Journal of Interconnection Networks

A Fully Distributed Location Management Scheme for Large PCS Networks*

Karunaharan Ratnam
Cisco Systems, 300 Apollo Drive, Chelmsford, MA 01824
kratnam@cisco.com

Ibrahim Matta
Dept. of Computer Science, 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

Received (3 March 2000)
Revised (7 February 2001)

This paper studies the effectiveness of a fully distributed location management scheme for PCS networks. Using analytical modeling and numerical simulation, we show that fully replicating location information is both appropriate and efficient for small PCS networks. Then, we extend the scheme in a hierarchical environment so as to reduce overhead and scale to *large* PCS networks. Through extensive numerical results, we show the superiority of our scheme compared to the current IS-41 standard.

Keywords: Location Management; PCS Networks; Centralized versus Distributed; Performance Analysis.

1. Introduction

One of the challenging tasks in a *Personal Communication Services* (PCS) network is to efficiently maintain the location of PCS subscribers who move around freely with their wireless unit (hereafter called mobile host or *mobile* for short). In North America, Telecommunications Industry Association's interim standard *IS-41* [1,2] is used for managing location information of subscribers and enabling them to send and receive calls and other services such as messaging and data service.

In a PCS network [3], every subscriber is registered with a home network, the *Home Location Register (HLR)* of which maintains the subscriber's current physical location. In IS-41, this physical location is the ID of the *Mobile Switching Center (MSC)* currently serving the subscriber. If the subscriber has roamed to another region then he/she has to register with the *Visitor Location Register (VLR)* that covers the new region. During registration, the VLR will contact the subscriber's

*This work was supported in part by NSF grants CAREER ANI-0096045 and MRI EIA-9871022.

HLR, and the HLR will update its database to reflect the new location of the subscriber. If the mobile has registered with some other VLR before, HLR will send a registration cancellation message to it.

In IS-41 [1], an incoming call is routed to the called subscriber as follows. The dialed call is received by the MSC in the home system. This MSC is called the *originating MSC*. If the mobile host is currently being served by the originating MSC (i.e. the mobile host is not roaming), then this MSC queries the HLR to obtain the registration status and feature information of the mobile host. After receiving the response from the 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 the circuit to terminate the call to the mobile host.

Figure 1 shows how a call is 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 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 originating MSC through the HLR. The originating MSC then routes 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 at the mobile.

Thus, HLR is a critical entity in the IS-41 location management system. There are many disadvantages to having a centralized location management scheme such as the scheme used in IS-41. One disadvantage is that since every location request as well as location registration are serviced through an HLR, in addition to the HLR being overloaded with database lookup operations [4], the traffic on the links leading to the HLR is heavy. This, in turn, increases the time required to establish a connection to a mobile host. Another 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. Thus, HLR is a single point of failure in the network.

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 to them by a long distance carrier. 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.

Many location management schemes and improvements to IS-41 have been proposed in recent years (e.g. [5,6]). In [7,8] we presented a novel approach for efficient location management by fully distributing the location information across Location Registers (LRs). These LRs replace the centralized VLRs and HLRs which are found in current PCS networks. *Thus, a unique feature of our proposed scheme is availability (fault-tolerance) by not having the concept of "home."* Since there are no HLRs or VLRs in this system, each LR maintains the location information of not only the mobiles that are local to it, but also of other mobiles in the network [7,8]. That is, the location information of all mobile hosts are fully replicated in all the

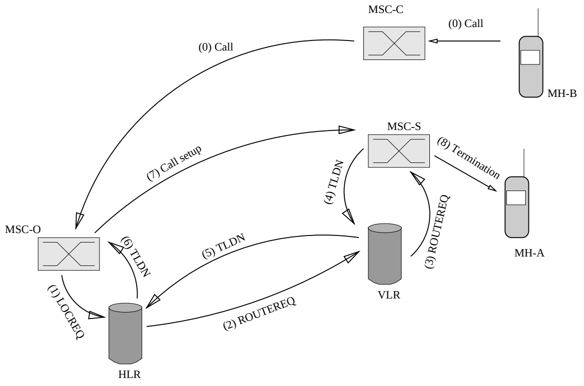

Figure 1: Call delivery in IS-41 and the tromboning problem

LRs. The LRs are distributed throughout the network. An LR may serve one or more MSCs just like the VLR in the PCS architecture. An LR/VLR could also co-exist with an MSC, and serves only that MSC.

LRs function as both the location registry for the 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. When a call request arrives at the local LR, this LR can directly contact the *serving* LR (cf. Figure 2), thus avoiding the tromboning problem present in the current IS-41 standard. Details of our scheme including correctness arguments can be found in [7,8]. In this paper, we analyze this fully distributed location management scheme. In order to scale to *large* PCS networks, we extend our scheme by organizing the LRs hierarchically so as to reduce the cost of updating location information.

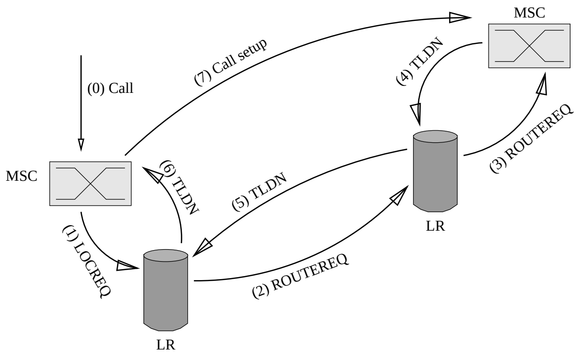

Figure 2: Call delivery in our fully distributed location management

The paper is organized as follows. Section 2 analyzes our location management scheme for flat (non-hierarchical) networks [7,8] and compare it to the IS-41 scheme.

