A Forwarding Strategy to Reduce Network Impacts of PCS

Ravi Jain, Yi-Bing Lin, Charles Lo† and Seshadri Mohan*
Applied Research, Bellcore, Morristown, NJ 07960

Abstract

We propose a per-user forwarding strategy for locating users who move from place to place while using Personal Communications Services (PCS). The forwarding strategy augments the basic location strategy proposed in existing standards such as GSM and IS-41, with the objective of reducing network signalling and database loads in exchange for increased CPU processing and memory costs. With the forwarding strategy, calls to a given user will first query the user's Home Location Register (HLR) to determine the first Visitor Location Register (VLR) which the user was registered at, and then follow a chain of forwarding pointers to the user's current VLR. This strategy is useful for those users who receive calls infrequently relative to the rate at which they change registration areas.

We use a reference PCS architecture and the notion of a user's call-to-mobility ratio (CMR) to quantify the costs and benefits of using forwarding and classes of users for whom it would be beneficial. We show that under a variety of assumptions forwarding is likely to yield significant net benefits for certain classes of users, in exchange (possibly) for a small increase in mean call setup time. For instance, under certain cost assumptions, for users with $CMR < 1$ forwarding can result in 20-60% savings over the basic strategy, with no increase in mean call setup time.

1 Introduction

We consider the problem of locating users who move from place to place while using Personal Communications Services (PCS). Previous studies [12, 15, 11, 13] have shown that, with predicted levels of PCS users, there will be significant loads upon the signalling network and network databases, and that these loads are dependent upon the data management strategies adopted. We present a user location strategy which has the potential to reduce these loads significantly. The strategy we discuss here is an *auxiliary* strategy, in that it augments the *basic* user location strategies proposed in standards such as the North American IS-41 cellular standard [5] and the European GSM standard for mobile communications [14, 18]. (For surveys see [8, 17].)

The strategy we present is the use of *per-user forwarding.* This strategy, like other auxiliary strategies [8], attempts to reduce the network signalling and database loads of the basic strategies in exchange for increased CPU processing and memory costs. Since technology trends are driving the latter costs down, deploying the forwarding strategy on a system-wide basis will become increasingly attractive. Once deployed, whether the forwarding strategy should be invoked for a particular user is a function of the user's mobility and communications patterns, as discussed below.

The *basic* user location strategies proposed in the IS-41 [5] and GSM [14, 18] standards use a two-tier system of Home Location Register (HLR) and Visitor Location Register (VLR) databases (see also [17]). The user performs a registration at the HLR every time the user changes registration areas, and deregisters at the previous VLR. The key observation we make is that, in many cases, it should be possible to avoid these 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; several techniques can be used to limit the maximum length of the chain. 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 registration areas. This idea attempts to exploit patterns in the call reception and mobility patterns of PCS users, and is called *per-user forwarding.*[1]

*Address correspondence to: Ravi Jain, Bellcore, 445 South St, Morristown, NJ 07960. Phone: 201-829-4756, Fax: 201-829-5888, E-mail: rjain@thumper.bellcore.com

†Author's present address: AirTouch Communications, Walnut Creek, CA 94596.

[1]We have previously designed an auxiliary strategy, called *per-user caching,* which attempts to exploit the calling an mo-

The idea of forwarding itself is not novel, as it has been used in various forms in computer systems software designs (e.g. [6]). Forwarding strategies have also been proposed for locating mobile users in [1, 7]. However, in this paper we propose a different forwarding strategy, concentrate on the issues which arise in its application to a PCS telecommunication signalling network, and attempt to understand the classes of users and the system parameters for which the strategy is likely to yield benefits.

The outline of this paper is as follows. In sec. 2, we describe the PCS network architecture that we assume. (Some of this background material in this paper is repeated from our earlier paper [10] in order to keep this paper self-contained). In sec. 3, we describe the forwarding strategy. A feature of the forwarding location strategy is that it is useful only for certain classes of PCS users, those meeting certain call and mobility criteria. We encapsulate this notion in the definition of the user's call-to-mobility ratio (CMR) in sec. 4. We then use this definition to quantify the costs and benefits of forwarding for different classes of users.[2] In sec. 5 we use our our PCS network reference architecture to determine, given various performance bottlenecks in the system, the threshold CMRs below which forwarding is beneficial. In sec. 6 we briefly discuss alternative schemes and implementation issues. Section 7 provides some conclusions.

It should be stressed at the outset that the assumptions we have used in this paper constitute one reference set of assumptions. A number of variations in the assumptions could be considered, including variations in the network architecture and in the auxiliary location strategy. The intent of this paper is to present the key ideas behind the forwarding auxiliary strategy and to develop a method for quantifying its costs and benefits. This method can then be applied to specific architectures and deployment scenarios as needed.

