Analysis of Caching-based Location Management in Personal Communication Networks*

Karunaharan Ratnam
Dept. of Elec. and Comp. Eng.
Northeastern University
Boston MA 02115
karu@ece.neu.edu

Ibrahim Matta†
Computer Science Dept.
Boston University
Boston, MA 02215
matta@cs.bu.edu

Sampath Rangarajan
Bell Laboratories
700 Mountain Avenue
Murray Hill, NJ 07974
sampath@research.bell-labs.com

Abstract

Personal communication networks support the delivery of communication services as the user moves from one region to another. When a mobile user/terminal receives a call, the network has to quickly determine its current location. The existing approach suffers from high delay in locating the mobile as it requires maintaining the current location in a stable storage that has to be always consulted to reach the mobile. To reduce this delay, many proposed schemes rely on caching the locations of mobiles, especially those which do not move too frequently. To measure mobility, the node that originates the call usually measures only *those movements that it sees between successive calls to that mobile. In this paper, we present a caching scheme based on fully disseminating the location updates of mobiles to* every *node so as to increase the chance that the cache entry points to the correct location of the mobile user. We analyze our full dissemination based scheme and compare it to other caching and non-caching based schemes.*

1 Introduction

One of the challenging tasks in a *Personal Communication Services* (PCS) environment is to efficiently maintain the location of PCS subscribers who move around freely with their wireless units (hereafter called mobile host or mobile for short). In North America, Telecommunications Industry Association's interim standard *IS-41* [1] is used for managing location information of the subscribers and enabling them to send and receive calls and other services, such as messaging and data service.

In IS-41 [1], an incoming call is routed to the called subscriber as follows. The call is initiated by dialing the subscriber's number. The dialed call is received by the Mobile Switching Center (MSC) in the home system. This MSC is called *originating MSC*. If the mobile host is currently being served by the orginating MSC (i.e. the mobile host is not roaming), then the originating MSC queries the Home Location Register (HLR) to obtain the registration status and feature information of the mobile host. After receiving the response from HLR, the originating MSC pages the mobile host. When the mobile host responds (i.e. subscriber accepts the call by pressing the proper button), the originating MSC sets up circuits to terminate the call to the mobile host.

Figure 1 shows how a call is typically delivered to a roaming mobile host. As before, when a call to a mobile is dialed, the call is first routed to the originating MSC. The originating MSC then sends a location request message to the HLR to find out the current location of the mobile. The HLR, in turn, sends a route request message to the Visitor Location Register (VLR) that is currently serving the mobile. The VLR then sends a route request message to the MSC that is currently serving the mobile. The serving MSC creates a Temporary Location Directory Number (TLDN) and returns it to the VLR. The TLDN is then passed back to the orginating MSC through the HLR. The originating MSC then route the call using this TLDN. When the serving MSC receives the call routed using the TLDN, it pages the mobile host. If the mobile responds, then the call is terminated to the mobile.

Thus, HLR is a critical entity in the IS-41 location management system. There are many disadvantages to this approach. One is that, since every location request

*This work was supported in part by NSF grants CAREER ANIR-9701988 and MRI EIA-9871022.

†This work was done while this author was with the College of Computer Science at Northeastern University.

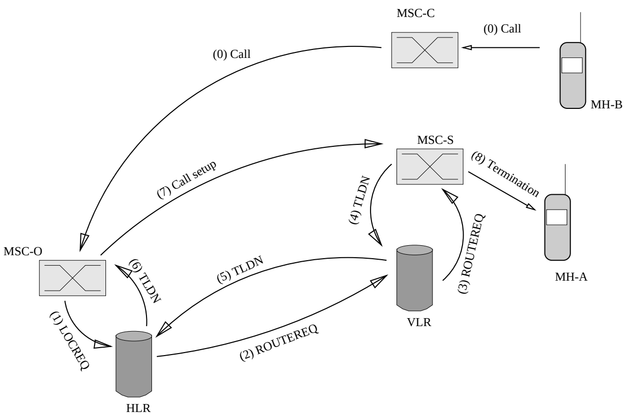

Figure 1. Call delivery in IS-41 and the tromboning problem

is serviced through HLR, the time required to establish a connection to a mobile host increases [7]. The other disadvantage is that, any HLR system failure causes all mobiles registered with the HLR to be unreachable even though mobiles may be roaming and away from the HLR region.

