Determination of Number of Probe Vehicles Required for Reliable Travel Time Measurement in Urban Network

KARTHIK K. SRINIVASAN AND PAUL P. JOVANIS

Several intelligent vehicle–highway system demonstration projects are currently assessing the feasibility of using probe vehicles to collect real-time traffic data for advanced traffic management and information systems. They have used a variety of criteria to determine the number of probes necessary, but few generalizable algorithms have been developed and tested. The described algorithm explicitly considers the time period for travel time estimation (e.g., 5, 10, or 15 min), the number of replications of travel time desired for each link during each measurement period (reliability criterion), the proportion of links to be covered, and the length of the peak period. This algorithm is implemented by using a simulation of the Sacramento Network (170 mi^2) for the morning peak period. The results indicate that the number of probe vehicles required increases nonlinearly as the reliability criterion is made more stringent. More probes are required for shorter measurement periods. As the desired proportion of link coverage in the network increases, the number of probes required increases. With a given number of probes a greater proportion of freeway links than of major arterials can reliably be covered. Probe vehicles appear to be an attractive source of real-time traffic information in heavily traveled, high-speed corridors such as freeways and major arterials during peak periods, but they are not recommended for coverage of minor arterials or local and collector streets or during off-peak hours.

Advanced traffic management and information systems (ATMISs) represent an important set of systems within the Intelligent Transportation System Program. ATMISs aim to achieve improved and coordinated traffic control, incident management, and vehicle routing within the network. ATMISs are characterized by the collection of real-time traffic data, responses to changes in traffic flow with traffic management strategies, areawide surveillance and detection systems, and integrated management of various functions.

Several technologies currently used for traffic surveillance for ATMISs include fixed-location surveillance mechanisms such as loop detectors, radar, and video image processing techniques. In contrast to these fixed-location techniques, vehicles can be used as traffic probes, experiencing travel times as they traverse various links of the network and transmitting point-to-point travel time to a traffic information center. Several demonstration projects for studying the feasibility of using probe vehicles to obtain real-time traffic data are under way. These include projects in Sydney, Australia (*1*); the ADVANCE project in Chicago (*2*); the Pathfinder project, a field trial of an in-vehicle real-time traffic information system in the Santa Monica Freeway Smart Corridor in Los Angeles (*3*); and the FAST-TRAC demonstration project in Michigan (*4*).

The cost of collecting travel time data depends on the number of probes required. There is therefore an interest in estimating the smallest number of probes required to measure link travel times in the network reliably and adequately. Various estimates of the required number of probe vehicles have been provided by different studies. In a study undertaken for the implementation of a travel time measurement system for the Sydney metropolitan area modeled with 200 nodes, a "majority" of 3,700 taxis was considered an adequate sample (*5*). In another study it was estimated that about 200 to 300 probe vehicles would provide sufficient coverage in the I-45 (North)–US-59 (Eastex) freeway corridor in Houston during the morning peak period (*6*). Boyce et al. (*7*) have estimated that about 5,000 probe vehicles are required to cover 60 percent of a study area in the northwest suburbs of Chicago for a measurement period of 10 min during the morning peak period.

The great disparity in the estimates from different studies is not surprising because they depend on specific network characteristics including link capacities and flow levels. None of these studies presents a procedure that can be used to obtain the number of probe vehicles required that accounts for both the reliability of travel times measured by probes and adequate coverage of the network by probes, although the procedure suggested by Boyce et al. (*8*) accounts for adequate coverage. In this paper a procedure that explicitly considers the following criteria is developed:

1. Reliability corresponds to the number of replications of travel times from probe vehicles for each link during each measurement period desired in the estimate of the number of probes.
2. Adequacy relates to the proportion of links to be sampled at least once (based on the reliability criterion) during the measurement period.

Several additional factors are considered in empirical testing, including trip length distribution and the mix of link classes (for freeways, arterials, etc.).

A general procedure for the estimation of the number of probe vehicles required is proposed. The implementation of this procedure by using a computer simulation is then discussed. The results are analyzed, and a discussion of its implications for real-time traffic information systems is presented. The findings of the study are summarized and recommendations for future research are presented in the last section.

METHODOLOGY FOR DETERMINING NUMBER OF PROBE VEHICLES REQUIRED FOR RELIABLE TRAVEL TIME MEASUREMENT

Reliable representation and an adequate sampling area are considered essential in the determination of the number of probes required.

K. K. Srinivasan, 6.302 ECJ Building, University of Texas, Austin, Tex. 78751. P. P. Jovanis, Institute of Transportation Studies, 2028 Academic Surge, University of California—Davis, Davis, Calif. 95616.

Definition and Formulation of Reliability Measures

The travel times measured by the probes must reliably represent link travel times in the network. Two measures of reliability are considered. The first measure defines reliability in terms of the probability of the absolute error (given by $|T_{plt} - \mu_{lt}|$, the difference between the travel time of probe vehicles and the mean travel time of all vehicles) not exceeding a threshold on ε_a. By using this criterion the sample size for reliability n_{lt} is shown (*9*) to be

$$n_{lt} = (ts_{lt}/\varepsilon_a)^2 \tag{1}$$

where t represents the t-value for a chosen reliability level (say 90 percent), and s_{lt} is the standard deviation of probe travel times.

May (*9*) observes that for traffic engineering travel time studies a sample size of between 50 and 100 is adequate. This is a substantial requirement, however, for probe vehicles during each measurement period for each link. For this measure the allowable variance on larger values of travel times is more stringent than that on smaller values of mean travel times. Furthermore, this measure ignores the possible dependence of the standard deviation on the mean travel time. Hence, its use is not recommended.

