Modeling Location Management in Personal Communication Services*

Derek Lam, Jan Jannink, Donald C. Cox, Jennifer Widom

Electrical Engineering & Computer Science Depts.

Stanford University, Stanford, CA 94305

October 18, 1995

Abstract

This paper presents a framework for modeling and evaluating the performance of location management schemes for a Personal Communication Services (PCS) network. The models take into account complex human behaviors, and are based on actual measurements of these behaviors. Real measurements are a crucial factor in achieving realistic performance evaluations. We have developed a powerful and highly configurable simulator based on the framework. Simulation results, showing the performance of an existing location management standard, IS-41, are presented in terms of database performance requirements (transactions per second) and network traffic load (messages per second).

1 Introduction

The goal of Personal Communication Services (PCS) [5] is to provide ubiquitous communication between users, independent of their locations and mobility patterns. The realization of this vision presents many challenging problems in network data management [8, 9, 25]. A key problem in this area is *location management*. Location management refers to accessing and maintaining user information for call routing purposes. Important per-user information, such as current location, authentication information, and billing information, are stored in *user profiles*. From an operational

*Research partially supported by the Center for Telecommunications and the Center for Integrated Systems at Stanford University, by the Anderson Faculty Scholar Fund, and by equipment grants from Digital and IBM Corporations.

perspective, location management relies on two functions: *profile lookups* and *profile updates*. A profile lookup occurs in any call between users:

- to access the caller's profile for authentication, and
- to access the callee's profile for location information and connection status.

A profile update may occur:

- to signal user equipment activation or deactivation,
- to signal user call connection, or
- to register user movement.

The performance of any location management scheme is a function of the underlying database architecture and the location management algorithms. Performance variables of interest are: profile lookup and update response times (a function of database loads and the amount of network signaling traffic), memory cost (a function of database size), and system equipment price. IS-41 [3, 18] and GSM [18, 19, 21] are the two location management standards for the current cellular telephony systems. They both use similar simple location management schemes. Previous studies [15] have shown that for projected numbers of PCS users, these schemes will incur a large increase in database loads over the current levels.

In recent years, many sophisticated location management schemes [1, 10, 11, 22] have been proposed to reduce profile lookup and update response times and signaling traffic. These methods utilize techniques such as data replication, caching, and hierarchical database schemes. It is beyond the scope of this paper to review all proposed schemes. However, it is important to note that the actual performance of these proposals depends strongly upon user behavior. For example, the merits of caching and data replication schemes are functions of user mobility and calling patterns. Furthermore, different schemes optimize their performance to different aspects of user behavior. As a result, accurate user behavior models are required when evaluating the performance of location management schemes.

Most studies in the literature have concentrated on theoretical analysis [2, 4, 20]. For tractability reasons, simple user behavior models were used. Some simulation studies have been done. Lo et. al. [15] estimated, both analytically and using Monte Carlo simulations, the database transaction volume required to support PCS based on certain assumptions about user calling patterns and a simple mobility model. This work gives good first-order estimates, but detailed simulations are required to obtain more accurate figures. Krishna et. al. [14] performed simulation studies on location management algorithms using simple call, mobility, and cost models. Their cost model gave performance figures only in relative units, which do not relate to real-life performance metrics for database and network loads.

In this paper, we present a framework for modeling and evaluating the performance of database architectures and location management algorithms. Our framework uses call and movement models aimed at capturing the complex patterns of user behavior. By analyzing actual call and vehicle traffic data, we have demonstrated the validity of our models and obtained realistic parameters for them. At the present time, large scale PCS networks have not yet been implemented. Our goal is to provide an accurate platform for evaluating the performance of various location management proposals and to assess their database transaction and network bandwidth requirements. We have developed a discrete event simulator, *Pleiades*, based on our framework. Simulation results showing the performance of an existing location management standard, IS-41, are presented in terms of database performance requirements (transactions per second) and signaling traffic load (messages per second). Even though in this paper we are focusing on PCS modeling, our framework can be adapted to model wireless data networks.

The rest of the paper is organized as follow. Section 2 gives a brief overview of PCS network architecture. Section 3 describes the IS-41 location management scheme. Section 4 presents our modeling framework. Section 5 shows how we obtain realistic model parameters by analyzing actual call and vehicle traffic data. Section 6 describes the architecture of our simulator. Section 7 presents simulation results for the IS-41 standard.

2 PCS Network Architecture

In PCS, mobile users communicate through portable handsets. The basic network architecture is shown in Figure 1. (Please note that all figures are included at the end of the paper.) It consists of a set of *radio ports* connected to a fixed wireline network through *mobile switching centers* (MSC). Radio ports are the communication service points for the portable handsets within their coverage areas. MSC's are the hardware interface between a group of radio ports and the wireline network. Databases of user profiles are also connected to the wireline network. The detailed operations of these databases and the connectivity between them are determined by the location management scheme. User communications take place by establishing call connections through radio ports, MSC's, and the wireline network.

3 Location Management in IS-41

IS-41 requires profiles to be stored permanently in a home database or *Home Location Register* (HLR). When a user moves out of the region maintained by his HLR, a copy of the user's profile is created in a visiting database or *Visiting*

Location Register (VLR) in the user's current region. In addition, the profile in the HLR is updated to reflect the user's current location. Profile lookups occur when a caller initializes a call to a callee. The caller's profile can be obtained from the local HLR or VLR. The caller's MSC searches for the callee's current location by first querying the local VLR. If the callee is not currently served by this VLR, then a location request is sent to the callee's HLR which might be at the same site. (A user's HLR is a function of the user identification number.) If the callee is not currently at home, the HLR needs to query the VLR at the callee's current location for the necessary location/routing information. Figure 2 shows the events that occur during the search for the callee's current location. It applies to the scenario where the callee is not at home and different VLR's are serving the caller and the callee. Figure 3 shows the profile update events that occur when a user moves into a region maintained by a new VLR.

4 Modeling Framework

We have developed an object-oriented modeling framework for realistic performance evaluations of sophisticated location management schemes. Our framework is divided into the following components: **Basic Topology Model**, **Location Management Modules**, **Call Model**, and **Movement Model**. The Basic Topology Model specifies the geographical and network topologies independent of location management schemes. Each Location Management Module implements a specific location management scheme using the structures defined in the Basic Topology Model. Call and Movement Models characterize call and movement behavior of individual users, respectively.

4.1 Basic Topology Model

The Basic Topology Model is composed of the following objects.

User represents a human user. A user object contains information describing the user's' current geographical location, the user's home location, and the database(s) currently containing a copy of the user profile. Location management modules use this information to perform profile lookups and updates. The user object also maintains user statistics in terms of number of calls and moves made by the user.

Site represents a geographical area. All Site objects together define the physical geography for user movements. A Site usually corresponds to the area covered by one profile database. A Site maintains statistics relating to number of site users and their combined call and movement activities.

Database represents any form of user profile database (HLR, VLR, etc.). A Database is often associated with a Site. (Exceptions are the higher-level databases of a hierarchical database architecture.) Each database object maintains access statistics relating to number of database reads and writes, database messages sent, and total cost of sending all database messages (in hop count).

