Location Tracking with Distributed HLR's and Pointer Forwarding

Yi-Bing Lin, *Senior Member, IEEE,* and Wen-Nung Tsai

***Abstract*— Location tracking operations in a personal communications services (PCS's) network are expensive. A location tracking algorithm called *pointer forwarding* has been proposed to reduce the location update cost. The key observation behind forwarding is that if users change PCS registration areas (RA's) frequently, but receive calls relatively infrequently, it should be possible to avoid registrations at the home-location register (HLR) database by simply setting up a forwarding pointer from the previous visitor-location register (VLR). Calls to a given user will first query the user's HLR to determine the first VLR, which the user was registered at, and then follow a chain of forwarding pointers to the user's current VLR.**

To reduce the "find" cost in call delivery, the PCS service provider may distribute HLR databases in the network. This paper integrates the concept of distributed HLR's with pointer forwarding, and the new scheme is referred to as the pointer forwarding with distributed HLR (PFDHLR). Since no registration to the HLR is performed in the pointer forwarding scheme when a user moves to the new locations, the cost of updating multiple HLR's is eliminated in PFDHLR. Our study indicates that PFDHLR may significantly reduce the mobility management cost compared with the single HLR approach.

***Index Terms*— Deregistration, distributed database mobility management, home-location register, personal communications services, registration, visitor-location register.**

I. Introduction

TO SUPPORT mobility in a personal communications services (PCS's) system, strategies have been proposed in protocols such as EIA/TIA IS-41 [1] and ETSI GSM MAP [2].

Two basic operations in PCS mobility management are *registration* (the process that a *portable* or *mobile phone* informs the system of its current location) and *location tracking* (the process that the system locates the portable). Location tracking is required when the network attempts to delivery a call to the mobile user. The mobility management strategies proposed in IS-41 and GSM are *two-level* strategies in that they use a two-tier system of home and visited databases. When a user subscribes the services to a PCS system, a record is created in the system's database called home-location register (HLR). When the mobile user visits a new registration area (RA), a temporary record for the mobile user is created in the visitor-location register (VLR) of the *visited system* (the new location). The VLR then sends a registration message to the HLR. Assume that the messages are delivered using the signaling system no. 7 (SS7) in the public switched telephone network (PSTN). To support nongeographic NPA-NXX-XXXX number in IS-41 Revision B, the VLR does not recognize the address of the HLR from the mobile identification number (MIN) of the portable. Instead, the registration message is forwarded to a signal transfer point (STP) through several SS7 network elements. The STP uses a table lookup technique called global title translation (GTT) to identify the HLR address. Then the message is forwarded from the STP to the HLR. A tutorial for PCS signaling using SS7 can be found in [3].

To deliver a call to a mobile subscriber from an *originating switch*, the HLR is queried to find the current VLR of the portable. Then the HLR sends a query to the VLR. The VLR returns a routable address called the temporary-location directory number (TLDN) to the originating switch (in the PSTN) through the HLR. Based on the TLDN, a trunk (voice circuit) is then set up from the originating switch to the portable. Note that the call-delivery procedure consists of two parts. In the first part (referred to as the **find** operation), the VLR is located, and the TLDN is returned to the originating switch. In the second part, the TLDN is used to set up a voice trunk to the portable. The portable is paged in the second part of call delivery. Like the registration process, several SS7 network elements are visited, and a GTT is required to access the HLR in the first part.

Studies [4]–[6] indicated that the message traffic due to PCS location tracking operations is significant. To reduce the location tracking cost, several algorithms have been proposed [6]–[12].

Pointer forwarding schemes proposed in [9] and [13] are based on the observation that in many cases, it should be possible to avoid the registrations at the HLR by simply setting up a forwarding pointer from the previous VLR. Calls to a given user will first query the user's HLR to determine the first VLR which the user was registered at and then follow a chain of forwarding pointers to the user's current VLR. This observation results in a strategy, which will be useful for those users who receive calls infrequently relative to the rate at which they change RA's. This idea attempts to exploit patterns in the call reception and mobility of PCS users.

The pointer forwarding scheme was studied in [9] by comparing its performance with the IS-41 scheme. The reader is referred to [9] for detailed operation-cost analysis and message-traffic analysis. It is interesting to note that when

Manuscript received November 20, 1995; revised August 27, 1996. This work was supported in part by the National Science Council under Contract NSC 86-2213-E-009-074.

The authors are with the Department and Institute of Computer Science and Information Engineering, National Chiao-Tung University, Taiwan, R.O.C. (e-mail: liny@csie.nctu.edu.tw).

Publisher Item Identifier S 0018-9545(98)00677-X.

