive is used in this study. The number following # is the link number.)

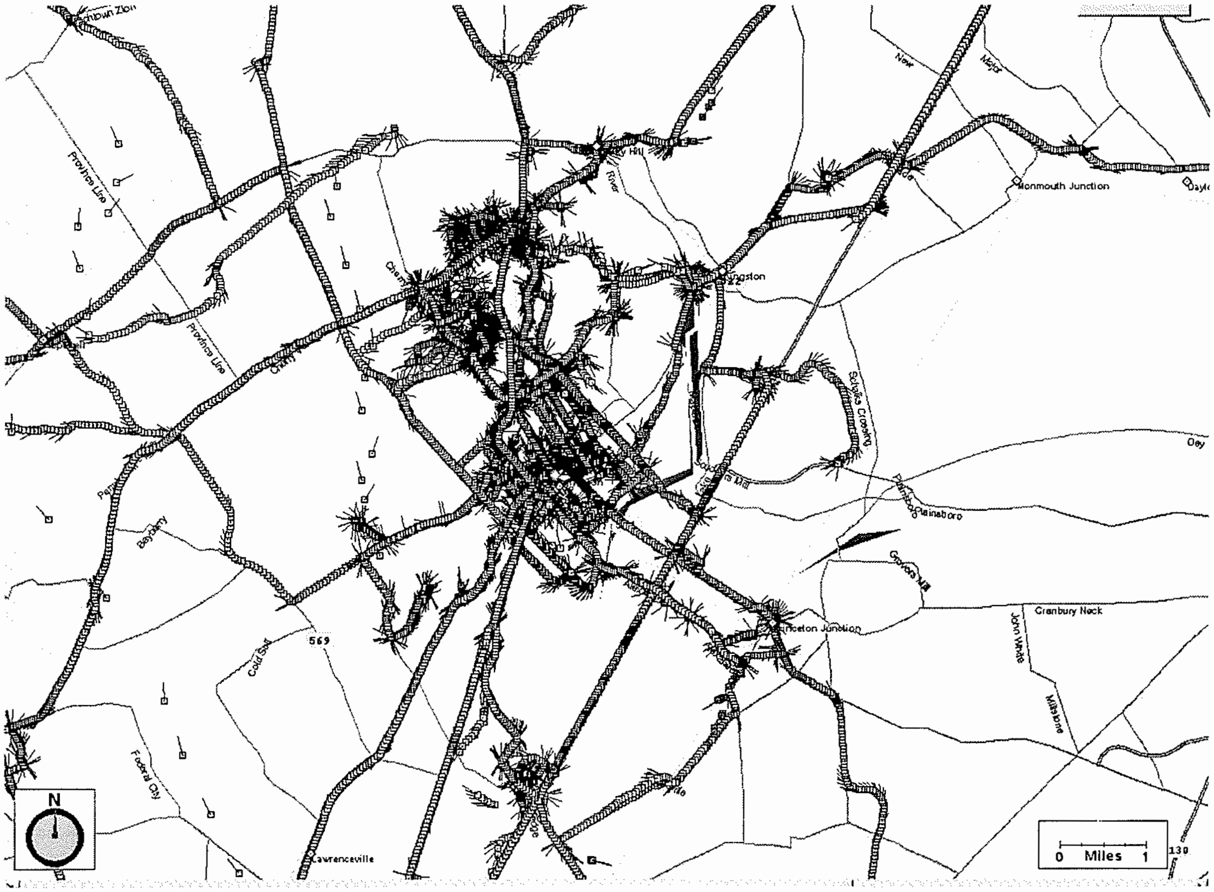

Figure 3. Traces of Probe Vehicles

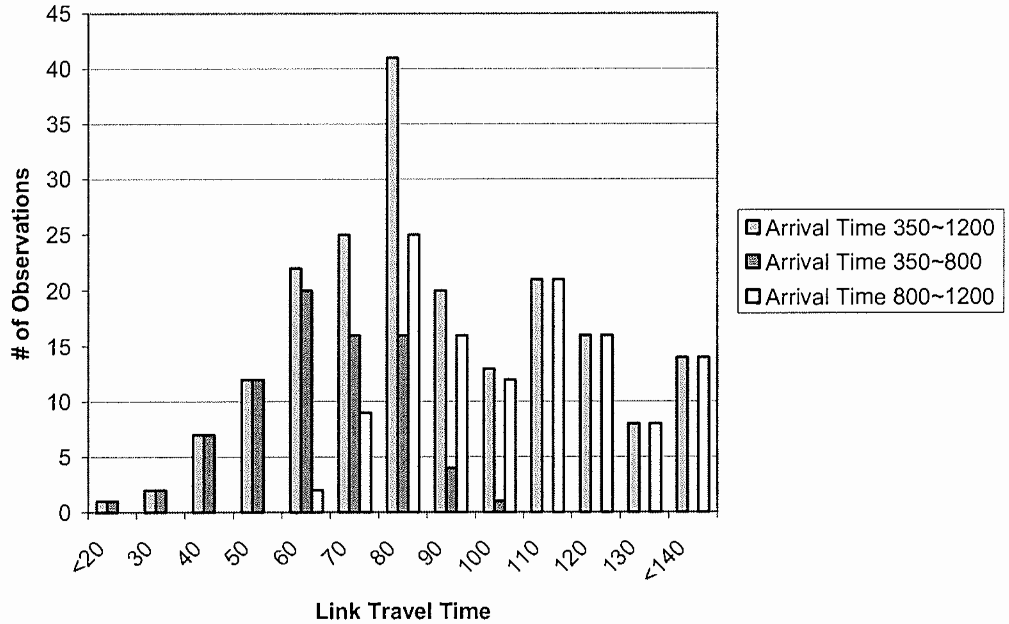

Figure 4. Travel Time Distribution of Link 11 in terms of Arrival Time

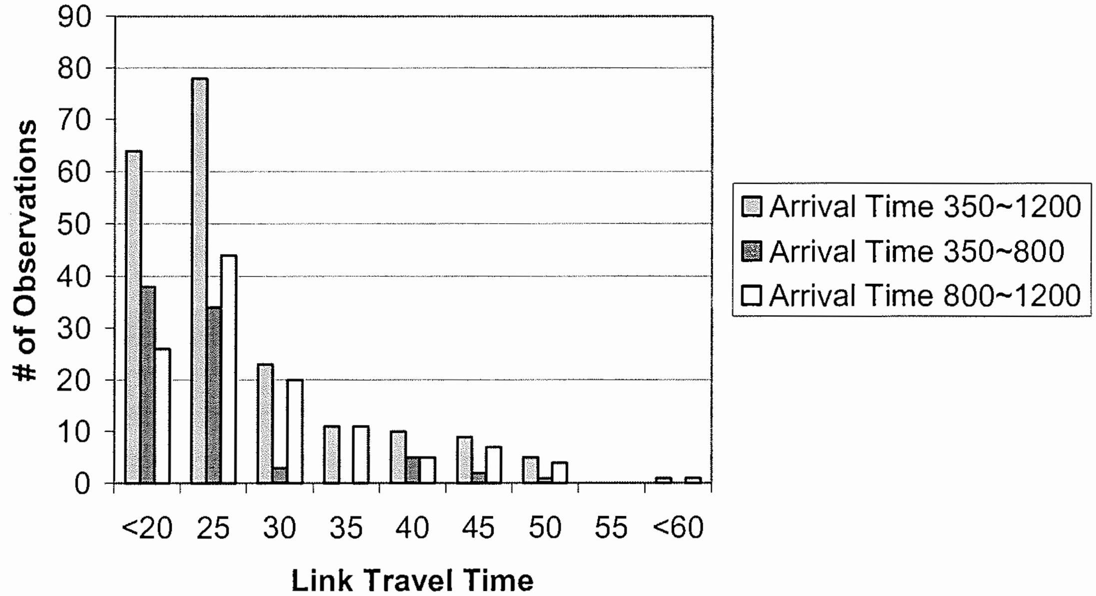

Figure 5. Travel Time Distribution of Link 12 in terms of Arrival Time

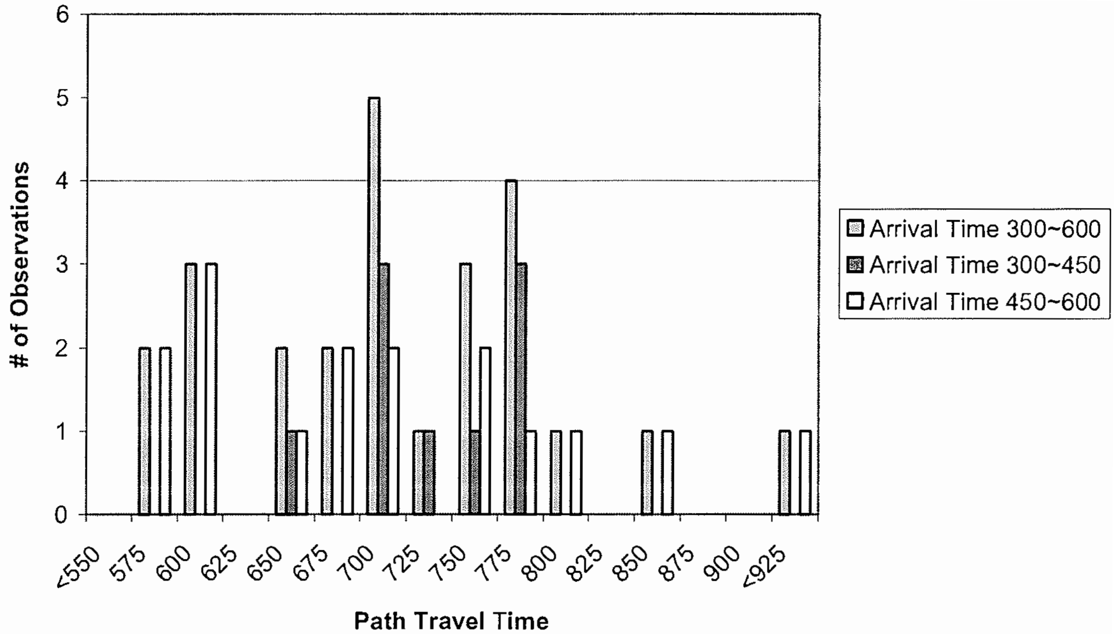

Figure 6. Northbound Path Travel Time Distribution

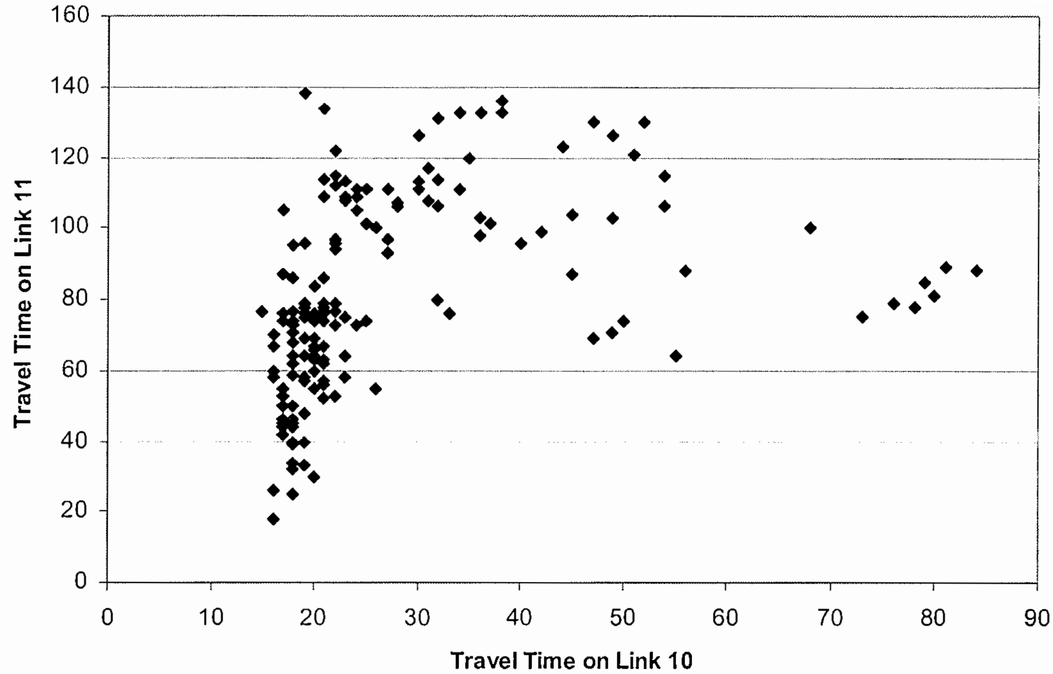