Link represents a direct communication link between two databases. It has a link cost describing the cost of sending a message through it. It maintains traffic statistics in terms of number of messages.

Geographical topology is defined by a *movement connectivity matrix* which specifies, for each Site, its neighbors and the probabilities of users crossing into each of them. Database network topology is derived from the Link objects. A fixed routing matrix is calculated to describe the minimum-cost communication path[1] between any pair of databases. During the simulation, database access and link statistics along this path are updated.

4.2 Location Management Modules

Each location management module performs profile lookups and profile updates specific to a particular location management scheme. Each module updates information in user objects and all other statistics according to the call and movement activities of the users and the particular scheme involved.

We have implemented location management modules to support the following existing and proposed location management schemes: HLR/VLR (IS-41 and GSM), distributed home database, centralized database, full replication, as well as hierarchical extensions to HLR/VLR. As mentioned earlier, in this paper we focus on modeling aspects so we consider only IS-41. The other schemes are discussed in [12].

4.3 Call Model

Our call model generates call traffic for each individual user. We have adopted a model that takes into account different classes of calls (e.g., short, long, retry), and real life behaviors such as a user calling more frequently a specific group of people. The model is divided into two parts: the Call Traffic Model and the Callee Distribution Model.

[1] The cost of a path is the sum of all immediate link costs along the path.

4.3.1 Call Traffic Model

The call traffic model describes how often individual users place calls to other people and characterizes the duration of each call. Extensive studies have been done on traffic in fixed telephone networks. However, very little is known about the traffic characteristics of future PCS networks. On fixed telephone networks, traffic is modeled very accurately by a Poisson arrival process and exponentially distributed call durations. Our call traffic model is an extension of this basic model to PCS. It generates call arrivals (i.e., calls initiated) for different classes of traffic and models changes to user behavior through time, as we now explain.

At any instant, the call traffic model associates a *CallType* with each user. Each CallType completely describes the call traffic behavior of a user for that instant. The CallType contains a parameter specifying the probability for a user to become unreachable due to turning off his handset and a list of *Call* objects to specify call traffic characteristics. Each Call corresponds to a class of call traffic. It is defined by the probability of occurrence of the traffic class, its call arrival rate, mean call duration and distribution, and an attribute field to indicate retry calls. (Retry calls are calls that a caller will retry if the original call is undeliverable.)

Changes to calling behavior, such as those between daytime and evening calling patterns, are modeled by changing CallType's as a function of simulation time. Each CallType has parameters that describe the length of its "active" period and the next CallType in the behavioral sequence. This feature is useful in investigating performance of location management schemes under daily cyclical load patterns as well as transitions into busy periods.

4.3.2 Callee Distribution Model

The callee distribution model characterizes how the callee is generated for each call. It is an important modeling issue because of its effects on performance evaluation, especially for schemes with caching or data replication. In the past, public telephone service providers were able to indirectly model callee distributions by measuring the cumulative call traffic into each service area. However, for PCS this approach no longer provides sufficient information because of user mobility.

We have developed a callee distribution model that models the behavior of each individual caller. It accounts for such real life behaviors as users calling a group of people (e.g., business associates, family, friends etc.) more frequently. In our model, each user is associated with its own callee list. When a call is generated for a user, the callee is selected either randomly from all users or from among the user's callee list according to a specified probability distribution. Various callee distributions can be simulated by different initializations of the callee lists and their probability distributions. Using actual measurement data, we have investigated various issues in setting up the callee lists. These results are discussed in detail in Section 5.1.2.

4.4 Movement Model

Our movement model characterizes user movements within the geography defined by the Basic Topology Model. Recall from Section 4.1 that the Basic Topology Model defines, for each site, its neighbors and the probabilities of crossing into them. A simple Markovian model could have been used to generate user movements based on the crossing probabilities (as in [4]). Instead, we have developed a more detailed model, which includes the Markovian model as a special case, because most proposed location management schemes optimize their performance for certain movement characteristics. For example, *return home* movements are important when studying schemes with home location registers (such as the current standards IS-41 and GSM). Similarly, *roundtrip* movements are important when studying schemes with user profile replication. Our model generates user movements with varying distances and velocities and includes roundtrip and return home movements.

Analogously to CallType's in the call model, at any instant each user is assigned a *MoveType* to characterize his movement behavior. A MoveType is composed of a number of *Move* objects. Each Move models a particular class of movement behavior by defining its probability of occurrence, mean movement velocity and distribution, mean number of site crossings and distribution, and its move attribute. The move attribute indicates whether a Move is a *simple move*, *roundtrip move*, *return home move* or a *stationary move* (i.e., no move). Simple moves refer to moves in which users roam around the geography according to the crossing probabilities. The only restriction is that, for moves with more than one site crossing, the users do not return to the same site between adjacent moves. (In other words, there are no movement loops of length two; these moves are modeled by roundtrip moves.) Roundtrip moves are moves in which the starting and the destination sites are the same. Return home moves return users to their home locations stored in user objects via the most direct path.

User behavioral changes are supported using the same technique as in the Call Model. Each MoveType has parameters that describe the length of its "active" period and the next MoveType in the behavioral sequence. By changing a user's MoveType as function of simulation time, we are able to simulate changes in user movement patterns, for example between morning rush hours and afternoon periods. Through survey and actual measurement data, we have investigated velocity and distance distributions of user movements. These results are discussed in Section 5.2.

5 Model Parameters

To achieve accurate modeling of user behavior, we need to use sensible parameters to set up our call and movement models. We have investigated current human behavior by analyzing actual call and vehicle traffic data. Based on these results, we were able to derive realistic model parameters for our simulations.

5.1 Call Model Parameters

Distributions of call arrival rates and call durations in fixed telephone networks are well documented in the literature. [15] reports that for current telephone usage, the mean call arrival rate and mean call duration during the busy hours are about 2.8 calls/hour/telephone set and 2.6 min/call, respectively. These results provide guidelines to us when specifying call arrival rates and call durations in our model.

Traffic volume patterns over a period of time, e.g., days or weeks, and callee distributions are important components of our call model. Unfortunately, these statistics are not well documented. We have obtained call traffic data from our local university telephone exchange to study existing behavioral patterns. While these data are not for actual PCS traffic, they provide to us insights into existing communication patterns. It is important to characterize traffic volume patterns, e.g., as a function of time of day, because we need to use associated mobility patterns for performance evaluations. The importance of characterizing callee distribution is obvious and, to the authors' knowledge, there has not been any previous research in this area.

5.1.1 Patterns in Call Traffic Volume

We have analyzed call traffic data [24] corresponding to a six-month period from our local university telephone exchange. This exchange serves the entire campus including university offices, student housing, and residential households. The data obtained has been preprocessed to protect the anonymity of the callers and callees. The data set contains caller identification (encrypted), callee identification (encrypted), time of call, and call duration for each call destined to locations outside of the local exchange. In this data set there are a total 19,592 distinct callers. While this constitutes a specialized sample of call traffic data, the particular composition of our campus allows us to infer to call traffic patterns in both business and residential settings.