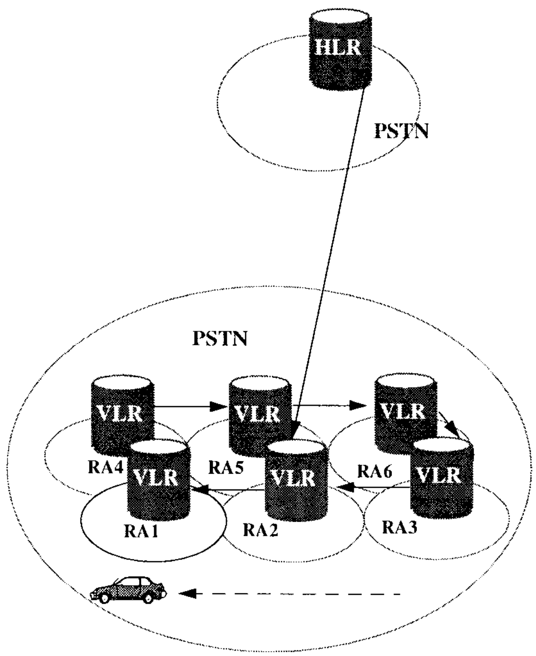

Fig. 1. Pointer forwarding with single HLR (the **registration** operation).

the call frequency is much lower than the portable move frequency, pointer forwarding may significantly outperform IS-41. When the call frequency is much higher than the portable move frequency, pointer forwarding behaves identically to IS-41.

This paper proposes a distributed HLR approach for pointer forwarding and compares its performance with the original (the single HLR) pointer forwarding scheme.

II. Pointer Forwarding with Single HLR

Pointer forwarding is a well-known technique used in the Andrew distributed file system [14]. This technique was proposed to reduce the location update cost in a PCS network [6], [7]. In this algorithm, no message is sent to update HLR when a portable moves to a new RA. Instead, a message is sent from the new RA directly to the old RA. On receipt of the message, the obsolete registration entry in the old RA is deleted, and a forwarding pointer to the new RA is created.

When an incoming call arrives, the forwarding pointers are traced to find the actual location of the portable.

There are several ways to manage the forwarding pointers. A simple algorithm is described as follows. When the portable moves to new RA's, the forwarding pointers are created as shown in Fig. 1. When the portable moves from RA2 to RA1, RA1 sends a message to RA2. Upon the receipt of the message, RA2 performs two tasks.

1) A forwarding pointer is created to point to RA1.
2) The portable is deregistered.

When a phone call arrives, the forwarding pointers are traced. (see Fig. 2). Also note that a global title translation is required at the network (e.g., in California) of the calling party, and the query message is then forwarded to the HLR (e.g., in New York).

III. Pointer Forwarding with Distributed HLR's

Because of the heavy signaling traffic generated by PCS location tracking, the HLR may become a bottleneck. To reduce the traffic to an HLR, one natural solution is to distribute the HLR function in several locations.

However, it is difficult to implement distributed HLR's in IS-41. For portable registration, it may be required to update some or all HLR's. Thus, extra traffic is generated for multiple HLR updates.

On the other hand, distributed HLR's can be efficiently implemented with pointer forwarding. Since the **registration** operation in pointer forwarding is done by sending a message from the new VLR to the old VLR, multiple HLR updates are eliminated. Thus the advantage of using distributed HLR for pointer forwarding is obvious from the aspect of database access delay.

The pointer forwarding with the distributed HLR's (PFDHLR) scheme is described as follows. The HLR's are distributed in remote PSTN's. A natural location for a distributed HLR is near by the STP that performs GTT in PFSHLR (see Fig. 4). The HLR's may point to different VLR's that the portable previously visited. Like forwarding pointer with single HLR (PFSHLR), the registration process only involves the old and new VLR's, and the HLR's are not updated.

When a call arrives from a particular PSTN, the HLR of the PSTN is queried as shown in Fig. 5. After the **find** operation is completed, the TLDN is returned from the current RA (e.g., RA1 in our example) to the HLR, and the HLR updates its pointer to the current RA of the portable (c.f., Fig. 3). Note that GTT is not required since we assume that the distributed HLR is near by or is collocated with the GTT STP. After the **find** operation, the VLR (RA1) returns the routable address back to the HLR, and the pointer of the distributed HLR is updated (i.e., it points to RA1) as shown in Fig. 6. After the **find** operation, the obsolete forwarding pointers (e.g., the pointers between RA6, RA5, RA4, RA3, RA2, and RA1) are *not* deleted. We assume that the VLR's have enough space to accommodate these pointers. Note that the current technologies are able to provide enough space to store these. Note that the GTT and an extra remote visit from the GTT STP to the HLR is avoided in PFDHLR (see Figs. 2 and 5).

IV. Performance Study

This section compares PFDHLR with PFSHLR from the network signaling aspect. Since the registration behaviors are the same for both schemes, it suffices to compare the **find** cost. Our performance study consists of two parts. The first part is to estimate the cost to query an HLR and the cost to traverse a pointer between two VLR's. The second part is to derive the numbers of pointers traced in both schemes.

To simplify our study, we make a simple cost estimate as follows.

1) The cost of the request from the calling party to the GTT STP in PFSHLR (see step 1 in Fig. 2) or to the distributed HLR in PFDHLR (see step 1 in Fig. 5) is normalized to one.

