Comparing Soft and Hard Handoffs

Yi-Bing Lin, Senior Member, IEEE, and Ai-Chun Pang

Abstract—This paper studies the soft-handoff mechanism and compares its performance with hard handoff. Our study indicates that although a handset may potentially consume extra radio links in soft handoff, the mechanism provides better opportunity to transfer the link successfully in the handoff procedure. Thus, by carefully planning the overlay areas of cells, soft handoff can outperform hard handoff.

Index Terms—Hard handoff, personal communications services, radio channel allocation, soft handoff.

I. INTRODUCTION

In a mobile communication network, a handset communicates with the outside world through the radio contact to a base station (BS). When a call arrives at a cell (i.e., the coverage area of a BS), the destination (or the originating) handset is connected if a channel is available. Otherwise, the call is blocked (this is referred to as a new call blocking). When a communicating handset moves from one cell to another, the channel in the old BS is released and a channel is required in the new BS. This process is called handoff. In mobile systems such as AMPS [1], global system for mobile communication (GSM) without macrodiversity [2], DECT [3], D-AMPS [4], and PHS [5], hard handoff is employed [6], [7]. In hard handoff, the old radio link is broken before the new radio link is established, and a handset always communicates with one BS at any given time. In the handoff procedure, the network needs to set up the new voice path for the handoff call. This setup time is referred to as the network response time \( t_{\text{net}} \). If the old radio link is disconnected before the network completes the setup, the call is forced terminated. Thus, even if idle channels are available in the new cell, a handoff call may fail if the network response time \( t_{\text{net}} \) for link transfer is too long. Note that a handoff failure may not necessarily cause a call drop. It is normally some time-out mechanism for the voice or signaling path which leads to a dropped call.

Some code-division multiple-access (CDMA) systems [8] and GSM with macrodiversity [2] utilize soft handoff where a handset may communicate with the outside world using multiple radio links through different BS’s at the same time. During handoff, the signaling and voice information from multiple BS’s are typically combined (or bridged) at the mobile switching center [9]. Similarly, voice and signaling information must be sent to multiple BS’s, and the mobile station must combine the results. In some soft-handoff systems, a handset may connect up to three or four radio links at the same time.

Thus, within the overlay area of cells, a handset can connect to multiple BS’s. During the process of dropping a failing link, the handset may communicate using other radio links. Thus, link transfer is not sensitive to the elapsed link-transfer time. Note that the soft-handoff link-transfer procedure may not be faster than that for hard handoff. However, soft handoff is not time critical as compared with hard handoff [10].

On the other hand, soft handoff degrades channel availability because a handset may consume multiple radio channels. Thus, it is desirable to investigate the performance of soft handoff and the tradeoff between hard and soft handoffs. In this paper, we present analytic and simulation models to study soft handoff and compare its performance with hard handoff. To strengthen the theme of our study, we do not consider the handoff prioritized schemes [11] that are seldom implemented in the commercial systems.

II. ANALYTIC MODEL

With minor modifications to the two analytic models we developed in [12] and [11], we compare the performance for hard and soft handoffs. For the reader’s benefit, we reiterate the models in [12] (for hard handoff) and [11] (for soft handoff) with new notation and new interpretation.

A. The Hard-Handoff Model

Fig. 1 illustrates the timing diagram for the hard-handoff model. In this figure, \( t_i \) represents the time that a handset can receive the signal from cell \( i \) (i.e., the time that the handset resides in cell \( i \)). Since the cells may overlay, the handset will enter the overlay area \( i \) before it moves from cell \( i \) to cell \( i + 1 \). Let \( \tau_i \) be the overlay time. Then \( t_i \) can be expressed as \( \tau_i = x_i + \tau_i \), where \( x_i \) is the time that the handset stays in the nonoverlay area of cell \( i \). In hard handoff, a communicating handset is switched from cell \( i \) to cell \( i + 1 \) at some point within \( \tau_i \). In Fig. 1, the handoffs occur at time \( w_i \), \( i = 0, 1, 2, \ldots \). From the viewpoint of the hard-handoff scheme, the residence time of the handset at cell \( i \) is \( y_i = w_i - w_{i-1} \). Let \( x_i \) be the nonoverlay period. If \( E[x_i] = (1/\eta) \) and \( E[\tau_i] = (1/\gamma) \), then

\[
E[y_i] = E[x_i] + E[\tau_i] = \frac{1}{\eta} + \frac{\eta + \gamma}{\eta \gamma}.
\]  

Manuscript received January 31, 1998; revised April 8, 1999. This work was supported in part by the National Science Council, R.O.C., under Contracts NSC-87-2213-E-009-013 and NSC88-2213-E009-079.