Figure 4 shows average call arrival volume patterns for a full 24 hour period. Each arrival value for a particular time is obtained by averaging over their corresponding values over each of the days in our data set. We have examined averages over all days, weekdays (Mondays – Fridays), and weekends (Saturdays and Sundays). We observe that there are essentially three periods of call activity during a typical weekday. The first corresponds to the late night period (12 a.m.–7 a.m.) when there is very little activity. The second is the peak period which spans the regular business hours (8 a.m.–4 p.m.). The last period corresponds to the off-peak period during the evening hours. One observation is that volume changes abruptly during the morning transition, but the evening transition is much more gradual. We also observed that weekend volume is significantly less than weekday. This is clearly due to the lack of business activity.

5.1.2 Callee Distributions

We model callee distributions by maintaining, for each caller, a list of the people they call most often with the probabilities of making calls to each of them. To assign probabilities, we need empirical information on the distributions of these probabilities and how they depend on the frequency rank of callees.

We have investigated empirical callee distributions according to *callee rank*. Callee rank is defined over a reference time period and is specific to a (caller,callee) pair. The rank k callee of a caller is the caller's k^{th} most frequently called person within the period. We have segmented our six months of call data into datasets according to different reference time periods. For example, to characterize callee distributions over a one month period, we divide our data into six one-month datasets. We have performed analysis for reference time periods of one day, one week, and one month (4 weeks). From each dataset j, we obtain, $f^i_{r,j}$, the relative frequency of caller i's calls destined to the rank r callee, by:

$$f^i_{r,j} = \begin{cases} \frac{C^i_{r,j}}{C^i_{tot,j}} & i = 1, \ldots, N_C, \quad j = 1, \ldots, D, \quad C^i_{tot,j} > 0 \\ 0 & C^i_{tot,j} = 0 \end{cases} \tag{1}$$

where N_C is the number of distinct callers from all datasets, N_D is the number of datasets for the current reference time period, $C^i_{r,j}$ is the number of calls made by caller i to his rank r callee in data set j, and $C^i_{tot,j}$ is the total number of calls made by caller i in data set j. Then we form an estimate of caller i's probability of calling his rank r callee, $\hat{P}^i_r$, by averaging over the datasets:

$$\hat{P}^i_r = \frac{1}{V^i} \sum_{j=1}^{N_D} f^i_{r,j} \tag{2}$$

where V^i is the number of datasets in which $C^i_{tot,j} \neq 0$. We calculate the mean call probability to the rank r callee, P_r, by averaging over $\hat{P}^i_r$. Figure 5 shows the relationship between P_r and callee rank for the three reference periods. Figure 6 is a log-log plot of P_r vs. callee rank. We observe that the mean probabilities obey a power or generalized

Time Scale	A	p	mean square error
1 Day	0.770	2.61	0.000014
1 Week	0.556	1.79	0.000039
1 Month	0.388	1.35	0.000024

Table 1: Fitted Power Law Parameters

Zipf's law at all three reference time periods:

$$P_r \simeq \frac{A}{r^p} \tag{3}$$

where A is the scaling parameter and p is the exponent parameter. A power relationship shows up as a straight line on a log-log plot. We have fitted our results to power laws. Table 1 shows the fitted parameters and the mean square errors of the fits. From our results, we found that callee list lengths of 5, 8, and 11 are sufficient to account for 90% of the calls over the time periods of one day, one week, and one month, respectively.

When initializing our callee lists, we would also like to know the distributions around the mean call probabilities. This is useful because we want to include in our simulations users that deviate from the "average" behavior. For example, there should be users who make calls only to one callee (i.e., $P_1 = 1.0$). Figure 7 shows the distributions of $\hat{P}_1^i$ for the three reference periods. We found that each empirical distribution can be approximated by a normal distribution with the same mean and variance. Figure 8 shows the cumulative distributions of $\hat{P}_1^i$ and their fits to normal distributions.

For the higher rank call probabilities, we look at the relative probabilities to the rank r callee, $\frac{\hat{P}_r^i}{\hat{P}_{r-1}^i}, r > 1$. By characterizing the relative probabilities, a recursive technique can be used to initialize the call probabilities in a callee list. We found that the distributions of $\frac{\hat{P}_r^i}{\hat{P}_{r-1}^i}$ can also be approximated by normal distributions. Figures 9 shows graphically the cumulative distributions of relative call probabilities for a few higher rank cases and their fits to normal distributions.

We have implicitly assumed in our call model that callee distributions are not dependent on call arrival characteristics. We have verified this assumption by observing that low correlation exists between callers' average call arrival rates and their observed call probabilities. Table 2 summarizes this result in terms of correlation coefficients between average call arrival rates of users (over the reference time period) and their respective call probabilities.

Time period	$\hat{P}_1$	$\frac{\hat{P}_2}{\hat{P}_1}$	$\frac{\hat{P}_3}{\hat{P}_2}$
1 Month	0.0532	0.0162	-0.0350
1 Week	0.2033	-0.1233	-0.1995
1 Day	0.1472	-0.1579	-0.1393

Table 2: Correlation between Average Call Arrival Rates and Call Probabilities

5.2 Movement Model Parameters

We have also investigated actual user movement behavior. The reports [7, 13] describe personal transportation behavior in the United States. based on a survey conducted throughout 1990. [6] contains vehicle traffic statistics for Europe. In addition, we have obtained actual movement statistics from [16] to correlate with survey statistics.

We have investigated time-of-day traffic volume patterns using data from [7, 16]. Figure 10 shows a summary of the results. In the figure, the patterns "survey commuter" and "survey non-commuter" are derived from [7] and the measured pattern "S.F. commuter" is from [16]. Both measured and surveyed data suggest that commuter movement patterns can be modeled using five MoveTypes. The first MoveType models the late night low activity period between midnight and 6 a.m. The second one models the morning rush hour, 6–9 a.m. The third one models daytime roaming activities during the business hours between 9 a.m. and 4 p.m. The fourth MoveType is for evening return home commutes that occur between 4 p.m. and 7 p.m. The fifth MoveType handles evening activities between 7 p.m. and midnight. The figure also suggests that non-commuters can be modeled essentially by three MoveTypes, corresponding to a late night low activity period (midnight–6 a.m.), a day time high activity period (6 a.m.–6 p.m.), and a evening moderate activity period (6 p.m.–midnight).

