Reducing Location Update Cost in a PCS Network

Yi-Bing Lin, *Senior Member, IEEE*

Abstract—**Location tracking operations in a personal communications services (PCS's) network are expensive. Several strategies have been proposed to reduce the location update cost. This paper studies a special case of a location tracking algorithm called the alternative location algorithm. This special case is referred to as the two location algorithm (TLA). An analytical model is proposed to compare the performance of TLA and the IS-41 protocol. Our study indicates that the performance of TLA is significantly affected by the user moving patterns and the call traffic. If the user mobility is higher than the call frequency or the user tends to move back to the previously visited registration areas, then TLA may significantly outperform IS-41. We also observe that the variance of the portable residence times in registration areas has impact on the performance of TLA (i.e., better performance is expected for larger variance).**

Index Terms—**Home location register, mobility management, personal communications services, roaming management, visitor location register.**

I. Introduction

IN a personal communications services (PCS's) system, the location of a called portable must be determined before the connection can be established. Due to the mobility of portables, a database called home location register (HLR) is required to store the location of a portable. The location record is modified when the portable moves to another location. A procedure to locate a portable is specified in IS-41 [1]. This procedure is referred to as the basic *find* operation in this paper. In IS-41, call termination consist of two parts. In the first part, the basic *find* operation is performed. In the second part, the trunk is set up for the conversation. The basic *find* operation is described below. We assume that the calling party is a wireline user.

- Step 1) The incoming call to a portable p is initiated at a switch called the *originating switch*.
- Step 2) The originating switch queries the HLR for the portable's current location.
- Step 3) The HLR queries the *visitor location register* (VLR) where p was last registered.
- Step 4) The VLR queries the *mobile switching center* (MSC) where p is located to determine whether p is capable of receiving the call. If so, the MSC returns a routable address *Temporary Local Directory Number* (TLDN) to the VLR. The VLR sends the TLDN back to the HLR.
- Step 5) The HLR forwards the TLDN to the originating switch.

After the originating switch has received the TLDN, it routes the call to the MSC where p is located.

We assume that every VLR serves a *registration area* (RA) which consists of one or more radio port coverage areas. When a portable moves from one RA to another, it registers at the VLR of the new RA, and its new location is reported to the HLR. This action is referred to as the *registration* operation. The *registration* operation may be followed by a deregistration (or cancellation) operation to remove the obsolete record in the old VLR. (This is referred to as *explicit deregistration* [2].) In IS-41, HLR sends a deregistration message to the old VLR. In [3], the deregistration message is sent directly from the new VLR to the old VLR. The deregistration operation may not be performed immediately after a *registration* operation. In *timeout deregistration*, the obsolete entries are cancelled periodically [2]. In *implicit deregistration*, no deregistration operation is performed. An old VLR record is replaced to accommodate a newly arrived portable when the resource such as TLDN is used up [2], [4]. The advantage of implicit deregistration is that no deregistration message is required to cancel an obsolete record. To simplify our analysis, implicit deregistration is assumed in this study.

Studies [5]–[7] indicated that the message traffic due to the *find* and *registration* operations is significant. When the frequency of the incoming calls is high with respect to the portable mobility (i.e., the rate that a portable moves to new RA's), the *location cache scheme* [8], [9] was proposed to reduce the traffic (i.e., to reduce the number of the *find* operations). When the call frequency is low with respect to the mobility, the number of the *registration* operations increases which significantly degrades the performance of the location cache scheme. Several algorithms such as the *pointer forwarding* algorithm [3], [7] and the *alternative location* algorithm (ALA) [10]–[12] have been proposed to reduce the registration cost.

This paper studies a special case of ALA called the *two location algorithm* (TLA). In TLA, we consider two locations to track the location of a portable. On the other hand, ALA allows more than two user locations in the profile. The performance of ALA has been studied in [10]–[12]. These studies focused on the cost analysis for the paging operations in call termination, and provided the mean analysis of the user mobility.

Manuscript received July 2, 1993; revised August 20, 1996; approved by IEEE/ACM Transactions on Networking Editor C.-L. I. This work was supported by the National Science Council under Grant NSC 86-2213-E-009-074.

The author is with the Department of Computer Science and Information Engineering, National Chiao Tung University, Hsinchu, Taiwan, R.O.C. (email: liny@csie.nctu.edu.tw).

Publisher Item Identifier S 1063-6692(97)01704-4.

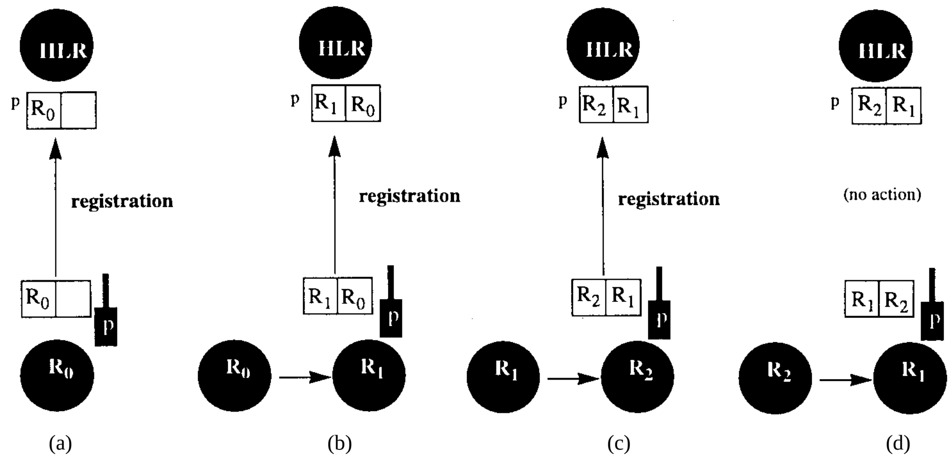


Fig. 1. The two location algorithm.

This paper analyzes the SS7 traffic cost of registration and call delivery for a basic scheme based on IS-41 and the TLA. We also consider the impact of the portable movement distributions, which provides important insight to complement the performance studies in [10]–[12].

TLA and ALA can be integrated with the cache scheme to further reduce the location update/tracking costs for all incoming call frequencies and mobility patterns.