2 PCS network architecture

PCS users receive calls via either wireless or wireline access. In general, calls may deliver voice, data, text, facsimile or video information. For our purposes, we define the location of a PCS user, as known by the wire-line network, as the *registration area* (RA) in which the user is located. For users attached directly to a wire-line network, the RA is defined as the point of attachment. For users attached via wireless links, the situation is described as follows. In order to deliver calls by wireless links, the geographical region covered by a PCS network is divided into radio port coverage areas, or *cells*. Each cell is primarily served by one radio *base station*, although a base station may serve one or more cells. The base station locates a user, and delivers calls to and from the user, by means of paging within the cell(s) it serves. Base stations are connected to the rest of the wire-line network by wireline links. An RA consists of an aggregation of cells, forming a contiguous geographical region.

bility patterns of a different population of users, namely those who receive calls frequently relative to the rate at which they change registration areas [10].

[2]As we will see in later sections, the network, and not the user, bears the burden of maintaining information about the user's CMR.

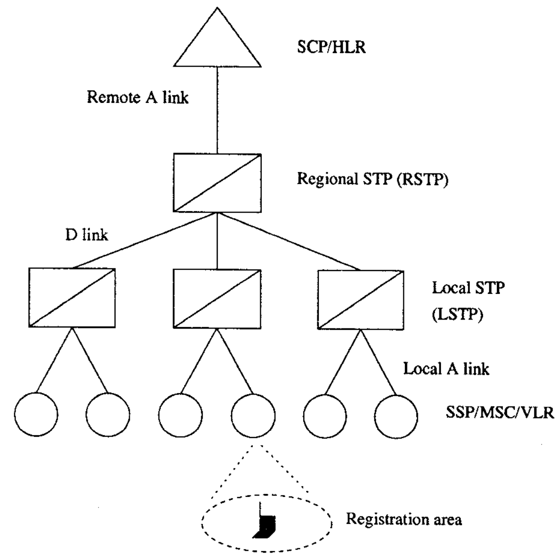

Figure 1: Reference CCS network architecture

We assume that a Common Channel Signalling (CCS) network, with a Signaling System No. 7 (SS7) protocol [16], as commonly used for telecommunications, is used to set up calls. Fig. 1 illustrates the reference signalling network assumed in this study.[3] The cells of the geographical region are served by base stations and are aggregated into RAs. The base stations of an RA are connected via a wire-line network to an end-office switch, or Service Switching Point (SSP). Each SSP serves a single RA. SSPs of different RAs are in turn connected to a two-level hierarchy of Signalling Transfer Points (STPs), comprised of a Regional STP (RSTP) connected to all Local STPs (LSTPs) in the region, which perform message routing and other SS7

[3]This architecture is meant as a reference architecture for a hypothetical geographical region, and is not necessarily the architecture corresponding to any particular implementation.

functions.[4] The RSTP is also connected to a Service Control Point (SCP), which is assumed to contain the functionality associated with a Home Location Register (HLR) database.

For simplicity, we assume that in functional terms the MSC is collocated with the SSP. In addition, a distinct Visitor Location Register (VLR) database is associated with each MSC. These assumptions are not unreasonable [10]. (In the rest of this paper, the terms switch or SSP are used interchangeably). Each switch is assumed to serve exactly one RA, which, in turn, may be comprised of one or more cells. This assumption is used to simplify the ensuing analysis, although in practice, each VLR may serve a number of RAs.

In this paper we do not address issues relating to the content of messages and other information transfer (e.g., for billing, etc.) which may occur during a call. Since we only perform a comparative analysis of the basic strategy with and without the auxiliary forwarding strategy it is assumed that message sizes are equal for different types of transactions (e.g., location request, registration and deregistration).

3 Per-user Forwarding

The basic idea behind per-user forwarding is to avoid unnecessary CCS message traffic, and updates of a user's HLR, if the user moves across RAs relatively frequently but receives calls relatively infrequently. Two operations which incur signalling are defined (we use the terminology of [10, 1]): *MOVE*, where the PCS user moves from one RA to another, and *FIND*, where the RA in which the PCS user is currently located is to be determined.

3.1 Basic user location procedures

We first present the *MOVE* and *FIND* procedures used in current PCS standards proposals such as IS-41 [5]; we call these the *BasicMOVE* and *BasicFIND* procedures. (Note that the *BasicMOVE* and *BasicFIND* procedures we present are simplifications of those in the standards, and only attempt to capture the major interactions between the HLR and VLR databases relevant to our purposes. The reader is referred to [5, 18] for their detailed specification.) We use the following pseudo-code to describe the *BasicFIND* and *BasicMOVE* procedures.