Numerical results are presented in Section 3, and simulation results in Section 4. Section 5 extends our scheme to hierarchical networks. Section 6 analyzes this hierarchical extension and numerical results are then presented in Section 7. Section 8 concludes the paper.

2. Performance Analysis for Flat Network

In this section, we analyze our recently proposed fully distributed (FD) location management scheme [7,8] and compare it with that of the IS-41 scheme. For simplicity, it is assumed that there is only one MSC per service area, and the LR/VLR is co-located with the MSC. Thus, we use LR to indicate an MSC/LR combination, and VLR to indicate MSC/VLR combination. In both schemes (fully replicated and IS-41), the total cost consists of *UPDATE cost* and *FIND cost*. The *UPDATE cost* covers all the costs involved in mobile host registration and location update. In the case of fully distributed location management, *UPDATE cost* also includes the cost involved in the dissemination of location information. The *FIND cost* covers all the costs involved in terminating a call to mobile host. In the case of IS-41, *FIND cost* consists of all the costs involved in the call termination as depicted in Figure 1. However, since we have assumed that VLR is co-located with the MSC, *FIND cost* basically consists of the cost of signaling between originating area VLR and HLR, and of signaling between HLR and serving VLR. In order to compare the cost efficiency of our FD scheme and the IS-41 scheme, we use the expected total cost incurred for a mobile host while it is in a single LR (or VLR) service area as the comparison metric. The total cost includes the *UPDATE cost* incurred for registering the mobile host when it moved into the LR (or VLR) service area, and the *FIND cost* incurred for *every* call terminated to the mobile host while it is in this service area and before it moves to another service area.

In IS-41, UPDATE involves the new VLR registering the MH with its HLR, and the HLR sending registration cancellation to the old VLR. Hence the *UPDATE cost* is given by:

$$\begin{aligned} UPDATE_{IS41} &= Cost(VLR_{new} \leftrightarrow HLR) + \\ &\quad Cost(HLR \leftrightarrow VLR_{old}) \end{aligned} \tag{2.1}$$

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 (cf. Figure 1):

$$\begin{aligned} FIND_{IS41}^{roam} &= Cost(VLR_{caller} \leftrightarrow VLR_{orig}) + \\ &\quad Cost(VLR_{orig} \leftrightarrow HLR) + \\ &\quad Cost(HLR \leftrightarrow VLR_{callee}) \end{aligned} \tag{2.2}$$

and for a non-roaming mobile host, the FIND cost is given by:

$$\begin{aligned} FIND_{IS41}^{local} &= Cost(VLR_{caller} \leftrightarrow VLR_{orig}) + \\ &\quad Cost(VLR_{orig} \leftrightarrow HLR) \end{aligned} \tag{2.3}$$

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

In our fully distributed (FD) scheme, new location information needs to be disseminated to all LRs in the network. For fault tolerance, we use *flooding* to implement full dissemination. For simplicity, assume that for any mobile host there is at most one location update propagating in the network at any point in time. Also assume normal operation of update (i.e. no message corruption). Let $Graph(V, E)$ be the topology, where V is the set of all nodes (LRs) and E is the set of all links. Let adj_v be the adjacency list of node v, and C_l be the average cost of a link between two adjacent (neighbor) MSCs/LRs. Then, the total update cost per move in the FD scheme is given by:

$$UPDATE_{FD} = C_l \left(1 + \sum_{v \in V} (|adj_v| - 1)\right) \tag{2.4}$$

Until the update about the mobile host is completed, some of the calls to the mobile may arrive to an old (incorrect) LR that is no longer serving the mobile host. Since the old LR is likely to be immediately updated after the move because of its proximity to the new LR, we assume that there will be at most one call forwarding involved from the old LR to the new LR. Then the FIND cost in the FD scheme is given by:

$$\begin{aligned} &FIND_{FD} = Cost\,(LR_{caller} \leftrightarrow LR_{callee}) + \\ &P\,(\text{call arrives to } LR_{old}) \times Cost\,(LR_{old} \leftrightarrow LR_{callee}) \end{aligned} \tag{2.5}$$

Now we need to find the probability that the call arrives to an old (incorrect) LR. We define the following (cf. Figure 3):

- $f_c(t)$: p.d.f. of the call arrival process to the mobile
- λ: average number of call arrivals per second
- t_r: amount of time the mobile spends in a service area
- $f_r(t)$: p.d.f. of the random variable t_r. It is assumed to be exponentially distributed with mean residence time $1/\mu$, i.e. $f_r(t) = \mu e^{-\mu t}$
- t_u: time taken for the location information to reach the LRs.
- $f_u(t)$: general distribution describing t_u. If the total number of LRs is M, after the mobile moves, let the time taken for the update to reach LR i be t_i, then $(t_1, t_2, ...t_M)$ describes t_u.

$$\begin{aligned} &P\,(\text{call arrives to } LR_{old}) = \\ &\frac{\text{Expected num. of calls arriving during } t_u}{\text{Expected num. of calls arriving during } t_r} \end{aligned} \tag{2.6}$$

$$\begin{aligned} &\text{Expected num. of calls arriving during } t_u = \\ &\int_0^{\infty} \left[\int_0^{t_r} \lambda\, t\, f_u(t)\, dt\right] f_r(t)\, dt \end{aligned} \tag{2.7}$$

$$\begin{aligned} &\text{Expected num. of calls arriving during } t_r = \\ &\int_0^{\infty} \lambda\, t\, f_r(t)\, dt = \frac{\lambda}{\mu} \end{aligned} \tag{2.8}$$