Figure 7. Scatter Plot for Observations on Link 10 and Link 11

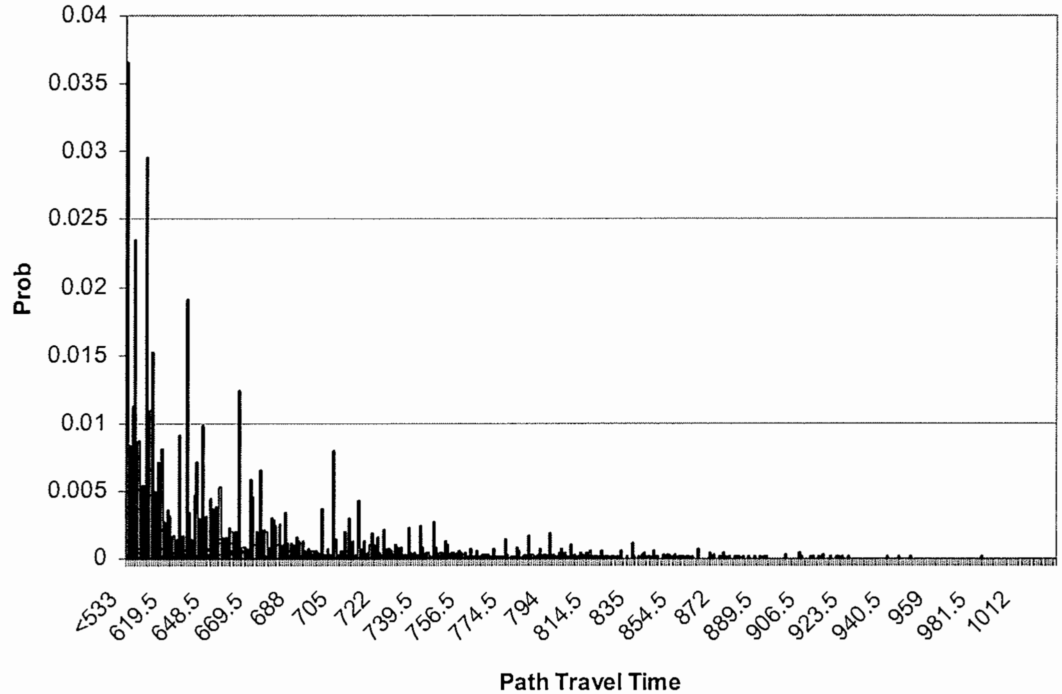

Figure 8. Path Travel Time Distribution derived from Empirical Conditional Link Travel Times

Table 1. Travel Time Analysis

	Link (# of observations)	Length (m)	Mean Time (s)	Standard Deviation	Coefficient of Variation	Shapiro-Wilk Test for Normality, P-value
North-bound Route	2 (370)	283.91	19.89	4.9	0.25	<0.0001
	3 (401)	227.35	17.79	8.46	0.48	<0.0001
	4 (111)	196.42	46.77	22.97	0.49	<0.0001