[4] In practice each STP actually consists of two STPs in a mated-pair configuration for redundancy [16]; for simplicity Fig. 1 only shows one of the two STPs of each mated pair.

BasicMOVE( ) {
    The mobile terminal detects that it is in a new registration area;
    The mobile terminal sends a registration message to the new VLR;
    The new VLR sends a registration message to the user's HLR;
    The HLR sends a registration cancellation message to the old VLR;
    The old VLR sends a cancellation confirm message to the HLR;
    The HLR sends a registration confirm message to the new VLR;
}

Notice that the *BasicMove* procedure involves notifying the HLR every time a user enters a new registration area. The registration information is used to find the user when a call is to be delivered to that user.

BasicFIND( ) {
    Call to PCS user is detected at local switch;
    if called party is in same RA **then** return;
    Switch queries called party's HLR;
    HLR queries called party's current VLR, *V*;
    VLR *V* returns called party's location to HLR;
    HLR returns location to calling switch;
}

3.2 Forwarding user location procedures

The forwarding procedures modify the basic procedures as follows. When a user moves from one RA to another it informs the switch (and VLR) at the new RA of the old RA from which it arrived. The switch at the new RA determines whether to invoke the basic *MOVE* or the forwarding *MOVE* strategy.

In *FwdMOVE*, the new VLR exchanges messages with the old VLR to set up a forwarding pointer from the old to the new VLR, but does not involve the user's HLR. (See Fig. 2). The message labeled $REG(RAa)$ in Fig. 2 is a registration message from the user terminal indicating that it has moved from the previous RA, *RAa*, to the current RA. The messages $REG(RAv)...REG(RAx)$ are similar. The message *REGPTR* is a message from the new VLR to the old VLR specifying that a forwarding pointer is to be set up; message *regptr* is the confirmation from the old VLR that this has been done.

A subsequent call to the user from some other switch will invoke the *FwdFIND* procedure, which queries the user's HLR as in the basic strategy, and ob-

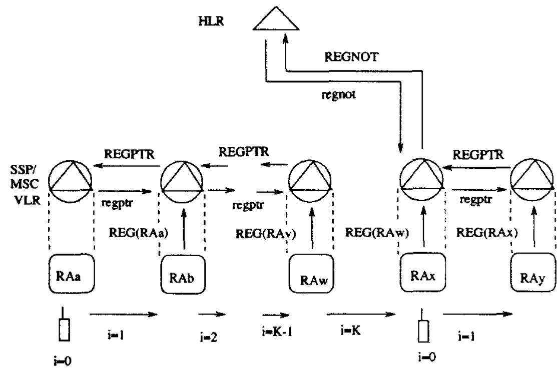

Figure 2: The *FwdMOVE* procedure

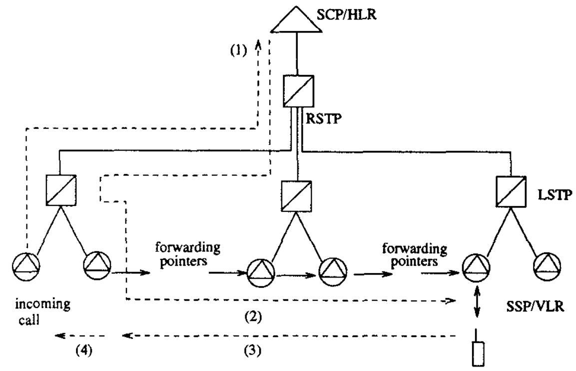

Figure 3: The *FwdFIND* procedure

tains an outdated pointer to the old VLR. The pointer from the old VLR is then followed to the new VLR, in order to determine the user's correct current location. (See Fig. 3). To ensure that the time taken by *FwdFIND*, and hence the call setup time, is bounded to a reasonable value, the length of the chain of forwarding pointers must be limited. This is accomplished by allowing the chain to grow to at most $K-1$ pointers during the *FwdMOVE* process. We describe the forwarding procedures using pseudo-code as follows.

FwdMOVE( ) {
    /* Initially, i is 0 */
    if ($i < K-1$) {
        User registers at new RA/VLR, passing id of former RA/VLR;
        New VLR deregisters user at old VLR;
        Old VLR sends ACK and user's
        service profile to new VLR;
        $i := i + 1$;
    }
    else {
        BasicMOVE;
        $i := 0$;
    }
}

FwdFIND( ) {
    Call to PCS user is detected at local switch
    if called is in same RA **then** return
    Switch queries called's HLR
    HLR responds to caller's switch with V_0
    Caller's switch queries V_0
    while (Queried VLR is not called's current VLR, V_c)
        VLR queries next VLR in pointer chain
        /* Now called's actual VLR has been found */
        /* Begin return phase */
    while (Queried VLR is not VLR V_0)
        VLR passes V_c to previous VLR in chain
    VLR V_0 send user location to caller's switch
}

4 Performance modeling of per-user forwarding

We characterize classes of users by their *call-to-mobility ratio* (CMR). The CMR of a user is the average number of calls *received* by a user per unit time, divided by the average number of times the user changes registration areas per unit time. Forwarding will only be applied to users meeting a certain CMR threshold.