The second measure of reliability (r) is defined as the minimum probability that the absolute value of relative error (ε_{lt}) is less than the maximum allowable relative error threshold (ε_{max}) (i.e., the percentage of time the relative error is less than maximum allowable relative error). The ε_{lt} of the probe travel times is defined as the ratio of the difference between the mean travel time of the probes and the overall mean travel time (of all vehicles traversing the link) to the overall mean travel time. Thus, relative error is given by

$$\varepsilon_{lt} = (T_{lt} - \mu_{lt})/(\mu_{lt}) \tag{2}$$

From the definition of reliability,

$$Pr\{|\varepsilon_{lt}| < \varepsilon_{max}\} \geq r \tag{3}$$

Reliability can now be formulated in terms of the number of probes required for a link l at time t and a measurement period of interest, tu:

Let μ_{lt} be the expected travel time (all vehicles),

Let μ_{elt} be the estimated mean travel time (all vehicles) based on historical travel times,

Let σ^2_{lt} be the travel time variance,

Let σ^2_{elt} be the estimated travel time variance based on historical travel times,

Let n_{lt} be the number of probe vehicles required to reliably measure link travel times,

Let T_{lt} represent the mean travel time experienced by n_{lt} probe vehicles,

Let s_{lt} represent the sample standard deviation of the travel times of the probe vehicles, and

Let n_{plt} represent the minimum number of probes required to reliably measure travel times during the measurement period; this n_{plt} is a lower-bound value for n_{lt}.

Substituting for the relative error in Equation 3 gives

$$Pr\{-\varepsilon_{max} < [(T_{lt} - \mu_{lt})/\mu_{lt}] < \varepsilon_{max}\} \geq r \tag{4}$$

Since T_{lt} is a sample mean of travel times of probe vehicles, an invocation of the central limit theorem implies asymptotically,

$$[(T_{lt} - \mu_{lt})/(\sigma_{lt}/\sqrt{n_{lt}})] \sim N(0,1) = Z \tag{5}$$

Substituting Equation 5 in Inequality 4,

$$Pr[-\varepsilon_{max}\mu_{lt}/(\sigma_{lt}/\sqrt{n_{lt}}) < Z < \varepsilon_{max}\mu_{lt}/(\sigma_{lt}/\sqrt{n_{lt}})] \geq r \tag{6}$$

$Z \sim N(0,1)$, $\Phi(x)$ is the cumulative distribution function of Z evaluated at x and Φ^{-1} is its inverse. By using symmetry of the normal distribution, Inequality 6 is rewritten as

$$\varepsilon_{max}\mu_{lt}/(\sigma_{lt}/\sqrt{n_{lt}}) \geq \Phi^{-1}[(1 + r)/2] \tag{7}$$

$$n_{lt} \geq \{\Phi^{-1}[(1 + r)/2]/[\varepsilon_{max}(\mu_{lt}/\sigma_{lt})]\}^2 = n_{plt} \tag{8}$$

For example, if the desired r is 95 percent and ε_{max} is 10 percent,

$$\begin{aligned} n_{plt} &= \{\Phi^{-1}(1.95/2)/[0.1(\mu_{lt}/\sigma_{lt})]\}^2 \\ &= [19.6/(\mu_{lt}/\sigma_{lt})]^2 \end{aligned} \tag{9}$$

Thus, the number of probe measurements on each link during each measurement is proportional to the coefficient of variation squared. The constant of proportionality depends on the desired r and ε_{max}. Since exact mean travel times and corresponding variances are not known, an estimate of the ratio μ_{lt}/σ_{lt} obtained by using historical data can be used in Equation 9, resulting in

$$n_{plt} \approx [19.6/(\mu_{elt}/\sigma_{elt})]^2 \tag{10}$$

It is assumed there are enough probes ($n_{lt} > 25$) for the central limit theorem to be applicable. This requirement may be satisfied in very heavily traveled corridors over long measurement periods (more than 10 min) but not in corridors where the traffic is light. If this requirement is violated, the exact mean travel time distribution of the probes, T_{lt}, may be determined by convolution or other statistical procedures. Let the probability density function of the mean travel time $T_{lt}(y)$ of all probes in link l during a measurement period starting at t be given by $g(\Theta, n_{plt}, y)$, where Θ is the vector of parameters determining the distribution of the travel time of an individual probe vehicle. Then Inequality 4 can be written as

$$Pr[-\mu_{lt}(1 - \varepsilon_{max}) < T_{lt} < \mu_{lt}(1 + \varepsilon_{max})] \geq r \tag{4a}$$

and simplified as

$$\int_{x_1}^{x_2} g(\Theta, np, y)dy \geq r \tag{4b}$$

where x_1 is equal to $-\mu_{lt}(1 - \varepsilon_{max})$, and x_2 is equal to $\mu_{lt}(1 + \varepsilon_{max})$.

For fixed values of ε_{max} and r, Inequality 4b can be solved for obtaining values of $n_{lt} \geq n_{plt}$. For many travel time distributions, this cannot be solved analytically. There is a trade-off between the central limit theorem approximation and the computational complexity in solving Inequality 4b, assuming that the distribution of T_{lt} can be found analytically.

Adequacy of Area Coverage by Probe Vehicles

The second major aspect for determining the total number of probe vehicles required is adequate area coverage. This is a design parameter of interest, because it is desirable to have reliable travel times on as many links as possible with a given number of probes. Area coverage can be defined in terms of the proportion of links in the network reliably covered by the probe vehicles. The travel times on certain link classes such as freeways and major arterials are arguably of a higher value than those on other road classes. Hence, area coverage can also be defined in terms of the proportion of links covered in each link class.

The objective, then, is to determine the number of probe vehicles N required in the network given that a desired proportion (p) of the links be covered reliably in a peak period of D minutes for a fixed measurement period of tu minutes.

The problem of adequate area coverage involves the consideration of spatial and temporal overlaps in the paths of the probes. For example, the probes stay on the network for different durations and may cover more than one link during a measurement period. Probes between different origins and destinations could traverse the same links during the same or different measurement periods. The adequate area coverage will therefore be influenced by the trip length distribution, the mix of link types such as freeways and arterials, and their speed differences, network characteristics, and probe trip characteristics such as origin, destination, and departure time. This problem of adequate area coverage of the probe vehicles is not tractable in a closed form.