II. The Two Location Algorithm

In the TLA, a portable p has a small built-in memory to store the addresses for the two most recently visited RA's. The record of p in the HLR also has an extra field to store the corresponding two locations. The first memory location (see Fig. 1) stores the most recently visited RA. The algorithm guarantees that the portable is in one of the two locations. Fig. 1 illustrates a sequence of p's movements $R_0 \rightarrow R_1 \rightarrow R_2 \rightarrow R_1$. When p joins the network, the location is stored in its memory, and a *registration* operation is required to modify the HLR record as shown in Fig. 1(a). When p moves to a new location, it checks if the new location is in the memory. If the new location is not found, the address for the RA p just left is kept, and the other address is replaced by the address for the new RA [see Fig. 1(b) and 1(c)]. A *registration* operation is required to make the same modification in the HLR record. If the address for the new location is already in the memory, no action is taken in the HLR record. In other words, no *registration* operation is performed [see Fig. 1(d); p moves back to the previously visited registration area R_1]. Note that in TLA, the portable always has the correct view of the "latest visited RA." On the other hand, the HLR may have an "incorrect view." In Fig. 1(a)–(c), both the portable and the HLR have the consistent view of the latest visited RA [which is R_0 in Fig. 1(a), R_1 in Fig. 1(b), and R_2 in Fig. 1(c)]. At Fig. 1(d), the HLR considers R_2 as the latest visited RA, which is incorrect.

When a phone call arrives, the two addresses are used to find the actual location of the portable. The order of the addresses selected to locate the portable affects the performance of the algorithm. If the portable is located in the first try (this is referred to as *location hit*), then the *find* cost for TLA is the same as the IS-41 algorithm. Otherwise, extra penalty incurs in TLA to locate the portable for the second try (this is referred to as *location miss*; note that the second try is always successful). After the second try, the HLR identifies the RA where the portable is visiting.

There are several alternatives to select the address for the first try. An obvious heuristic is to select the location randomly. Another heuristic is the *latest-RA-first* strategy where the latest visited RA address (in the HLR's view) is selected. We have observed that the latest-RA-first strategy is better than the random selection algorithm in all cases studied in this paper. We only present the results for the latest-RA-first strategy.

III. Movements Between Two Phone Calls

This section studies the patterns of the incoming calls and the portable movement. We derive the probability $\alpha(K)$ that a portable moves across K RA's between two phone calls. Assume that the incoming calls to a portable are a Poisson process, and the time the portable resides in an RA has a general distribution. We show that for a portable, different moving patterns result in different $\alpha(K)$ distributions.

Consider the timing diagram in Fig. 2. Let t_c be the interval between two consecutive phone calls to a portable p. Suppose that the portable resides in an RA R_0 when the previous phone call arrived. After the phone call, p visits another K RA's, and p resides in the jth RA for a period t_{M_j} $(0 \leq j \leq K)$. Let t_m be the interval between the arrival of the previous phone call and the time when p moves out of R_0. Let $t_{c,i}$ be the interval between when the portable enters R_i and when the next phone call arrives. Let t_{M_i} be independent identically distributed random variables with a general distribution $F_m(t_{M_i})$, the density function $f_m(t_{M_i})$ and the Laplace–Stieltjes Transform

$$f_m^*(s) = \int_{t=0}^{\infty} e^{-st} f_m(t)\, dt. \tag{1}$$

Note that it is possible that the next phone call arrives while the previous call is still in progress. The phenomenon is called *busy line* effect. In this case, the portable cannot initiate/accept the next call and the assumption of Poisson process for calls

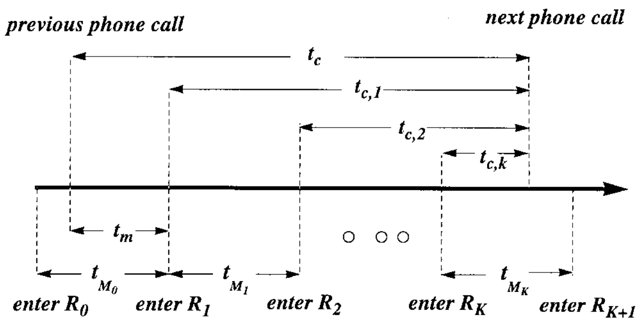


Fig. 2. The timing diagram.

is no longer valid. The *busy line* effect was studied in [13]. To simplify and strengthen our results, we assume that the busy line effect is insignificant. This assumption is justified when the inter-call arrival times are more than 10 times longer than the call holding times.

Let $f_c(t)$ and $r_m(t)$ be the density function of t_c and t_m, respectively. Let $E[t_c] = 1/\lambda_c$ and $E[t_{M_i}] = 1/\lambda_m$. Since the incoming phone calls are a Poisson process, we have

$$f_c(t) = \lambda_c e^{-\lambda_c t}.$$

From the memoryless property of the exponential distribution, $t_{c,\,i}$ have the same exponential distributions as t_c for all i. Also from the random observer property [14], [15]

$$\begin{aligned} r_m(t) &= \lambda_m \int_{\tau=t}^{\infty} f_m(\tau)\, d\tau \\ &= \lambda_m [1 - F_m(t)] \end{aligned} \tag{2}$$

(the detailed derivation can be found in [15, pp. 67–69], and is referred to as the *excess life theorem*). The Laplace–Stieltjes Transform for the t_m distribution is

$$\begin{aligned} r_m^*(s) &= \int_{t=0}^{\infty} e^{-st} r_m(t)\, dt \\ &= \int_{t=0}^{\infty} e^{-st} \lambda_m [1 - F_m(t)]\, dt \\ &= \frac{\lambda_m}{s} - \int_{t=0}^{\infty} e^{-st} \lambda_m F_m(t)\, dt \\ &= \frac{\lambda_m}{s} \\ &\quad - \left[\frac{\lambda_m}{s} e^{-st} F_m(t) \Big|_{t=0}^{\infty} - \frac{\lambda_m}{s} \int_{t=0}^{\infty} e^{-st} f_m(t)\, dt \right] \\ &= \frac{\lambda_m}{s} \left[1 - \int_{t=0}^{\infty} e^{-st} f_m(t)\, dt \right] \\ &= \frac{\lambda_m}{s} [1 - f_m^*(s)]. \end{aligned} \tag{3}$$