From the same sources we have obtained statistics relating to mode of transportation, travel distance, and travel time statistics for various movement types and their percentages of occurrence. In our movement model, we keep track of distances in number of site crossings. We estimate our distance parameters by dividing average travel distance with site diameter. We have derived Table 3 using the data in [7] and a site diameter of five miles. It describes the percentage of occurrence for each category of movements, and its average move distance and velocity. We have used these numbers to specify movement behavior of in our movement model. In particular, we based commuter movements during the morning and evening rush hours using the "To/From Work" figures. "Work-Related Business" figures are part of the movement characteristics of commuters during the business hours. "Personal Business" and

Trip Purpose	Percentage of Trips	Average Trip Length (miles)	Average Trip Length site crossings	Average Velocity mi/hr	Average Velocity sites/hr
To/From Work	20.2	10.65	2	31.3	6.3
Work-Related Business	1.4	28.20	6	81.3	16.3
Personal Business	52.9	6.74	1	28.3	5.7
Social/Recreation/Other	25.3	11.53	2	39.2	7.8
Vacation	0.2	218.22	(44)	261.5	(52.3)

Table 3: Movement Statistics

"Social/Recreational/Other" statistics are used to specify non-commuter movements, and commuter movements during the non-business hours. Vacation statistics need special attention because mobile users will most likely turn off their equipment during such a move. We have omitted these statistics from the metropolitan area simulation described in Section 7. This omission will not affect our results since we are simulating a metropolitan area and the volume of this type of move is low.

6 Simulator

Our discrete event simulator is based on the modeling framework described above. It is written in C++ with approximately 5000 lines of code. Each location management module is implemented in about 100 lines of code.

Our simulator architecture is that of a discrete event handler, centered around an earliest time first priority queue. Figure 11 depicts the architecture of the simulator. Initially, the simulator constructs the geographical and network topologies in the Basic Topology Model according to a simulation script. The simulator then establishes a population of users, each of which has a specified movement model and call model, from a set also defined in the simulation script.

Having constructed the simulation geography, the simulator enters a cycle of event generation and event application. The event generator generates call and move events according to the length of a time window, the population size, and the level of activity. Each event is enqueued at its proper time. Once a full window of events has been enqueued, the event handler dispatches and transcribes them into simulation statistics according to a specified location management scheme. The statistics generated relate to the movement and calling activities of the users at each time

Area Code region	Counties	1990 Population	Simulated subregions	Border subregions
North Bay (707)	Solano, Napa, Sonoma	839,408	8	4
Peninsula (415)	S. F., San Mateo	1,603,978	14	7
East Bay (510)	Alameda, Contra Costa	2,082,914	17	7
South Bay (408)	Santa Clara, San Jose	1,497,577	11	2
Totals	9 counties	6,023,877	50	NA

Table 4: Population Figures

window. The statistics are collected in an output file for analysis at the end of the simulation.

7 Simulation Results

Using the framework and the simulator described above, we have investigated the performance of the IS-41 standard in terms of database and network signaling loads. We performed our simulation on a geography that models the San Francisco Bay Area, which is composed of four area codes. The transportation behavior in the area is such that there is a significant amount of traffic between these area codes, particularly during rush hours. Figure 12 is a map of the Bay Area. Regions corresponding to different area codes are represented by different shades in the figure; bridges, ferries and public transportation are also included. Figure 13 [23] is an overlay map that shows the relationship between our simulation model and the actual geography of the Bay Area. Site objects for the Basic Topology Model are represented by polygons in the figure. Databases and their network topology are represented by data centers and their corresponding networks.

In our simulation, Site objects were used to represent urbanized land areas. We specified connectivities between Sites to model actual geographical transportation routes such as highways, bridges and roads. Using actual traffic volume statistics from [16], we estimated movement between area codes (see Table 5). We then fine-tuned our geographical connectivity parameters to produce similar large scale movement behavior. Movements on a small scale, for example between adjacent sites, are handled by specifying movement classes as described in Section 5.2.

We simulated a 24-hour period for a population of 3,025,000 Users using a HP 9000/755/99 workstation with 512 MB RAM in an UNIX environment. We populated Sites with Users according to census information from [17]. Table 4 is calculated using data from [17] and it shows the population in the four area codes. We divided our

Border crossing	Formula for traffic estimate between area codes	Peak	Total	pop. %
707 ⇒ 415	.5*(North Bay ferries) + .725*(Golden Gate bridge & ferries)	19,200	47,500	5.6%
707 ⇒ 510	2*(Richmond bridge)	12,300	48,500	5.7%
510 ⇒ 415	Richmond bridge + Bay bridge + San Mateo bridge + Dumbarton bridge + BART + .25*(Golden Gate bridge & ferries)	106,500	318,500	15.2%
510 ⇒ 408	San Mateo bridge + Dumbarton bridge	27,500	63,000	3.0%
415 ⇒ 408	Bay bridge + Caltrain	45,000	140,000	9.3%
Totals	$(707 \Rightarrow 415) + (707 \Rightarrow 510) +$ $(510 \Rightarrow 415) + (510 \Rightarrow 408) + (415 \Rightarrow 408)$	210,500	617,500	10.2%
Non-Peak Dirs.	.667*Totals	140,000	416,000	6.8%

Table 5: Movement Estimates

population of Users into 43% "commuters" and 57% "non-commuters". The proportion between the two groups is derived from the national average in [7] and the peak-to-total traffic figure for the Bay Area in Table 5. We specified the calling and movement behaviors of each group according to the models previously discussed. In Figure 10, we have shown movement traffic recorded from our simulation. The figure shows traffic patterns for all Users and for "commuters" alone. It can be seen that both patterns agree with surveyed and measurement data. Also, we have setup parameters in our movement models so that the total number of area code crossings in our simulation matches the number suggested by Table 5 to within 0.5%.

Figure 14 shows database read and write activities in transactions per second (TPS) throughout the simulation. Each access rate value on the plot is calculated from statistics collected from a fifteen minute simulation window. Since there is over an order of magnitude difference between database read and write activities, a log scale is used on the y-axis to aid in comparing their relative levels. In our results, we observed peak rate for lookups at 4,761 TPS, for updates at 743 TPS, and for their combined total at 5,310 TPS. These peak rates occured at 12:25 p.m., 3:15 p.m., and 1 p.m. in our simulated day, respectively.

Figure 15 shows network activities for location management. It shows the number of database messages sent per second and the total number of hop counts per second throughout the simulation. Total number of hop counts is defined to be the total number of network hops travelled by all messages within a time window. Again, the values in the figure are calculated from statistics collected every simulation window. The units on the y-axis is operations per second with which we use to refer to both the number of messages per second and the number of hop counts per second. From our results, we observed a peak signaling bandwidth of 4,408 messages per second and 12,950 message-hop per second at 1 p.m. in our simulated day.

8 Conclusions

We have presented a framework for modeling and evaluating the performance of location management schemes in PCS. The framework achieves its accuracy through realistic user behavior models that are based on measurements and surveys of actual human activities. A software system, capable of simulating a large population of users, was developed based on this framework. We have demonstrated the functionality of our software through a detailed simulation of the San Francisco Bay Area. At a time when large scale wireless communication networks have not yet been implemented, our work provides a practical way to evaluate the performance of various location management proposals and to assess database transaction and network bandwidth requirements for providing PCS. From the simulation results, we have shown that these requirements can be high for the simple location management scheme, IS-41. These results suggest that research on more efficient location management schemes is necessary.

Our framework and simulation software are not limited to PCS environments. They can be adapted to model wireless data networks by configuring the various model parameters and developing new modules to perform location management related functions in the corresponding data protocols.