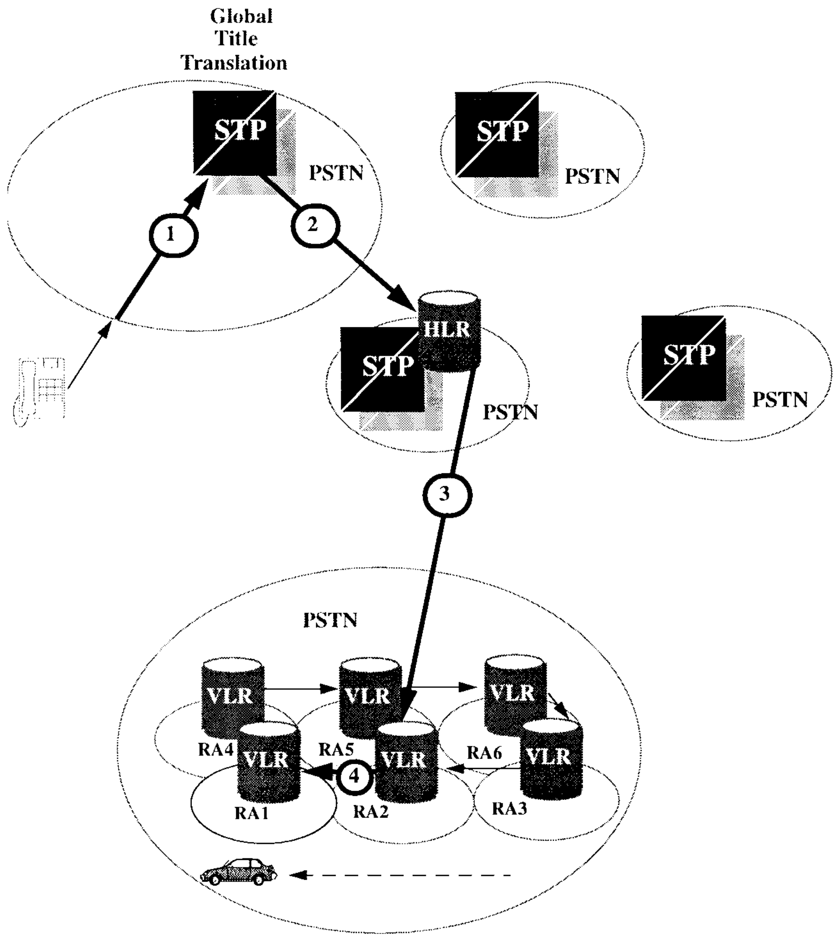

Fig. 2. Pointer forwarding with single HLR (the **find** operation).

2) The cost of sending a message from the GTT STP to the HLR in PFSHLR is one (see step 2 in Fig. 2).
3) The cost of querying the first VLR from the HLR in both PFSHLR (see step 3 Fig. 2) and PFDHLR (see step 2 in Fig. 5) is one.
4) The cost of traversing a pointer from a VLR to another is δ (see step 4 in Fig. 2 and steps 3 and 4 in Fig. 5). Since the old and the new VLR's are likely to be next to each other, $\delta \ll 1$ can be expected (see the cost analysis in [9]). In our study, $\delta = 0.5$ and 1 are considered.

The expected numbers $E[k]$ of pointers traced in both PFSHLR and PFDHLR are derived as follows. Suppose that between two consecutive incoming calls the portable moves to new RA's K times, and the number of forwarding pointers traced to find the actual location is k. The probability $\alpha(K)$ of K moves between two incoming calls was derived in [15]

$$\alpha(K) = \begin{cases} 1 - \frac{1 - f_m^*(\lambda_c)}{\theta}, & K = 0 \\ \frac{1}{\theta}[1 - f_m^*(\lambda_c)]^2 [f_m^*(\lambda_c)]^{K-1}, & K > 0 \end{cases} \quad (1)$$

where $\theta = \frac{\lambda_c}{\lambda_m}$ is the call-to-mobility ratio [9], where λ_c is the call-arrival rate, and λ_m is the RA moving rate. In the equation, f_m^* is the Laplace–Stieltjes transform of the portable residence time in an RA. For our purpose, we assume that the portable residence times have a Gamma distribution. The Gamma distribution is selected for two reasons. First, the Gamma distribution does not have a specific distribution shape, and it has the desirable property to fit an arbitrary distribution by setting appropriate parameters [16]. Second, the Gamma distribution has a simple Laplace–Stieltjes transform format, which simplifies the calculation of (1). With mean $1/\lambda_m$ and variance V, the Gamma Laplace–Stieltjes Transform is expressed as

$$f_m^*(s) = \left(\frac{\lambda_m \gamma}{s + \lambda_m \gamma}\right)^\gamma \quad \text{where} \quad \gamma = \frac{1}{V\lambda_m^2}.$$

We use $\alpha(K)$ and a two-dimensional (2-D) random walk model to determine the expected number $E[k]$ of the VLR pointers traversed in the **find** operation. In a 2-D random walk, the portable p may move to one of the four neighbor RA's [c.f., Fig. 7(a)]. For simplicity, we assume that p moves to one of the neighbor RA's with probability 0.25. During the K movements between two call arrivals, the portable may revisit an RA several times, and $k \leq K$. For example, $K = 32$ and $k = 4$ in Fig. 7(b). Clearly, in the PFSHLR scheme, what we have here is not a chain of the pointers, but a tree of pointers.