The probability $\alpha(K)$ that p moves across K RA's between two phone calls is derived in two cases. For $K = 0$

$$\begin{aligned} \alpha(0) &= \Pr[t_c \le t_m] \\ &= \int_{t_m=0}^{\infty} \int_{t_c=0}^{t_m} \lambda_c e^{-\lambda_c t_c} r_m(t_m)\, dt_c\, dt_m \\ &= r_m^*(s)|_{s=\lambda_c} \\ &= 1 - \frac{1 - f_m^*(\lambda_c)}{\rho} \end{aligned} \tag{4}$$

where $\rho = \lambda_c/\lambda_m$ is the *call-to-mobility ratio* defined in [8]. For $K \ge 1$

$$\begin{aligned} &\alpha(K) \\ &\quad = \Pr[t_m + t_{M_1} + \cdots + t_{M_{K-1}} < t_c \le t_m + \cdots + t_{M_K}] \\ &\quad = \Pr[t_c > t_m] \left(\prod_{i=1}^{K-1} \Pr[t_{c,\,i} > t_{M_i}] \right) \Pr[t_{c,\,K} \le t_{M_K}]. \end{aligned} \tag{5}$$

$\Pr[t_{c,\,i} > t_{M_i}]$, $\Pr[t_{c,\,K} < t_{M_K}]$, and $\Pr[t_c > t_m]$ are derived as follows:

$$\begin{aligned} &\Pr[t_{c,\,i} > t_{M_i}] \\ &\quad = \int_{t_{M_i}=0}^{\infty} \int_{t_{c,\,i}=t_{M_i}}^{\infty} \lambda_c e^{-\lambda_c t_{c,\,i}} f_m(t_{M_i})\, dt_{M_i}\, dt_{c,\,i} \\ &\quad = f_m^*(\lambda_c). \end{aligned} \tag{6}$$

From (6),

$$\begin{aligned} \Pr[t_{c,\,K} \le t_{M_K}] &= 1 - \Pr[t_{c,\,K} > t_{M_K}] \\ &= 1 - f_m^*(\lambda_c). \end{aligned} \tag{7}$$

From (4),

$$\begin{aligned} \Pr[t_c > t_m] &= 1 - \Pr[t_c \le t_m] \\ &= \frac{1 - f_m^*(\lambda_c)}{\rho}. \end{aligned} \tag{8}$$

From (4)–(8), we have

$$\alpha(K) = \begin{cases} 1 - \dfrac{1 - f_m^*(\lambda_c)}{\rho}, & K = 0 \\ \dfrac{1}{\rho}[1 - f_m^*(\lambda_c)]^2 [f_m^*(\lambda_c)]^{K-1}, & K > 0. \end{cases} \tag{9}$$

We assume that the portable residence times have a Gamma distribution. Most studies use the exponential distribution for its simplicity. Exponential is a special case of Gamma. In this paper, the Gamma distribution is selected for three reasons. Firstly, 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]. Secondly, the Gamma distribution has a simple Laplace–Stieltjes Transform format which simplifies the calculation of (9). Thirdly, with Gamma, the impact of variance of the residence times can be conveniently studied. With mean $1/\lambda_m$ and variance V, the Laplace–Stieltjes Transform of a Gamma random variable is expressed as

$$f_m^*(s) = \left(\frac{\lambda_m \gamma}{s + \lambda_m \gamma} \right)^{\gamma}$$

where

$$\gamma = \frac{1}{V \lambda_m^2}.$$

Fig. 3 plots $\alpha(K)$ for Gamma portable residence time distributions with different Variance values. The figure indicates that the α distribution is significantly affected by the variance of the residence times. The next section will show that the variance of the distribution may have a significant impact on the performance of TLA.

IV. Performance of the Location Update Strategies

This section studies the performance for TLA and IS-41. We first estimate the costs of the *find* and the *registration*

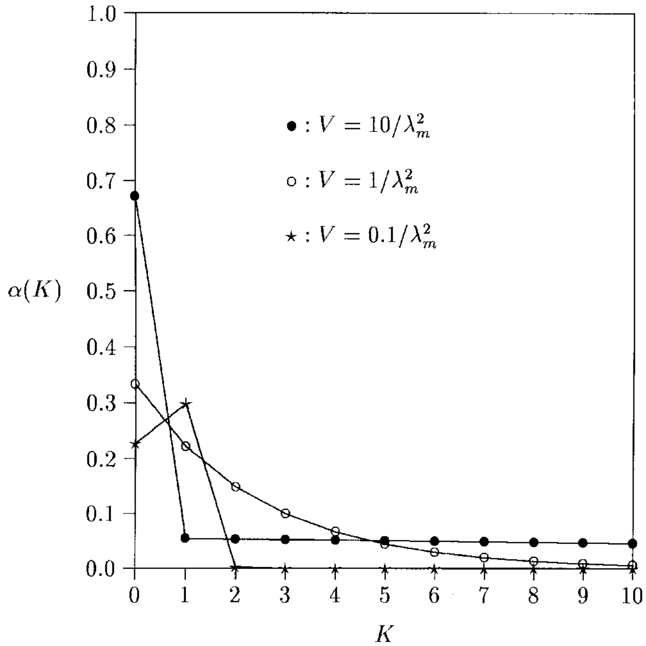


Fig. 3. $\alpha(K)$ for Gamma portable residence time distribution with different variance values ($\lambda_c = 0.5\,\lambda_m$.

operations, then derive the number of operations performed in the two algorithms.

For a particular PCS system, economic analysis is required to estimate the costs for the *find/registration* operations. For the demonstration purpose, we give a simple cost analysis as follows.