Suppose a user crosses several RAs between two consecutive phone calls. If the basic user location strategy is used, the *BasicMOVE* procedure is called every time the user moves to a new RA. If the forwarding strategy is used, the HLR is updated only every Kth move, with forwarding pointers being set up for all other moves.

We calculate C_B and C_F, the total $MOVE$ and $FIND$ costs of maintaining user information and locating the user between two consecutive RA crossings for the basic and forwarding strategies, respectively. Let M and F represent the costs of a single invocation of *BasicMOVE* and *BasicFIND* respectively, and let M' and F' represent the average costs of an invocation of *FwdMOVE* and *FwdFIND*. Then, since there are p call arrivals per $MOVE$, we have

$$C_B = M + pF \quad (1)$$

$$C_F = M' + pF' \qquad (2)$$

We will be interested in two measures of performance: the ratio C_F/C_B, which reflects the average benefit of forwarding in terms of reduced total costs, and the ratio F'/F, which reflects the average additional cost paid during call setup for those users to whom call forwarding is applied.

In the next section we will compare the costs of the forwarding strategy with that of the basic strategy for specific cost model based upon the reference architecture of Fig. 1. For the moment we denote costs for various operations used in the forwarding strategy as follows.

S = *Cost of setting up a forwarding pointer between VLRs during a* FwdMOVE.

T = *Cost of traversing a forwarding pointer between VLRs during a* FwdFIND

Here S includes the cost of both the $REGPTR$ and *regptr* messages, and T includes the cost of traversing the forwarding pointer in both directions. Then, the average cost of a *FwdMOVE* is simply the weighted average of the cost of setting up forwarding pointers and the cost of performing a *BasicMOVE*. Every K moves require $K-1$ pointer creations, so the expected pointer creation cost per move is $(K-1)S/K$. Every K moves require one *BasicMOVE* operation, and the cost per move is M/K. Thus, for $K > 0$,

$$M' = \frac{K-1}{K}S + \frac{M}{K} \qquad (3)$$

We evaluate the cost of *FwdFIND* as follows. When a *FwdFIND* occurs only half the number of messages need to be exchanged between the HLR and VLRs compared to the *BasicFIND* procedure. In the worst case, a call which arrives to the user after the user has moved $K-1$ times will incur a cost of traversing $K-1$ pointers. We also make the simplifying assumption that calls arrive to the mobile user uniformly and deterministically as the user moves, so that on average a *FwdFIND* will involve traversing $(K-1)/2$ pointers. Finally, we assume that returning the called party's actual location to the calling switch also costs T. Thus, the cost per *FwdFIND* becomes (assuming $K > 0$)

$$F' = \frac{F}{2} + \frac{KT}{2} \qquad (4)$$

From Eq. (1-4),

$$\frac{C_F}{C_B} = \frac{pF + pKT}{2(M + pF)} + \frac{(K-1)S + M}{K(M + pF)} \qquad (5)$$

$$\frac{F'}{F} = \frac{1}{2} + \frac{KT}{2F} \qquad (6)$$

In order to evaluate the objective functions C_B/C_F and F'/F we need values for S, T, M and F. As a first approximation, without performing any detailed analysis of the operations involved, we observe that *BasicMOVE* and *BasicFIND* involve the same number of messages between HLR and VLR databases, so we set $M = F$. We also assume that the cost of setting up a forwarding pointer is the same as the cost of traversing it, i.e., we set $S = T$. Finally, we normalize $M = 1$, and $S = \delta$, with $\delta < 1$. Then

$$\frac{C_F}{C_B} = \frac{[1 + (K-1)\delta](2 + Kp) + pK\delta}{2K(1+p)} \qquad (7)$$

$$\frac{F'}{F} = \frac{1}{2} + \frac{K\delta}{2} \qquad (8)$$

Thus we see from Eq. 8 that $F' < F$ as long as $K < 1/\delta$. In Figs. 4 and 5 we plot C_F/C_B and F'/F as functions of p and K for various values of δ and $K > 1$. It is clear that forwarding can be of substantial benefit, and can be used to reduce total $FIND$ and $MOVE$ costs by upto 30-50% for certain classes of users by choosing appropriate values of K, but that the benefits are highly dependent upon the value of δ.

