Teletraffic Modeling for Personal Communications Services *

Derek Lam, Donald C. Cox, Jennifer Widom
Electrical Engineering & Computer Science Depts.
Stanford University, Stanford, CA 94305
{dlam, dcox}@wireless.stanford.edu, widom@cs.stanford.edu

Abstract

This paper presents a realistic teletraffic modeling framework for Personal Communications Services. The framework captures complex human behaviors and has been validated through analysis of actual call and mobility data. Using the proposed framework, a large-scale simulation was performed on a model of the San Francisco Bay Area. Simulation results showing the performance of IS-41 are presented.

1 Introduction

The goal of Personal Communications Services (PCS) [3] is to provide integrated communications (e.g., voice, video, and data) between nomadic subscribers independent of time, locations, and mobility patterns. The market for various wireless communications services is growing rapidly. In the United States, for example, there are more than 30 million cellular customers and this number is continuing to grow at an exponential rate. Other parts of the world, such as Europe and Asia, are experiencing the same tremendous growth in the demands for wireless communications services. There is evidence supporting this growth from the widespread use of the European Global System for Mobile Communications (GSM) and the reported explosive growth of the Japanese Personal HandyPhone System (PHS). It is, therefore, not far-fetched to envision a future PCS network that needs to support a large number of mobile subscribers scattered over a vast geographical region — a continent, or perhaps the world!

Teletraffic models are an invaluable tool for network planning and design. They are useful in areas such as network architecture comparisons, network resource allocations, and performance evaluations of protocols. Traditional traffic models have been developed for wireline networks. These models predict the aggregate traffic going through telephone switches. As such, they do not include subscriber mobility or callee distributions and therefore need modifications to be applicable for modeling PCS traffic. Since a general model for PCS traffic does not yet exist, most researchers resort to adding their own ad-hoc mobility models to the traditional wireline models. These ad-hoc mobility models seldom reflect actual movement patterns. They will unlikely be able to describe adequately the range of subscriber behaviors that will appear on a PCS network — one that covers a large geography. Mobility models are required to describe movement behavior at different scales. Another aspect of the ad-hoc approach is that callee distribution is usually neglected. Callee distribution is an important modeling aspect and we will show evidence supporting this in our simulation results.

In this paper, we discuss the PCS traffic modeling framework that was developed for our research in data management for wireless communications networks. Our traffic models are based on call traffic data, airplane passenger traffic data, and personal transportation surveys. We have developed a call traffic model

*Research partially supported by the Center for Telecommunications and the Center for Integrated Systems at Stanford University, and by equipment grants from Digital, IBM, and Pacific Bell Corporations.

that takes into account callee distributions. Using techniques and results from transportation research, we have developed three mobility models to characterize movements at different scales: within a metropolitan area, within a national area, and at the international level. Another unique aspect of our framework is that it models time-varying behavior which allows for investigations into transient, peak, and average performance. Unlike other synthetic models, we expect that our framework can provide a platform for generating realistic simulation results. Based on our framework, we have developed a large scale software simulator, *Pleiades*. Using Pleiades, we illustrate the importance of our models for data management research by comparing them with other simpler models commonly used. We also present simulation results on the performance of IS-41, the current data management standard in the United States, using our models for the San Francisco Bay Area. Finally, based on this modeling research, we have generated the **S**tanford **U**niversity **M**obile **A**ctivity **TRA**ces (SUMATRA). Our aim is for these traces to be used widely for wireless research and serve as a common platform for comparing research results. SUMATRA is available for the general community on the world wide web at **http://db.stanford.edu/sumatra**.

2 Motivation

Our motivation for PCS traffic modeling research comes from our work in scalable and efficient data management schemes for wireless communications networks. PCS presents many challenging problems in data management [10, 12, 32]. The PCS network stores important per-user information, such as current location, authentication information, and billing information, in *user profiles*. Data management refers to accessing and maintaining the information in user profiles. For example, during call setup, among other tasks, the network needs to access the callee's profile for location information and the caller's profile for authentication information. Also, the network registers user movements by updating location information in user profiles. The performance of any data management scheme is a function of the underlying database architecture, protocol, and algorithms. Performance variables of interest are: profile lookup and update response times, memory cost, and system equipment cost. Previous studies [20, 21] have shown that for projected numbers of PCS subscribers, existing data management standards, IS-41 [6, 22] and GSM [23], will incur a large increase in database loads over the current levels. In recent years, many sophisticated data management schemes [13, 14, 15, 26] have been proposed to reduce profile lookup and update response times and signaling traffic. These methods utilize techniques such as data replication and caching. It is beyond the scope of this paper to discuss these schemes. However, it is important to note that actual performance of these proposals depends strongly on subscriber behavior. In other words, the merits of caching and data replication schemes are functions of mobility and calling patterns. As a result, traffic models based on realistic calling and mobility patterns are a critical aspect of performance evaluation.

3 PCS Network Architecture

In a PCS network, mobile subscribers communicate through portable handsets. The basic network architecture is shown in Figure 1. It consists of a set of *radio ports* (i.e., basestations) 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 data management scheme. Communications take place by establishing call connections through radio ports, MSC's, and the wireline network.