The basic *find* operation consists of two parts. The first part includes the interactions between the originating switch and the HLR [Steps 1), 2), and 5) in the *find* operation; see Fig. 4(a)]. We assume that the PCS messages are sent through the signaling system no. 7 (SS7) network in the public switched telephone network (PSTN). (A tutorial for PCS signaling using SS7 can be found in [17].) Upon the receipt of the dialed number (the mobile identification number (MIN) of the portable), the originating switch is not capable of associating this number with the physical address of the HLR. Instead, it forward the message to a signal transfer point (STP). 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. (GTT is needed because nongeographic numbering is assumed [1], [18].) Note that several SS7 network elements may be visited before the STP and the HLR are reached. The second part includes the interactions between the HLR, the VLR, the MSC, and the portable [Steps 3) and 4) in the *find* operation; see Fig. 4(a)]. For systems such as PACS [2], in the second part of the *find* operation the MSC pages the portable to make sure that the portable is in its area before the VLR sends the TLDN to the HLR. The design of the PACS alerting protocol allows a large paging capacity with low cost (the alerting capacity of PACS [19] is ten times larger than DECT [20] and 1000 times larger than CT2 [21], [22]). Note that for other systems such as AMPS [23], the paging cost is high and the portable is not paged during the VLR query. The portable is paged during trunk setup.

Note that the global title translation (and visits to several extra STP's) is performed in Part 1 but is not performed in Part 2, and the paging operation performed in Part 2 is not performed in Part 1. Without loss of generality, we normalize the cost of the basic *find* operation to one. We assume that the cost for Part 1 is δ, where $0.2 \le \delta \le 0.8$. For systems with low paging cost such as PACS, a large δ (e.g., $0.6 \le \delta \le 0.8$) is appropriate. For systems with high paging cost such as AMPS, a low δ (e.g., $0.2 \le \delta \le 0.4$) is expected. (A detailed paging cost analysis can be found in [10]–[12].)

In TLA, if the HLR selects the right location at the first time (a location hit), then the TLA *find* operation is exactly the same as the basic *find* operation. If the first try fails (a location miss), then the HLR tries the second location [see Steps 3a, 3b, 3, and 4 in Fig. 4(b)]. The extra cost (for Steps 3a and 3b) is the same as the cost for Part 2 of the basic *find* operation.

Fig. 4(c) illustrates the *registration* operation. Similar to Part 1 of the basic *find* operation, a global title translation is required to access the HLR. Also, a deregistration operation is required to delete the obsolete record in the old VLR. To simplify our analysis, we assume that the cost for registration is the same as the cost for Part 1 of the basic *find* operation (without considering the authentication procedure [6]).

Based on our simple cost analysis we have the following estimation.

1) The cost for the basic *find* and the TLA *find* (location hit) operations are one.
2) The cost for the *registration* is δ where $0.2 \le \delta \le 0.8$.
3) The cost for the TLA *find* (location miss) operation is $1 + (1 - \delta)$.

The numbers of operations for TLA and IS-41 are derived as follows. Suppose that the portable moves across K RA's between two phone calls. Let n_1 be the number of **registration** operations performed among the K moves. From (9), the cost n_1 for the IS-41 algorithm is

$$\begin{aligned} n_{1,\,\text{IS-41}} &= \sum_{j=0}^{\infty} j\alpha(j) \\ &= \sum_{j=1}^{\infty} \left(\frac{j}{\rho}\right)[1 - f_m^*(\lambda_c)]^2[f_m^*(\lambda_c)]^{j-1} \\ &= \frac{1}{\rho} \end{aligned}$$

and the cost is

$$c_{1,\,\text{IS-41}} = \frac{\delta}{\rho}. \tag{10}$$

Consider TLA. Suppose that the portable moves across K RA's between two phone calls. The conditional probability $\Pr[I = i|K]$ that i *registration* operations are performed among the K moves has a Bernoulli distribution:

$$\Pr[I = i|K] = \binom{K}{i} \theta^{K-i}(1-\theta)^i$$

where θ is the probability that when a portable p moves, the new RA address is in the portable memory [i.e., the probability that Fig. 1(d) occurs]. Since there are only two addresses stored in a portable, θ is the probability that p moves back

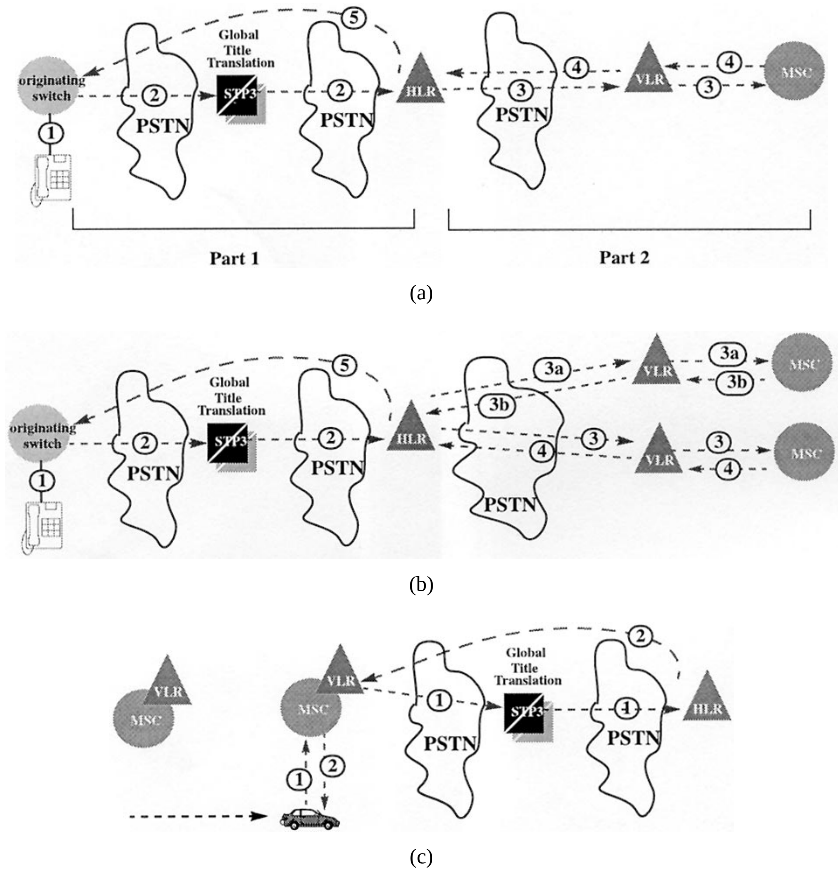


Fig. 4. Cost estimations for different operations. (a) The basic find operation or the MLA find operation (location bit) (b) The MLA find operation (location miss). (c) The registration operation.

to the RA where it came from.