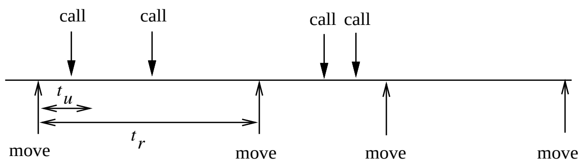

Figure 3: Diagram depicting the parameters

The solution to equation (2.7) depends on the distribution $f_u(t)$, and it can generally be solved by numerical techniques. Here, we solve it for the simple case of exponentially distributed t_u with mean $1/\alpha$.

$$\text{Expected num. of calls arriving during } t_u = \frac{\lambda}{\alpha} - \lambda\mu\left(\frac{1}{\alpha(\alpha+\mu)} + \frac{1}{(\alpha+\mu)^2}\right) \tag{2.9}$$

$$P\,(\text{call arrives to } LR_{old}) = \frac{\mu}{\alpha} - \mu^2\left(\frac{1}{\alpha(\alpha+\mu)} + \frac{1}{(\alpha+\mu)^2}\right) \tag{2.10}$$

Table 1 shows the cost incurred due to incorrect location information as the update is propagating over the network to all LRs, for a 400-LR network and λ of 10 calls per hour. Given $1/\alpha$ is the average update duration (assuming exponentially distributed duration), this cost is given by:

$$(\lambda/\alpha)\ P\,(\text{call arrives to } LR_{old})\ \ Cost\,(LR_{old} \leftrightarrow LR_{callee}) \tag{2.11}$$

Assuming $Cost\,(LR_{old} \leftrightarrow LR_{callee})$ equals 1 msec, the table shows that the cost due to outdated location information is negligible. We henceforth estimate the FIND cost in the FD scheme as:

$$FIND_{FD} = Cost\,(LR_{caller} \leftrightarrow LR_{callee}) \tag{2.12}$$

Table 1: Cost due to outdated location information

1 / α	**P(call arrives to incorrect LR)**	**Average Cost**
1 sec	2.77×10^{-5}	7.69×10^{-11}
10 sec	2.77×10^{-4}	7.69×10^{-9}
60 sec	1.66×10^{-3}	2.76×10^{-7}

3. Numerical Results for Flat Network

The above analysis provides a way to compare the cost of using the two alternative location management schemes: IS-41 and our fully distributed location management scheme. In this section, we numerically compare the cost by making reasonable assumptions on network deployment and parameters. We assume an $N \times N$ network topology as shown in Figure 4. For this network topology, the average distance

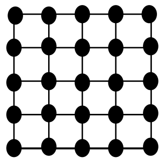

Figure 4: A 5×5 network topology ($M = 25$ service areas)

between any two nodes is given by $1.333(N/2)$. As mentioned earlier, it is assumed that VLR (or LR) is co-located with the MSC. We further assume here that HLR is co-located with the MSC, thus each node in Figure 4 represents a VLR/LR, HLR and MSC. Hence we assume that the link cost between VLR (LR), HLR, and MSC of the same service area is negligible. Table 2 shows our cost assumptions and parameters.

Table 2: Cost Assumptions and parameters for flat network

Link	**Average cost**	**Justification**
Single hop	C_l	Cost of a single hop
$LR_{old} \leftrightarrow LR_{callee}$	C_l	A mobile's move generally involves neighboring LRs
$VLR_{callee} \leftrightarrow HLR$ $VLR_{caller} \leftrightarrow HLR$	$1.33(N/2)C_l$	Average cost to an HLR
$LR_{caller} \leftrightarrow LR_{callee}$ $LR_{caller} \leftrightarrow LR_{old}$	$1.33(N/2)C_l$	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	

As mentioned earlier, the comparison merit is the expected total cost incurred for a mobile host while it is in a single LR (or VLR) service area. The cost we compute here assumes that the mobile host is away from its home service area (i.e. roaming), and is given by:

$$\begin{aligned} &\text{TotalCost}_{IS41} = \text{UPDATE}_{IS41} + \\ &\text{Expected num. of calls per move} \times \text{FIND}^{roam}_{IS41} \qquad (3.13) \\ &\text{TotalCost}_{FD} = \text{UPDATE}_{FD} + \\ &\text{Expected num. of calls per move} \times \text{FIND}_{FD} \qquad (3.14) \end{aligned}$$

where λ/μ is the expected number of calls per move, and $\text{UPDATE}_{IS41}, \text{FIND}^{roam}_{IS41}$, UPDATE_{FD} and FIND_{FD} are given by equations (2.1), (2.2), (2.4) and (2.12), respectively.