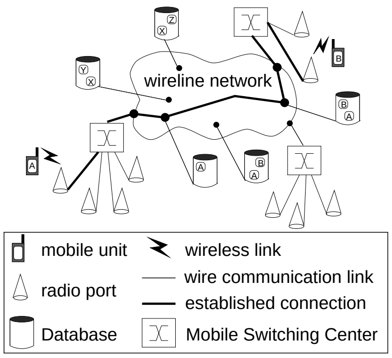

Figure 1: PCS Network Architecture

4 PCS Traffic Modeling Framework

We have developed a general modeling framework for our data management research. The framework includes the modeling of call and movement traffic, data management activities, and the signaling network. In this paper , we focus on the call and mobility traffic models. Interested readers can refer to [17, 18] for details of the general modeling framework. The traffic modeling framework is composed of the following objects.

Site represents a geographical area. All Site objects together define the physical geography for subscriber movements.

User represents a human subscriber.

The framework can also be divided into the following components: **Topology Model**, **Call Model**, and **Movement Model**. The Topology Model specifies the geographical topology or connectivities between *Sites*. The Call Model and Movement Model describe, respectively, how *Users* make calls to one another and how they move through the geography defined by the Topology Model. In our framework, we have decided to independently model call and movement behaviors. Correlations between call and movement activities are indirectly modeled through our time-of-day traffic analysis.

4.1 Call Model

Our call model generates call traffic for each individual user. The model is divided into two parts: the **Call Traffic Model** and the **Callee Distribution Model**. We have corroborated our models using encrypted call traffic data [28] from our local university telephone exchange. This exchange serves the entire campus including university offices, student housing, and faculty and staff residential households.

4.1.1 Call Traffic Model

The call traffic model describes how often individual users place calls to other users and characterizes the duration of each call. Very little is known about the traffic characteristics of future PCS networks. However, on fixed telephone networks, traffic is modeled accurately. Reference [8] gives an overview of existing call traffic models. For current telephone usage, according to [21], the mean call arrival rate and the mean call duration during busy hours are 2.8 calls/hour and 2.6 min/call, respectively. Our call traffic model generates call arrivals (i.e., calls initiated) for different classes of traffic and models time-varying user behavior. Each call traffic class is characterized by its probability of occurrence, call arrival rate, mean call duration, and distribution.

We have investigated time-of-day call traffic volume patterns because we need to use corresponding mobility patterns in performance evaluations. Figure 2 shows the traffic volume patterns derived from our

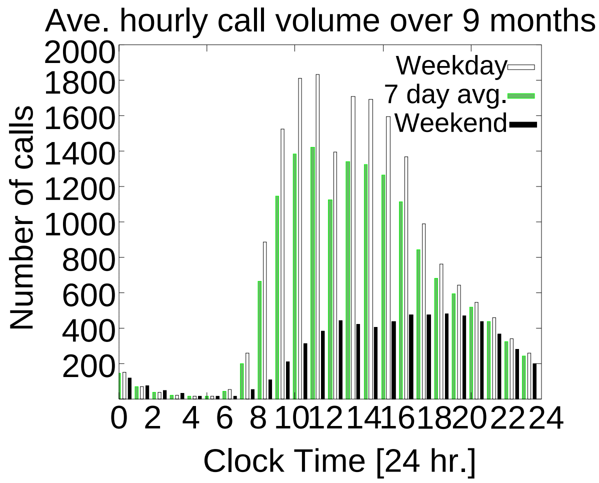

Figure 2: Average Number of Call Arrivals Per Hour

call traffic data in [28]. 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. These patterns along with call arrival rates in [21] provide guidelines for us when specifying the parameters in our Call Traffic Model.

4.1.2 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 effect on performance evaluation, especially for schemes with caching or data replication.

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 and 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. To obtain reasonable parameters, we have investigated empirical probability distributions using the notion of *callee rank*. The rank k callee of a caller is the caller's k^{th} most frequently called person within a reference period. For each caller i in [28], we calculate the call probability to the rank r callee ($\hat{P}_r^i$) over the periods of 1 day, 1 week, and 1 month. We observe that the mean call probability to the rank r callee, $\bar{P}_r$, can be modeled using a power or generalized Zipf's law at all three reference time periods: $\bar{P}_r \simeq \frac{A}{r^p}$, where A is the scaling parameter and p is the exponent parameter. Table 1 shows the fitted parameters and mean square errors of the fits. Figure 3 and Figure 4 are linear and log-log plots of $\bar{P}_r$ versus callee rank for the three reference periods.

We have investigated distributions around $\bar{P}_r$ because we want to include in our model callers that deviate from the "average" behavior. Figure 5 shows the distributions of first rank call probabilities ($\hat{P}_1^i$) derived from [28] for the three reference periods. We modeled each empirical distribution with a truncated

Time Scale	A	p	mean sq. error
1 Day	0.778	2.61	0.000010
1 Week	0.574	1.84	0.000028
1 Month	0.383	1.34	0.000030