In a PCS network, we expect that the mobility of a portable exhibits *spatial locality* (i.e., a portable tends to revisit RA's). To capture this phenomenon, the movement of a portable is modeled by a *2-D random walk with reflecting barriers*. We assume that between two call arrivals, the mobility of a portable is restricted in a square region with size S. Fig. 8

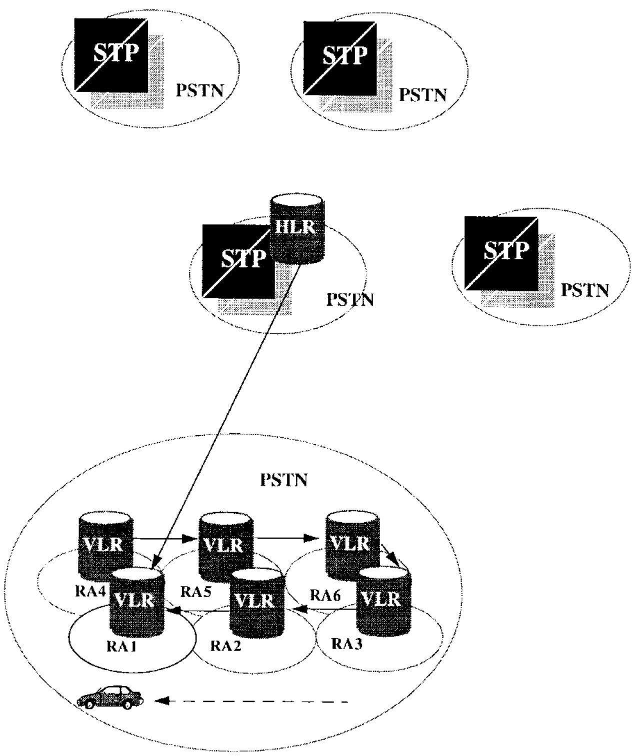

Fig. 3. Pointer forwarding with single HLR (after the **find** operation).

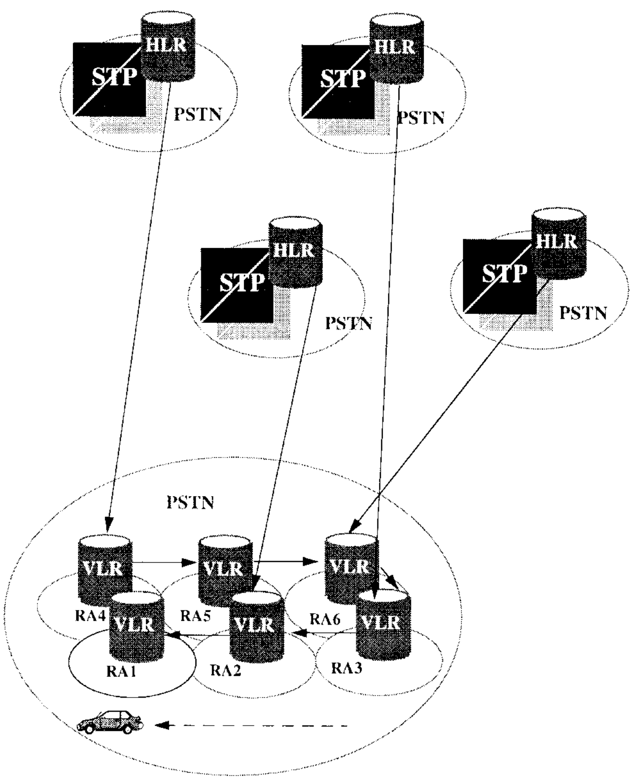

Fig. 4. Pointer forwarding with distributed HLR's (the **registration** operation).

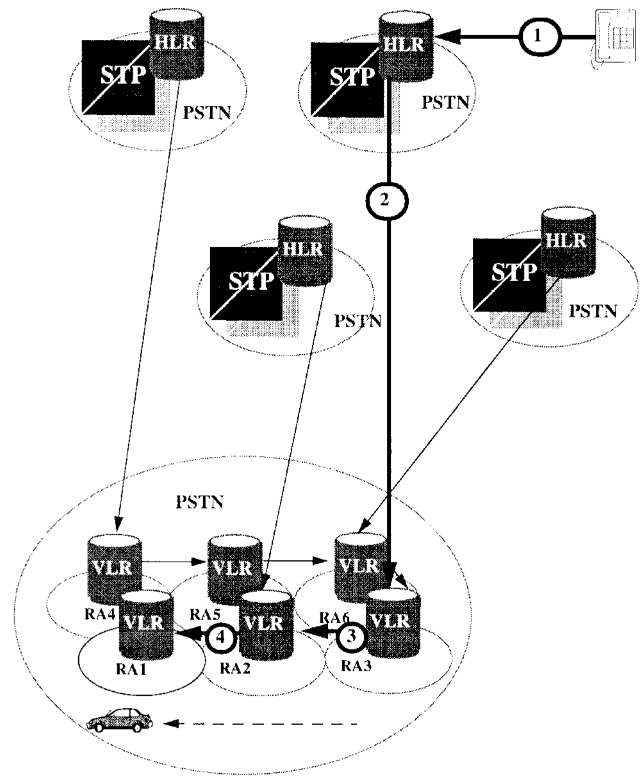

Fig. 5. Pointer forwarding with distributed HLR's (the **find** operation).