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

Publisher Item Identifier S 0018-9545(00)03667-7.
Since the radio link between the BS and the handset is broken before it is connected in hard handoff, the link transfer may fail due to long network response time even if radio channels are available in the new BS. The following assumptions are used in the model.

- The call arrivals to/from a handset are a Poisson process.
- The net new call arrival rate to a cell is $\lambda_o$.
- The mobile residence time $y_i$ in a cell $i$ has an exponential distribution with the density function
  
  \[ f_m(y_i) = \theta e^{-\theta y_i}. \]

This assumption will be relaxed to accommodate general residence time distribution in Appendix A.

- The call holding time $t_c$ is exponentially distributed with the mean $1/\mu$.

The output measures are:

- $\lambda_b$ handoff call arrival rate to a cell;
- $p_o$ new call blocking probability;
- $p_r$ probability that a handoff call is blocked because no radio channel is available;
- $\alpha_h$ probability that a hard-handoff call is blocked because the network response time $t_nrt$ is too long;
- $p_f$ forced termination probability or the probability that a handoff call is blocked because no radio channel is available or because the network response time is too long;
- $p_{nc}$ call incompletion probability.

As mentioned before, a handoff call is forced terminated if the network response time is too long (with probability $\alpha_h$) or no channel is available (with probability $p_r$). Since a nonprioritized scheme is considered, $p_r = p_o$ and

\[ p_f = 1 - (1 - \alpha_h)(1 - p_o) = 1 - (1 - \alpha_h)(1 - p_r). \]  

From [12], we have

\[ \lambda_b = \frac{\theta(1 - p_o)\lambda_o}{\mu + \theta[1 - (1 - \alpha)(1 - p_o)]}. \]  

The channel occupancy time of a call in a cell is the minimum of the remaining call holding time (note that the call holding time for a handoff call has the same distribution as a new call because of the memoryless property of the exponential distribution) and the remaining cell residence time. Thus, the channel occupancy time is also exponentially distributed with rate $\mu + \theta$. The net traffic to the system is $\lambda_b + \lambda_h$. Let $c$ be the number of channels in a cell. The hard-handoff scheme can be modeled by an $M/M/c/c$ system and from the Erlang-B formula

\[ p_r = p_o = \left[ \frac{\lambda_o + \lambda_h \theta}{(\mu + \theta)(\mu + \theta + \gamma)} \right] \left( \sum_{i=1}^{c} \frac{(\lambda_o + \lambda_h \theta)^i}{(\mu + \theta)^i i!} \right)^{-1}. \]  

The probability $p_{nc}$ can be obtained by assigning an initial value for $\lambda_h$ and by iterating (4) and (3) until the $\lambda_h$ value converges. From [12], the call incompletion probability is derived as

\[ p_{nc} = p_o + \frac{\theta(1 - p_o)[1 - (1 - \alpha)(1 - p_o)]}{\mu + \theta[1 - (1 - \alpha)(1 - p_o)]}. \]  

From [12], we have

\[ \lambda_b = \frac{\theta(1 - p_o)\lambda_o}{\mu + \theta[1 - (1 - \alpha)(1 - p_o)]}. \]  

The probability $p_{nc}$ can be obtained by assigning an initial value for $\lambda_h$ and by iterating (4) and (3) until the $\lambda_h$ value converges. From [12], the call incompletion probability is derived as

\[ p_{nc} = p_o + \frac{\theta(1 - p_o)[1 - (1 - \alpha)(1 - p_o)]}{\mu + \theta[1 - (1 - \alpha)(1 - p_o)]}. \]  