Table 1: Fitted Power Law Parameters

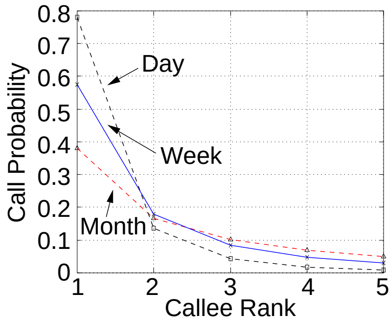

Figure 3: Mean Call Probability vs. Callee Rank

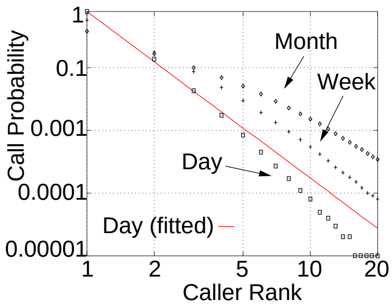

Figure 4: Log-Log Plot of Mean Call Probability vs. Callee Rank

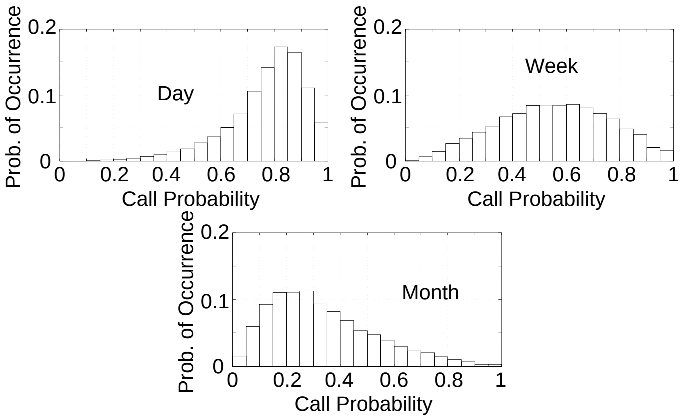

Figure 5: Distributions of Call Probabilities to First Rank Callee (for three reference time periods)

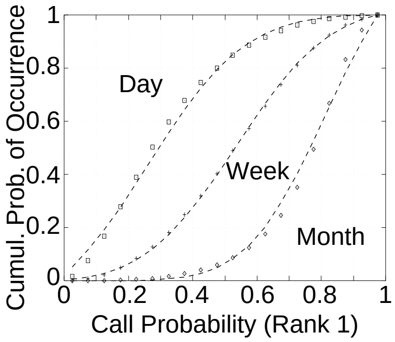

Figure 6: Cumulative Distribution of Call Probabilities to First Rank Callee

normal distribution. Figure 6 shows the cumulative distributions of $\hat{P}_1^i$ and their fits to our model. For the higher rank call probabilities, we looked at the relative probabilities to the rank r callee, $\hat{P}_r^i/\hat{P}_{r-1}^i, r > 1$. By characterizing the relative probabilities, we can then obtain higher rank call probabilities using a recursive procedure. We also modeled the distributions of $\hat{P}_r^i/\hat{P}_{r-1}^i$ by truncated normal distributions. Figure 7 shows graphically the cumulative distributions of relative call probabilities for a few higher rank cases and

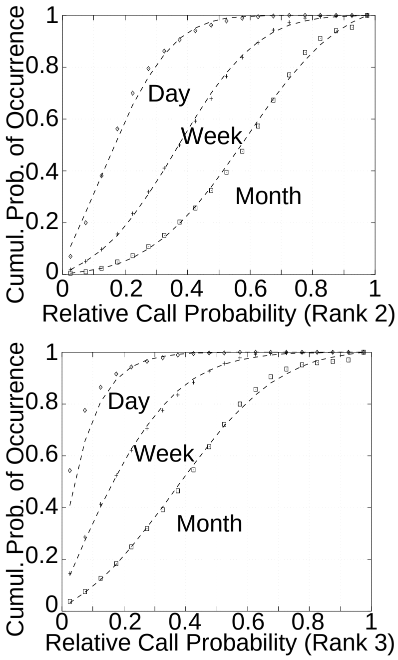

Figure 7: Cumulative Distribution of Relative Call Probabilities to Higher Rank Callees

their fits to our model.

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 periods and their respective call and relative call probabilities.

Time period	$\hat{P}_1$	$\hat{P}_2/\hat{P}_1$	$\hat{P}_3/\hat{P}_2$
1 Month	-0.0009	0.0014	0.0018
1 Week	0.0016	0.0004	0.0003
1 Day	0.0033	-0.0011	-0.0027

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

4.2 Mobility Models

Before discussing the mobility models we developed, we first review the common approaches to modeling human movements.

Fluid Model In a fluid model [8, 19, 21], traffic flow is conceptualized as the flow of a fluid. It describes macroscopic movement behavior. One of the simplest fluid models describes the amount of traffic flowing out of a region to be proportional to the population density within the region, the average velocity, and the length of the region boundary. For a circular region with a population density of ρ, an average velocity of v, and region diameter of L, the average number of site crossings per unit time, N, is:

$$N = \rho \pi L v \tag{1}$$

One of the limitations of this model is that it describes aggregate traffic and therefore is hard to apply to situations where individual movement patterns are desired, for example when evaluating network protocols or data management schemes with caching. Another limitation comes from the fact that since average population density and average velocity are used, this model is more accurate for regions containing a large population such as the case in [21].

Markovian Model The Markovian (or random walk) model [1] describes individual subscriber movements. In this model, a subscriber will either remain within a region or move to an adjacent region according to a transition probability distribution. One of the limitations of this approach is that there is no concept of *trips* or consecutive movements through a series of regions. *Trip* is again an important modeling aspect when considering replication and caching data management schemes mentioned previously.