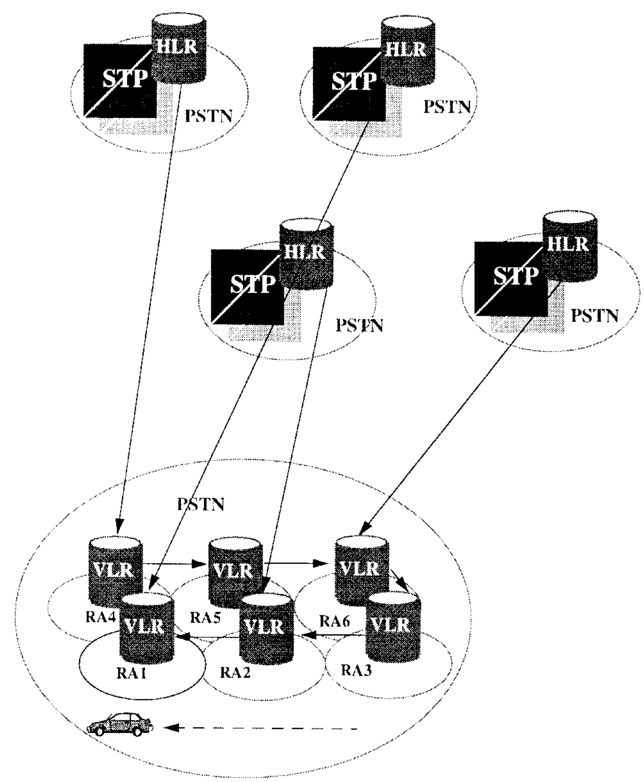

Fig. 6. Pointer forwarding with distributed HLR's (after the **find** operation).

shows a region of size $S = 25$. If the portable is in a interior RA of the region (e.g., R_2), then it may move to one of the four directions with the equal probability 0.25. For a boundary RA such as R_1, if the portable moves horizontally, then it must move to the right. In other words, a portable can only move to one of the three neighbors with the routing probabilities 0.25 (up), 0.5 (right), and 0.25 (down). Similarly, for a portable at R_3, the routing probabilities are 0.5 (left) and 0.5 (down). It

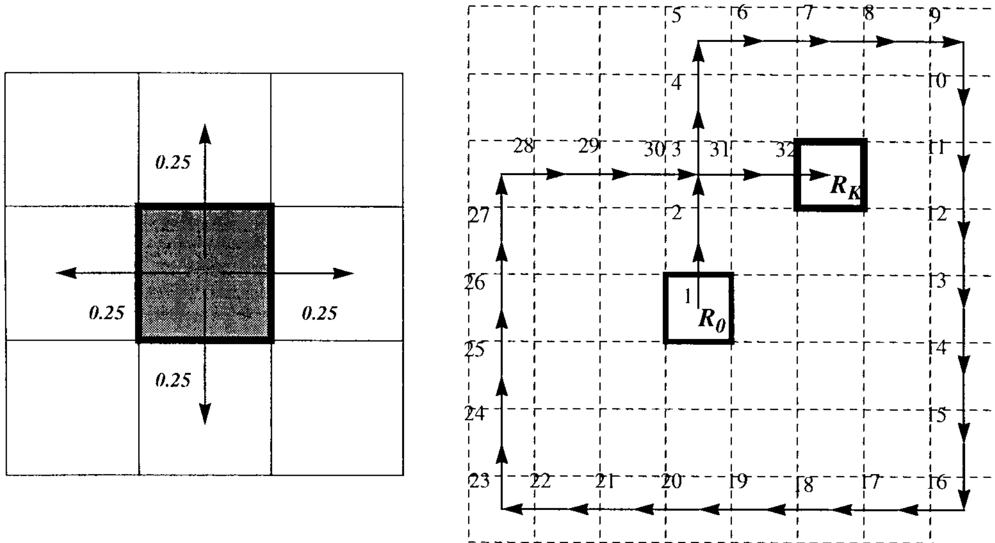

Fig. 7. The 2-D random walk model.

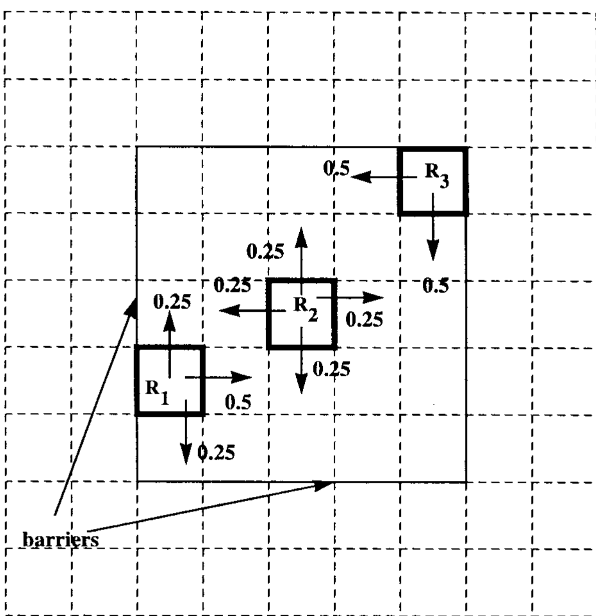

Fig. 8. The random walk model with reflecting barriers.