The expected number of calls per move is often referred to as the *call-to-mobility ratio* [5]. Figure 5 shows the total cost for IS-41 and fully distributed schemes in a 5×5 network for both uniformally and exponentially distributed location update duration. C_l is taken to be 1. As expected, when the call-to-mobility ratio increases our fully distributed scheme offers better cost performance. For larger networks, the fully distributed scheme introduces heavy update cost, which increases the total cost as can be seen in Figures 6 and 7. In the case of 10×10 network, FD outperforms IS-41 when the call arrival rate is higher than 4 calls per hour compared to 2 calls per hour in the 5×5 network. And, for the 20×20 network, in the range of reasonable

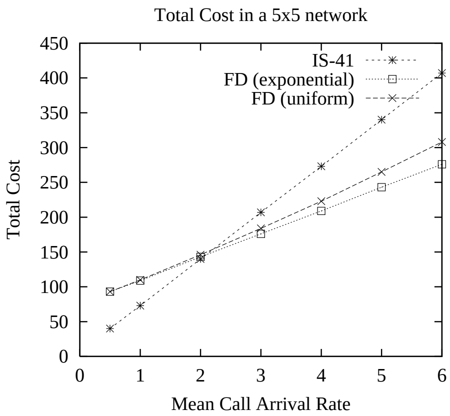

Figure 5: Total cost versus call arrival rate (5×5 network)

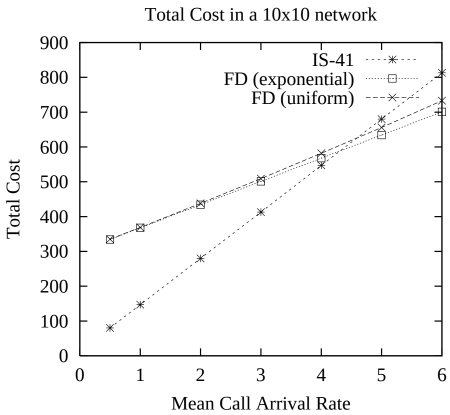

Figure 6: Total cost versus call arrival rate (10×10 network)

call arrival rates, IS-41 outperforms FD. In Section 5, we extend our FD scheme so as to scale to larger networks.

4. Simulation for Flat Network

The fully distributed location management scheme and IS-41 scheme are simulated using CSIM—a commercial process oriented discrete event simulator for the C programming language [9]. CSIM provides structures such as Processes, Facilities, Storage units, Events, Mailboxes, Data Collection structures, etc., to model the interaction of different entities of a simulation. For example, modeling of a single server and queue can be done by declaring the server as a CSIM facility, and the arriving customer as a CSIM process. Arrival of the customer is modeled by a CSIM event that can be scheduled to occur at a future time. The events are stored in a global event list and serviced in the order of increasing event time. CSIM maintains a logical clock, and each event advances the clock by the time difference between the last event and the current event. If the server is free, the customer next in the queue gets the service from the server by reserving the facility, and releases it after

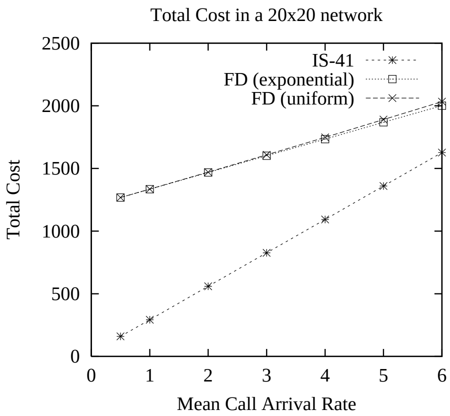

Figure 7: Total cost versus call arrival rate (20×20 network)

obtaining the service. Using CSIM data collection structures, one can retrieve the current statistics (e.g., average queuing delay, throughput, etc).

In our simulation, network nodes—wired nodes and mobile hosts—are modeled as facilities. The communication between the nodes are modeled by exchanging messages between CSIM mailboxes associated with each node. Each LR maintains different mailboxes for wired communication with other LRs and wireless communication with mobile hosts. The mailboxes are further divided into those that handle mobility related messages, and that handle call related messages. Hence there are four mailboxes associated with each LR. The mobile hosts maintain only two mailboxes for wireless communication with their local LR, one handles mobility related messages, and the other handles call related messages.

We assume identical per-hop transmission delays. A table of the distances between the wired network nodes, in terms of number of hops, is passed to the simulation from the input script. Our simulation implementation does not perform actual hop-by-hop routing of packets, instead we use the above table of distances to approximately simulate the delay involved in transmitting a packet from source to destination. Queuing delay is also simulated, but it is simulated at the destination node only. It is assumed that there are no queue overflows (i.e. infinite buffers), and messages are delivered in FIFO order. The message processing delay is assumed to be exponentially distributed. Table 3 lists the parameters used in the simulation.

Table 3: Simulation parameters

Parameter	**Value**	**Distribution**
Message processing time (wired or wireless)	1 ms	Exponential
LR/HLR database access time	5 ms	Exponential
Per-hop transmission delay (wired or wireless)	50 ms	Fixed value
Simulation time	20,000 sec	
Call inter-arrival time	1 every 30 min	Exponential
Number of mobile hosts	250 per LR/HLR	
Mobile host stop-or-go threshold	0.90	

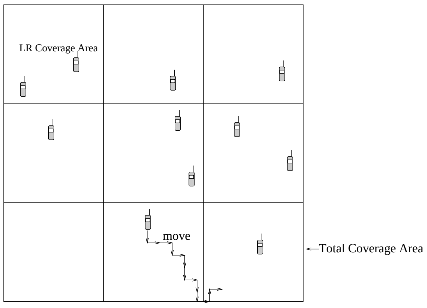

Figure 8: Sample 3×3 network topology showing the LR coverage area and user mobility pattern

There is no single model to capture the mobility of users. In fact, user mobility depends on many factors: terrain, time of day, mode of transportation, to name a few. These give rise to diverse mobility models. In this simulation, we assume a user moves in a zig-zag fashion. That is, the total coverage area map is overlayed with grid lines, and the user moves from one grid point to another. On each move, the direction of movement is decided by the outcome of a coin toss. If outcome is a head, movement is in north-south direction, otherwise movement is in east-west direction. However, if the movements in the past are, say in north or east directions, then the future move is also in north or east direction depending on the outcome of the coin toss. But, the direction is reversed when the user goes over the boundary of the total coverage area (see Figure 8) to ensure that the user stays within the boundaries. The amount of time a mobile host spends at a point before deciding to move is computed from the user configurable average residence time a mobile host spends in an LR service area. Another parameter, stop-or-go, is added to the characteristic of the mobile host to capture stop and go behavior. That is, when the time is up for the next move, a random number is drawn from a uniform distribution. If the mobile host stop-or-go threshold is larger than the random number, the mobile host moves to the next point. Otherwise, it stays in the same point until the next move.