Gravity Model Gravity models have been used to model human movement behavior in transportation research. They have been applied to regions of varying sizes, from city models [2, 11] to national and international models [7, 27]. There are many variations among the gravity models and it is not possible to describe all of them here. In its simplest form, the amount of traffic $T_{i,j}$ moving from region i to region j is:

$$T_{i,j} = K_{i,j} P_i P_j \tag{2}$$

where P_i is the population in region i, and $\{K_{i,j}\}$ are parameters that have to be calculated for all possible region pairs (i,j). In this form, the model describes aggregate traffic and therefore suffers from some of the same limitations as the Fluid Model. However, a variation of this model can be applied to describe individual movements directly if we interpret P_i as the "attractivity" of region i and $T_{i,j}$ as the probability of movements between i and j. In this approach, the parameters $\{P_i\}$ also have to be calculated from traffic data in addition to $\{K_{i,j}\}$. The advantage of the gravity model is that frequently visited locations can be modeled easily since they are simply regions with large attractivity. The main difficulty with applying the gravity model is that many parameters have to be calculated and therefore it is hard to model a geography with many regions.

Mobility Traces Mobility traces indicate current movement behavior and are more realistic than mobility models. However, traces for large population sizes and large geographical area are hard to come by. Researchers in the wireless local area network area, for example in [29], have collected movement traces for performance evaluation purposes. However, because of their applications, these traces are usually restricted to within an in-building environment or a small region. Also, they are generated from a small section of the population. For our mobility research, we have focused on larger areas, such as for a metropolitan area, because we feel that it is at this level where there will be significant network and database activities. Another limitation of mobility traces is that without a mathematical framework it is difficult to use them to predict future behavior for network planning purposes. In contrast, the other models described above can, for example, be used to predict movement behavior as population increases.

Our mobility model characterizes user movements within the geography defined by our Topology Model. We have developed a hierarchy of mobility models for movements at different scales: *Metropolitan Mobility Model*, *National Mobility Model*, and *International Mobility Model*. The Metropolitan Mobility Model describes subscriber movements within a metropolitan area. It is a generalized version of the Markovian Model that includes varying trip lengths and velocities. The National and International Mobility Models describe aggregate movement behavior at the national and international levels, and are variations of the gravity model discussed above. All these models have been derived using actual traffic data and official transportation surveys.

4.2.1 Metropolitan Mobility Model (METMOD)

The Metropolitan Mobility Model (METMOD) describes subscriber movements within a metropolitan area. It is a detailed model that includes the Markovian model as a special case. Each *Site* object is used to describe a small region of a metropolitan area. The geographical connectivity between Sites is modeled by the Topology Model described previously. Furthermore, probabilities for moving into adjacent Sites are specified by the *movement connectivity matrix*. Each element (i, j) in the matrix is the conditional probability that during a move a subscriber in site i will move into site j. Our model generates movement trips corresponding to different classes of mobility behavior: *simple move*, *roundtrip move*, *return home move*, and *stationary move*. Each of the movement classes is characterized by its probability of occurrence, mean velocity and distribution, mean number of site crossings and distribution. *Return home* movements are important when studying data management schemes with home location registers (such as IS-41 and GSM). *Roundtrip* movements are important when studying schemes with caching or data replication.

We have investigated actual human movement behavior using survey results from [5, 9, 16] and actual movement statistics from [24]. Figure 8 is a summary of the time-of-day traffic volume patterns we obtained from [9, 24]. From the data in [9], we have derived statistics (see Table 3) 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 then represent each trip purpose in Table 3 as a movement class with their appropriate mean move velocity and distance.

We have applied METMOD to a geographical model of the San Francisco Bay Area. Simulation results from this model are presented in the Simulation Results section.

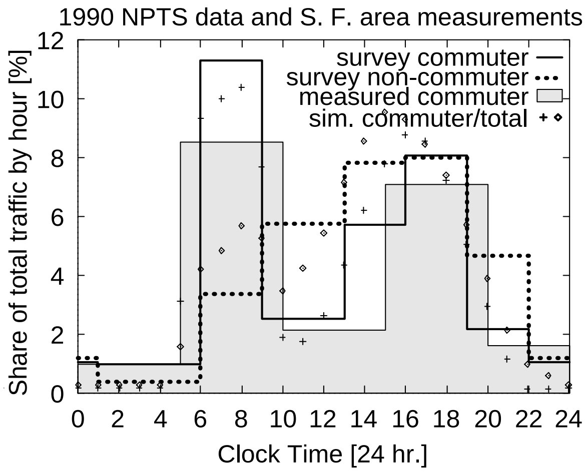

Figure 8: Estimated Commuter Pattern from Traffic Measurements

Trip Purpose	% of Trips	Average Trip Length (mi)	Average Velocity (mi/hr)
To/From Work	20.2	10.65	31.3
Work-Related	1.4	28.20	81.3
Personal	52.9	6.74	28.3
Social/Other	25.3	11.53	39.2
Vacation	0.2	218.22	261.5

Table 3: Movement Statistics

4.2.2 National Mobility Models (NATMOD)

The National Mobility Model (NATMOD) characterizes movement behavior between metropolitan areas in the United States. Each Site object now represents a metropolitan area. NATMOD characterizes traffic volume flowing between two Sites as a function of the population in each Site and the distance separating them.