The number n_1 for TLA is

$$n_{1,\,\mathrm{TLA}} = \sum_{K=0}^{\infty}\sum_{i=0}^{K} i\ \Pr[I=i|K]\alpha(K) = \frac{(1-\theta)}{\rho}$$

and the cost is

$$c_{1,\,\mathrm{TLA}} = \frac{(1-\theta)\delta}{\rho}. \tag{11}$$

Let n_2 be the number of *find* operations performed to locate a portable. The number n_2 and the cost c_2 for IS-41 are

$$n_{2,\,\text{IS-41}} = 1$$

and

$$c_{2,\,\text{IS-41}} = 1. \tag{12}$$

For TLA, the number $n_{2,\,\mathrm{TLA}}$ is derived in the Appendix. The probability ω that the HLR has a location miss in a call termination is given in (21)

$$\omega = \frac{1-f_m^*(\lambda_c)}{\rho}\left\{1 - \frac{1-\theta+\theta^2 f_m^*(\lambda_c)+[\theta f_m^*(\lambda_c)]^2}{1-[\theta f_m^*(\lambda_c)]^2}\right\}.$$

The number n_2 and the cost c_2 for TLA is

$$n_{2,\,\mathrm{TLA}} = 1+\omega$$

and

$$c_{2,\,\mathrm{TLA}} = 1+(1-\delta)\omega. \tag{13}$$

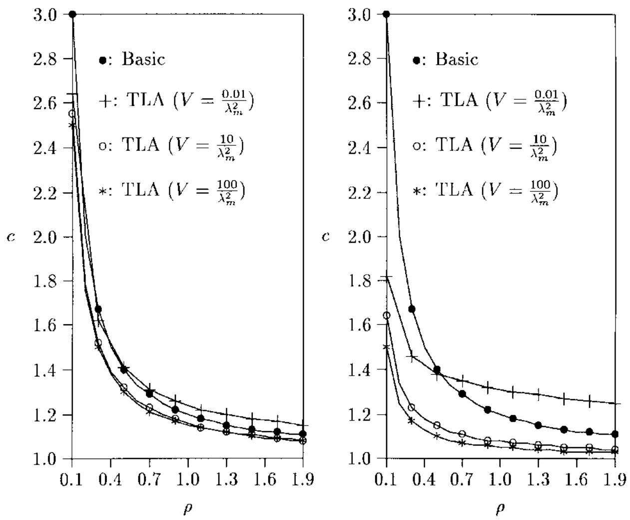
Define $c = c_1 + c_2$ as the net cost of *registration* and *find* operations performed between two phone calls, then from (10)–(13) we have

$$c_{\text{IS-41}} = \frac{\delta}{\rho}+1$$

$$c_{\mathrm{TLA}} = \frac{(1-\theta)\delta}{\rho}+1+(1-\delta)\omega. \tag{14}$$

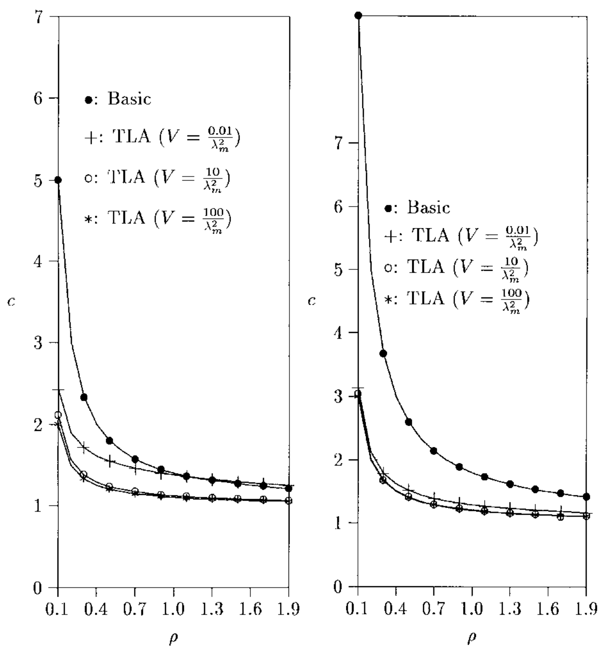
Figs. 5 and 6 plot $c_{\text{IS-41}}$ and c_{TLA}. These figures indicate that when the call-to-mobility ratio ρ is small, TLA outperforms IS-41 even if the user movement exhibits low locality (θ is small) and the *registration* cost is low. For example, when $\rho = 0.1$, $\theta = 0.25$, and $\delta = 0.2$, TLA reduces 16% of the cost over IS-41 [see Fig. 5(a)]. When the user movement exhibits high locality ($\theta = 0.75$), TLA significantly outperforms IS-41 when ρ is small. The improvements of TLA over IS-41 are 31% to 50% [for a low registration cost $\delta = 0.2$; see Fig. 5(b)], 51% to 58% [for a moderate registration cost $\delta = 0.4$; see Fig. 6(a)], and 65% to 67% [for a large registration cost $\delta = 0.8$; see Fig. 6(b)].

When ρ is large, the performance of TLA is significantly affected by the portable's RA residence time distribution (specifically, the variance V of the Gamma distribution). For a reasonable large variance (e.g., $V > 10/\lambda_m^2$), TLA outperforms IS-41 for the ranges of all input parameters considered in our study (when ρ is large, 5% to 20% improvements are observed in Figs. 5 and 6). On the other hand, IS-41 outperforms TLA when V and δ are small, and ρ is large



(a) (b)

Fig. 5. The costs for IS-41 and TLA. (a) $\Theta = 0.25, \delta = 0.2$. (b) $\Theta = 0.75, \delta = 0.2$.