9 Research Directions

The performance and requirements of various location management schemes are under investigation using our simulator. Some particular issues of interest are the performance of caching and data replication schemes under realistic user behavior conditions. Our simulator will serve as the platform for developing and testing more efficient location management schemes.

10 Acknowledgments

We would like to thank Susan Phillips and Stanford Communication Services for providing the call traffic data and Nesrin Basoz for the vehicle traffic data. We also thank the other members of the Pleiades group — Bora Akyol, Wiburt Labio, and Narayanan Shivakumar — for numerous useful discussions.

References

[1] V. Anantharam, M. L. Honig, U. Madhow, and V. K. Wei. Optimization of a database hierarchy for mobility tracking in a personal communications network. *Performance Evaluation*, pages 287–300, 1994.

[2] B. Awerbuch and D. Peleg. Concurrent online tracking of mobile users. *Computer Communication Review*, 21(4):221–33, 1991.

[3] EIA/TIA IS-41.3 (Revision B). *Cellular Radiotelecommunications Intersystem Operations*, July 1991.

[4] A. Bar-Noy and I. Kessler. Mobile users: To update or not to update? In *Proceedings of INFOCOM '94 Conference on Computer Communications*, volume 2, pages 570–576, June 1994.

[5] D. C. Cox. Wireless personal communications: What is it? *IEEE Personal Communications*, pages 20–35, April 1995.

[6] ECMT. *European transport trends and infrastructural needs.* OECD publications, 2 Rue Andre-Pascal, 75775 Paris CEDEX 16, France, 1995.

[7] P.S. Hu and J. Young. 1990 NPTS databook: Nationwide personal transportation survey. Technical report, Federal Highway Administration, November 1993.

[8] T. Imielinski and B.R. Badrinath. Data management for mobile computing. *SIGMOD Record*, 22(1):34–39, March 1993.

[9] B. Jabbari, G. Colombo, A. Nakajima, and J Kulkarni. Network issues for wireless communications. *IEEE Communications Magazine*, 33(1):88–99, January 1995.

[10] R. Jain and Y.-B. Lin. An auxiliary user location strategy employing forwarding pointers to reduce network impacts of PCS. In *Proceedings of ICC '95 — International Conference on Communications*, pages 740–744, 1995.

[11] R. Jain, Y.-B. Lin, C. Lo, and S. Mohan. A caching strategy to reduce network impacts of PCS. *IEEE Journal on Selected Areas in Communications*, 12(8):1434–44, October 1994.

[12] J. Jan. Data management for user profiles in wireless communications systems.

[13] R. Kitamura. *Time-of-Day Characteristics of Travel: An Analysis of 1990 NPTS Data*, chapter 4. Federal Highway Administration, February 1995.

[14] P. Krishna, N. H. Vaidya, and D. K. Pradhan. Location management in distributed mobile environment. In *Proceedings 1993 IEEE Workshop on Advances in Parallel and Distributed Systems*, pages 83–88, October 1993.

[15] C. N. Lo, R. S. Wolff, and R. C. Bernhardt. An estimate of network database transaction volume to support universal personal communications services. In *First International Conference on Univeral Personal Communications, ICUPC '92*, pages 09.03/1–6. IEEE, 1992.

[16] Metropolitan Transportation Commission. 1990 commute summary. Lotus 123 format spreadsheet, August 1990. S.F. bay area transportation.

[17] UpClose Publishing. 1990 census demographic summaries. http://www.upclose.com/upclose/, 1994. 1990 population figures for nine counties in S.F. bay area.

[18] S. Mohan and R. Jain. Two user location strategies for personal communications services. *IEEE Personal Communications*, pages 42–50, First Quarter, 1994.

[19] M. Mouly and M.-B. Pautet. *The GSM System for Mobile Communications*. Palaiseau, France, 1992.

[20] G.P. Pollini, S Meier-Hellstern, and D. J. Goodman. Signaling traffic volume generated by mobile and personal communications. *IEEE Communication Magazine*, pages 60–65, June 1995.

[21] M. Rahnema. Overview of the GSM system and protocol architecture. *IEEE Communications Magazine*, pages 93–100, April 1993.

[22] N. Shivakumar and J. Widom. User profile replication for faster lookup in mobile environments. In *To appear in 1st ACM International Conference on Mobile Computing and Networking, Berkeley, CA.*, 1995.

[23] U.S. Geological Survey. Urban Land Use for Central California, 1990, 1995. http://geo.arc.nasa.gov/usgs/images/urban2.gif.

[24] Telephone call data set, 3/95–9/95, October 1995. Encrypted ID numbers.

[25] P. E. Wirth. Teletraffic implications of database architectures in mobile and personal communications. *IEEE Communications Magazine*, pages 54–59, June 1995.