We have obtained domestic air passenger traffic data [4] maintained by the Department of Transportation. The data analyzed contain domestic air passenger origin and destination traffic between all major U.S. commercial airports for two three-month periods (10/91-12/91 and 10/93-12/93). Our assumption is that commercial flights are the major mode of transportation between metropolitan areas.

Using dataset [4], we have derived traffic volume between the ten largest metropolitan areas in the United States. These metropolitan areas are chosen from their population in 1995 [30]. We model the traffic volume using a variation of the gravity model. Instead of developing a detailed gravity model that could model the traffic very accurately, our approach is to find a sufficiently accurate model with a small number of parameters to allow efficient simulation. In order to do this, we have assumed symmetric traffic flow, i.e., the traffic volumes between any two metropolitan areas are the same along both directions. We note that, for our data in the year 1991, there is an average of less than 2% fluctuation between the two traffic directions. Another interpretation of this is that we are modeling only the average traffic volume between two regions. We have experimented with many different variations of the gravity model. The following model gave the best results.

$$T^*_{i,j} = \frac{m_i m_j P_i P_j}{d_{i,j}^{\gamma_i+\gamma_j}} \tag{3}$$

where $T^*_{i,j}$ is equal to $0.5(T_{i,j}+T_{j,i})$ and $T_{i,j}$ is the traffic between region i and region j, P_i is the reported population for region i, and $d_{i,j}$ is the distance between the largest airports in regions i and j. $\{m_i\}$ and $\{\gamma_i\}$ are parameters that have to be calibrated. Comparing to Equation 2 in Section 4.2, we note that the product $(m_i/d_{i,j}^{\gamma_i}\ m_j/d^{\gamma_j})$ is equivalent to $K_{i,j}$. With the assumption of symmetric traffic flow, we were able to separate the dependence of $K_{i,j}$ to i and j, and obtain a model that has $2N_R$ parameters where N_R is the number of regions. We calculated the parameters $\{m_i\}$ and $\{\gamma_i\}$ twice. For the first calculation, we used the three months of movement data in 1991 and population estimates for the metropolitan areas in 1991 [30]. We performed the second calculation using traffic and population data in 1993. Both calculations were done by minimizing the mean normalized root-mean-square (RMS) difference using the simplex method. Mathematically speaking, if $T^*_{i,j}$ is the actual average traffic flow between region i and region j and $\hat{T}^*_{i,j}$ is the average traffic flow predicted by the calibrated model, then the normalized RMS difference, e_i^{rms}, is:

$$e_i^{rms} = \sqrt{\frac{1}{N_R}\sum_j \left(\frac{\hat{T}^*_{i,j} - T^*_{i,j}}{\sum_j T^*_{i,j}}\right)^2} \tag{4}$$

where N_R is the total number of regions. We note that $T^*_{i,j} = T^*_{j,i}$ and $T^*_{i,i} = 0$, and similar relations also hold for $\hat{T}^*_{i,j}$. The mean normalized RMS difference is the average of e_i^{rms}. Table 4 shows the calculated parameters for each region and their e_i^{rms}. Columns 2 to 4 show results from the first calculation, while columns 5 to 7 show results from the second calculation. For all these results, $\{d_{i,j}\}$ are measured in miles and the values of $\{T^*_{i,j}\}$ and $\{P_i\}$ are measured in number of persons and thousands of persons, respectively.

The two sets of parameters in Table 4 indicate how much the model parameters can vary over time. We have examined whether the parameter variations are significant. Table 5 shows the results. Column 2 reports the the change in e_i^{rms} if we were to predict 1991 traffic using the calculated parameters for 1993 and the 1991 population figures. Similarly, column 3 shows the change in e_i^{rms} when we use the calculated parameters for 1991 to predict the 1993 traffic. We observe that the changes in $\{e_i^{rms}\}$ are small.

Regions	1st. Calc. (1991)			2nd Calc. (1993)		
	$ln(K_i)$	γ_i	e_i^{rms}	$ln(K_i)$	γ_i	e_i^{rms}
New York + area	-10.7	-0.0629	0.0142	-11.1	-0.124	0.0150
L.A. + area	-7.55	0.351	0.0082	-7.06	0.435	0.0116
Chicago Metro.	-8.57	0.145	0.0171	-8.70	0.104	0.0182
Wash. + Balt.	-7.69	0.305	0.0162	-8.28	0.205	0.0182
S.F. + Bay area	-6.61	0.393	0.0140	-6.88	0.369	0.0171
Philadelphia + area	-11.8	-0.154	0.0739	-11.8	-0.177	0.0639
Boston + area	-6.75	0.455	0.0201	-6.47	0.490	0.0167
Detroit + area	-6.87	0.504	0.0148	-7.53	0.395	0.0165
Dallas + area	-8.71	0.0758	0.0207	-9.76	-0.0646	0.0235
Houston + area	-2.93	0.987	0.0167	-0.194	1.13	0.0197