Simulation Results

The performance of IS-41 and fully distributed (FD) location management schemes are compared based on the time to locate a mobile host (i.e. FIND time) in each scheme. Note that here the FIND time includes delays due to the dissemination of location update messages over the network. Results in this section are taken for a network where the LRs (or HLR/VLRs in the case of IS-41) are located at the cross-points of a mesh as shown in Figure 4. We simulated two networks with different sizes—25 service areas and 100 service areas—and the average number of mobile hosts per service area is 250. Simulating a large-scale network is impractical. Even for a small network (i.e. 25 service areas), there are 6250 mobile hosts and 25 LRs. The memory (main and swap) required to simulate 6250 mobile hosts in CSIM and the total time required to complete the simulation is too large.[†] Hence, we scaled down the simulation by 1:10. That is, in the simulation, each simulated mobile host represents 10 physical mobile hosts. Hence, the simulated mobile host has to move 10 times faster and receive 10 times more calls than a physical mobile

[†]The simulation was run on a dual 500MHz Pentium III Xeon processor machine with 1MB cache and 2GB of memory running Linux 2.0.36 kernel.

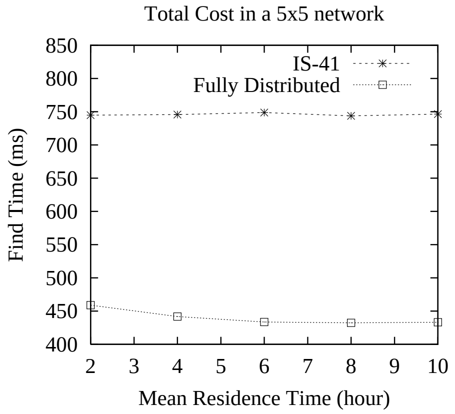

Figure 9: FIND time in a 5×5 network

host. This is done by dividing the average residence time by 10 and multiplying the average call arrival rate by 10.

As shown in Figure 8, each LR covers a square geographical area, and there are no coverage overlaps. The coverage area is assumed identical for all LRs. Mobile hosts are randomly spread across the total coverage area, and the total number of mobile hosts in the simulation is the number of service areas times 25.

Figure 9 shows the total time taken to locate a mobile host (i.e. FIND time) using fully distributed location management and IS-41 location management for a network with 25 service areas. The FIND time is mainly affected by the network delay. User mobility rate in the IS-41 scheme has minimal effect on FIND time which stayed around 745 ms for this network. This means that the load offered to the HLR is negligibly small, and the HLR processing delay, which includes database access time, queuing and message processing time, has little effect on the FIND time. In the case of FD, the FIND time increases with the user mobility rate due to the increase in the load caused by location information updates flooded in the network. When the average residence time is 2 hours, FIND time is 460 ms. The FIND time gradually drops with the residence time and reaches 433 ms when the residence time is 10 hours.

The effect of user mobility on the FIND time is even more prominent in a larger network. In Figure 10, we plot the FIND time for IS-41 and FD schemes for a network of 100 service areas. The total number of mobile hosts in this network is 25,000. Again, the FIND time using IS-41 remains almost constant at 1.43 seconds. Note that, for mesh networks, the average distance between any two nodes is proportional to the square root of the size of the network. This explains why the FIND time in a 100-service area network is twice that of a 25-service area network. It also means that the effect of HLR processing delay is still insignificant.

Since there are 4 times more mobile hosts in a 10×10 network than in the 5×5 network, each LR has to process 4 times more location update information when using FD. Consequently, LR processing delay is a significant contributor to the increased FIND time especially at high mobility rates. When the residence time is 2 hours, the FIND time in a 10×10 network is 1825 ms compared to the FIND time of 460 ms in the case of a 5×5 network. As the residence time increases, the number of location updates reduces. Hence, at higher residence times, contribution from LR processing delays on FIND time is less, and the effect of per-hop transmission delays dominates. In this case, as expected, for a residence time of 10 hours, the

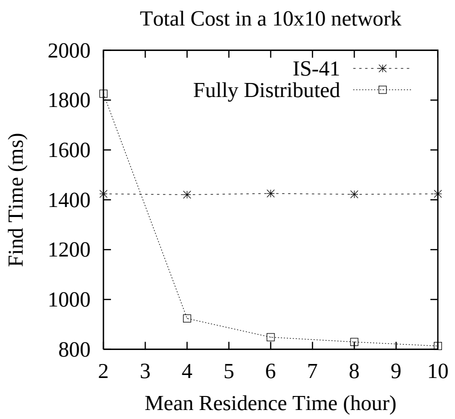

Figure 10: FIND time in a 10×10 network

FIND time in a 10×10 network is almost twice that of a 5×5 network.