(a) (b)

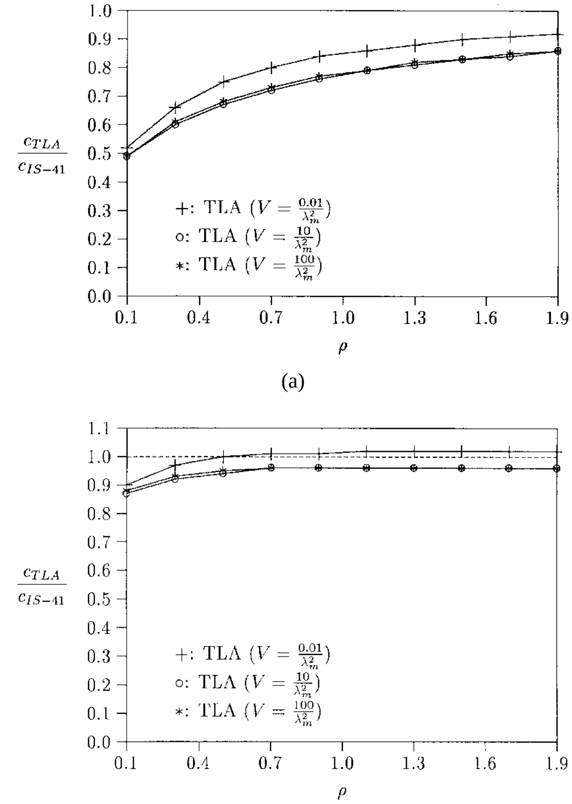
Fig. 6. The costs for IS-41 and TLA. $\Theta = 0.75, \delta = 0.4$. (b) $\Theta = 0.75, \delta = 0.4$.

(see Fig. 5). From (21), it is not difficult to see that ω is a decreasing function of V. Consider the cases when $V \to \infty$ and $V \to 0$:

$$\lim_{V\to\infty} f_m^*(\lambda_c) = \lim_{\gamma\to\infty}\left(\frac{\gamma}{\rho+\gamma}\right)^\gamma = 1 \Rightarrow \lim_{V\to\infty}\omega = 0$$

$$\lim_{V\to 0} f_m^*(\lambda_c) = 0 \Rightarrow \lim_{V\to 0}\omega = \frac{\theta}{\rho}. \tag{15}$$

From (13), $c_{2,\text{TLA}}$ also is a decreasing function of V and the performance of TLA degrades as the variance V of the



(b)

Fig. 7. The effects of the call-to-mobility ratio ρ. $\Theta = 0.6, \delta = 0.6$. (b) $\Theta = 0.2, \delta = 0.2$.

RA residence time distribution decreases (see Figs. 5 and 6). When δ is small, $c_{2,\text{TLA}}$ in (13) has more effect on the cost c_{TLA}, and IS-41 has a better chance to outperform TLA when V is small.

Fig. 7 indicates that $c_{\text{TLA}}/c_{\text{IS-41}}$ is either an increasing functions of ρ [see Fig. 7(a)] or a convex function [see the curve for $V = 0.01/\lambda_m^2$ in Fig. 7(b)], $\lim_{\rho\to\infty} c_{\text{TLA}}/c_{\text{IS-41}} = 1$, the value will approach to one). From (21), ω is a decreasing function of ρ, and

$$\lim_{\rho\to\infty} f_m^*(\lambda_c) = 0 \Rightarrow \lim_{\rho\to\infty}\omega = 0.$$

Since

$$c_{\text{IS-41}} - c_{\text{TLA}} = \frac{\delta\theta}{\rho} - (1-\delta)\omega.$$

We have

$$\lim_{\rho\to\infty}(c_{\text{IS-41}} - c_{\text{TLA}}) = 0. \tag{16}$$

That is, for all θ and δ values, TLA is the same as IS-41 when ρ is large (i.e., $\lim_{\rho\to\infty} c_{\text{TLA}}/c_{\text{IS-41}} = 1$ in Fig. 7). In this case, there are many call terminations between two portable movements. For these call terminations the HLR always has location hits (except for the first one that may or may not have a location miss), the basic *find* operation is performed in TLA.

When $\rho \to 0$, we have

$$\lim_{\rho\to 0} f_m^*(\lambda_c) = 1 \Rightarrow \lim_{\rho\to 0}\rho\omega = 0$$

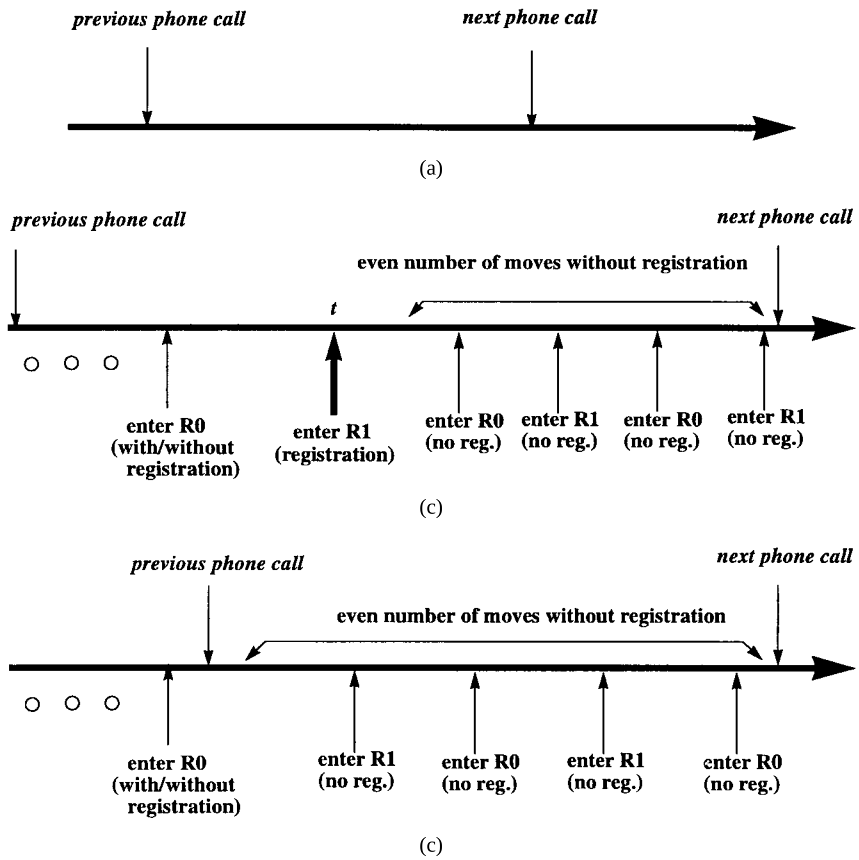


Fig. 8. Situations when the HLR has the correct view of the latest visited RA. (a) Case I: No move between two phone calls. (b) Case II: The last registration followed by an even number of moves. (c) Case III: No registration between two phone calls.