Table 4: NATMOD Calculated Parameter Values and Residues

Regions	1991 Traffic $e_i^{rms,1993} - e_i^{rms,1991}$	1993 Traffic $e_i^{rms,1991} - e_i^{rms,1991}$
New York + area	0.0030	0.0046
L.A. + area	0.012	0.0076
Chicago Metro.	-0.0002	0.0038
Wash. + Balt.	0.0075	0.0069
S.F. + Bay area	0.0091	0.0076
Philadelphia + area	-0.0088	0.0087
Boston + area	-0.0025	0.0033
Detroit + area	-0.0003	0.0010
Dallas + area	0.0037	-0.0001
Houston + area	0.0094	-0.0028

Table 5: NATMOD Sensitivity to Parameter Variations

Countries	K_j				
	1990	**1991**	**1992**	**1993**	**1994**
Canada	0.103	0.0918	0.0935	0.0950	0.0924
United Kingdom	0.0351	0.0321	0.0368	0.0390	0.0395
Japan	0.0146	0.0143	0.0154	0.0152	0.0155
Mexico	0.0199	0.0198	0.0197	0.0196	0.0203
Germany	0.0121	0.0117	0.0132	0.0136	0.0132
France	0.0121	0.0109	0.0127	0.0123	0.0130
Netherlands	0.0216	0.0233	0.0259	0.0310	0.0342
South Korea	0.0072	0.0073	0.0083	0.0089	0.0097
Dom. Republic	0.0511	0.0438	0.0477	0.0500	0.0504
Jamaica	0.1511	0.1374	0.1314	0.1431	0.1466

Table 6: INTMOD Calculated Parameters for 1990 - 1994

The significance of NATMOD is that it is a realistic and reasonable model for movements between metropolitan areas. We have shown parameter values and their ranges for current inter-metropolitan movements. In addition, we have also shown that NATMOD is relatively insensitive to parameter variations and therefore is reasonable to be used in estimating traffic volumes in the future or for other geography. Another interesting aspect of NATMOD is that it associates only two parameters, m_i and γi, per Site and therefore permitting more convenient and efficient simulations.

4.2.3 International Mobility Model (INTMOD)

The International Mobility Model (INTMOD) characterizes movement behavior between the U.S. and ten other countries. Each Site object in INTMOD represents a country. From [25] we obtained air passenger departure and arrival traffic between U.S. and other countries from 1990 to 1994. We also used a variant of the gravity model following the same methodology as in NATMOD. However, in this case, due to the limited amount of data, a fully-interconnected movement model between all country pairs is not possible. We have tried to model movements between USA and other countries using the following model.

$$T_j^* = K_j P_{usa} P_j \tag{5}$$

T_j^* is the average traffic flow between USA and country j, K_j is the calculated parameter for country j, and P_j is the population of country j (P_{usa} is the USA population). T_j^* is found by averaging over the arrival and departure traffic. From [30], we obtained five years of population figures and proceed to determine parameters for these five years. We note that because of the simpler model, K_j is simply equal to $T_j^*/P_{usa}P_j$ and hence there is no residue. Table 6 shows the calculated parameters. We also investigated how the accuracy of this model changes over time. Table 7 shows the normalized RMS difference when we use the 1990 parameters to predict traffic between 1991 to 1994.

Compared to NATMOD, the INTMOD gravity model is missing the inverse dependence to distance. One of the reasons for not using the same gravity model as NATMOD was explained previously by the lack of traffic data. Another justification is that there is generally some amount of uncertainty in defining distances between countries. Defining the distance between USA and Canada, two large territories, is a good example. In any case, our goal for INTMOD is to have a simple, easy to use, and realistic model for international movement traffic. Similar to NATMOD, we have presented current parameter values and

Traffic Year	e_{usa}^{rms}
1990	0
1991	0.0109
1992	0.00900
1991	0.0102
1994	0.0125

Table 7: INTMOD Variations of e_{usa}^{rms} over Time

their ranges, and have shown that INTMOD parameters remain relatively constant over a period of five years.

5 Simulation Results

We have developed a discrete event simulator, Pleiades, based upon the framework described above. Pleiades contains several modules to emulate the functions of various data management schemes, such as IS-41, GSM, and new novel proposals. The architecture of Pleiades is shown in Figure 9; further details on Pleiades can be found in [17].

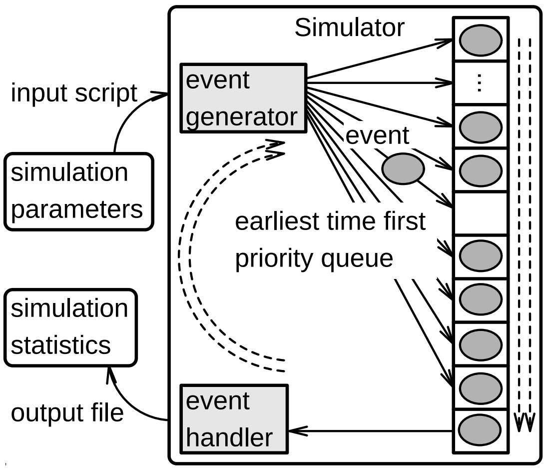

Figure 9: Simulator Architecture

We now show that our callee distribution and mobility models are critical to performance evaluation. Table 8 shows three sets of simulation parameters for three different models. The proposed callee model for set 1 and 3 is our callee distribution described previously in Section 4.1.2. Tables 9 and 10 are summaries of simulation results for two data management schemes, IS-41 and centralized database. In the centralized database scheme, all user profiles are stored in one single centralized database. If the caller and the callee are not in the same registration zone, then a profile lookup is required at the centralized database. We observe that using our proposed models, significantly different results were obtained in the key performance measures. We note that, in the centralized database scheme, global average update rates are independent of the models because we need to update the same number of profiles per user movement.