Boyce et al. (*8*) have suggested the following procedure for estimating the number of probes required: sample N probe trips at random from the pool of all vehicle trips and assign them to the network by using a static user-optimal route choice model for the peak period. For a fixed measurement period tabulate the number of links traversed by at least one probe and plot that number against the total number of probes N.

Although this approach is attractive regarding the handling of adequate area coverage, it has some serious limitations. It ignores the reliability requirement by assuming that if a link is traversed by at least one probe vehicle the link travel times experienced by the probe vehicles reliably represent the travel times on the link. The use of a static, user-optimal route choice model constrains the optimal route to be fixed throughout the peak period resulting in possibly nonoptimal routing.

The following procedure is proposed. An appropriate mechanism for sampling the probe vehicle trips from the pool of all vehicle trips must be adopted. Random selection or suitably weighted selection of origins, destinations, and departure times of probe vehicle trips may be considered. Instead of the static user-optimal route assignment, a stochastic or dynamic route assignment process is suggested. Several possible criteria could be used for the route assignment model. However, it is not clear whether there exists a unique routing strategy that can always achieve the minimum number of probes. Further investigation is required to examine this issue.

The estimate of the number of probes would vary with the route assignment criterion, depending on the context of usage. For instance, when the traffic controller has no control over the routing decisions, origins and destinations, and paths of the vehicles, a random selection of probes is reasonable and it is presumed that users take the paths with the shortest times. In this case the need to specify desired link traversal proportions for different link classes is obviated. However, if the goal is to maximize link coverage with a given fleet of probes, then criteria such as the minimum capacity-to-links ratio may be used, in which the capacity on a path is the maximum number of probes required on any link on that path. Therefore, the vehicles would be assigned to paths that have the maximum number of links covered by the minimum number of probes (n). Because the controller has some control over the routing decision of the probes, fewer probes will be required. However, here it might be necessary to consider prespecified requirements for link coverage proportion for each link class.

Therefore, the use of a generic stochastic or dynamic route is suggested. A more realistic route choice and selection mechanism will lead to accurate estimates of the number of probes required.

General Heuristic Algorithm for Estimating Number of Probe Vehicles Required in Network for Reliable Travel Time Estimation

The following general algorithm is proposed:

1. Determine n_{plt}, the minimum number of probe vehicles required from the reliability criterion (Equation 10 or Inequality 4b) on each link l during each measurement period tu.
2. Solve a stochastic or dynamic route assignment model for the network.
3. Sample N probe vehicle trips from the pool of all vehicle trips occurring in the network during the peak period.
4. Assign these N probe vehicle trips by using the route assignment model from Step 2.
5. Determine the proportion of links p_t covered reliably by probes (links in which the number of probes in the measurement period starting at time t is at least n_{plt}). Average the link coverage p_t over all measurement periods in the peak period to obtain the average proportion of link coverage p.

This simulation is repeated for a given N and measurement period tu to obtain a confidence interval for p. This entire simulation is repeated for different values of N until the desired proportion of the links p_0 is covered, that is, $p \geq p_0$. If different desired proportions are specified for each link class, the algorithm continues until the coverage requirement of each link class is satisfied.

SIMULATION MODEL FOR IMPLEMENTING ALGORITHM

The general algorithm is implemented by using a simulation procedure. It is assumed that the traffic controller has no control over the origins and destinations and routes of the probes but can monitor their paths as they traverse the network. The general algorithm has been simplified as follows because of the lack of reliable and detailed data, such as the means and variances of travel times for all of the links for measurement periods as short as 5 min.

May (*9*) notes that the coefficient of variation (σ_{lt}/μ_{lt}) is in the range of 0.08 to 0.17, which corresponds to n values of 2 to 11 (for $r = 95$ percent and $\varepsilon_{max} = 10$ percent). This is used as a basis for the reliability criterion; n values of 1, 2, 3, 5, and 10 are used in the simulation. It is assumed that μ_{lt}/σ_{lt} is constant across all links and for all the time periods.

In the simulation probe trips are selected at random with respect to origin, destination, and departure time. If some origins, destinations, and departure times in a city are more attractive than others, then randomization can be performed by giving suitable weights to result in a more representative selection of probe vehicle trips.

A stochastic shortest-path-route assignment model is used with the algorithm. The rationale is that a majority of the users will choose paths that minimize their travel times. Hence, random selection of origins and destinations and shortest-path-route assignment would lead to a representative selection of probe vehicle trips from the pool of all trips.

The route assignment procedure is based on normally distributed travel times. McShane and Roess (*10*) note that normal distributions are often fitted to observed travel times, speeds, and delays. For this distribution the procedure for determining n_{plt} becomes exact even for small samples. Alternatively, travel times can be generated from normally distributed speeds. Convolution procedures can be used if some other analytical distribution is assumed.

The Sacramento network, with an area of about 170 mi^2, is used for the simulation. This network has 174 nodes and 248 links (73 freeway and 175 nonfreeway links). Only major arterials and freeways are included in the network, and a 2-hr peak period is considered.

The simulation model is used to calculate p for given values of N, n, and tu. The proportion of freeways and arterials covered by the probes is also computed. This simulation is repeated for various combinations of n of 1, 2, 3, 5, and 10 vehicles per link per measurement period and tu of 5, 10, 15, and 20 min, respectively, for N values of 500, 1,000, 1,500, and so forth. For each combination of design parameters five replicates are performed to obtain a confidence interval for p.