and

$$\lim_{\rho\to 0}(c_{\text{IS-41}} - c_{\text{TLA}}) = \lim_{\rho\to 0}\frac{1}{\rho}[\delta\theta - (1-\delta)(\rho\omega)] = \infty. \tag{17}$$

That is, for all δ values and all $\theta \neq 0$, TLA is much better than IS-41 when ρ is small. In this case, the portables move without receiving any phone call. It is clear that the move cost for TLA is always cheaper than IS-41 if $\theta > 0$. When $\theta = 0$,

$$\lim_{\theta\to 0}(c_{\text{IS-41}} - c_{\text{TLA}}) = 0.$$

That is, if a portable never moves back to the previously visited RA, then TLA behaves exactly the same as IS-41 does.

Our analysis indicated that TLA may underperform IS-41 when the variance V of the RA residence times is small. In this case, it is desirable to disable TLA when the cost of the TLA is higher than IS-41. Heuristics similar to the ones proposed in [8] and [9] can be used to dynamically disable/enable TLA.

V. Conclusion

This paper studied the performance of a PCS location tracking strategy called the TLA. TLA is a variation (and a special case) of the alternative location algorithm proposed in [10]. We studied how the incoming call rate, the portable mobility, and the portable locality affect the performance of TLA. Our study indicated that TLA significantly outperforms IS-41 when the call-to-mobility ratio is low (i.e., the user moves more frequently than to receive calls) or when the registration cost is large. We also showed that the variance V of the portable residence times (in registration areas) may affect the performance of TLA when the registration cost is low. The performance of TLA degrades when V is small.

TLA can be easily implemented by modifying the existing IS-41 based systems. The modifications are done in two places. Firstly, one more field should be added to the portable location record in the HLR to hold the address of the second portable location. A few lines of code are added to the HLR registration/call delivery procedures to exercise TLA. The HLR modifications can be easily done by using the user friendly tools with graphical inputs such as ISCP SPACE System [24]. Secondly, a few lines of code should be added to the micro kernel of the protable to exercise TLA. Note that we may need to slightly modify the SS7 TCAP message format, but no hardware modifications are necessary.

Our future research directions include modeling the general alternative location strategy by combining our model and the models proposed in [10]–[12], and comparing TLA with the pointer forwarding scheme [3]. Also, it is possible that IS-41 may outperform TLA in certain conditions. We are developing an adaptive mechanism (similar to the one described in [8]) to dynamically switch between IS-41 and TLA based on the user mobility.

Appendix
Derivation for ω

We first consider the probability $(1-\omega)$ that the HLR has the correct view of the latest visited RA when a call arrives (i.e., a location hit occurs and the *find* cost for TLA is the same as IS-41).

If the HLR has the correct view of the current RA of the portable when a call arrives, one of the following situations holds.

1) Case I: The portable has not moved since the last call arrival [see Fig. 8(a)]. After a call termination, the HLR always identifies the current RA of the portable. Thus, location hit occurs when the second call arrives. The probability ω_1 of Case I is

$$\omega_1 = \alpha(0). \tag{18}$$

2) Case II: Suppose that the portable visits $K > 0$ RA's between two phone calls and *registration* operations are performed during some of the K moves. If the HLR has the correct view of the latest RA at the next call arrivals, then during the K moves, the last *registration* operation performed in TLA is followed by even number of portable movements (and no *registration* operations are performed in these movements) as shown in Fig. 8(b). Note that at time t in Fig. 8(b), the HLR's view of the current RA is the same as the portable (which is R_1). After four moves (without any registration), the portable moves back to R_1, and the HLR's view for the latest visited RA is correct. Let ω_2 be the probability of Case II, and $\omega_2(K)$ be the conditional probability of Case II when the portable makes K moves between two phone calls. Since the probability of a move with registration is $(1-\theta)$ and the probability of a move without registration is θ, we have

$$\begin{aligned}\omega_2(K) &= (1-\theta)\sum_{i=0}^{\lfloor (K-1)/2\rfloor} \theta^{2i}\\ &= \frac{1-\theta^{2\lfloor (K-1)/2\rfloor+2}}{1+\theta}, \qquad \text{for } K>0.\end{aligned}$$

Note that $\omega_2(2i+2) = \omega_2(2i+1)$ for $i \geq 0$. The probability ω_2 is

$$\begin{aligned}\omega_2 &= \sum_{K=1}^{\infty} \omega_2(K)\alpha(K)\\ &= \sum_{i=0}^{\infty} \omega_2(2i+1)[\alpha(2i+1)+\alpha(2i+2)]\\ &= \frac{[1-f_m^*(\lambda_c)]^2[1+f_m^*(\lambda_c)]}{\rho(1+\theta)}\\ &\quad \times \left\{\sum_{i=0}^{\infty}[1-\theta^{2(i+1)}]\, f_m^*(\lambda_c)^{2i}\right\}\\ &= \frac{[1-f_m^*(\lambda_c)]^2[1+f_m^*(\lambda_c)]}{\rho(1+\theta)}\\ &\quad \times \left\{\frac{1}{1-[f_m^*(\lambda_c)]^2} - \frac{\theta^2}{1-[\theta f_m^*(\lambda_c)]^2}\right\}\\ &= \frac{[1-f_m^*(\lambda_c)]^2[1+f_m^*(\lambda_c)]}{\rho(1+\theta)}\\ &\quad \times \left\{\frac{1-\theta^2}{\{1-[f_m^*(\lambda_c)]^2\}\{1-[\theta f_m^*(\lambda_c)]^2\}}\right\}\\ &= \frac{[1-f_m^*(\lambda_c)](1-\theta)}{\rho\{1-[\theta f_m^*(\lambda_c)]^2\}}.\end{aligned} \tag{19}$$