Using Pleiades, we have investigated the performance of IS-41 on a geography that models the San Francisco Bay Area, which is composed of four area codes. Figure 10 is a map of the Bay Area. Regions

Parameters	Parameter Set 1	2	3
# users	16200		
Geography	9x9 square grid		
Callee Model	Proposed 8 callees	Random	Proposed 8 callees
Move Model	25% no move 25% simple 25% roundtrip 25% return home	25% no move 25% simple 25% roundtrip 25% return home	Markov

Table 8: Simulation Parameters for Model Comparisons

Perf. vars. (Global)	Set 1	Percentage Difference (Set2-Set1)/Set1	(Set3-Set1)/Set 1
ave. lookup rates (per sec)	26.02	21.1%	18.2%
ave. update rates (per sec)	42.74	$\approx$ 0%	19.8%
ave. message rates (per sec)	52.00	21.1%	34.5%
ave. msg-hop count per sec	112.9	49.8%	31.8%

Table 9: Simulation Results and Percentage Difference for IS-41

Perf. vars. (Global)	Set 1	Percentage Difference (Set2-Set1)/Set1	(Set3-Set1)/Set 1
ave. lookup rates (per sec)	20.65	17.0%	12.3%
ave. update rates (per sec)	18.47	$\approx$ 0%	$\approx$ 0%
ave. message rates (per sec)	31.51	9.93%	46.7%
ave. msg-hop count per sec	68.54	40.4%	46.3%

Table 10: Simulation Results and Percentage Difference for Centralized Database Scheme

corresponding to different area codes are represented by different shades in the figure; bridges, ferries and public transportation are also included. Figure 11 [31] is an overlay map that shows the relationship between our simulation model and the actual geography of the Bay Area.

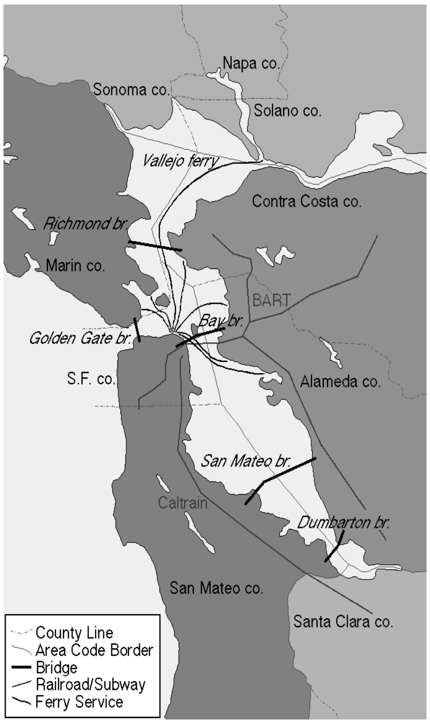

Figure 10: Map of the San Francisco Bay Area

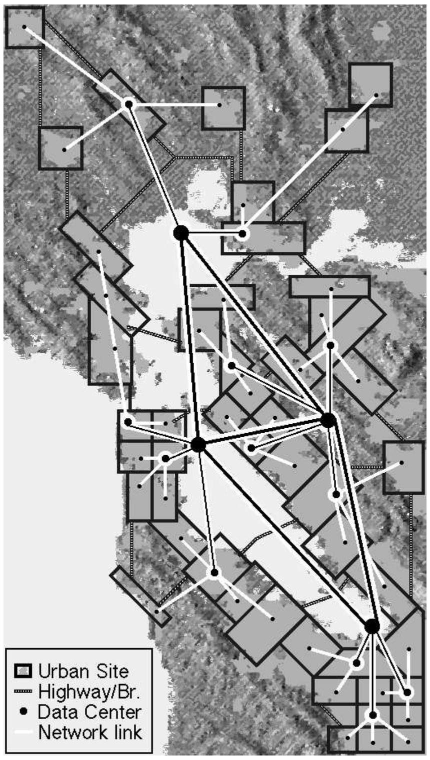

Figure 11: Overlay of Simulation and Network Topologies

Figure 12 and Figure 13 show systemwide database and network activities throughout the simulation. Each data point on these plots is calculated by averaging over the 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. We note that in the following summary, peak lookup and update rates occur at different times. This is revealed only through our detailed time-varying models and suggests possible optimizations in the utilizations of network resources. The following summarizes the results.

- We observe peak access rate for lookups at 4,746 TPS, for updates at 741 TPS, and their combined total at 5,304 TPS. These peak rates occurred at 12:45 p.m., 3:15 p.m., and 1 p.m. in our simulated day, respectively. Table 11 shows lookup, update, and total access rates at these peak times.
- We observe a peak signaling rate of 4,401 messages per second and 12,721 message-hops per second at 1 p.m. in our simulated day.