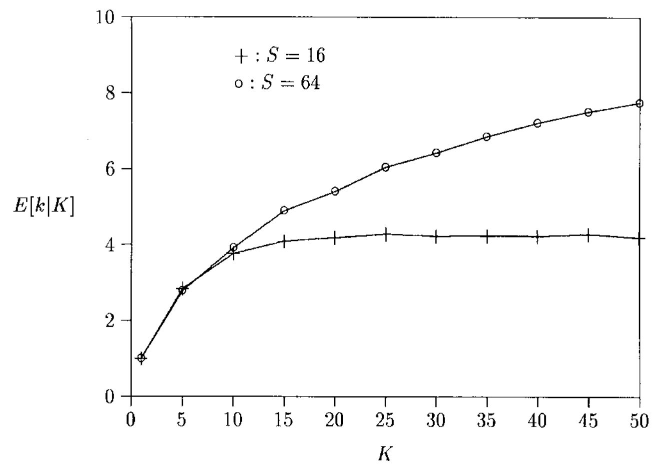

Fig. 9. The expected number $E[k \mid K]$.

is difficult to derive the number k analytically. Instead, for a given K, we computed the expected number $E[k \mid K]$ by using the simulation technique. Fig. 9 plots $E[k \mid K]$ for different K values. The figure indicates that $E[k \mid K]$ does not increase as fast as K. When the portable exhibits high locality (e.g., $S = 16$ or the portable moves among the 16 RA's between two phone calls), $E[k \mid K]$ does not increase for $K > 15$. The expected values $E[k]$ are expressed as

$$E[k] = \sum_{i=0}^{\infty} E[k \mid K]\alpha(K).$$

Fig. 10 plots $E[k]$ for different θ values. The figure indicates the following.

1) The variance V of the user residence time distribution has an effect on $E[k]$. The larger the variance V, the smaller the number $E[k]$. Consider the same mean residence time $\frac{1}{\lambda_m}$. With large V, a large number of small K values (particularly, $K = 0$) and a large number of large K values are observed. A very small K value results in a small k value. A very large K value does not necessarily result in a large k value (see Fig. 9—for $S = 16$ and $K > 15$, k does not increase as k increases). Thus, a smaller $E[k]$ is expected for a larger V. This phenomenon also was observed in [9].
2) User locality has a significant effect on $E[k]$ when θ is small (i.e., when the portable moves more frequently than the call arrivals). This is consistent with our intuition. When $\theta \to \infty$, the impact of user locality can be ignored. In this case, the HLR always points to the current VLR, and $E[k] = 0$ for all S and V values.

We assume that there are N remote PSTN's in our study. To compare PFSHLR and PFDHLR, the expected number of k is

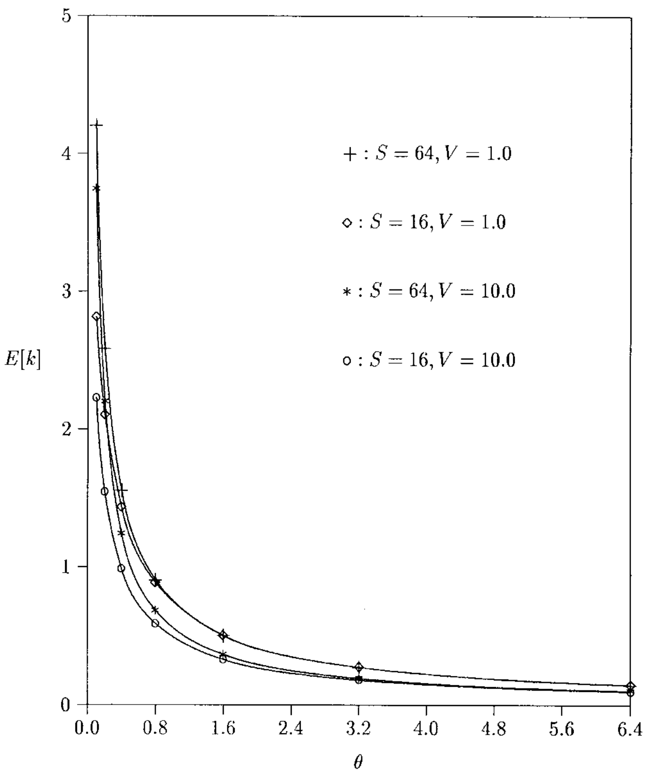

Fig. 10. The expected number $E[k]$.