5 Forwarding threshold analysis

In this section we compare the costs of the forwarding strategy with that of the basic strategy for a specific cost model based upon the reference architecture of Fig. 1. Thus instead of treating the relative costs of setting up pointers and performing a *BasicMOVE* simply in terms of δ, we examine the relative costs incurred by each network element for these operations.

To quantify this further, assume costs for traversing various network elements as follows. Let L and R represent the costs of processing and routing a message by an LSTP and RSTP, respectively. Let H_Q and V_Q be the cost of querying the HLR and a VLR respectively. Let A_l, D, and A_r be the cost of transmitting a message over a local A-link, a D-link and the remote A-link, respectively.

From Fig. 1, the cost for *FwdFIND* depends upon whether the originating switch shares the same LSTP as the switch/VLR indicated by the called's HLR (which in turn may not necessarily reflect the called's actual location). If not, the originating switch must query the pointed VLR via the RSTP. Also, if the called party moves outside its LSTP area frequently, searching for its current location will also involve following forwarding pointers via the RSTP frequently.

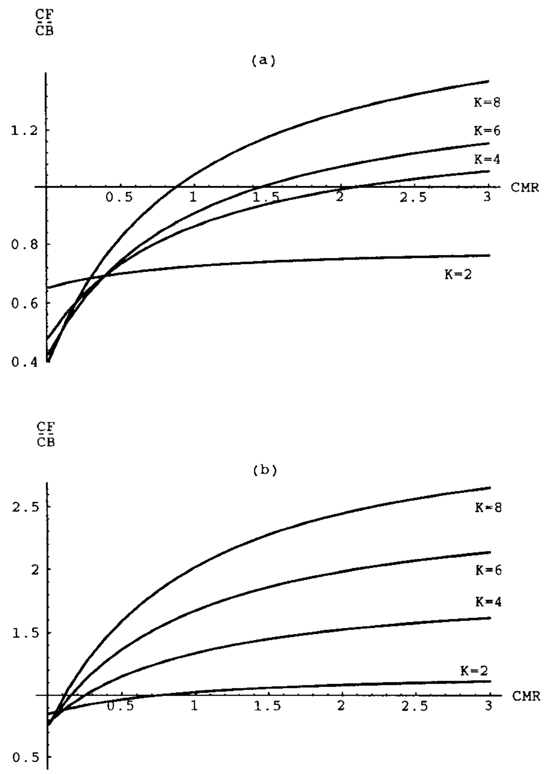

Figure 4: Relative total $MOVE$ and $FIND$ cost of forwarding: (a) $\delta = 0.3$, (b) $\delta = 0.7$

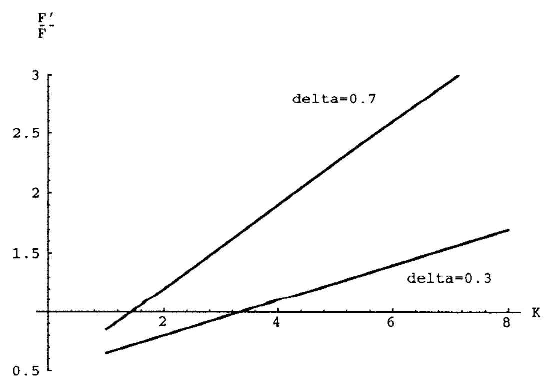

Figure 5: Relative call setup time for *FwdFIND* and *BasicFIND*,for $\delta = 0.3$ and $\delta = 0.7$

To quantify this behavior we introduce some additional parameters. Later we will comment on how to estimate them. Let

t = *Prob (called user's location, according to HLR, is in caller's LSTP area)*

s = *Prob (any given* MOVE*by the user is within its current LSTP area)*

j = *number of forwarding pointers traversed to locate a user, per* FwdFIND

As before, we denote the average cost of a *FwdFIND* by F' and that of a *BasicFIND* by F; the former can be calculated by evaluating the *FwdFIND* algorithm line-by-line, as follows.

$$
\begin{aligned}
F' &= A_l + L + D + R + A_r + H_Q \\
&+ A_l + L + D + R + A_r \\
&+ 2A_l + L + (1-t)(L + 2D + R) + V_Q \\
&+ j[2A_l + L + (1-s)(L + 2D + R) + V_Q] \\
&+ j[2A_l + L + (1-s)(L + 2D + R)] \\
&+ 2A_l + L + (1-t)(L + 2D + R) \qquad (9)
\end{aligned}
$$

Similarly, the cost F can be derived for *BasicFIND*. It is straightforward to show that the difference is:

$$
\begin{aligned}
\Delta F &\equiv F' - F \\
&= (2 + 4j)A_l + 2j(2 - s)L \\
&+ [2(1 - 2t) + 4j(1 - s)]D \\
&+ [2j(1 - s) - 2t]R - 2A_r + jV_Q \qquad (10)
\end{aligned}
$$