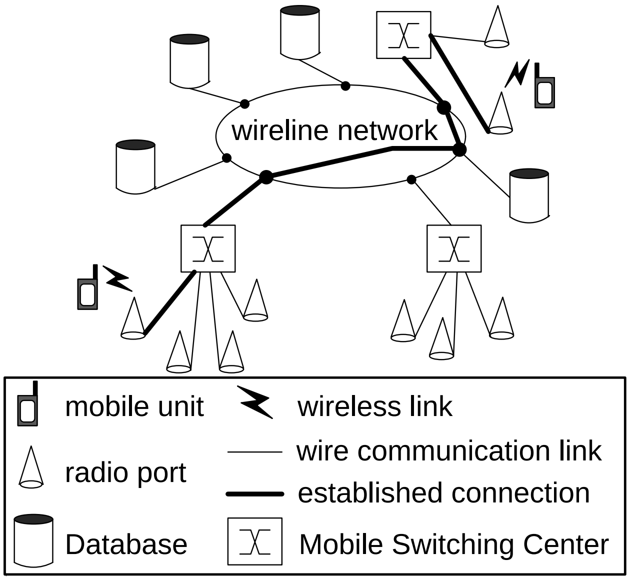

Figure 1: PCS Network Architecture

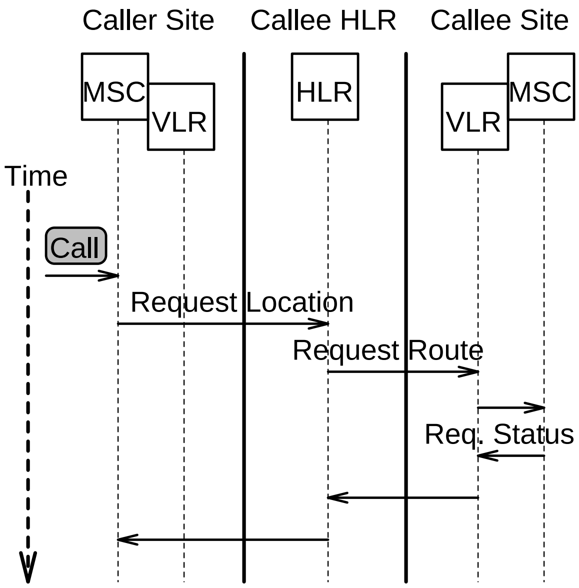

Figure 2: IS-41 Callee Profile Lookup

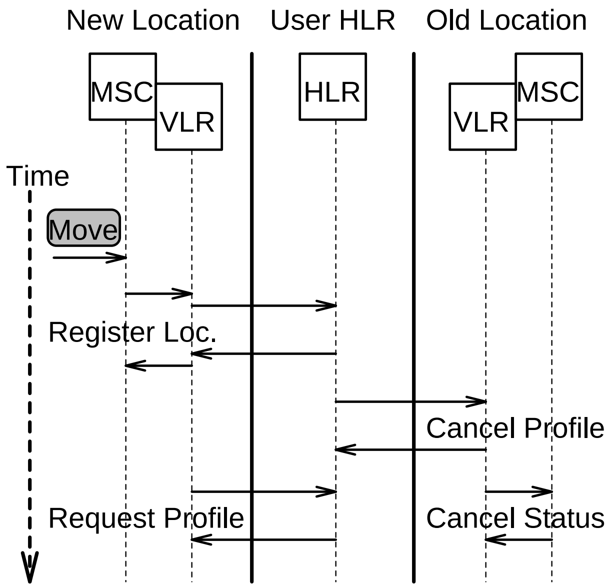

Figure 3: IS-41 User Profile Update

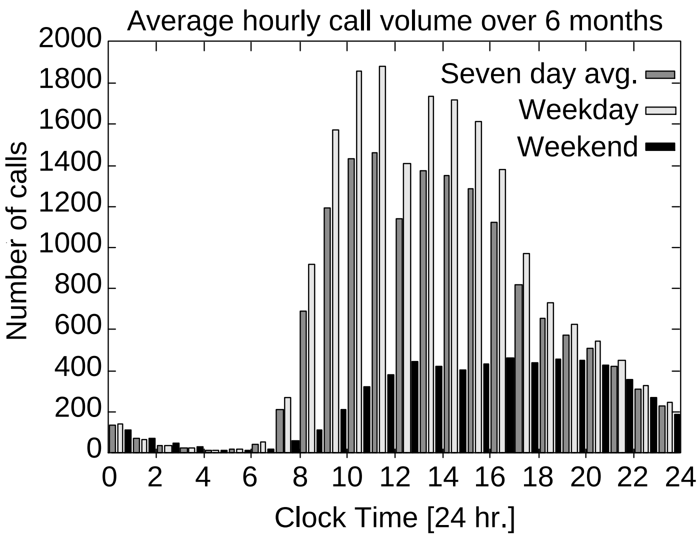

Figure 4: Average Number of Call Arrivals Per Hour

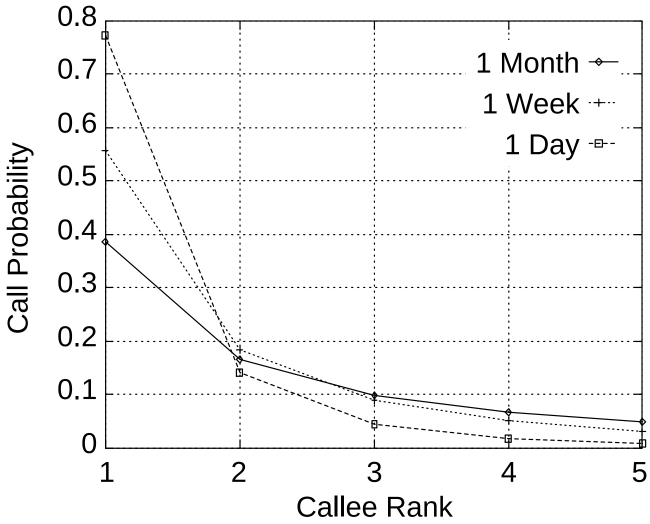

Figure 5: Mean Call Probability vs. Callee Rank

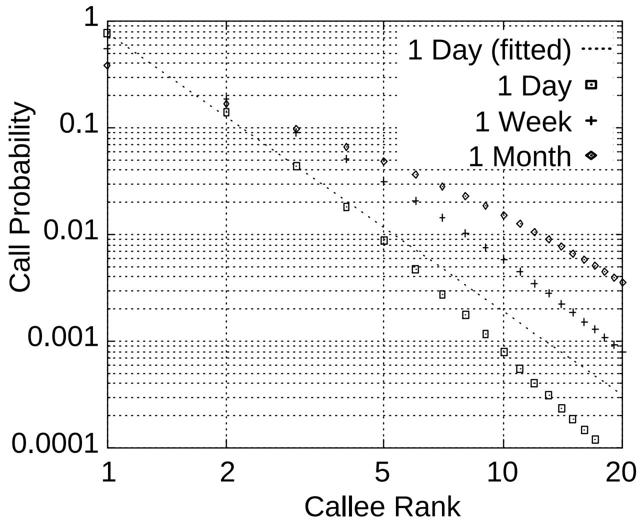