There is also another disadvantage, which is generally referred to as *tromboning problem.* Consider the situation depicted in Figure 1. The subscriber MH-A's home MSC is MSC-O, and MH-A is currently roaming and being served by MSC-S. Another mobile MH-B, which is currently being served by MSC-C, makes a call to MH-A. MSC-C and MSC-S are geographically closer to each other, and connected by the local exchange carrier. But, MSC-O (the home MSC of MH-A) is geographically far away from both MSC-C and MSC-S and connected by a long distance carrier. Then, routing the call from MH-B to MH-A involves two long distance legs, one between MSC-C and MSC-O, and the other between MSC-O and MSC-S. The latter leg is used twice, first to obtain the TLDN, and then to provide the voice/data connection.

Furthermore, if a subscriber is called several times from the same switch, the call processing procedure (cf. Figure 1) has to be repeated each time a call is terminated to the subscriber. In this case, the call is routed through unnecessary links several times. Avoiding these extra links will not only decrease the overall (signaling) cost but also decrease the delay in establishing the call. This could be achieved if the current location of the callee is stored or *cached* by the frequent caller.

Many location management schemes and improvements to IS-41 have been proposed in recent years (e.g. [2, 4, 6, 3, 11]). In this paper, we are concerned with improvements to IS-41 that are based on *caching* the location of mobiles so as to improve call delivery and cost.

Our Contribution

In [8, 10], we presented a fully distributed location management scheme. Because of its distributed nature, we replace the centralized registers VLRs and HLRs with location registers (LRs). Since there are no HLRs or VLRs in this system, each LR maintains the location information of not only mobiles that are local to it, but also of other mobiles in the network. Thus, this is a protocol where the location information of the mobile hosts are *fully replicated* in all the LRs. The LRs are distributed throughout the network. An LR serves one or more MSCs just like the VLR in the PCS architecture (cf. Figure 1). An LR could co-exist with an MSC, and serve only that MSC.

LR functions as both location registry of local mobile hosts as well as the lookup directory for the location of other mobile hosts. The type of location information maintained for a mobile host depends on whether the mobile is local to the LR or not. For local mobile hosts, LR maintains the id of the MSC that is currently serving the mobile. For mobile hosts that are not local, LR maintains the id of the LR where the mobile host currently resides. When a mobile registers with an LR, the new location information is disseminated to all other LRs in the network. This dissemination is carried out in parallel through the whole network so that the new location is very quickly updated at all LRs.

In this paper, we modify our fully distributed location management scheme to implement caching. The location information is fully disseminated to all the location registers (LRs). However, the location information is *not* fully replicated, but *cached* based on per user call and mobility pattern. Full replication requires the location information be maintained in a stable storage at each LR, and the access to stable storage is generally slow compared to the access to the cache. Hence, caching instead of replication expedite the processing of disseminated location update information. As in IS-41, the location information of a mobile is still maintained on stable storage of HLR. We analyze our full dissemination based caching strategy. We show that, under certain network, traffic and mobility conditions, if the mobile host service area change is promptly known to all the nodes, the cached location information can be effectively validated, and the performance of our caching scheme with full location information dissemination is more cost efficient than the basic IS-41 scheme or a per user caching scheme where not all service area changes by mobiles are known to all nodes.

The paper is organized as follows. Section 2 describes traditional caching-based location management schemes, proposed to improve the basic IS-41 standard. Section 3 presents our caching with full dissemination scheme. Section 4 analyzes the various caching-based schemes as well as the basic IS-41 standard. Numerical results are presented in Section 5. Due to lack of space, we refer the reader to [9] for more results. Section 6 concludes the paper.

2 Caching Strategy for PCS

An improvement to IS-41, called *per user caching scheme*, is proposed in [2, 6, 4]. The authors observe that, when a particular MSC receives a large number of calls to a particular mobile that belongs to a different home system, the signaling and database lookup cost involved in setting up the call can be significantly reduced by caching the location information (i.e. a mapping of mobile host id and the identity of the serving VLR) at the calling MSC. Each time when a call is attempted, the cached information is checked first. Since the access time in looking up an entry in the cache memory is very short (order of microseconds), checking the cached information for every call does not significantly affect the performance of the MSC. If the mobile's VLR information is cached (i.e. cache hit), then that VLR is contacted directly. If that VLR is still serving the mobile, the HLR need not be contacted to route the call, thus reducing connection establishment delay and avoiding unnecessary HLR database lookup delay. However, if the subscriber has moved out of that VLR's region, the HLR need to be contacted after sending a connection request to the wrong VLR. In that case, the connection establishment time will be longer than that of the basic IS-41. If the VLR information for the particular mobile is not cached (i.e. cache miss), then the normal IS-41 call setup procedure is followed.