To compute $p_0$ and $p_r$, the soft-handoff scheme can be modeled by a Markov process with states $s(n)$, where $n \geq 0$ represents the number of busy channels as in [11]. Fig. 3 illustrates the Markov process. When the process is in state $s(n)$ (for $0 \leq n < c$), $n$ channels are busy. The effective call traffic to a cell at $s(n)$ is $\lambda_b + \lambda_h$ [and the process moves from $s(n)$ to $s(n + 1)$ with this rate]. Since a busy channel is released with the rate $\mu + \omega$, the process moves from $s(n)$ to $s(n - 1)$ (for $0 < n \leq c$) with the rate $\eta(n)(\mu + \omega)$.

When the process is in $s(c + j)$, where $j \geq 0$, all channels are busy, and $j$ handoff calls are looking for the second links. When a call arrives at state $s(c + j)$, the call is dropped immediately if it is a new call. On the other hand, if the call is a handoff call, then it is trying to connect to the second link before it leaves the overlay area. Thus, the process moves from $s(c + j)$ to $s(c + j)$. The soft-handoff model.

B. The Soft-Handoff Model

For the demonstration purpose, we assume that a handset can connect up to two radio links in a CDMA system. Fig. 2 illustrates the timing diagram for the soft-handoff model. The notations $t_i$, $\tau_i$, and $x_i$ are the same as that in Fig. 1. In soft handoff, a communicating handset at cell $i$ utilizes one channel during the nonoverlay period $x_i$ and is looking for a second radio link from cell $i + 1$ during $\tau_i$. Suppose that the second link is found at time $t_i$, then the channel occupancy time of the handset at cell $i + 1$ is the minimum of $\tau_i + x_i$ (in Fig. 2) and the remaining call holding time. Assume that $t_i$ is exponentially distributed, then from the memoryless property, $\tau_i$ also has the same distribution as $t_i$, i.e., it is exponentially distributed with mean

\[ \frac{1}{\omega} = \frac{1}{\eta} + \frac{1}{\gamma} + \frac{1}{\gamma} = \frac{\gamma + 2\eta}{\gamma \eta} \]

\[(nonexponential t_i are considered in Appendix A). For a fixed period, the number of calls visited by a handset is independent of the handoff schemes and the moving rate of a handset in soft handoff is $\theta$ as expressed in (1). Let $\alpha_s$ be the probability that a soft-handoff call is blocked because the network response time is too long. Unlike the hard handoff, it is apparent that $p_f < p_o$ in this scheme. Following the same reasoning in the previous section, $p_f$, $\lambda_b$, and $p_{nc}$ for soft handoff are similar to (2), (3), and (5) and can be expressed as

\[ p_f = 1 - (1 - \alpha_s)(1 - p_r) \]

\[ \lambda_b = \frac{\theta(1 - p_o)\lambda_o}{\mu + \theta[1 - (1 - \alpha_s)(1 - p_r)]} \]  

\[ p_{nc} = p_o + \frac{\theta(1 - p_o)[1 - (1 - \alpha_s)(1 - p_r)]}{\mu + \theta[1 - (1 - \alpha_s)(1 - p_r)]}. \]

Fig. 2. The timing diagram for the soft-handoff model.
Fig. 3. The Markov chain.

The process leaves the system with the rate $j(\mu + \gamma)$, and the process moves from $s(c + j)$ to $s(c + j - 1)$ with rate $c(\mu + \omega) + j(\mu + \gamma)$ for $j > 0$.

Let $\pi_i$ be the steady-state probability for $s(i)$. Then

$$
\pi_i = \begin{cases}
\frac{(\lambda_0 + \lambda_1)^i}{2(\mu + \omega)^i} \pi_0, & i \leq c \\
\frac{(\lambda_0 + \lambda_1)^{i-c}}{c(\mu + \omega)^c} \prod_{1 \leq k \leq n-c} [c(\mu + \omega) + j(\mu + \gamma)] \pi_0, & i > c.
\end{cases}
$$