Figure 6: Log-Log Plot of Mean Call Probability vs. Callee Rank

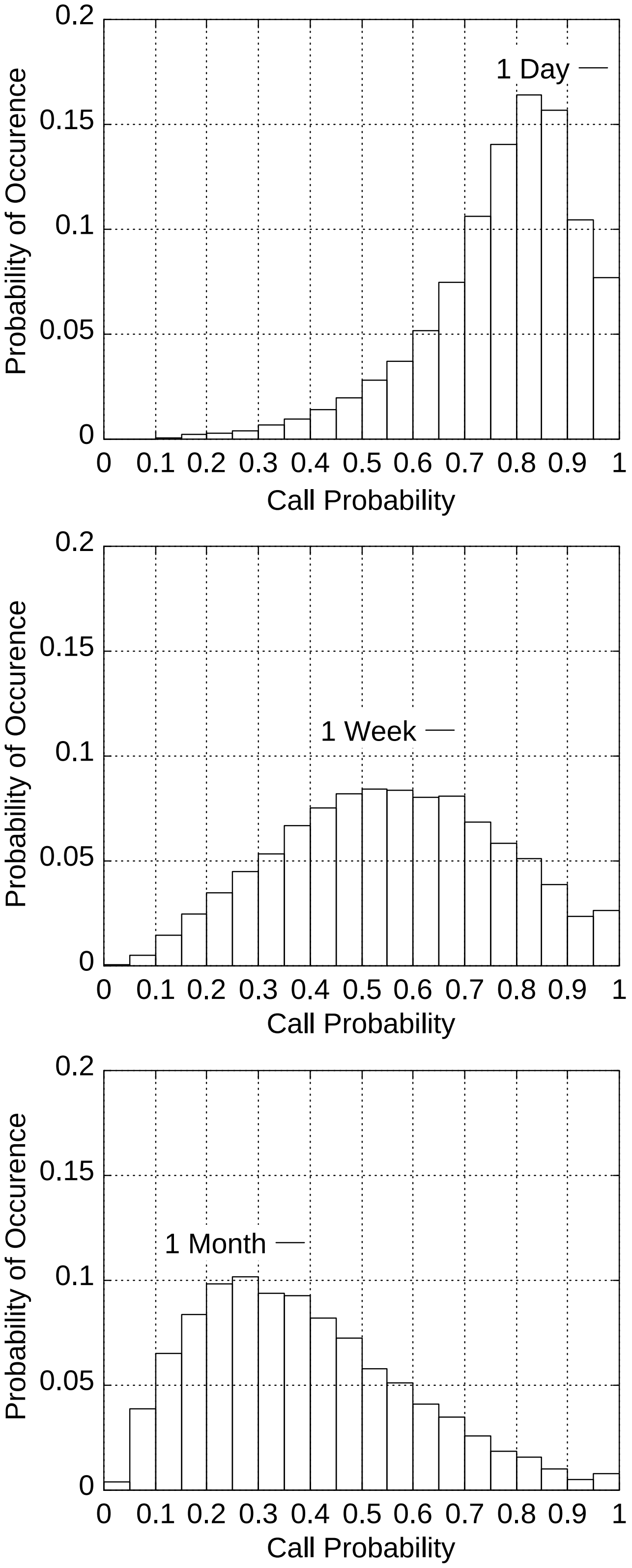

Figure 7: Distributions of Call Probabilities to First Rank Callee (for three reference time periods)

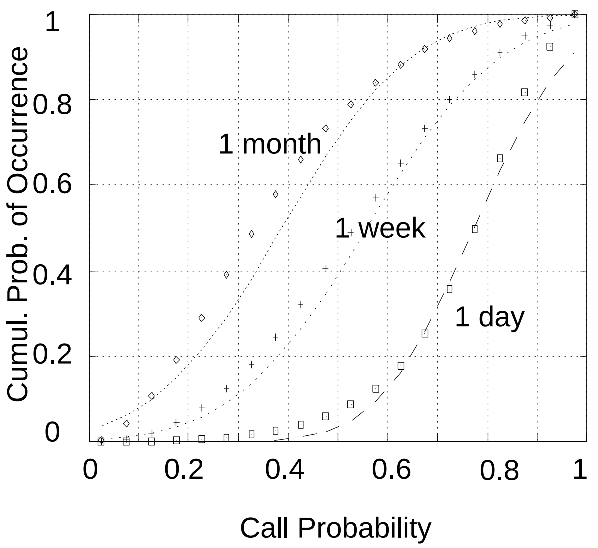

Figure 8: Cumulative Distribution of Call Probabilities to First Rank Callee

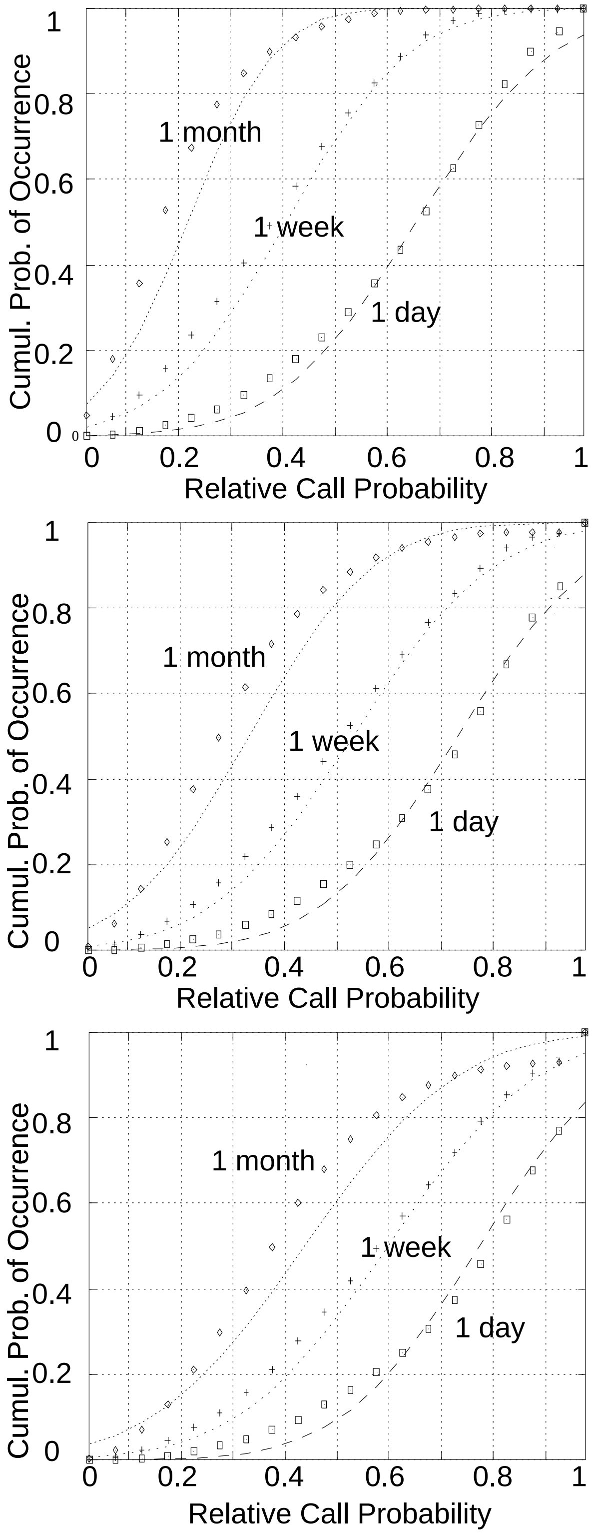

Figure 9: Cumulative Distribution of Relative Call Probabilities to Higher Rank Callees