Taking into account these trade-offs, the location information of a subscriber has to be cached at an MSC only if the mobile changes its service area (i.e. area served by a VLR) less frequently than it receives calls from that MSC. The authors use a figure of merit called *local call-to-mobility ratio* (LCMR) in caching the location information. LCMR is the ratio between the number of calls originating from an MSC to the number of times the mobile changes its service area as *seen* by that MSC. The location information of a subscriber is cached at an MSC, if the LCMR maintained for the subscriber at the MSC is larger than a threshold derived from the link and database access cost of the network. Note that, the *true* call-to-mobility ratio which takes into account *all* the service area changes by the mobile host whether it is seen by a particular MSC or not, would yield more valid (correct) cache entries and lower overall cost than LCMR.

A cache entry has to be removed from the cache after it is deemed less useful. Because cache memory is becoming less expensive, capacity (in terms of number of mobiles that are cached) may not be an issue. However, the longer the time a mobile has not been called from an MSC, the higher the chance that the corresponding cache entry, if cached, is not correct (i.e. points to a wrong location for the mobile host). In [6], a cache invalidation approach called *T-threshold scheme* is proposed. In this scheme, a cache entry is invalidated after a threshold amount of time T since its last usage.

3 Our Approach: Caching Strategy with Full Dissemination

In our fully distributed scheme [8, 10], the location information is fully distributed to all the location registers (LRs). In this paper, the location information is *not* fully replicated, but *cached* based on per user call and mobility pattern. Since the location information is no longer replicated, it must be maintained at some stable storage. Similar to IS-41, we also call the stable storage HLR, mainly because it provides the same functionality.

The role of LR now is to provide location registry services for visiting mobile hosts. That is, LR provides the same functions as the IS-41 VLR. Since we assume LR is co-located with the MSC, the caching functionality is delegated to LR so that MSC can be viewed as just switching calls to mobile hosts. When a call originates for a mobile, the current location of the mobile is found by first querying the local LR to check the cache. If the location information is not cached in the local LR or the cached location information turns out to be incorrect, only then the HLR is contacted.

When the mobile moves from one LR region to another LR region, this location information is updated to the HLR and disseminated to all the LRs in the network. LRs *selectively* cache this location information. The decision to cache the location of a mobile host could be made based on any of the schemes discussed in [2, 4]. In these schemes, the caching decision is based on the local call-to-mobility ratio, which is the ratio between the number of calls to a mobile host from a VLR area and the number of times the mobile host has changed its location as seen by that VLR. The VLR will be able to see a change in the location of a mobile host only when a call is originated for that mobile from the VLR area. Hence, even if the mobile host has changed several VLR areas between two consecutive calls from a VLR area, the VLR will count the mobility of that mobile host as just one movement. Since the location information is fully disseminated in our scheme, all LRs can keep track of the *true* mobility of *any* mobile host. Hence, LRs can cache the location information efficiently, and yield higher cache hit rates. Also, if the location information is not disseminated (as in [2, 4]), the cache is more likely to have stale entries compared to our scheme. If an LR uses this stale entry to contact a mobile host, it will send a connection request to a wrong LR first, and wait for the reply from that LR before it can contact the HLR of that mobile to receive the current (correct) location. In this case, the cost of contacting a wrong LR negates the gain of caching the location information. Since full dissemination yields more accurate cache entries, we expect our scheme to yield lower connection establishment times. Given that dissemination time of location information is small, we henceforth assume that a cache entry for a mobile (if it exists) is always correct.

4 Analysis of Caching Strategies

We assume that each service area is served by a single MSC, and the VLR/LR is co-located with each MSC. We refer to the IS-41 based per user caching as *IS41-Cache scheme*, and our caching with full location information dissemination scheme as *FD-Cache*. Our performance measure is the expected total cost defined as the sum of the UPDATE cost and the FIND cost, where the UPDATE cost is the cost involved in updating the location information of a mobile user as the user moves to a service area, and the FIND cost is the cost of locating and establishing a connection to the mobile user before the user moves to another service area. Since the total cost is derived for a single service area crossing, only one location update contributes to the UPDATE cost. However, the FIND cost will have contribution from all the calls that are terminated to the mobile host while it is in that service area. Here we assume that the mobile host is away from its home service area (i.e. roaming).