We now calculate the average cost of a *FwdMOVE*, taking into account that one out of every K *FwdMOVE* is the same as a *BasicMOVE*. Also, let V_U and H_U denote the costs of updating a VLR and an HLR, respectively.

$$
\begin{aligned}
M' &= \frac{K-1}{K}[V_U + 2A_l + L \\
&+ (1-s)(L + 2D + R) + V_U + 2A_l + L \\
&+ (1-s)(L + 2D + R)] + \frac{M}{K} \qquad (11)
\end{aligned}
$$

Note that in this expression, we assume that the user's profile and other necessary information are part of the acknowledgment received by the new VLR. We denote the benefit of *FwdMOVE* over *BasicMOVE* as ΔM, and it can be shown that

$$
\begin{aligned}
\Delta M &\equiv M - M' = \frac{K-1}{K}[2sL \\
&+4sD + 2(1+s)R + 4A_r + H_U] \quad (12)
\end{aligned}
$$

A *FwdMOVE* is never more costly than a *BasicMOVE*. For there to be a net benefit, the excess cost of

Dom. Term	ΔM	ΔF	Use Forward--ing?
A_l	0	$(2+4j)A_l$	Never
L	$\frac{2(K-1)sL}{K}$	$2j(2-s)L$	see text
D	$\frac{4(K-1)sD}{K}$	$(2-4t+4j$ $-4js)D$	see text
R	$\frac{2(K-1)(1+s)R}{K}$	$2(j-js$ $-t)R$	see text
A_r	$\frac{4(K-1)A_r}{K}$	$-2A_r$	Always
H_U	$\frac{(K-1)H_U}{K}$	0	Always
V_Q	0	jV_Q	Never

Table 1: Evaluation of Forwarding for Various Dominant Cost Terms

FwdFIND over *BasicFIND*, over all calls received by the called party, has to be less than the benefit of *FwdMOVE* relative to *BasicMOVE*, i.e., the called party should receive calls relatively infrequently. Since, to a given called party the ratio of *FwdFIND*'s per *FwdMOVE* is CMR, we require that $\Delta M - p\Delta F > 0$, i.e., $p < \Delta M / \Delta F$; this can be evaluated using equations (10) and (12) above.

In order to get an intuitive understanding for this formula, we consider situations for which only one of the cost terms in equations (10) and (12) above dominates, and calculate the costs and benefits of *FwdMOVE* and *FwdFIND*. (See Table 1).

In three of the situations in Table 1, it is not clear if forwarding is worthwhile or not. To investigate this, it is necessary to make additional assumptions, namely to quantify s and j. As before, we assume that on average $j = (K-1)/2$.

To estimate s, we need a model of the called's mobility across LSTP areas. We use the following simple model. Assume that an LSTP area consists of $m \times m$ RA's arranged in a square, and each RA is itself a square. Users are assumed to be uniformly distributed throughout the LSTP area. Furthermore, each time a user leaves an RA, one of the four sides is crossed with equal probability. Then,

Prob[user crosses an LSTP area boundary]
= Prob[user is in border RA of LSTP area]
× Prob[user crosses LSTP area boundary]
$= 1/m$

Thus, under these assumptions, $s = 1 - 1/m$. We now try to estimate a value for s for a typical regional telephone company region. There are 160 LATA over

Dom. Term	ΔM	ΔF	Max CMR
L	$\frac{1.74(K-1)L}{K}$	$1.13(K-1)L$	$\frac{1.54}{K}$
D	$\frac{3.48(K-1)D}{K}$	$(1.74+0.26K$ $-4t)D$	see text
R	$\frac{3.74(K-1)R}{K}$	$(0.13K-0.13$ $-2t)R$	see text

Table 2: Evaluation of Forwarding With Simplifying Assumptions

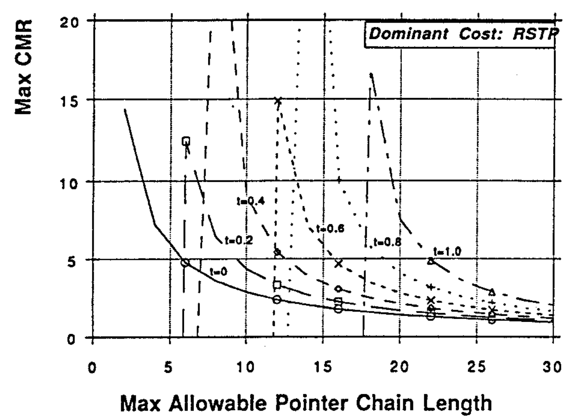

Figure 6: Max. CMR vs. length of pointer chain when RSTP cost dominates

seven regional telephone company regions, and typically there are 1250 SSP per region [4]. Assuming one LSTP per LATA, the number of SSPs per LSTP is $\frac{1250\times7}{160}$. Now, assuming each SSP corresponds to a single RA, $m = \sqrt{1250 \times 7/160} \approx 7.4$, and therefore $s \approx 0.87$. Using these estimates for j and s, we can investigate the cases when L, D or R is the dominant cost. (See Table 2).