Due to these reasons, for large networks when the average residence time is low, IS-41 is better suited. However, if the residence time is higher, using FD, the FIND time can be greatly improved over IS-41.

5. Hierarchical Location Management

Our fully distributed location management scheme requires that new location information about all mobiles be disseminated to all the LRs in the network. As the size of the network grows, location information dissemination not only consumes a significant portion of the network bandwidth but also consumes significant portion of LR resources to process large number of update messages. In addition, the gain of employing full dissemination diminishes with the size of the network. That is, for a large network, it is impractical to have a location management scheme based on full location information dissemination. Full location information dissemination can be avoided by logically arranging LRs in a hierarchical fashion—a tree structure as in [10] or a cluster-supercluster arrangement as in [11]. The idea here is to divide the LRs into hierarchy of clusters, and confine location information dissemination to within the clusters as much as possible. This section analyzes the performance of our fully distributed location management in a hierarchical environment and assesses its applicability and benefits. *Our scheme is different from other hierarchical schemes (e.g., [6]) in that its goal is not only to reduce the overhead of location management, but also to uniquely provide high availability through (selective) replication of location information.*

5.1. *Proposed Hierarchical Location Management*

Figure 11 shows the conceptual arrangement of the LRs in a hierarchical network under our proposed scheme. The proposed approach uses a distributed location management. The mobile hosts are not associated with a home location register like in IS-41. Each LR maintains the location information of all the mobiles that are currently being served in the subtree rooted from the LR. It *also* maintains the location of the mobiles that belong to the subtree rooted from its sibling LRs. Note here that the subtree rooted from a leaf node contains only that leaf node. If a mobile host is being served by one of the descendants of an LR, then the LR

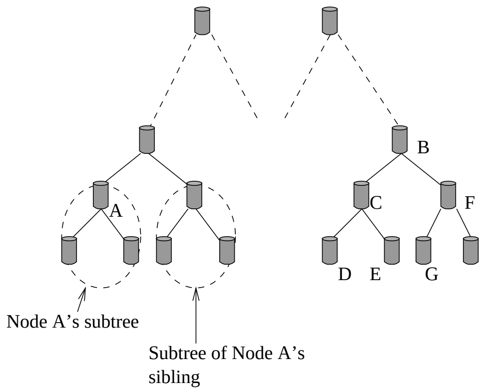

Figure 11: Conceptual diagram showing the hierarchical arrangement

maintains the ID of its immediate child LR, whose subtree contains the mobile host, to track the mobile host. Referring to Figure 11, if a mobile host is in the service area of LR **D**, then location information in LR **C** for the mobile host would point to LR **D**, but the location information in LR **B** for the same mobile host would point to LR **C**. For the mobile hosts that reside in the subtree of a sibling, the LR maintains its sibling's ID to track the mobile host. That is, location information in LR **F** for that mobile host served by LR **D** would be LR **C**. This way the location information of a mobile host is only maintained by the following LRs: serving LR of the mobile host, sibling LRs of the serving LR, ancestor LRs of the serving LR, and sibling LRs of the ancestors. That is, location information of the mobile host being served by **D** are maintained only in the LRs **D**, **E**, **C**, **F**, **B**, and so on. LRs **A** and **G** do not maintain the location information for that mobile host.

Tracking the LR serving a mobile host involves traversing the LR tree hop-by-hop until the serving LR is reached. If the location entry for a mobile host does not exist in an LR, then the tracking request is forwarded to the LR's parent LR. In this way the tracking request traverses the tree upwards until the LR which has the location information for the mobile host is reached. That LR forwards the tracking request to the LR pointed to by the location information. Here, location tracking traverses laterally. From there, it traverses downwards until the LR currently serving the mobile host is reached. For example, if **G** were to track the LR of a mobile host being served by **D**, **G** forwards the tracking request to **F**. **F** forwards the request to **C**, which forwards it to **D**. This information is returned back to **G**.

5.2. *Registration and Location Update Algorithm*

Mobile hosts identify their current LR by the periodic beacon message broadcasted by the base stations. If the mobile host receives a beacon message with a different service area than its currently registered service area, it registers with the new LR serving the area. The registration message contains the id of the mobile host. This registration message is propagated to the serving LR of the area. Upon receiving the registration message, in addition to sending registration confirmation back to the mobile host, the LR also sends a location update message to other LRs in the dissemination list. The dissemination list of an LR contains all its sibling LRs and the parent LR [12].

6. Analysis for Hierarchical Network

In this section, we try to answer the question of when our hierarchical location management system is cost efficient compared to IS-41 and the flat fully distributed location management evaluated in Section 2. Here we analyze a two-level hierarchy as shown in Figure 12. Note here that this analysis can be extended in a straightforward way to higher levels of hierarchy as well.

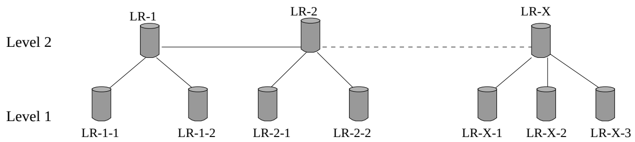

Figure 12: Conceptual diagram showing a two-level hierarchical arrangement used in the analysis

Now, if a mobile host moves across level-1 LRs belonging to the same level-2 LR, henceforth called *level-1 move*, then the cost of updating the move is the cost of distributing the location update to all the LRs in that cluster only. We assume here that location information is carried reliably. Then, instead of disseminating location updates to other LRs using flooding as in Section 2, they can be efficiently disseminated to all the LRs in the dissemination list over a spanning tree rooted at the new level-1 LR that is currently serving the mobile host. Then the cost is given by:

$$\mathrm{UPDATE}_{H-level1} = C_{level1} \times (M_{level1} - 1) \tag{6.15}$$

Here, C_{level1} is the average cost of the link connecting two adjacent level-1 LRs, and M_{level1} is the average number of LRs in a level-1 cluster.