RESULTS FROM SIMULATION MODEL

The simulation is used to examine the nature of relationships between the design parameters and the number of probes required and to identify possible design trade-offs, including the following:

1. The effect of the reliability requirement (n) on N.
2. The effect of measurement period tu on N.
3. How the proportion of links traversed by the probe vehicles influences N.
4. Given a number of probes, analysis of how the proportion of link coverage by the probes varies with link classes.

The plot between N and n (Figure 1) indicates that as the requirement of n (on each link per measurement period) increases, N increases almost linearly for each value of tu. This suggests that if the variance in travel times on a link is small, then fewer probes are adequate to reliably represent travel times.

The plot of N versus tu (Figure 2) for a p value of .6 and n values of 1, 2, 3, 5, and 10 indicates that as the measurement period decreases the number of probes increases nonlinearly. Interestingly, there is a "knee" at a tu of about 10 min, particularly for high values of n. So measurement periods of less than 10 min in duration carry a premium in the requirement for n. For measurement periods longer than 10 min the relationship between N and tu is nearly linear.

The narrow joint confidence intervals (90 percent) for an n value of three probes per link per measurement period and a tu of 10 min (Figures 3 and 4) indicate that the point estimates of N have small variations around their mean values.

As the desired proportion of link traversals p increases, more probes (N) are required. The proportion of links traversed increases at a decreasing rate with increasing N. There appears to be a knee beyond which large increases in N are required for a moderate increase in p. The level of N at which the knee occurs depends on other design parameters (such as n).

The results also indicate that the freeways have significantly (statistically) higher link traversals than the arterials for given values of N and tu. For small values of N the probe vehicles are attracted to freeways because of their higher speeds, resulting in higher p values. As the number of probes increases the p value for freeways increases rather slowly. For nonfreeways, however, with an increasing value of N, the value of p increases at a decreasing rate and finally levels off to a saturation value. This is natural because the user-optimal route choice model results in the drivers choosing the faster routes (freeways) over the slower routes (arterials). However, some drivers may prefer marginally slower routes, which have smaller variances in travel times than the high-speed routes (*11*). Hence, it appears viable to use probes in heavily traveled corridors such as freeways instead of major arterials.

For reliability requirements with n values of 1, 3, and 10 probes per link per measurement period (r = 95 percent; σ/μ_{tt} = 0.06, 0.09, and 0.16; ε_{max} = 0.1; tu = 10 min; p = 80 percent) about 1,200, 3,500, and 10,000 probe vehicles, respectively, are required. If the reliability requirement is not considered, the estimated number of probe vehicles could be underestimated by a factor between 2 (n = 2) and 8 (n = 10). An underestimate in the number of probes results in a reduced reliability of the travel times or a reduced proportion of link coverage for a given reliability level.

For an n value of three probes per link per measurement period (p = .8) about 6,500, 3,500 and 2,500 probes are required for measurement periods of 5, 10, and 15 min, respectively. The use of a 5-min measurement period requires substantially more equipped vehicles than a 10-min measurement period, and hence, 10 min seems a reasonably appropriate measurement period in travel time monitoring systems with probes.

For a reliability requirement of three probes per link during each measurement period of 10 min, the total number of probes required is about 900 for a 40 percent link traversal proportion, 1,200 for a 60 percent link traversal proportion, and 3,500 for an 80 percent link traversal proportion.

The adequacy criterion of 80 percent overall link coverage is reasonable since most of the freeways (97 percent) are covered and 74 percent of nonfreeways are covered. A 60 percent overall link coverage results in 90 percent of freeways and but only 49 percent of nonfreeways being covered. This is because as the speed declines, each probe covers fewer links in any measurement period; hence, more probes are required.

An empirical model is fitted to the simulation results to study the simultaneous variation between N and other design factors including p, tu, and n. The linear regression on the logarithms of the variables of interest is performed:

$$\log N = a + b \log n + c \log p + d \log tu + \varepsilon \qquad (11)$$

where $\varepsilon \sim iid\ N(0,\sigma^2)$. The calibrated model is

$$\log N = 9.10 + 1.16 \log p + 0.8 \log n - 0.7 \log tu$$

This yields

$$N = 8,947(p)^{1.16}(n)^{0.8}(tu)^{-0.7}$$

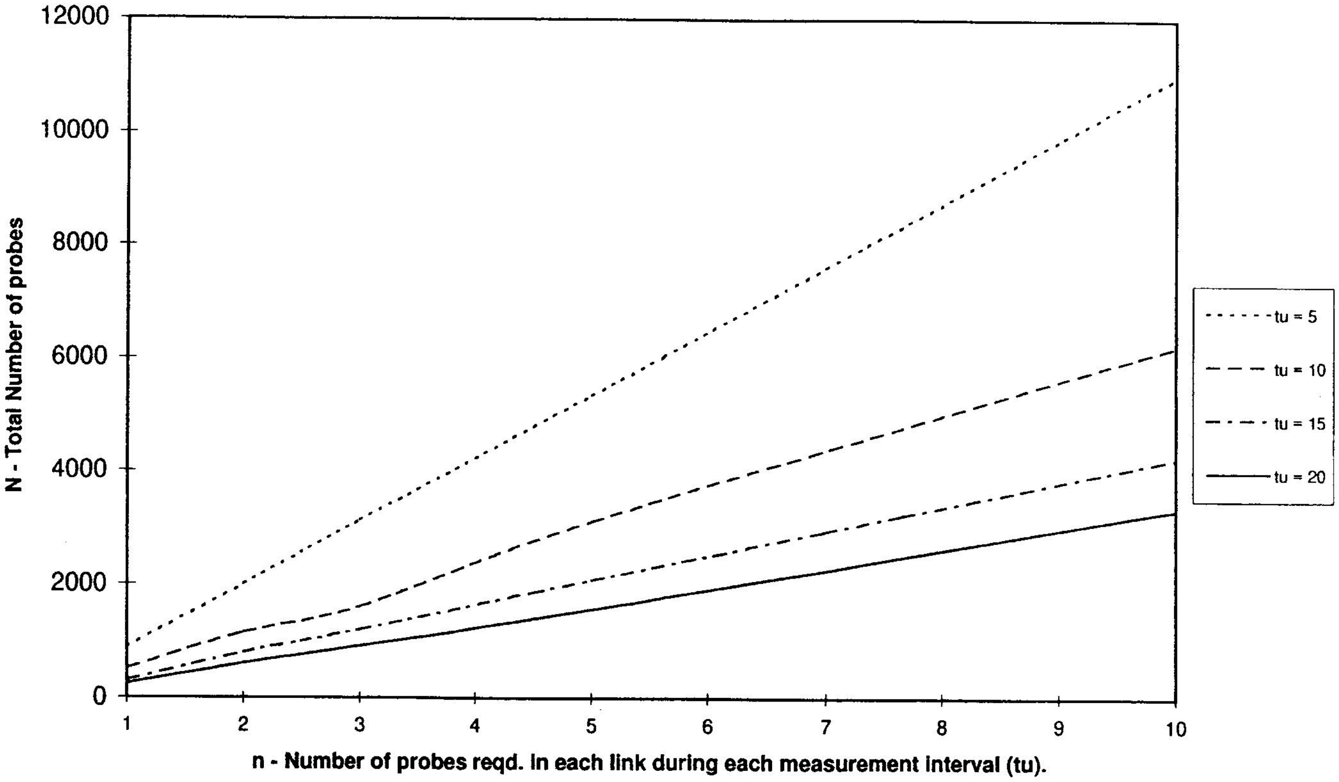

FIGURE 1 Plot of N versus n (p = .6).

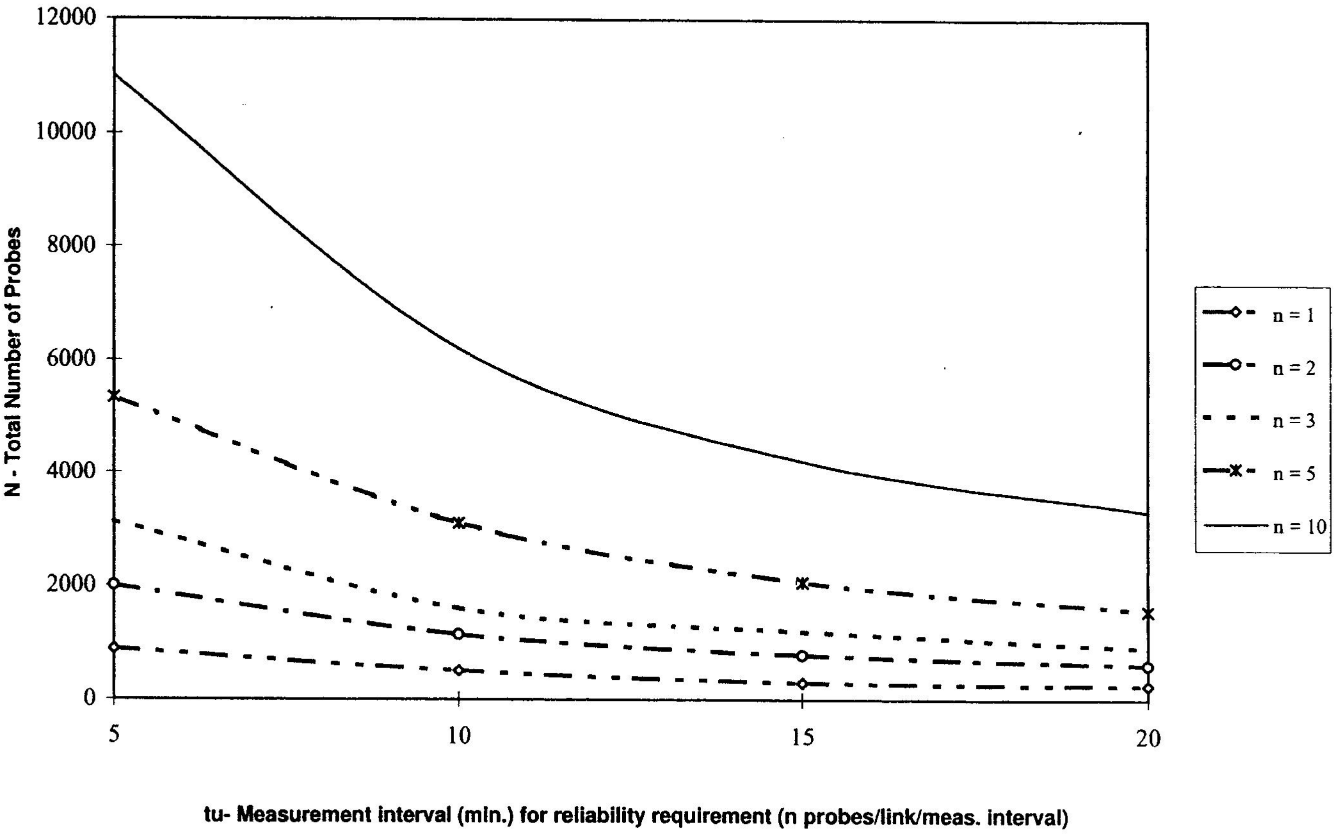

FIGURE 2 Plot of N versus tu (p = .6).