In Fig. 6 we plot the maximum CMR for which forwarding is beneficial as t and K are varied when the RSTP is the dominant cost. (The plot for the case where the D-link is dominant is similar, since traversing a D-link implies traversing an RSTP; it is omitted due to lack of space.)

Let K_0 denote the value of K for which the curves in Figs. 6 cross the x-axis. Excluding values of t for which K_0 is not defined, we see that when $K < K_0$, the maximum CMR is negative. Since the benefit per *FwdMOVE* is never negative, this implies that forwarding is worthwhile regardless of a user's CMR in these cases. Informally this is because for small K, it is unlikely that either a *FwdMOVE* or a *FwdFIND* will traverse the RSTP, so that on average both operations result in a benefit.

As t decreases, i.e., the called party's location as known by the HLR gets further from the caller's RA, K_0 decreases. Informally, this is because as t decreases it is increasingly likely that a *FwdFIND* will leave the LSTP area, thus imposing traffic on the D-links and RSTP. For $K > K_0$, the maximum CMR decreases with K. This is because the average benefit to the RSTP per *FwdMOVE* increases with K, but relatively slowly; as K gets larger, the fraction of the benefit which is lost by using a *BasicMOVE* every Kth move becomes smaller. On the other hand, the average cost to the RSTP per *FwdFIND* increases linearly with K.

From Fig. 6 choosing $K \leq 5$ results in net benefits regardless of the user's CMR (except if $t = 0$). This is the case even if the user tends to receive calls from a fairly wide geographical area (t as low as 0.2). Fig. 6 also shows that, if the RSTP is the dominant cost, it is worthwhile to use forwarding if $CMR < 0.5$, regardless of the value of K chosen and assumptions about the user's call reception patterns.

6 Discussion

6.1 User terminal capabilities required

The *FwdMOVE* procedure requires that the user's equipment inform the new VLR of the number of RA crossings it has performed, which uses this to determine whether a basic move is required.[5] It also requires that the user's equipment inform the new VLR of the registration area from which the user just arrived. We stress three points related to this process:

1. The user terminal equipment must have the capability of storing this information. We observe that such a capability is implicitly or explicitly required in the terminal registration procedure specified in several proposed standards documents [3, 2, 18].

2. The user terminal equipment need not use SS7 address information for this purpose. Typically the radio port periodically broadcasts a registration area identifier which could be used.

3. Current air interface protocols do not support the relaying of this information to the network. Thus new protocol elements (new messages or message elements) would have to be designed to support this.

[5]Note that in IS-41 this information is provided to the VLR for security purposes anyway [5].

6.2 Alternative forwarding strategies

Numerous variations of forwarding can be envisioned. One possibility is that when a caller tries to locate a user, the local VLR is first queried to see if a forwarding pointer is available for this user; this would eliminate a query of the called user's HLR. A second possibility is that once a *FwdFIND* procedure locates the actual VLR of the called user, this information is cached at the caller's VLR, for use in subsequent call deliveries to the same user. A third possibility is to prevent loops from being created in the forwarding chain by ensuring that, during *FwdMOVE*, the VLR at the new RA is checked to see if an outgoing pointer already exists for this user; if so, it is deleted and no forwarding pointer from the old RA need be set up. Yet another possibility is that during *FwdFIND*, once the user's actual current location has been found, the information is returned directly to the calling switch, rather than retracing the pointer chain. We have investigated an alternative forwarding strategy in detail in [9].

6.3 Estimating the users' CMR

Here we sketch some methods of estimating the CMR. A simple and attractive policy is to not estimate CMR on a per-user basis at all. For instance, if it is known that the average CMR over all users in a PCS system is low enough, then forwarding could be used at every SSP to yield net system-wide benefits.

One possibility for deciding about forwarding on a per-user basis is to calculate running estimates of users' CMRs at each SSP. We have discussed two schemes for doing so in [10]. Another possibility is to maintain information about a user's calling and mobility pattern at the HLR. It is easy to envision numerous variations on this idea. In general, schemes for estimating the CMR range from static to dynamic, and distributed to centralized. Evaluating their effectiveness depends upon numerous factors whose discussion is outside the scope of this paper.