If a mobile host moves across level-2 LRs, henceforth called *level-2 move*, then the cost of updating the move is the cost of distributing the location update in the new cluster plus the cost of updating all the level-2 LRs to point to the new level-2 LR plus the cost of distributing "delete" message to all the LRs in the old cluster.

$$\begin{aligned} \mathrm{UPDATE}_{H-level2} &= 2 \times C_{level1} \times (M_{level1} - 1) \\ &\quad + C_{level2} \times (M_{level2} - 1) \end{aligned} \tag{6.16}$$

Assuming $P_{local-move}$ is the probability that a mobile host move is across the LRs in level-1, the update cost in the hierarchical system is given by:

$$\begin{aligned} &\mathrm{UPDATE}_H = P_{local-move} \times \mathrm{UPDATE}_{H-level1} + \\ &(1 - P_{local-move}) \times \mathrm{UPDATE}_{H-level2} \end{aligned} \tag{6.17}$$

The find cost (location tracking cost) of a mobile depends on whether the call is from a mobile host in the local cluster or not. The find cost for a call from a local cluster is given by:

$$\begin{aligned} \mathrm{FIND}_{local-call} &= \mathrm{Cost}\,(LR_{caller} \leftrightarrow LR_{callee}) \\ &= \mathrm{Cost}\,(LR_{local} \leftrightarrow LR_{local}) \end{aligned} \tag{6.18}$$

If the call is from a mobile in another cluster (henceforth called a remote-call), then the calling party LR (a.k.a. LR_{caller}) needs to contact its parent LR (a.k.a. $LR_{caller-level2}$) to track the callee. $LR_{caller-level2}$ will contact the callee level-2 LR (a.k.a. $LR_{callee-level2}$), which in turn will contact the currently serving LR of the callee (a.k.a. LR_{callee}). Hence the find cost of a remote-call is given by:

$$\begin{aligned}
&\mathrm{FIND}_{remote-call} = \mathrm{Cost}\,(LR_{caller} \leftrightarrow LR_{caller-level2}) \\
&+ \mathrm{Cost}\,(LR_{caller-level2} \leftrightarrow LR_{callee-level2}) \\
&+ \mathrm{Cost}\,(LR_{callee-level2} \leftrightarrow LR_{callee})
\end{aligned} \tag{6.19}$$

Let $P_{local-call}$ be the probability that the call that arrived is from a mobile in the local cluster. Then the find cost in the hierarchical network is given by:

$$\begin{aligned}
\mathrm{FIND}_H \;&=\; P_{local-call} \times \mathrm{FIND}_{local-call} \\
&\quad + (1 - P_{local-call}) \times \mathrm{FIND}_{remote-call}
\end{aligned} \tag{6.20}$$

Following the same method of analysis as in Section 2, given λ is the call arrival rate to a mobile and $1/\mu$ is the mean of the (exponentially distributed) residence time of the mobile in a service area, the total cost of the location management in the hierarchical network is given by:

$$\begin{aligned}
\mathrm{TotalCost}_H \;&=\; \mathrm{UPDATE}_H + \\
&\quad\ \text{Expected num. calls per move} \times \mathrm{FIND}_H \\
&=\; \mathrm{UPDATE}_H + \frac{\lambda}{\mu} \times \mathrm{FIND}_H
\end{aligned} \tag{6.21}$$

Table 4: Cost Assumptions and parameters for hierarchical network

Parameter	**Value**	**Comment**
N_{LR}	N_{LR}	Total number of LRs
N_c	N_c	Number of clusters
C_{level1}	C_{level1}	Cost of the link connecting adjacent level-1 LRs
C_{level2}	$\sqrt{\frac{N_{LR}}{N_c}}C_{level1}$	Cost of the link connecting adjacent level-2 LRs
$Cost\,(LR_{local} \leftrightarrow LR_{local})$	$1.33(\sqrt{N_{LR}/N_c}/2)C_{level1}$	Average cost between any two level-1 LRs
$Cost\,(LR_{caller} \leftrightarrow LR_{caller-level2})$ $Cost\,(LR_{callee} \leftrightarrow LR_{callee-level2})$	$1/2\sqrt{N_{LR}/N_c}C_{level1}$	Average distance between the level-1 LR and its parent LR
$Cost\,(LR_{caller-level2} \leftrightarrow LR_{callee-level2})$	$1.33(N_c/2)C_{level2}$	Average cost between two level-2 LRs

7. Numerical Results for Hierarchical Network

We consider a mesh topology as shown in Figure 13. If the total number of LRs in the network is N_{LR}, and the number of clusters is N_c, then the average number of LRs in a cluster is N_{LR}/N_c. The level-2 LR is placed along with the level-1 LR at the center of the cluster. If there is no single center LR, then the level-2 LR co-locates with one of the four center LRs. Assuming the cost of the link connecting two adjacent LRs is proportional to the distance between the LRs, parameters C_{level1} and C_{level2} are related by $C_{level2} = \sqrt{N_{LR}/N_c}\ C_{level1}$. The cost between any of the level-1 LR and its level-2 LR is given by $1/2\ \sqrt{N_{LR}/N_c}\ C_{level1}$ for large values of