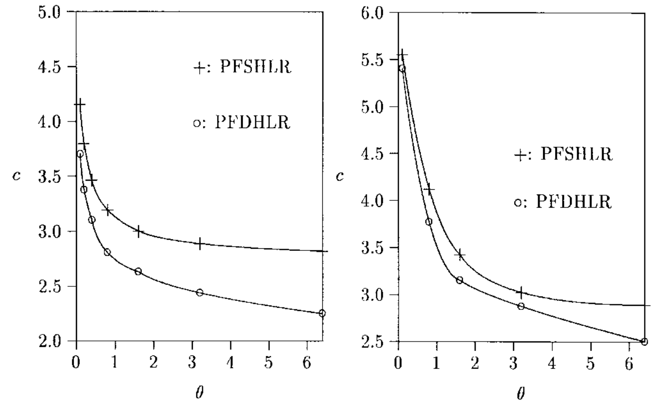

Fig. 11. The comparison of the **find** cost of PFSHLR and PFDHLR. ($S = 36$ and $N = 4$).

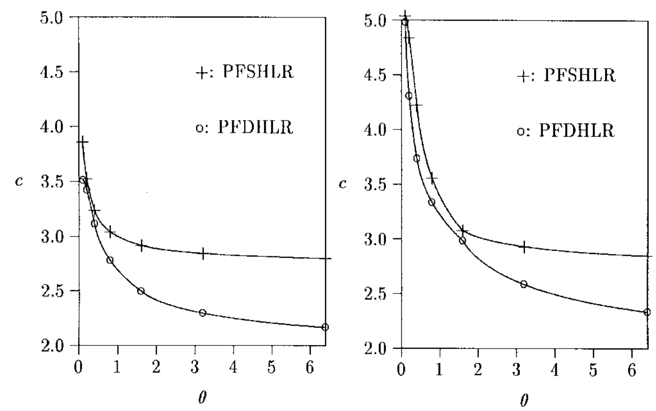

Fig. 12. The comparison of the **find** cost of PFSHLR and PFDHLR. $S = 36$ and $N = 4$.

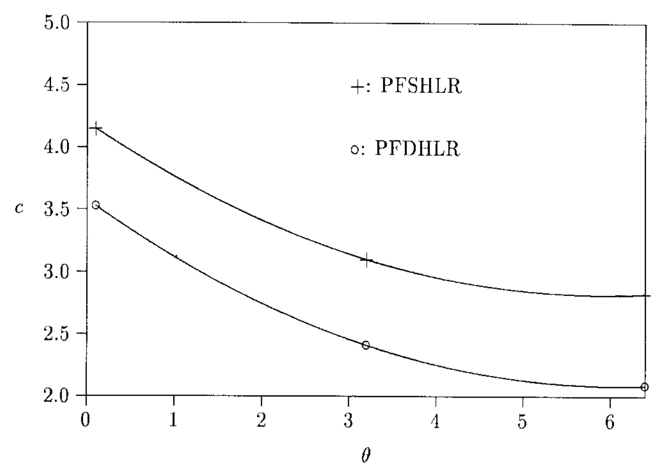

Fig. 13. The comparison of the **find** cost of PFSHLR and PFDHLR. $S = 16$, $V = 1.0$, $\delta = 0.5$, and $N = 4$.

reexpressed as $E[k \mid \theta]$, a function of θ (because α is affected by θ). In PFSHLR, the **find** cost c_S (where "S" stands for "single HLR" approach) for a call delivery is

$$c_S = 3\left(\frac{N-1}{N}\right) + 2\left(\frac{1}{N}\right) + E[k \mid \theta]\delta \qquad (2)$$

$$= 3 - \frac{1}{N} + E[k \mid \theta]\delta \qquad (3)$$

where δ is the cost of traversing a pointer from a VLR to another. We assume that the calling parties are uniformly distributed in the N areas. In (2), $(N-1)/N$ calls are from remote PSTN's, and step 1 in Fig. 2 is required. In this case, the cost to access the first VLR is three (the cost for steps 1, 2, and 3 in Fig. 2). On the other hand, $1/N$ calls are from the PSTN of the HLR, and step 1 in Fig. 2 is not required. In this case, the cost to access the first VLR is two (the cost for steps 2 and 3 in Fig. 2). The VLR pointer traversal cost is $E[k \mid \theta]\delta$. For PFDHLR, the **find** cost c_D (where "D" stands for "distributed HLR" approach) for a call delivery is

$$c_D = 2 + E\left[k \,\middle|\, \frac{\theta}{N}\right]\delta. \qquad (4)$$

Note that the call to mobility ratio for PFSHLR is $\lambda/\lambda_m = \theta$. Since the call-arrival rate for PFDHLR is λ/N, its call to mobility ratio is $\lambda/(N\lambda_m) = \theta/N$. Thus, the expected number of pointer traversals in PFSHLR is $E[k \mid \frac{\theta}{N}]$, and the pointer traversal cost is $E[k \mid \frac{\theta}{N}]\delta$ in (4). The PFSHLR cost to access the first VLR is two in (4) as indicated in Steps 1 and 2 of Fig. 6. Figs. 11–13 plot the cost for PFSHLR and PFDHLR with different parameters.