3) Case III: Suppose that the portable visits $K > 0$ RA's between two phone calls and no *registration* operation is performed during the K moves. If the HLR has the correct view of the latest visited RA at the next call arrival, then K must be an even number as shown in Fig. 8(c). In this figure, the HLR has the correct view of the latest visited RA (which is R_0) immediately after the previous call termination. After four moves without registration, the portable moves back to R_0 when the next call arrives. The probability ω_3 of Case III is

$$\begin{aligned}\omega_3 &= \sum_{i=1}^{\infty}\theta^{2i}\alpha(2i)\\ &= \frac{[1-f_m^*(\lambda_c)]^2}{\rho f_m^*(\lambda_c)}\sum_{i=1}^{\infty}[\theta f_m^*(\lambda_c)]^{2i}\\ &= \frac{[1-f_m^*(\lambda_c)]^2\theta^2 f_m^*(\lambda_c)}{\rho\{1-[\theta f_m^*(\lambda_c)]^2\}}.\end{aligned} \tag{20}$$

From (18) to (20), the probability ω that the HLR has a location miss in a call delivery is

$$\begin{aligned}\omega &= 1-\omega_1-\omega_2-\omega_3\\ &= \frac{1-f_m^*(\lambda_c)}{\rho}\\ &\quad \times\left\{1-\frac{1-\theta+\theta^2 f_m^*(\lambda_c)+[\theta f_m^*(\lambda_c)]^2}{1-[\theta f_m^*(\lambda_c)]^2}\right\}.\end{aligned} \tag{21}$$

References

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

[2] Bellcore, "Generic criteria for version 0.1 wireless access communications systems (WACS) and supplement," Bellcore, Tech. Rep. TR-INS-001313, Issue 1, 1994.

[3] 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.

[4] Y.-B. Lin and A. Noerpel, "Implicit deregistration in a PCS network," *IEEE Trans. Veh. Technol.*, vol. 43, pp. 1006–1010, 1994.

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

[6] K. S. Meier-Hellstem and E. Alonso, "The use of SS7 and GSM to support high density personal communications," presented at the IEEE ICC, 1992.

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

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

[9] Y.-B. Lin, "Determining the user locations for personal communications networks," *IEEE Trans. Veh. Technol.*, vol. 43, pp. 466–473, 1994.

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

[11] G. P. Pollini and S. Tabbane, "The intelligent network signaling and switching costs of an alternate location strategy using memory," *IEEE Trans. Veh. Technol.*, vol. 42, 1993.

[12] H. Xie, S. Tabbane, and J. D. Goodman, "Dynamic location area management and performance analysis," *IEEE Trans. Veh. Technol.*, vol. 42, 1993.

[13] Y.-B. Lin and W. Chen, "Impact of busy lines and mobility on call blocking in a PCS network," *Int. J. Commun. Syst.*, vol. 9, pp. 35–45, 1996.

[14] W. Feller, *An Introduction to Probability Theory and Its Applications, Vol. I.* New York: Wiley, 1966.

[15] S. M. Ross, *Stochastic Processes.* New York: Wiley, 1983.

[16] N. L. Johnson and S. Kotz, *Distributions in Statistics: Continuous Multivariate Distributions.* New York: Wiley, 1972.

[17] Y.-B. Lin and S. K. DeVries, "PCS network signaling using SS7," *IEEE Personal Commun. Mag.,* pp. 44–55, June 1995.
[18] EIA/TIA, "Cellular intersystem operations–Baseline text draft 3 9-2-94 (IS-41 Rev. C)," EIA/TIA, Tech. Rep. PN-2991, 1994.
[19] A. R. Noerpel, L. F. Chang, and Y.-B. Lin, "PACS contention algorithm for initial access and handoff," in *Proc. 5th WINLAB Workshop,* 1995, pp. 217–226.
[20] ETSI, Digital European Cordless Telephone Common Interface, Version 05.03, 1991.
[21] Radio Advisory Board of Canada, *CT2Plus class 2: Specification for the Canadian Common Air Interface for Digital Cordless Telephony, Including Public Access Services, Annex 1 to Radio Standards Specification 130,* 1992.
[22] ETSI, *Common Air Interface Specification to be ued for the Interworking Between Cordless Telephone Apparatus in the Frequency Band 864.1 MHz, Including Public Access Services, I-ETS 300 131:1990,* June 1991.
[23] ANSI/EIA/TIA, "Mobile station-land station compatibility specification," EIA/TIA, Tech. Rep. 553, 1989.
[24] Bellcore, "ISCP SPACE System Service Creation Guide," Bellcore, Tech. Rep. BR-780-003-124, 1993.

Yi-Bing Lin received the B.S.E.E. degree from National Cheng Kung University, R.O.C., in 1983, and the Ph.D. degree in computer science from the University of Washington, Seattle, in 1990.

Between 1990 and 1995, he was with the Applied Research Area at Bell Communications Research, Morristown, NJ. In 1995, he was appointed Full Professor of the Department of Computer Science and Information Engineering, National Chiao Tung University, Hsinchu, Taiwan, R.O.C. In 1996, he was appointed as Deputy Director of Microelectronics and Information Systems Research Center, NCTU. His current research interests include design and analysis of personal communications services network, mobile computing, distributed simulation, and performance modeling. He is an Associate Editor of the *ACM Transactions on Modeling and Computer Simulation*, a Subject Area Editor of the *Journal of Parallel and Distributed Computing*, an Associate Editor of the *International Journal in Computer Simulation*, an Associate Editor of IEEE NETWORKS, an Associate Editor of S*imulation* magazine, an Area Editor of *ACM Mobile Computing and Communication Review*, and a columnist of *ACM Simulation Digest*. He is a member of the Editorial Board of the International Journal of Communications, a member of the Editorial Board of Computer Simulation Modeling and Analysis, Program Chair for the 8th Workshop on Distributed and Parallel Simulation, General Chair for the 9th Workshop on Distributed and Parallel Simulation. Program Chair for the 2nd International Mobile Computing Conference, the publicity chair of ACM Sigmobile, Guest Editor for the ACM/Baltzer WINET special issue on Personal Communications, and Guest Editor for the IEEE TRANSACTIONS ON COMPUTERS special issue on Mobile Computing.