Basic IS-41 Scheme (no caching): In IS-41, UPDATE involves the new VLR registering the mobile with its HLR, and the HLR sending registration cancellation to the old VLR. Hence the UPDATE cost is given by:

$$UPDATE_{IS41} = Cost(VLR_{new} \leftrightarrow HLR) + Cost(HLR \leftrightarrow VLR_{old}) \quad (1)$$

Assuming the time to register with the HLR is very short (i.e. the probability that a location request to the HLR falls during the registration time is negligible), the FIND cost of a roaming mobile is given by:

$$FIND_{IS41} = Cost(VLR_{caller} \leftrightarrow HLR) + Cost(HLR \leftrightarrow VLR_{callee}) \quad (2)$$

where VLR_{caller} is the MSC/VLR where the call is generated, HLR is the *home* MSC of the roaming mobile host, and VLR_{callee} is the MSC/VLR that is currently serving the roaming mobile host.

Thus, the expected total cost in the IS-41 scheme is given by:

$$\text{TotalCost}_{IS41} = \text{UPDATE}_{IS41} + \rho \times \text{FIND}_{IS41} \quad (3)$$

where ρ is the expected number of calls per move. UPDATE_{IS41} and FIND_{IS41} are given by equations (1) and (2), respectively.

Now we need to find the expected number of calls a callee receives at the current VLR. In general, the proportion of calls a customer receives from each service area is not uniformly distributed. In fact, an actual calling trace [11] collected over a period of six months shows that, on average, more than half the calls each customer receives on a daily basis are from a single caller. The trace further shows, on a daily basis, a customer receives 90% of the calls from his/her top 5 callers.

Let us rank each VLR service area (i.e. switch) with respect to each customer based on the percentage of calls the customer receives from that service area. Let N_T be the total number of service areas in the network. The customer receives the highest number of calls from service area ranked 0, and the lowest number of calls from service area ranked $N_T - 1$. Let f_k denote the proportion of calls one customer receives from a service area with rank k. Let us also assume that the call arrivals from each service area is a Poisson process, and the mean arrival rate from service area ranked k is λ_k. Then, combining the call arrivals to the mobile host from every service area, the call inter-arrival time is exponentially distributed with mean call arrival rate λ given by $\lambda = \lambda_0 + \lambda_1 + \cdots + \lambda_{N_T-1}$. Let us define the following parameters:

- t_c: the call inter-arrival time.
- $f_c(t)$: the (exponential) probability distribution function of call inter-arrival times.
- λ: the mean call arrival rate, i.e. $f_c(t) = \lambda e^{-\lambda t}$.

- λ_k: the mean call arrival rate from switch ranked k, given by $\lambda_k = f_k \lambda$.
- t_r: the residence time the mobile spends in a service area.
- $f_r(t)$: the (exponential) probability distribution function of mobile residence times.
- $1/\mu$: the mean residence time, i.e. $f_r(t) = \mu e^{-\mu t}$.
- t_h is the residual residence time of the mobile host seen by an arriving call.

Then, from the residual life time property [5], the probability distribution function of the residual residence time t_h is exponential, and $f_h(t) = \mu e^{-\mu t} = f_r(t)$. Thus, the number of expected calls from switch ranked k per move is $\frac{\lambda_k}{\mu}$. Substituting in equation (3), the total cost in the IS-41 scheme is given by:

$$\begin{aligned} \text{TotalCost}_{IS41} &= \text{UPDATE}_{IS41} + \\ &\quad \sum_{k=0}^{N_T-1} \frac{\lambda_k}{\mu} \text{FIND}_{IS41} \\ &= \text{UPDATE}_{IS41} + \frac{\lambda}{\mu} \text{FIND}_{IS41} \end{aligned} \tag{4}$$

IS41-Cache Scheme: The expected total cost in the IS41-Cache scheme is given by:

$$\begin{aligned} \text{TotalCost}_{IS41-Cache} &= \text{UPDATE}_{IS41-Cache} \\ &+ \rho \times \text{FIND}_{IS41-Cache} \end{aligned} \tag{5}$$