7 Conclusions

Depending upon the length of pointer chains allowed, and the cost of setting up a pointer relative to a basic *MOVE*, for users with $CMR < 1$ and stable call and move patterns, forwarding can result in cost reductions of 20-60% over the basic user location strategy, with little or no penalty in call setup time.

If the HLR or the remote A-link in an SS7 architecture is the performance bottleneck, forwarding is useful regardless of user call and mobility patterns. The reduced load on the HLR and the remote A-link is obtained in exchange for increased load on the VLRs and local A-link; if these elements are the bottleneck, forwarding should not be used. Thus the forwarding strategy serves to redistribute the network load to lower levels of the signalling network where excess capacity may be more likely to exist.

When the D-link, RSTP, and LSTP are the performance bottlenecks the benefits of forwarding typically depend upon user call and mobility patterns. For reasonably low CMRs ($CMR < 0.5$), forwarding can provide benefits regardless of user call and mobility patterns. If pointer chains are kept short ($K < 5$), forwarding is beneficial up to much higher ceilings on CMR. Thus it appears that when the D-link, RSTP or the LSTP are the bottlenecks, forwarding (with appropriate choice of K) could potentially provide benefits to large classes of users.

There are several issues deserving further study with respect to deployment of the forwarding strategy, such as the effect of alternative PCS architectures and integration with other auxiliary strategies (e.g. caching [10]). We have performed a detailed analysis of an alternative forwarding strategy in [9].

Acknowledgements. We thank numerous colleagues at Bellcore, in particular T. Ikuenobe, A. Knapp, S. Levenson and J. Rizzo, for many useful discussions, and the anonymous referees for their helpful comments.

References

[1] B. Awerbuch and D. Peleg. Concurrent online tracking of mobile users. In *Proc. SIGCOMM Symp. Comm. Arch. Prot.*, Oct. 1991.

[2] Bellcore. Generic criteria for Version 0.1 Wireless Access Communications Systems (WACS). Technical Advisory TA-NWT-001313, Issue 1, Bellcore, July 1992.

[3] Bellcore. Network and Operations Plan for access services to Personal Communications Services (PCS) providers. Special Report SR-TSV-002402, Bellcore, Aug. 1992.

[4] Bellcore. Switching system requirements for interexchange carrier interconnection using the integrated services digital network user part (ISDNUP). Technical Reference TR-NWT-000394, Bellcore, Dec. 1992.

[5] EIA/TIA. Cellular radiotelecommunications intersystem operations. Technical Report IS-41 (Revision B), EIA/TIA, July 1991.

[6] R. J. Fowler. The complexity of using forwarding address for decentralized object finding. In *Proc. SIGCOMM Symp. Comm. Arch. Prot.*, 1986.

[7] J. Ioannidis, D. Duchamp, and G. Q. Maguire Jr. IP-based protocols for mobile internetworking. In *Proc. SIGCOMM Symp. Comm. Arch. Prot.*, pages 234–245, 1991.

[8] Ravi Jain. A classification scheme for user location strategies in personal communications services systems. Aug. 1993. Submitted for publication.

[9] Ravi Jain and Yi-Bing Lin. An auxiliary user location strategy employing forwarding pointers to reduce network impacts of PCS. 1994. Submitted for publication.

[10] Ravi Jain, Yi-Bing Lin, Charles Lo, and Seshadri Mohan. A caching strategy to reduce network impacts of PCS. *IEEE J. Sel. Areas Comm.*, Oct. 1994.

[11] C. Lo, S. Mohan, and R. Wolff. A comparison of data management alternatives for personal communications applications. Technical Memorandum TM-ARH-022333, Bellcore, June 1993.

[12] C. N. Lo, R. S. Wolff, and R. C. Bernhardt. An estimate of network database transaction volume to support personal communications services. In *Proc. Intl. Conf. Univ. Pers. Comm.*, 1992.

[13] Charles Lo and Richard Wolff. Estimated network database transaction volume to support wireless personal data communications applications. In *Proc. Intl. Conf. Comm.*, May 1993.

[14] Edgar Lycksell. GSM system overview. Technical report, Swedish Telecom. Admin., Jan. 1991.

[15] K. Meier-Hellstern and E. Alonso. The use of SS7 and GSM to support high density personal communications. In *Proc. Intl. Conf. Comm.*, 1992.

[16] A. R. Modaressi and R. A. Skoog. Signalling System No. 7: A tutorial. *IEEE Comm. Mag.*, pages 19–35, July 1990.

[17] S. Mohan and R. Jain. Two user location strategies for PCS. *IEEE Pers. Comm. Mag.*, 1(1), First Quarter 1994. (Premiere issue).

[18] Michel Mouly and M. B. Pautet. *The GSM System for Mobile Communications.* M. Mouly, 49 rue Louise Bruneau, Palaiseau, France, 1992.