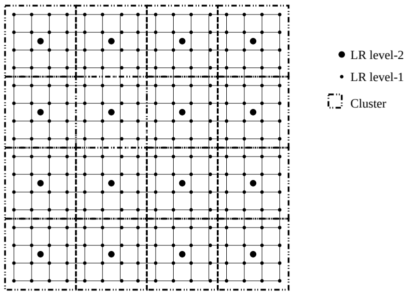

Figure 13: Mesh deployment of Location Registers

$\sqrt{N_{LR}/N_c}$ (greater than 4). Table 4 summarizes values of the parameters involved in the equation for total cost. In the following numerical results, C_{level1} is taken to be 1, $P_{local-move} = 90\%$ and $P_{local-call} = 1/N_c$.

Figure 14 shows the total cost versus call arrival rate for IS-41, our flat FD and hierarchical FD schemes. Both FD schemes implement full dissemination over a spanning tree rather than by flooding. The (two-level) hierarchical FD scheme performs the best.

Figure 15 illustrates how hierarchical FD scales compared to flat FD and IS-41 as the network size increases for $\lambda = 3$ and $\mu = 0.1$. The network size is varied from 50×50 to 200×200, where number of clusters (N_c) equals to 10, 20, 30, and 40, respectively. The results suggest that our hierarchical FD scales much better than IS-41.

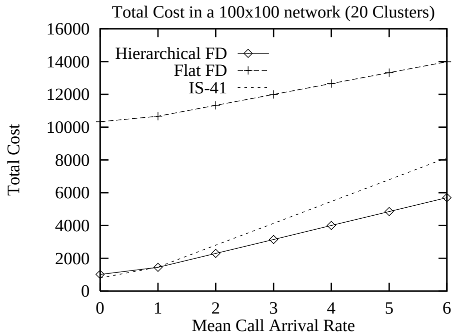

Figure 14: Total cost in a 100×100 network

8. Conclusions

As shown by analytical and numerical performance analysis, our fully distributed (FD) location management scheme is more suitable than IS-41 management. The fully distributed location management not only reduces the overall system cost, but also reduces the call establishment latency and increases the availability of the system. The hierarchical implementation of our FD scheme allows for scaling to large PCS networks while still providing high availability.

Our FD location management approach is general and can make use of other scalability and fast location lookup techniques [13], besides hierarchically organizing the location registers. For example, location information can be disseminated only to

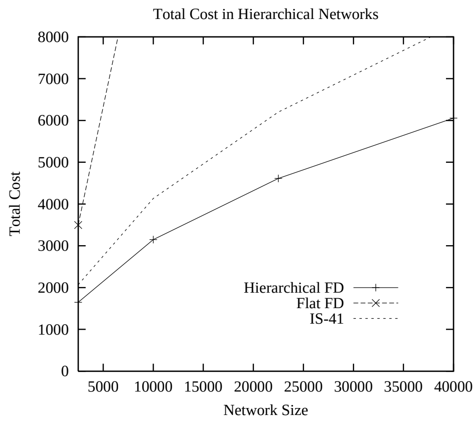

Figure 15: Total cost in hierarchical networks

most frequent callers, who can *cache* it in memory so as to speed up their location lookups. In our future work, we will investigate such extensions and study the tradeoffs between availability and cost/overhead.

References

1. "Cellular radiotelecommunications intersystem operations," December 1990.
2. Michael D. Gallagher and Randall A. Snyder, *Mobile Telecommunications Networking with IS-41*, McGraw-Hill, Inc., 1997.
3. Jan Thörner, *Intelligent Networks*, Artech House, Inc., 1994.
4. C. N. Lo, R. S. Wolff, and R. C. Bernhards, "An estimate of network database transaction volume to support personal communication services," in *IEEE First international conference on Universal Personal Communications*, 1992, pp. 236–241.
5. Harry Harjono, Ravi Jain, and Seshadri Mohan, "Analysis and simulation of a cache based auxiliary location strategy for PCS," in *IEEE Conf. Networks for Personal Communication*, March 1994.
6. Joseph S. M. Ho and Ian F. Akyildiz, "Local anchor scheme for reducing location tracking costs in PCNs," *IEEE/ACM Transaction in Networking*, October 1996.
7. Sampath Rangarajan, Karunaharan Ratnam, and Anton T. 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.
8. Karunaharan Ratnam, Sampath Rangarajan, and Anton T. Dahbura, "An efficient fault-tolerant location management protocol for personal communication networks," *IEEE Transactions on Vehicular Technology*, To appear. Available at www.ece.neu.edu/personal/karu/papers/ieeetvt.ps.gz.
9. "CSIM18-the simulation engine," Mesquite Software, Inc.
10. J. Z. Wang, "A fully distributed location registration strategy for universal personal communication system," in *IEEE Journal on Selected Areas in Communications*. IEEE, Oct. 1993, vol. 11, pp. 850–860.
11. J. Westcott and G. Lauer, "Hierarchical routing for very large networks," in *Proc. of IEEE INFOCOM*, 84.
12. K. Ratnam, I. Matta, and S. Rangarajan, "A fully distributed location management

scheme for large PCS," Tech. Rep. BU-CS-1999-010, Computer Science Department, Boston University, Boston, MA 02215, August 1999.

13. K. Ratnam, I. Matta, and S. Rangarajan, "Analysis of caching-based location management in personal communication networks," in IEEE ICNP '99, 1999.