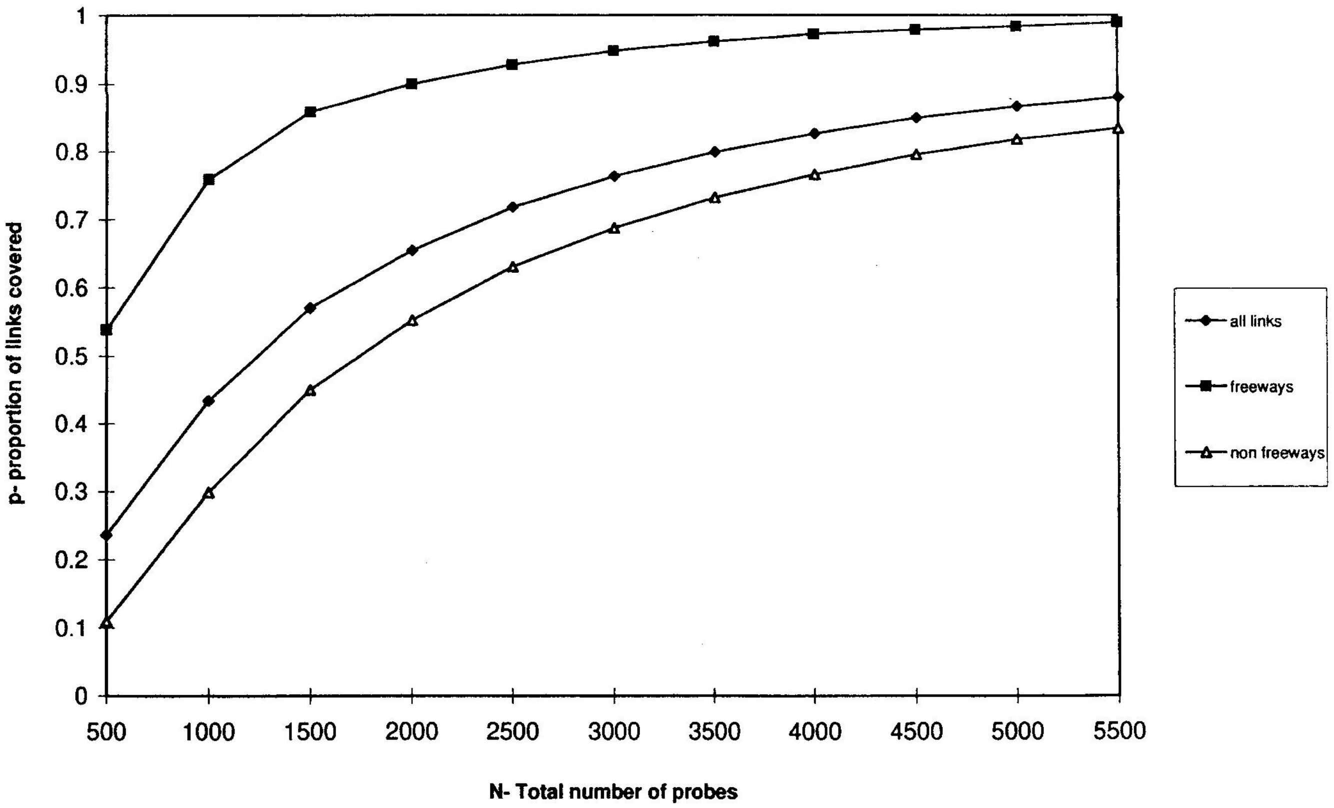

FIGURE 3 Plot of p versus N (measurement period = 10 min; n = 3).

All the variables are highly significant ($p < .0001$) and have very small standard errors of estimates, and the model appears to be a good fit (adjusted $R^2 = 0.83$). The diagnostic analysis indicated that the major assumptions of normality and heteroscedasticity are satisfied. The model, however, is likely to contain some amount of serial correlation because of both the randomization and assignment processes.

This model suggests that the variation between N and n is nonlinear and increasing but at a decreasing rate, which was not clear from the N-versus-n plots. The model is consistent with the plots, because it suggests a nonlinear and decreasing relationship between N and tu and a nonlinear relationship between N and p that increases at an increasing rate.

The simulation model was run for values of n of 4, tu of 10, n of 6, and tu of 15 (not used in the calibration) and various values of N to determine p values. The regression model applied to the prediction data set yielded a high R^2 value of 0.86. Thus, the empirical model is a good approximation of the simulation results. However, the constant is likely to be network specific and will be influenced by the percentage of freeways, arterials, and collector streets in the network.

The scenario-based analysis (i.e., for a range of p and tu values) can be used to select reasonable values of the measurement period (from N-versus-tu curves), proportion of link coverage (from the N-versus-p curve), and the number of probe vehicles required for a given reliability requirement (n). This also permits a sensitivity analysis of the estimate of the number of probe vehicles when design parameters are varied (in particular, the coefficient of variation could vary with time of day, season, etc.). Other networks of a similar scale can be simulated with appropriate n values.

Three simplifying assumptions in the simulation are the normality of travel time distribution, constant μ/σ ratios over all links, and constant μ/σ ratios over all measurement periods. Because a constant μ/σ ratio over time is used, it is considered that this ratio is the average ratio $[(\mu/\sigma)_{AVE}]$ during the peak period. This results in an n_{AVE} value that is between the upper and lower bounds for the reliability requirement. If the proportion of time that the μ/σ ratio is close to the average ratio is low, then the bounds of n values can be suitably weighted to get a better estimate of N. This assumption is therefore not very restrictive.

Let the μ/σ ratios be different for freeways and nonfreeways, resulting in N_f and N_{nf}, respectively. The lower value would result in an overestimate, whereas the larger value would result in an underestimate of the number of probes. On the basis of the proportion of freeway links to nonfreeway links, an appropriate weighting strategy could be used to obtain an estimate that is arbitrarily close to the actual number of probe vehicles necessary.

In the absence of actual travel time data it is not clear how the normal travel time distribution assumption would affect the simulation results. If the normal approximation is untenable, then the appropriate distribution (obtainable from traffic engineering studies or from ATMIS data covering only a portion of the network) could be used for the assignment procedure. The effect of this assumption is not likely to be serious with respect to N versus n and tu, since they are input parameters for the simulation. However, this assumption is likely to affect the relationship between N and p, because p is determined by the stochastic assignment procedure.