	5 (121)	177.82	53.41	26.55	0.5	0.0011
	6 (176)	215.87	33.59	18.23	0.54	<0.0001
	7 (172)	372.93	18.65	3.99	0.21	<0.0001
	8 (168)	386.36	36.32	55.87	1.54	<0.0001
	9 (179)	343.59	18.44	4.73	0.26	<0.0001
	10 (214)	394.24	26.69	13.77	0.52	<0.0001
	11 (202)	421.49	82.21	27.31	0.33	0.0076
	12 (201)	397.98	24.1	7.74	0.32	<0.0001
	13 (235)	460.94	32.33	12.75	0.39	<0.0001
	14 (140)	410.13	100.93	81.93	0.81	<0.0001
South-bound Route	17 (59)	617.58	107.53	59.08	0.55	0.0018
	18 (155)	410.13	31.32	12.39	0.4	<0.0001
	19 (180)	460.94	29.19	8.95	0.31	<0.0001
	20 (200)	397.98	31.49	11.12	0.35	<0.0001
	21 (154)	421.49	55.45	22.25	0.4	0.0018
	22 (161)	394.24	18.68	2.5	0.13	<0.0001
	23 (176)	343.59	19.1	3.82	0.2	<0.0001
	24 (140)	386.36	18.21	2.87	0.16	<0.0001
	25 (144)	372.93	19.87	5.3	0.27	<0.0001
	26 (68)	215.87	40.19	20.25	0.5	0.0003
	27 (21)	177.82	59.76	33.79	0.57	0.0364
	28 (25)	196.42	55.56	40.09	0.72	0.0009
	29 (293)	227.35	17.82	9.92	0.56	<0.0001
	30 (278)	283.91	18.47	5.05	0.27	<0.0001