Since caching does not affect the IS-41 move update cost, we have:

$$\text{UPDATE}_{IS41-Cache} = \text{UPDATE}_{IS41} \tag{6}$$

Let P_k denote the probability that the location information cached for a mobile host at service area (i.e. switch) ranked k is correct. Thus, P_k defines the cache *correctness* ratio. Alternatively, at steady state, P_k denotes the probability that the mobile has *not* moved since the last call from switch k.

$$\begin{aligned} P_k &= \text{Prob}(t_c < t_h) \\ &= \int_{t_c=0}^{\infty} f_c(t_c) \int_{t_h=t_c}^{\infty} f_h(t_h) dt_h dt_c \\ &= \int_{t_c=0}^{\infty} \lambda_k e^{-\lambda_k t_c} \int_{t_h=t_c}^{\infty} \mu e^{-\mu t_h} dt_h dt_c \\ &= \frac{\lambda_k}{\lambda_k + \mu} \end{aligned} \tag{7}$$

Let C_B be the FIND cost for a call in the basic IS-41 scheme (i.e. without caching), and C_H be the FIND cost for a call to a cached mobile host in the IS41-Cache scheme when the cached location information is correct. A net cost saving is achieved if

$$P_k C_H + (1 - P_k)(C_H + C_B) < C_B \tag{8}$$

Thus, a net cost saving is achieved by locally maintaining (caching) the location information if P_k is larger than a threshold P_k^T given by:

$$P_k > P_k^T = \frac{C_H}{C_B} \tag{9}$$

The costs C_H and C_B are given by:

$$\begin{aligned} C_H &= Cost(VLR_{caller} \leftrightarrow VLR_{callee}) \\ C_B &= Cost(VLR_{caller} \leftrightarrow HLR) + \\ &\quad Cost(HLR \leftrightarrow VLR_{callee}) \end{aligned}$$

Let $LCMR_k$ be the local call-to-mobility ratio at switch ranked k. This ratio is given by:

$$\text{LCMR}_k = \frac{\lambda_k}{\mu} \tag{10}$$

However, a switch sees a mobile host movement only when a call is terminated to the mobile from that switch. That is, the switch does not see the true mobility of the mobile host. In order make a decision whether to cache a location update, a local measurement of cache correctness ratio (i.e. fraction of time the cache contains correct location information when accessed) is used in [2, 6, 4]. That is, the location information of a mobile host is cached in a switch k if the *measured* cache correctness ratio P_k^M is larger than P_k^T. Considering long-term average, $P_k^M \approx P_k$.
Then, the FIND cost for a call to the mobile host from switch ranked k is given by:

$$\text{FIND}_{IS41-Cache}^k = \begin{cases} C_H + (1 - P_k) C_B & \text{if } P_k^M \geq \frac{C_H}{C_B} \\ C_B & \text{if } P_k^M < \frac{C_H}{C_B} \end{cases} \tag{11}$$

Then, the total cost in the IS41-Cache scheme is given by:

$$\begin{aligned} \text{TotalCost}_{IS41-Cache} &= \text{UPDATE}_{IS41-Cache} \\ &+ \sum_{k=0}^{N_T-1} \frac{\lambda_k}{\mu} \text{FIND}_{IS41-Cache}^k \end{aligned} \tag{12}$$

where $\text{UPDATE}_{IS41-Cache}$ and $\text{FIND}_{IS41-Cache}^k$ are given by equations (6) and (11), respectively.

FD-Cache Scheme: In FD-Cache, location updates of the mobile host are disseminated to all LRs. However, the time required to complete a location information dissemination is small compared to the call interarrival time. Hence, the likelihood of a cache lookup pointing to a wrong LR is negligibly small. Thus electing to cache the location information of a mobile host always yields cost benefit compared to not caching the location information. However, caching involves processing and memory usage. Thus, the number of mobile host location entries maintained in a cache is limited. A number of caching policies can be employed at each switch to decide whether to cache a mobile host. In this paper, we cache a mobile host at switch k if the average rate λ_k of calls terminated to the mobile host exceeds a threshold rate λ^T.