Therefore, it appears that the assumptions will not adversely affect the quality of the results. However, the assumptions used must be validated. The empirical model fits well, indicating that the qual-

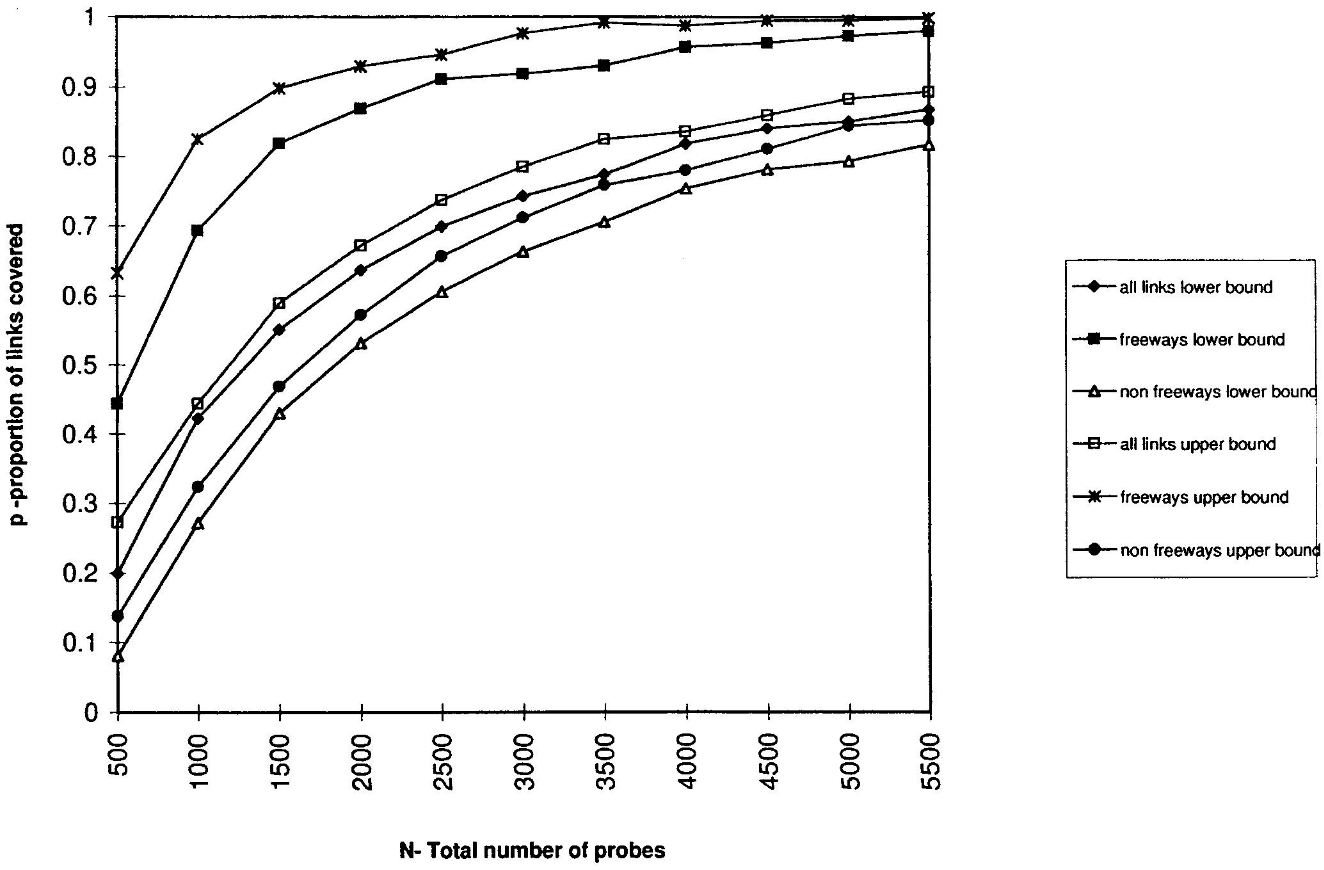

FIGURE 4 Ninety percent joint confidence intervals for p versus N plots (measurement period = 10 min; n = 3).

itative relationship between N versus n, p, and tu is reasonably valid. Because a range of n values between 1 and 10 vehicles per link per measurement period is used, the actual estimate is likely to be bounded by these endpoints. The use of reasonably approximate n values of about two to five would yield realistic estimates of the number of probes required.

IMPLICATIONS OF SIMULATION RESULTS FOR REAL-TIME TRAFFIC INFORMATION SYSTEMS

The system configuration required for monitoring the probes could be used to provide drivers with real-time information.

Many studies have noted that as the fraction of drivers with information increases, the efficiency of the system goes down. Mahmassani and Herman (*12*) observed that there is an optimal fraction of drivers who should receive route guidance advice. Mahmassani and Jayakrishnan (*13*) argue that about 20 to 25 percent of the drivers should be provided with information to have a maximum reduction in network travel times. It is therefore of interest to compare the estimate of the number of probe vehicles in terms of the optimal fraction of vehicles that should be instrumented. From the simulation, for a link coverage proportion of 80 percent with a 10-min measurement period and reliability requirement of three probe vehicles per link per measurement period, about 3,500 probe vehicles are required for the Sacramento network. This amounts to less than 5 percent of vehicles traveling the network during a 2-hr peak period and is well below the optimal percentage of drivers who should receive information. The provision of real-time route guidance information to all the equipped vehicles would not result in a significant deterioration of network performance in terms of travel time savings.

It is clear that the freeways require a smaller number of probe vehicles than the arterials for a given p. This is clear evidence of the concentration phenomenon discussed by Ben-Akiva et al. (*14*). Analogously, it can be inferred that the number of probe vehicles required in minor arterials and collector streets would be higher for a given link traversal proportion. Therefore, the use of vehicles as probes will not be effective for determining link travel times on minor arterials and collector streets. During the off-peak periods, with fewer vehicles and probes on the road, a greater proportion of probe vehicles would be required.