Table 2. Correlations between Links on Northbound Route

	Between Links (# of Observations)	Sample Mean	Sample Covariance	F-statistic p-value
Northbound Route	2&3 (355)	(19.86, 17.84)’	$\begin{pmatrix} 23.28 & -2.19 \\ -2.19 & 67.77 \end{pmatrix}$	0.3009
	3&4 (107)	(18.92, 47.22)’	$\begin{pmatrix} 65.7 & -37.1 \\ -37.1 & 528.42 \end{pmatrix}$	0.03975
	4&5 (85)	(45.41, 58.67)’	$\begin{pmatrix} 396.34 & -184.18 \\ -184.18 & 738.44 \end{pmatrix}$	0.001432
	5&6 (81)	(54.11, 30.31)’	$\begin{pmatrix} 761.4 & -178.79 \\ -178.79 & 344.89 \end{pmatrix}$	0.001412
	6&7 (131)	(33.92, 18.76)’	$\begin{pmatrix} 355.83 & -27.07 \\ -27.07 & 15.09 \end{pmatrix}$	1.405×10^{-5}
	7&8 (144)	(18.75, 36.01)’	$\begin{pmatrix} 16.12 & 5.84 \\ 5.84 & 2779.69 \end{pmatrix}$	0.7429
	8&9 (139)	(34.8, 18.71)’	$\begin{pmatrix} 2479.39 & 29.98 \\ 29.98 & 25.64 \end{pmatrix}$	0.1633
	9&10 (177)	(18.47, 27.36)’	$\begin{pmatrix} 22.46 & -9 \\ -9 & 201.98 \end{pmatrix}$	0.07592
	10&11 (163)	(27.5, 80.79)’	$\begin{pmatrix} 224.07 & 146.26 \\ 146.26 & 720.81 \end{pmatrix}$	1.794×10^{-6}
	11&12 (171)	(82.96, 23.92)’	$\begin{pmatrix} 791.75 & 2.22 \\ 2.22 & 52.48 \end{pmatrix}$	0.8876
	12&13 (157)	(23.16, 33.25)’	$\begin{pmatrix} 56.07 & -6.93 \\ -6.93 & 204.71 \end{pmatrix}$	0.3937
	13&14 (94)	(26.97, 110.78)’	$\begin{pmatrix} 49.92 & -88.18 \\ -88.18 & 6749.57 \end{pmatrix}$	0.1858

Table 3. Correlations between Links on Southbound Route

	Between Links (# of Observations)	Sample Means	Sample Covariance	F-statistic p-value
South-bound Route	17&18 (28)	(86.36, 28.64)'	$\begin{pmatrix} 2347.65 & -127.76 \\ -127.76 & 91.79 \end{pmatrix}$	0.1335
	18&19 (132)	(31.38, 30.14)'	$\begin{pmatrix} 158.79 & -1.82 \\ -1.82 & 80.42 \end{pmatrix}$	0.7182
	19&20 (157)	(29.56, 31.14)'	$\begin{pmatrix} 83.15 & 9.9 \\ 9.9 & 114.28 \end{pmatrix}$	0.2058
	20&21 (107)	(31.35, 55.04)'	$\begin{pmatrix} 132.81 & 87.58 \\ 87.58 & 591.53 \end{pmatrix}$	0.00105
	21&22 (141)	(54.24, 18.93)'	$\begin{pmatrix} 478.46 & -12.65 \\ -12.65 & 6.38 \end{pmatrix}$	0.006301
	22&23 (136)	(18.75, 19.41)'	$\begin{pmatrix} 6.54 & 3.63 \\ 3.63 & 16.64 \end{pmatrix}$	3.346×10^{-5}
	23&24 (118)	(19.18, 18.48)'	$\begin{pmatrix} 17.82 & -0.49 \\ -0.49 & 9.24 \end{pmatrix}$	0.8347
	24&25 (125)	(18.21, 19.78)'	$\begin{pmatrix} 8.09 & -0.36 \\ -0.36 & 26.45 \end{pmatrix}$	0.7586
	25&26 (50)	(18.4, 41.26)'	$\begin{pmatrix} 11.39 & 23.73 \\ 23.73 & 407.18 \end{pmatrix}$	0.01322
	26&27 (7)	(25, 77)'	$\begin{pmatrix} 60.33 & 56.33 \\ 56.33 & 704.67 \end{pmatrix}$	0.5533
	27&28 (21)	(59.76, 61.81)'	$\begin{pmatrix} 1141.59 & -889.8 \\ -889.8 & 1671.66 \end{pmatrix}$	0.001626
	28&29 (19)	(53.68, 22.16)'	$\begin{pmatrix} 1723.56 & 70.83 \\ 70.83 & 218.81 \end{pmatrix}$	0.6617
	29&30 (264)	(18.07, 18.5)'	$\begin{pmatrix} 99.71 & 10.29 \\ 10.29 & 25.16 \end{pmatrix}$	0.0008218

Table 4. Comparison of Path Travel Time Distribution

	Mean	Variance
Observed Distribution	697.84	7160.54
Derived Distribution from Independent Links	511.12	12690.63
Derived Distribution from Correlated Links under Normal Assumption	519.55	11524.06
Derived Distribution from Correlated Links using empirical distribution	643.18	43501