In our experiments, $0.1 \leq \theta \leq 6.4$ are considered. Fig. 11 considers the residence time with low variance (i.e., $V = 1$), where $S = 36$ and $N = 4$. The figure indicates that for $\delta = 0.5$, up to 20% improvement can be expected from PFDHLR over PFSHLR. For $\delta = 1.0$, up to 10% improvement can be expected.

Fig. 12 considers the residence time with high variance (i.e., $V = 10$), where $S = 36$ and $N = 4$. The figure indicates that for $\delta = 0.5$, up to 22% improvement can be expected from PFDHLR over PFSHLR. For $\delta = 1.0$, up to 18% improvement can be expected. That is, with larger V, the advantage of PFDHLR becomes more significant.

Fig. 13 considers the residence time with high locality (i.e., $S = 16$, that is, the portable only moves across 16 RA's between two call arrivals). The figure indicates that for $\delta = 0.5$, up to 28% improvement can be expected from PFDHLR over PFSHLR.

V. Conclusion

Due to the heavy signaling traffic generated by PCS location tracking, the HLR may become a bottleneck. To reduce the traffic to an HLR, one natural solution is to distribute the HLR function in several locations. However, it is difficult to implement distributed HLR's in protocols such as IS-41 and GSM. For portable registration, it may be required to update all HLR's. Thus, extra traffic is generated for multiple HLR updates.

On the other hand, distributed HLR's can be efficiently implemented with pointer forwarding. Since the **registration** operation in pointer forwarding is done by sending a message from the new VLR to the old VLR, multiple HLR updates are eliminated. Thus the advantage of using distributed HLR for pointer forwarding is obvious from the aspect of database access delay.

One potential problem of pointer forwarding with distributed HLR (PFDHLR) is that long pointer chain may be traversed to locate a portable. Our study indicated that in the network traffic aspect, PFDHLR outperforms its single HLR counterpart with up to 28% improvement for the range of θ under this study.

Acknowledgment

The authors would like to thank J. R. Cruz and the anonymous reviewers for their valuable comments.

References

[1] EIA/TIA, "Cellular radio-telecommunications intersystem operations," EIA/TIA Tech. Rep. IS-41 Revision B, 1991.

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

[3] Y.-B. Lin and S. K. DeVries, "PCS network signaling using SS7," *IEEE Personal Commun. Mag.*, pp. 44–55, June 1995.

[4] C. N. Lo *et al.*, "An estimate of network database transaction volume to support universal personal communication services," in *8th ITC Specialist Seminar on Universal Personal Telecommunications*, 1992.

[5] K. S. Meier-Hellstem and E. Alonso, "The use of SS7 and GSM to support high density personal communications," in *IEEE ICC*, 1992.

[6] T. Imielinski and B. R. Badrinath, "Querying in highly mobile distributed environment," in *Proc. 18th VLDB Conf.*, 1992.

[7] R. Jain, Y.-B. Lin, C. N. Lo, and S. Mohan, "A caching strategy to reduce network impacts of PCS," *IEEE J. Select. Areas Commun.*, vol. 12, no. 8, pp. 1434–1445, 1994.

[8] Y.-B. Lin, "Determining the user locations for personal communications networks," *IEEE Trans. Veh. Technol.*, vol. 43, no. 3, pp. 466–473, 1994.

[9] R. Jain and Y.-B. Lin, "Performance modeling of an auxiliary user location strategy in a PCS network," *ACM-Baltzer Wireless Networks*, vol. 1, no. 2, pp. 197–210, 1995.

[10] S. Tabbane, "Comparision between the alternate location strategy and the classic location strategy," WINLAB Tech. Rep. WINLAB-TR-37, Rutgers Univ., New Brunswick, NJ, 1992.

[11] G. P. Pollini and S. Tabbane, "The intelligent network signaling and switching costs of an alternate location strategy using memory," in *IEEE VTC*, 1993.

[12] H. Xie, S. Tabbane, and J. D. Goodman, "Dynamic location area management and performance analysis," in *IEEE VTC*, 1993.

[13] S. Rao, B. Gopinath, and D. Kurshan, "Optimizing call management of mobile unit," in *IEEE PIMRC*, 1992.

[14] M. Satyanarayanan, J. H. Howard, D. N. Nichols, R. N. Sidebotham, A. Z. Spector, and M. J. West, "The ITC distributed file system: Principles and design," in *Proc. 10th ACM Symp. Operating Systems Principles*, Dec. 1985, pp. 35–50.

[15] Y.-B. Lin, "Reducing location update cost in a PCS network," *IEEE/ACM Trans. Network.*, vol. 5, no. 1, pp. 25-33, 1997.

[16] N. L. Johnson, *Continuous Univariate Distributions—1*. New York: Wiley, 1970.

Yi-Bing Lin (S'80–M'96–SM'96), for a photograph and biography, see this issue, p. 57.

Wen-Nung Tsai received the B.S. degree in 1977 and the M.S. degree in computer science in 1979, both from National Chiao Tung University, Taiwan, R.O.C. He is currently working toward the Ph.D. degree at Northwestern University.

He is a Lecturer in the Department of Computer Science and Information Engineering, National Chiao Tung University. His current research interests include design and analysis of personal communications services network, mobile computing, distributed computing, parallel processing, and network security.