Even though the fraction of probe vehicles is small, the estimated number of probe vehicles could still produce congestion in communication networks. This effect of the probe vehicles on the communication network needs further investigation.

CONCLUSIONS

Probe vehicles have been proposed as an attractive source of real-time travel times. They can directly measure congestion, delays, and incident clearance times.

A heuristic algorithm has been developed for determining the number of probe vehicles required to estimate link travel times during a peak period, considering both reliability and adequacy requirements. This generic algorithm is the first to comprehensively consider design requirements, such as the proportion of links covered, the number of repeated observations of link travel time, and the length of the measurement period. It is also flexible because the choice of specific procedures such as route assignment and selection of probes from a pool of all trips is left to the user.

A simulation model was developed to test the algorithm for the Sacramento network. This model can be used in several other networks by suitably modifying the design parameters. The simulation is used to analyze the interrelationship between the design parameters n, tu, p, and N. The variation between N and n is nonlinear and increasing, but at a decreasing rate with increasing n. There appears to be a nonlinear and decreasing relationship between N and tu, and the relationship between N and p is nonlinear and increasing, but the rate decreases as N increases. The narrow confidence intervals for the plot between N and p indicate reasonably accurate point estimates. The plots between N versus tu and N versus p indicate the presence of knees, which can be used to identify trade-offs between the design parameters tu and p. Furthermore, with a given number of probes a substantially greater proportion of freeway links than arterials can reliably be covered.

It appears that a substantial number of probe vehicles would be required to estimate link travel times if all the classes of links including minor arterials and local and collector streets are to be covered adequately. However, the heavily traveled high-speed routes, such as the major arterials and freeways, require a much smaller fraction of probe vehicles. For a 10-min measurement period and an adequacy requirement that 80 percent of major arterials and freeway links be reliably covered with a minimum of three probe vehicles in each link during each measurement period, about 3,500 probe vehicles are required for a 2-hr peak period for the Sacramento network. These probe vehicles constitute less than 5 percent of the total peak period volume. If this fraction of drivers were provided with information, it would not result in a significant deterioration of network performance.

The use of vehicles as probes is an attractive option for collecting real-time information only in heavily traveled corridors and high-speed roads, such as freeways and major arterials, and during specific periods, such as the peak period. However, probes cannot be used as a stand-alone source of travel time information, especially during off-peak periods and on lightly traveled corridors and low-speed roads, such as local and collector streets and minor arterials. It is necessary to explore other sources of travel time data for these.

To minimize the cost of instrumenting the probe vehicles and to minimize possible problems from the simultaneous use of limited communication links, it is desirable to use as few probes as possible. This could be achieved by using the report by exclusion principle, in which the probe vehicles report only abnormally large deviations from the mean travel times. Further research is required to determine the effect of report by exclusion on the estimate of the required number of probe vehicles.

REFERENCES

1. Aplin, N. Overview of IVHS Activities in Australia. *IVHS America Annual Proceedings*, 1993, pp. 286–291.
2. Ligas, J., S. F. Bowcott, D. Boyce, P. Nelson, J. Schofer, F. Koppelman, and A. Kirson. An Update on ADVANCE. *IVHS America Annual Proceedings*, 1993, pp. 621–626.
3. Clelland, A., M. Wendtland, S. Tedesco, F. Mammano, and G. Endo. Evaluation Results and Lessons Learned from the Pathfinder Operational Test. *IVHS America Annual Proceedings*, 1993, pp. 448–457.
4. Barbaresso, J., and J. L. Grubba. Overview of the FAST-TRAC IVHS Program. *IVHS America Annual Proceedings*, 1993, pp. 630–638.
5. Longfoot, J. An Automatic Network Travel Time System (ANTTS). *Proc., Vehicle Navigation and Information Systems Conference*, Vol. 2, 1991, pp. 1053–1062.
6. Balke, K., W. McCasland, S. Z. Levine, and C. L. Dudek. Collection and Dissemination of Real-Time Travel Time and Incident Information with In-Vehicle Communication Technologies. *Proc., Vehicle Navigation and Information Systems Conference*, Vol. 1, 1991, pp. 77–82.
7. Boyce, D., A. Kirson, and J. Schofer. Design and Implementation of ADVANCE: The Illinois Dynamic Navigation and Route Guidance Program. *Proc., Vehicle Navigation and Information Systems Conference*, Vol. 1, 1991, pp. 415–426.
8. Boyce, D., J. Hicks, and A. K. Sen. In-Vehicle Navigation System Requirements for Monitoring Link Travel Times in a Dynamic Route Guidance System. Presented at 70th Annual Meeting of the Transportation Research Board, Washington, D.C., 1991.
9. May, A. D. *Traffic Flow Fundamentals*. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1990.
10. McShane, W. R., and R. P. Roess. *Traffic Engineering*. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1990.
11. Abdel-Aty, M. A., R. Kitamura, and P. P. Jovanis. Investigating the Effect of Travel Time Reliability Using Repeated-Measurement Stated Preference Data. In *Transportation Research Record 1435*, TRB, National Research Council, Washington, D.C., 1993, pp. 39–45.
12. Mahmassani, H., and R. Herman. Interactive Experiments for the Study of Trip Maker Behavior Dynamics in a Congested Network. In *New Developments in Dynamic and Activity Based Approaches*, 1990, pp. 272–298.
13. Mahmassani, H., and R. Jayakrishnan. System Performance and User Response under Real-Time Information in a Congested Network. *Transportation Research A*, Vol. 25, No. 5, 1991, pp. 293–308.
14. Ben-Akiva, M., A. DePalma, and I. Kaysi. Dynamic Network Models and Driver Information Systems. *Transportation Research A*, Vol. 25, No. 5, 1991, pp. 251–266.

Publication of this paper sponsored by Committee on Intelligent Transportation Systems.