Since $\pi_0 + \pi_1 + \ldots + \pi_n + \ldots = 1$, we have

$$
\pi_0 = \sum_{n=c}^{\infty} \frac{(\lambda_0 + \lambda_1)^{n-c}}{c(\mu + \omega)^c} \prod_{1 \leq k \leq n-c} [c(\mu + \omega) + j(\mu + \gamma)] ^{-1}.
$$

Since a new call is blocked when the system is in state $s(n)$ (where $n \geq c$) at its arrival, the originating call blocking probability is

$$
p_0 = \sum_{n=c}^{\infty} \pi_n. \quad (8)
$$

Following the technique we developed in [11], the probability $p_0$ is derived as follows. Suppose that a handoff call $C_t$ arrives at time $t$ when the cell is in state $s(n)$ (where $n = c + j$) and the call leaves the overlay area at time $t + \tau$. Let $\tau_c$ be the remaining call holding time of $C_t$ at time $t$ (i.e., the cell will be completed at time $t + \tau_c$). From the memoryless property, $\tau_c$ has the same exponential distribution as $t_c$. Consider the $c + j$ outstanding calls that arrive at the cell earlier than $C_t$. Suppose that among these $c + j$ calls, the first call leaves the system (either completes, expires, or leaves the cell) at time $t + t_j$. Then the density function for $t_j$ is

$$
f_j(t_j) = [c(\mu + \omega) + j(\mu + \gamma)] e^{-[c(\mu + \omega) + j(\mu + \gamma)]t_j}. \quad (9)
$$

If $t_j < \tau$, then at time $t + t_j$, $C_t$ sees $c$ handsets in conversations and $j - 1$ handoff calls looking for the second links. Now consider the first call that leaves the system among these $c + j - 1$ calls (excluding $C_t$). Suppose that the call leaves the system at time $t + t_j + t_{j-1}$. Because of the memoryless property of the call occupancy distribution and the overlay time distribution, $t_{j-1}$ has the density function $f_{j-1}$ as expressed in (9). Let $T_j = t_0 + \ldots + t_j$. For a call $C_t$ arriving at state $s(n)$ ($n = c + j, j \geq 0$), the probability that $C_t$ is blocked is

$$
Pr\{\tau < T_j \text{ and } \tau > \tau_c | s(c + j)\}
$$

$$
= \int_{t_j=0}^{\infty} \ldots \int_{t_0=0}^{\infty} \int_{t_0+t_1+t_2+\ldots+t_j=0}^{\infty} \int_{\tau_c=0}^{\tau} \gamma e^{-\gamma \tau} e^{-\mu \tau} \prod_{k=1}^{j} f_k(t_k) \ d\tau_c \ d\tau \ d\tau_1 \ d\tau_2 \ldots \ d\tau_j
$$

$$
= \frac{(j+1)\gamma}{c(\mu + \omega) + (j+1)(\mu + \gamma)}.
$$

Thus, the probability $p_0$, (that no radio resource is available for a handoff call) is

$$
p_0 = \sum_{0 \leq j < \infty} Pr\{\tau < T_j \text{ and } \tau > \tau_c | s(c + j)\} \pi_{c+j}
$$

$$
= \sum_{0 \leq j < \infty} \frac{(j+1)\gamma \pi_{c+j}}{c(\mu + \omega) + (j+1)(\mu + \gamma)}. \quad (10)
$$

By using the same iterative procedure described in the previous sections, $\lambda_h, p_0, p_f$, and $p_{nc}$ can be obtained.

**C. Derivation for $\alpha_h$ and $\alpha_s$**

Suppose that $\tau_i$ and $t_{nrtt}$ are exponentially distributed with rates $\gamma$ and $\beta$, respectively (nonexponential distributions will be considered in Appendix A). In soft handoff, let $\tau_i^s$ be the period between the handset connects to the new cell and when the handset leaves the overlay area. Then from the Markov model in Fig. 3, $\tau_i$ and $\tau_i^s$ have the same exponential distribution. Thus, we have

$$
\alpha_s = Pr\{t_{nrtt} > \tau_i^s\}
$$

$$
= \int_{\tau_i^s=0}^{\infty} \int_{t_{nrtt}=0}^{\infty} \beta e^{-\beta t_{nrtt}} \gamma e^{-\gamma \tau_i^s} \ dt_{nrtt} \ d\tau_i^s
$$