The total cost in the FD-Cache scheme is given by:

$$\text{TotalCost}_{FD-Cache} = \text{UPDATE}_{FD-Cache} + \rho \times \text{FIND}_{FD-Cache} \quad (13)$$

The location update information should clearly be disseminated efficiently for FD-Cache to be most effective. Full dissemination could be achieved efficiently by propagating location update information over a logical spanning tree. In this case, the UPDATE cost in FD-Cache is given by:

$$UPDATE_{FD-Cache} = (|V| - 1)C_l \quad (14)$$

where C_l is the average cost of a link, and V is the set of all nodes (switches) in the network.

The FIND cost is given by:

$$\text{FIND}^k_{FD-Cache} = \begin{cases} Cost\,(LR_{caller} \leftrightarrow LR_{callee}) & \text{if } \lambda_k \geq \lambda^T \\ Cost(LR_{caller} \leftrightarrow HLR) + \\ \quad Cost(HLR \leftrightarrow LR_{callee}) & \text{if } \lambda_k < \lambda^T \end{cases} \quad (15)$$

Thus, the total cost of the FD-Cache scheme is given by:

$$\text{TotalCost}_{FD-Cache} = \text{UPDATE}_{FD-Cache} + \sum_{k=0}^{N_T-1} \frac{\lambda_k}{\mu} \, \text{FIND}^k_{FD-Cache} \quad (16)$$

where $\text{UPDATE}_{FD-Cache}$ and $\text{FIND}^k_{FD-Cache}$ are given by equations (14) and (15), respectively.

5 Numerical Results

In this section, we present numerical results to compare the cost benefits of the schemes analyzed in Section 4. We assume a grid $N \times N$ network topology. For this network topology, the average distance between any two nodes is given by $1.333(N/2)$. As mentioned earlier, it is assumed that VLR (or LR) and HLR are co-located with the MSC (i.e. each node represents a VLR/LR, HLR and MSC). Table 1 shows our cost assumptions and parameters.

We assume the probability f_k that an incoming call is from a particular service area k (i.e. switch ranked k) is a linearly decreasing function of k. Since it is unlikely that a user will receive calls from each and every service area in the network, we assume that the user has non-zero probability of receiving calls from only n service areas, according to function f_k defined in Table 1. The highest value of a is $2/n$, in which case the probability of a call arrival from service area ranked $n-1$ or higher is zero. If the user's call probability is identical for all n service areas, then a is equal to $1/n$. In this case, $f_k = 1/n$ (i.e. uniform distribution). In the cost comparison that follows, a is assumed to be equal to $2/n$, and n is assumed to be $N^2/3$ (i.e. third of the total nodes in the network).

In IS-41 caching, the location information of a mobile is cached at a switch if the cache correctness ratio P^M is larger than the caching threshold C_H/C_B, which is equal to 1/2 given the parameters in Table 1. Since we consider long-term average, P^M is equal to the probability given by equation (7). In our full dissemination based caching (FD-Cache), the local caching threshold λ^T is varied from 0.01 to 0.5.

In Figure 2, we compare the various schemes in terms of the FIND cost, i.e. the cost of locating a mobile. For both networks with 25 and 100 service areas, FD-Cache outperforms all other schemes for all caching threshold values. Figure 3 shows the total cost of location management (i.e. FIND plus UPDATE) for the various schemes. Caching is more beneficial for higher average call arrival rates. Furthermore, FD-caching is more cost efficient when the number of service areas is small. In FD-Cache, the higher the value of the caching threshold λ^T is, the lower the likelihood of caching is, and consequently the total cost is higher for higher values of caching threshold.

When the average call arrival rate is low, caching does not significantly improve cost performance. Varying the caching threshold λ^T significantly affects the FD-Cache cost only when the average call arrival rate λ is high. The reason is that as the caching threshold increases, the number of calls that do not benefit from caching increases. Thus, when the mean call arrival rate is high, the performance of FD-Cache degrades (or total cost increases) as the contribution to the total cost from calls that do not benefit from caching becomes significant.