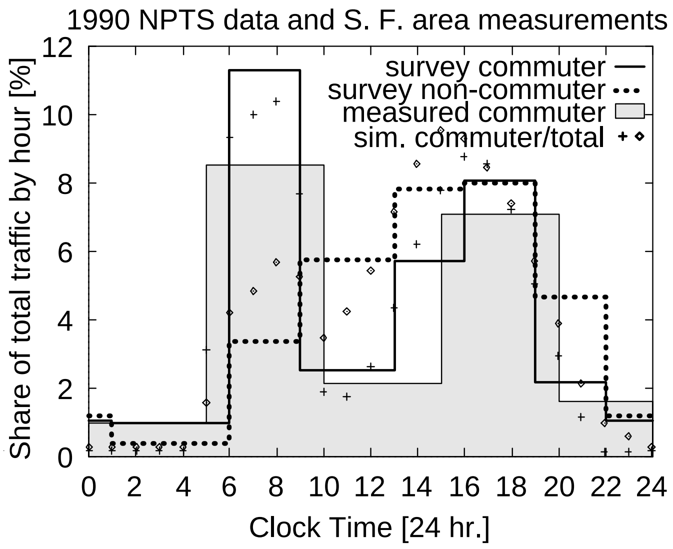

Figure 10: Estimated Commuter Pattern from Traffic Measurements

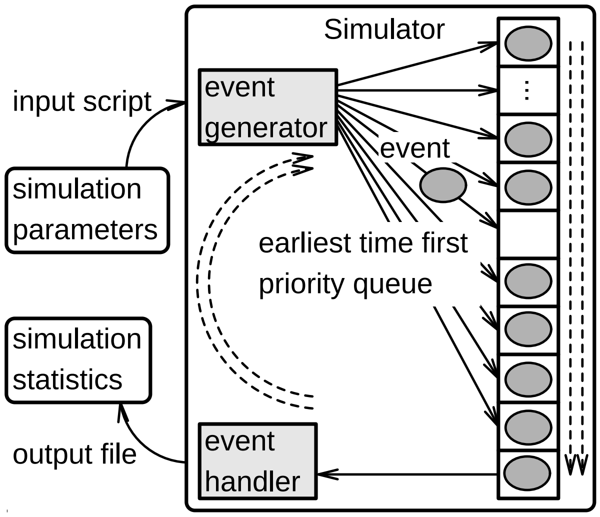

Figure 11: Simulator Architecture

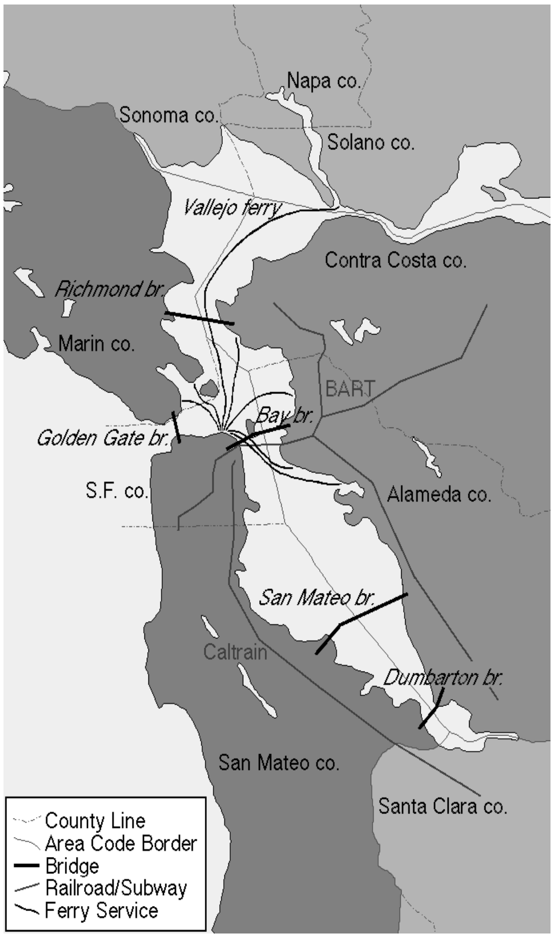

Figure 12: Map of the San Francisco Bay Area

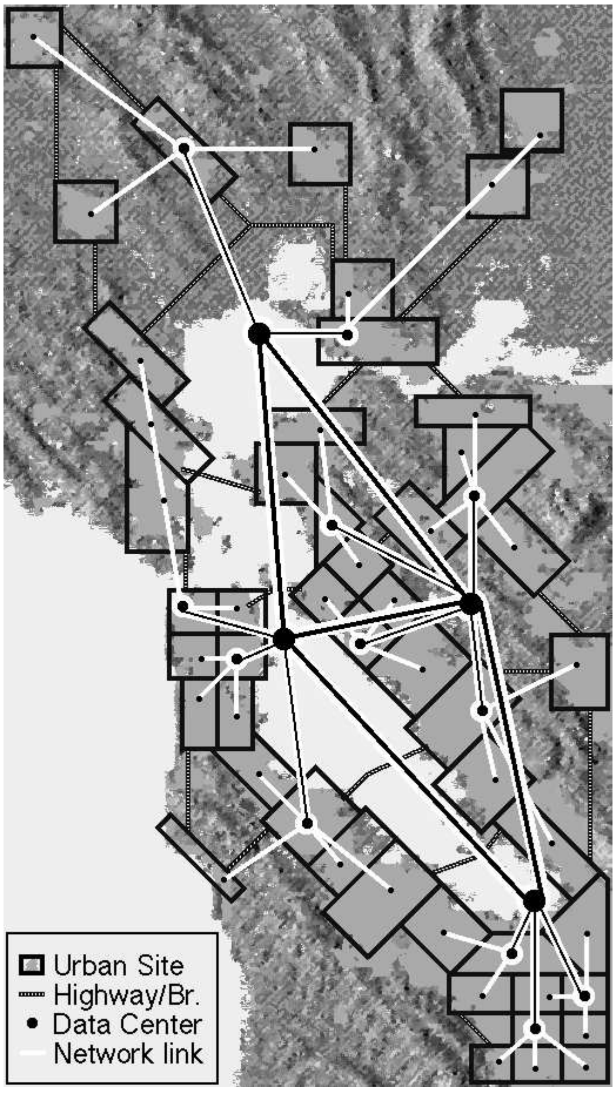

Figure 13: Overlay of Simulation and Network Topologies

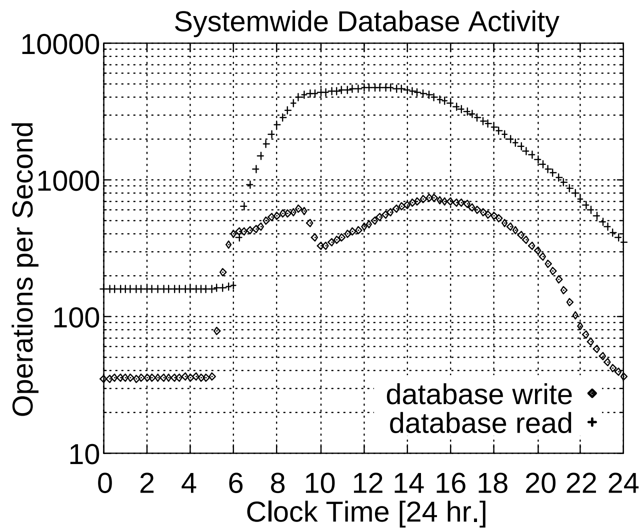

Figure 14: Database Access Rate

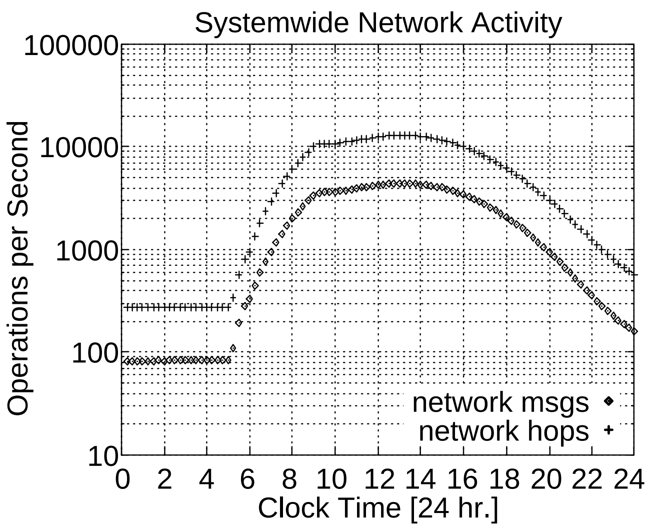

Figure 15: Number of Database Messages