$$
= \frac{\gamma}{\gamma + \beta}. \quad (11)
$$

**TABLE I**

<table>
<thead>
<tr>
<th>$\tau_c$</th>
<th>$0.05\mu$</th>
<th>$0.075\mu$</th>
<th>$0.10\mu$</th>
<th>$0.125\mu$</th>
<th>$0.15\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_0 = 3\mu$ (Simulation)</td>
<td>16.62%</td>
<td>16.63%</td>
<td>9.04%</td>
<td>7.26%</td>
<td>6.22%</td>
</tr>
<tr>
<td>$\lambda_0 = 4\mu$ (Simulation)</td>
<td>16.49%</td>
<td>11.67%</td>
<td>9.05%</td>
<td>7.26%</td>
<td>6.19%</td>
</tr>
<tr>
<td>$\lambda_0 = 6\mu$ (Simulation)</td>
<td>16.60%</td>
<td>11.62%</td>
<td>8.90%</td>
<td>7.33%</td>
<td>6.19%</td>
</tr>
<tr>
<td>Analysis</td>
<td>16.61%</td>
<td>11.76%</td>
<td>9.09%</td>
<td>7.41%</td>
<td>6.23%</td>
</tr>
</tbody>
</table>
Note that $\alpha_{hs}$ is independent of $\lambda_o$, $c$, and $\eta$. Table I lists the $\alpha_{hs}$ values obtained from simulation experiments (described in Appendix A) and the values computed from (11). The table indicates that (11) is consistent with the simulation experiments.

For hard handoff, the handoff procedure is initiated when the signal of the new link is better than the old link. Thus, we assume that $E[\tau_4] = 2\gamma + \beta^2$.

The analytic model is validated against a simulation model described in Appendix A.

Fig. 4 plots the $p_{hs}$ curves obtained from the analytic model (the dashed curves) and the simulation model (the solid curves). The figure indicates that the analytic and the simulation results are consistent.
III. NUMERICAL EXAMPLES

This section uses some numerical examples to illustrate the effects of the overlay time $\tau_i$, the mobility $\eta$, and the network response time $t_{\text{rt}}$ on output measures such as $p_{\text{st}}$, $p_{\text{f}}$, and $p_{\text{nc}}$.

1) The Effect of the Overlay Time: Fig. 5 illustrates the soft-handoff output measures $p_{\text{st}}$, $p_{\text{f}}$, and $p_{\text{nc}}$ as functions of the exponential overlay time $\tau_i$, where $\eta = 0.5/\mu$ and the network response time is exponentially distributed with mean $0.01/\mu$. Fig. 5(b) shows that $p_{\text{f}}$ decreases as $\tau_i$ increases (the longer the overlay time, the higher the probability that the second radio link is successfully connected to the handset). Fig. 5(a) shows that $p_{\text{st}}$ increases as $\tau_i$ increases (since handoff calls have better opportunity to obtain radio channels as the overlay time increases, the new call attempts are more likely to be blocked). Fig. 5(c) plots the $p_{\text{nc}}$ curves. We first note that $p_{\text{st}}$ [in Fig. 5(a)] and $p_{\text{f}}$ [in Fig. 5(b)] are two major factors [see (7)] that determine $p_{\text{nc}}$. For $\lambda_o \leq 5 \mu \lambda_o$, $p_{\text{st}}$ slightly increases as $\tau_i$ increases. On the other hand, $p_{\text{f}}$ significantly decreases as $\tau_i$ increases. Thus, $p_{\text{nc}}$ decreases as the overlay time increases. On the other hand, when the offered load is large (e.g., $\lambda_o = 6\mu$), $p_{\text{st}}$ significantly increases as $\tau_i$ increases. Since $p_{\text{f}}$ significantly decreases as $\tau_i$ increases, the net effect is that $p_{\text{nc}}$ decreases then increases as $\tau_i$ increases.