Link	Average cost	Justification
Single hop	C_l	Cost involved over a single hop
$VLR/LR \leftrightarrow HLR$	$1.33(N/2)C_l$	Average cost to an HLR
$VLR/LR_{caller} \leftrightarrow VLR/LR_{callee}$	$1.33(N/2)C_l$	Average cost between any two nodes
Mean residence time ($\frac{1}{\mu}$)	10 hours	
Mean call arrival rate (λ)	$0.5 - 6$ calls per hour	
f_k: prob. that an incoming call arrived from switch k	$f_k = \begin{cases} a - \frac{2(na-1)}{n(n-1)}k & \text{if } k < n \\ 0 & \text{otherwise} \end{cases}$ f_k is a linearly decreasing function of rank k, where n is the maximum number of nodes with non-zero call arrival probability	

Table 1. Cost assumptions and parameters

The UPDATE cost due to full dissemination is proportional to the number of service areas N (cf. equation (14)). Hence, when N is large and the average call arrival rate is low, the UPDATE cost dominates the total cost of FD-Cache. This is why FD-Cache is less cost effective than both IS-41 and IS41-Cache at low mean call arrival rates and large networks. In summary, FD-Cache is most effective for smaller caching threshold values (or equivalently, with larger cache sizes), on smaller networks, and at higher call arrival rates.

6 Conclusion

In this paper, we analyzed caching schemes used to quickly determine the current location of a mobile user in personal communication networks. We analyzed a caching strategy where the location of users who move frequently is cached. This mobility is measured locally by each node based on *only* those movements that it sees between successive calls to that mobile. To improve the mobility measure, we proposed a caching scheme based on fully disseminating the location updates of mobiles to *every* node so as to increase the chance that the cache entry points to the correct location of the mobile user. We confirmed the premise of our full dissemination based caching scheme through extensive numerical results.

We are currently investigating the various schemes using a detailed simulation model. Our preliminary results show that the performance improvement in the *delay* to find a mobile significantly offset the performance loss due to the *message overhead* of fully disseminating location updates. We are also investigating hierarchical based solutions to improve scalability. We will report these results in a future paper.

References

[1] Cellular radiotelecommunications intersystem operations, December 1990.

[2] H. Harjono, R. Jain, and S. Mohan. Analysis and simulation of a cache based auxiliary location strategy for PCS. In *IEEE Conf. Networks for Personal Communication*, March 1994.

[3] J. Ho and I. Akyildiz. Local anchor scheme for reducing location tracking costs in pcns. *IEEE/ACM Transaction in Networking*, October 1996.

[4] R. Jain, Y-B. Lin, C. Lo, and S. Mohan. A caching strategy to reduce network impacts of PCS. 12:1434–1444, October 1994.

[5] L. Kleinrock. *Queueing Systems*, volume I and II. New York: Wiley, 1976.

[6] Y.B. Lin. Determining the user locations for personal communications networks. In *IEEE Transaction on Vehicular Technology*, volume 43, pages 466–473, 1994.

[7] C. Lo, R. Wolff, and R. Bernhards. An estimate of network database transaction volume to support personal communication services. In *IEEE First international conference on Universal Personal Communications*, pages 236–241, 1992.

[8] S. Rangarajan, K. Ratnam, and A. Dahbura. A fault-tolerant protocol for location directory maintenance in mobile networks. In *Proc. of the 25th International Symposium on Fault-Tolerant Computing*, June 1995.

[9] K. Ratnam, I. Matta, and S. Rangarajan. Analysis of caching-based location management in personal communication networks. Technical Report NU-CCS-99-02, College of Computer Science, Northeastern University, Boston, MA 02115, May 1999. www.cs.bu.edu/fac/matta/publications.html

[10] K. Ratnam, S. Rangarajan, and A. Dahbura. An efficient fault-tolerant location management protocol for personal communication networks. *IEEE Transactions on Vehicular Technology*. To appear.

[11] N. Shivakumar, J. Jannink, and J. Widom. Per-user profile replication in mobile environments: Algorithms, analysis, and simulation results. *ACM/Baltzer Journal of Mobile Networks and Applications*, 2(2):129–140, 1997. www-db.stanford.edu/~widom/pubs.html.