6 Conclusions

In this paper, we have presented a realistic framework for modeling teletraffic in PCS. The framework incorporates realistic behavior models that have been corroborated using measurements and surveys of

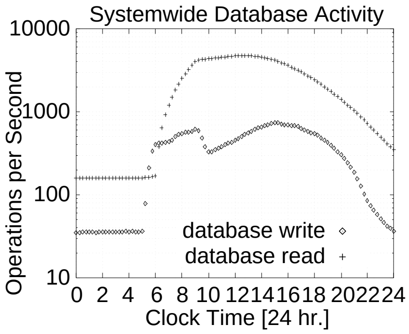

Figure 12: Database Access Rate

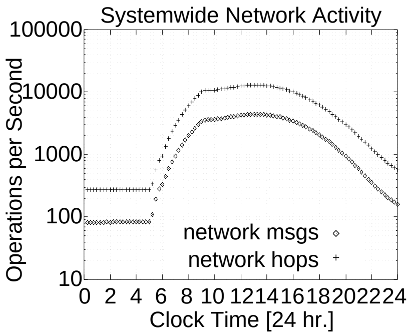

Figure 13: Number of Database Messages

Perf.vars (Global)	Simulated Times		
	12:45 p.m.	13:00 p.m.	15:25 p.m.
lookup rate (per sec)	4,745.9	4,745.8	4,002.6
update rate (per sec)	529.5	558.3	741.2
total rate (per sec)	5,375.4	5,304.1	4,743.8

Table 11: Access Rate at Three Selected Simulation Times

actual human activities. We have compared simulations results with some commonly used models and showed that the framework produced significantly different results for our applications in data management research. We have developed a software system capable of simulating a large population of users and have presented results of a detailed 24-hour simulation of the San Francisco Bay Area.

References

[1] 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, jun 1994.

[2] C. Ben, R. Bouchard, and E. S. Jr. Clyde. An evaluation of simplified procedures for determining travel patterns in a small urban area. In *Computer Simulation of Human Behavior*, chapter 20, pages 400–430. John Wiley & Sons, Inc., 1971.

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

[4] Origin and destination survey, T-100 domestic market data. Department of Transportation. Dataset for periods 10/91-12/91 and 10/93-12/93.

[5] ECMT. *European Transport Trends and Infrastructural Needs.* OECD publications, 2, Rue Andre-Pascal, 75775 Paris CEDEX 16, France, 1995.

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

[7] L. Fridstrom and H Thune-Larsen. An econometric air travel demand model for the entire conventional domestic network: the case of Norway. *Transportation Research*, 23B(3):213–223, June 1989.

[8] V.S. Frost and B. Melamed. Traffic modeling for telecommunications networks. *IEEE Communications Magazine*, 32(3):70–81, March 1994.

[9] P. S. Hu and J. Young. *1990 NPTS Databook: Nationwide Personal Transportation Survey.* Federal Highway Administration, November 1993.

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

[11] Bouchard. R. J. and C. E. Pyers. Use of gravity model for describing urban travel. *Highway Research Record*, 88:1–43, 1965.

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

[13] 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.

[14] 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.

[15] J. Jannink, D. Lam, N. Shivakumar, J. Widom, and D. C. Cox. Efficient and flexible location management techniques for wireless communication systems. Technical report, Stanford University, Computer Science Dept., May 1996. To appear in MOBICOMM'96.

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

[17] D. Lam, J. Jannink, D. C. Cox, and J. Widom. Modeling location management for Personal Communications Services. Technical report, Stanford University, Computer Science Dept., October 1995. Long version of [18].

[18] D. Lam, J. Jannink, D. C. Cox, and J. Widom. Modeling location management for Personal Communications Services, September 1996. To appear in ICUPC'96.

[19] K. K. Leung, W.A. Massey, and W. Whitt. Traffic models for wireless communication networks. *IEEE Journal on Selected Areas in Communications*, 12(8):1353–1364, October 1994.

[20] C. N. Lo and R. S. Wolff. Estimated network database transaction volume to support wireless personal communications services. In *International Conference on Communications, ICC '93*. IEEE, May 1993.

[21] 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.

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

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

[24] 1990 commute summary. Metropolitan Transportation Commission, Lotus 123 format spreadsheet, August 1990. S.F. Bay area transportation measurements.

[25] International air traffic between United States and other countries. National Transportation Statistics 1996: Table17,18.

[26] N. Shivakumar and J. Widom. User profile replication for faster lookup in mobile environments. In *1st ACM International Conference on Mobile Computing and Networking (MOBICOM'95)*, pages 161–169, Berkeley, California, 1995.

[27] P. B. Slater. International migration and air travel: global smoothing and estimation. *Applied Mathematics and Computation*, 53(2):225–234, March 1993.

[28] Telephone call traffic data set, 11/94–7/95. S. Phillips, personal communication, October 1995. Encrypted ID numbers.

[29] C.-K. Toh. Performance evaluation of crossover switch discovery algorithms for wireless ATM LANs. In *Proceedings of INFOCOMM'96*, pages 1380–1387, March 1996.

[30] International population data. U.S. Bureau of Census, International Data Base.

[31] 1990 urban land use in central California. U.S. Geological Survey, http://geo.arc.nasa.gov/usgs/images/urban2.gif, 1995. Urbanization in S.F. Bay area.

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