2) The Effect of the Nonoverlay Period: Fig. 6 plots $p_{\text{nc}}$ against the nonoverlay period $E[x_i] = 1/\eta$ where $\lambda_o = 6\mu$ and $E[t_{\text{rt}}] = 0.01/\mu$. This figure indicates that $p_{\text{nc}}$ is more sensitive to $\tau_i$ for large $\eta$ than small $\eta$. In other words, when the user mobility is large, the cell overlay area layout significantly affects the performance of soft handoff. Fig. 7 plots $p_{\text{nc}}$ against the variance $\text{Var}[x_i]$ of $x_i$ with the normal distribution. We observe that when $[x_i] < 1/\mu^2$, $p_{\text{nc}}$ is insensitive to the change of the variance of the nonoverlay period. On the other hand, when $\text{Var}[x_i] > 1/\mu^2$, $p_{\text{nc}}$ is very sensitive to $\text{Var}[x_i]$.

3) The Effect of Network Response Time: Fig. 8 plots $p_{\text{nc}}$ as a function of the network response time $t_{\text{rt}}$. In this figure, $\eta = \mu$ and $\lambda_o = 6\mu$. Fig. 8(a) shows the effect of the exponential $t_{\text{rt}}$ with various mean values. If the network response time is zero, then $p_{\text{nc}} = 1$ and handoff always fails. In this case, $p_{\text{nc}}$ increases as the overlay time increases. However, when $t_{\text{rt}}$ is nonzero, $p_{\text{nc}}$ decreases then increases as $\tau_i$ increases [this phenomenon was explained in Fig. 5(a)]. The effect of $t_{\text{rt}}$ on $p_{\text{nc}}$ is similar to the effect of $\tau_i$. That is, $p_{\text{nc}}$ is more sensitive to $\tau_i$ for large $t_{\text{rt}}$ than small $t_{\text{rt}}$. Fig. 8(b) demonstrates how the variance $\text{Var}[t_{\text{rt}}]$ of $t_{\text{rt}}$ with the normal distribution affects the system performance, where $\lambda_o = 6\mu$, $\eta = \mu$, and $E[t_{\text{rt}}] = 0.01/\mu$. The curves indicate that $p_{\text{nc}}$ decreases as $\text{Var}[t_{\text{rt}}]$ decreases. When $\text{Var}[t_{\text{rt}}] \geq 1$, the variance of the
distribution only has an insignificant effect on $p_{nc}$. On the other hand, when $\Var[\tau_{nrt}] < 1$, $p_{nc}$ significantly decreases as $\Var[\tau_{nrt}]$ decreases.

4) Soft Versus Hard Handoffs: Fig. 9 compares the call incompletion probabilities for soft handoff (the solid curves) and for hard handoff (the dashed curves). In this figure, (11) and (12) are used to compute $\alpha_s$ and $\alpha_h$, respectively. The figure indicates that in the ranges of the input parameters we considered, soft handoff outperforms hard handoff.

IV. Conclusion

This paper proposed an analytic model and a simulation model to study the performance of soft handoff. Our study indicated that the handoff network response time, the mobility of the user and the overlay time significantly affect the performance of soft handoff. Furthermore, we observed that the call incompletion probability can be significantly affected by the variances of the network response time and the nonoverlay time. Under the ranges of the input parameters we considered, soft handoff may significantly outperform hard handoff. Our study provides guidelines to determine the degree of the overlay among cells.

APPENDIX I

THE SIMULATION MODEL

This Appendix describes a simulation model to investigate the performance of CDMA soft handoff. In the simulation experiments, the PCS system consists of 64 BS's connected as an $8 \times 8$ wrapped mesh [13]. In the simulation model, a handset resides in the nonoverlay area of a cell for a period $\tau_{i}$, then moves to the overlay area of one of the four neighboring cells (selected with equal routing probabilities) for a period $\tau_2$, and finally moves to the nonoverlay area of the new cell. In the simulation, a cell $i$ is modeled as a cell object with data structure $C(i)$ to represent the number of idle channels at the cell.

The simulation model consists of four types of events. Let $\text{clock}$ be the system clock. The event types are described below.