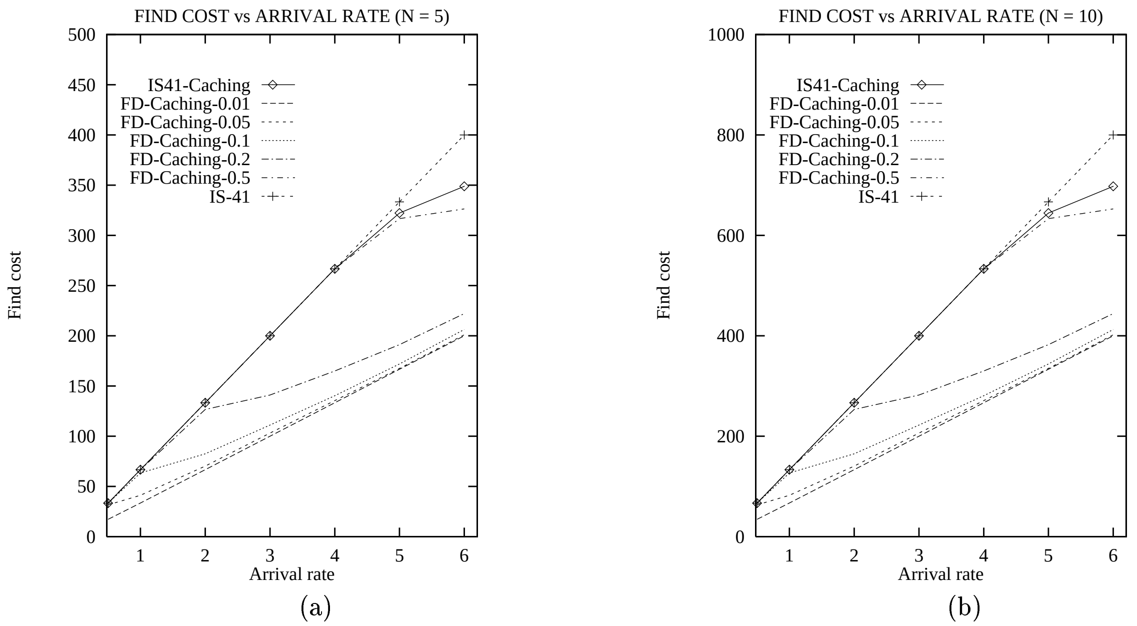

Figure 2. (a) Find cost for 25 service areas, (b) Find cost for 100 service areas.

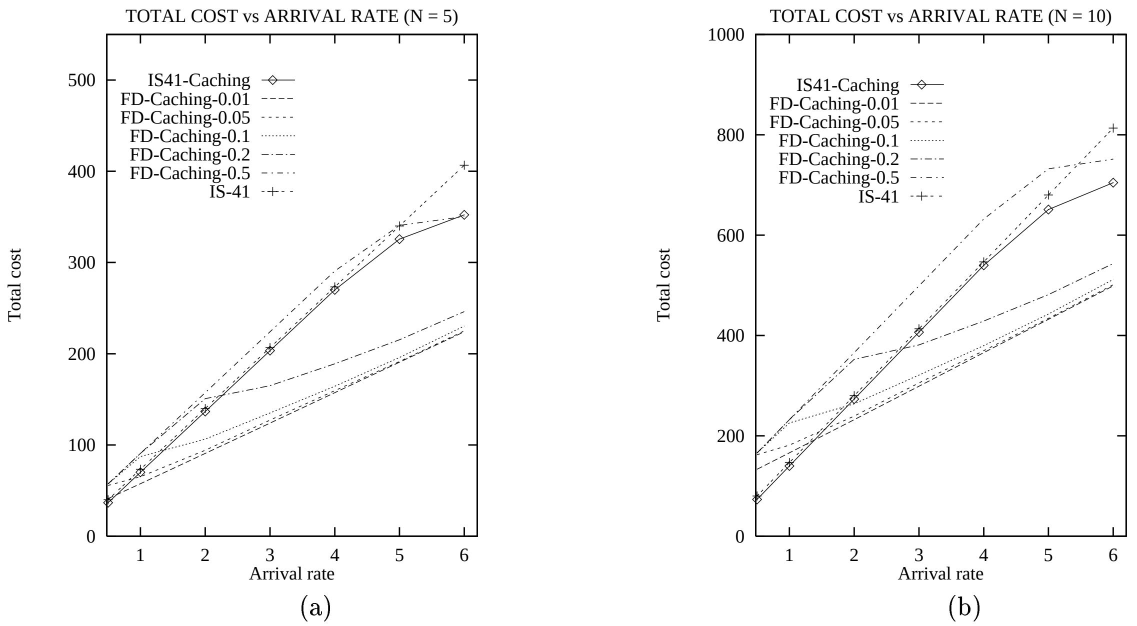

Figure 3. (a) Total cost for 25 service areas, (b) Total cost for 100 service areas.