- **ARRIVAL** event represents a new call arrival at a cell $i$. There are two cases.
  - $C(i) > 0$, In this case, $C(i)$ is decremented by one, and the call holding time $t_c$ and the nonoverlay time $x_i$ for this call are generated. If $t_c > x_i$, then generate a **COMPLETION** event with timestamp $\text{clock} + t_c$. Otherwise, generate an **OVERLAY** event with timestamp $\text{clock} + x_i$ (with the destination cell where the handset is moving, and the overlay period $\tau_2$).
  - $C(i) = 0$, The call is blocked. Update the call blocking statistics ($p_{bc}$ and $p_{nc}$).

The simulation generates the next **ARRIVAL** event according to the call arrival rate $\lambda_i$.

- **OVERLAY** event represents that a handset in conversation moves into the overlay area between the old cell $i$ and the new cell $j$.
  - $C(j) > 0$, the soft-handoff procedure is exercised. Let $t_c^i$ be the remaining call holding time and $\tau_{i}$ be the overlay period. Decrement $C(j)$ by one. If $t_c^i < \tau_{i}$, then generate a **COMPLETION** event with timestamp $\text{clock} + t_c^i$. Otherwise, generate the handoff network response time $\tau_{nrt}$. Generate a **RELEASE** event with timestamp $\text{clock} + \tau_{nrt}$.
  - $C(j) = 0$, The handset should continue to try until a radio channel in the new cell is available. This part can be implemented by generating another **OVERLAY** event if the handset is still in the overlay area at the next try time, or a **RELEASE** event if the next try occurs after the handset leaves the overlay area ($\tau_{nrt}$ is set $\infty$ to indicate soft failure due to shortage of radio channel).

- **RELEASE** event represents various situations described below.
  - $\tau_{nrt} = \infty$: The call is forced terminated because no radio channel is available. $C(i)$ is incremented by one, and the output statistics ($p_{tr}$ and $p_{nc}$) are updated.
  - $\tau_{i} < \tau_{nrt}$: The call is forced terminated because long network response time. $C(i)$ and $C(j)$ are incremented by one, and the output statistics ($\alpha_s$ and $\alpha_h$) are updated.
  - $\tau_{i} \geq \tau_{nrt}$: The second link is successfully connected. Let $t_c^j$ be the remaining call holding time. Generate the nonoverlay period $x_j$ at cell $j$. If $t_c^j < x_j$, then generate a **COMPLETION** event with timestamp $\text{clock} + t_c^j$. Otherwise, generate an **OVERLAY** event as described before.

- **COMPLETION** event represents the completion of a call at cell $i$. If the completion occurs at the nonoverlay area of cell $i$, then $C(i)$ is incremented by one. If the completion occurs at the overlay area between cell $i$ and cell $j$, then both $C(i)$ and $C(j)$ are incremented by one.

In the simulation experiments, the PCS system consists of 64 BS's connected as an $8 \times 8$ wrapped mesh [13]. The call arrivals to a cell form a Poisson process with arrival rate $\lambda_i$. The call holding times $t_c$ are exponentially distributed with mean $1/\mu$. The periods $x_i$, $\tau_{i}$, and $\tau_{nrt}$ are generated from exponential or
normal random number generators. In the experiments, every BS has ten channels. In each simulation experiment, 500,000 incoming calls are simulated to ensure that the simulation results are stable.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable comments.

REFERENCES


Yi-Bing Lin (S’80–M’96–SM’96) received the B.S.E.E. degree from National Cheng Kung University, Taiwan, R.O.C., in 1983 and the Ph.D. degree in computer science from the University of Washington, Seattle, in 1990. From 1990 to 1995, he was with the Applied Research Area at Bell Communications Research (Belcore), Morristown, NJ. In 1995, he was appointed Professor at the Department of Computer Science and Information Engineering (CSIE), National Chiao Tung University (NCTU), Hsinchu, Taiwan. In 1996, he was appointed Deputy Director of the Microelectronics and Information Systems Research Center, NCTU. Since 1997, he has been Chairman of the CSIE, NCTU. His current research interests include design and analysis of a personal communications services network, mobile computing, distributed simulation, and performance modeling. He was a Guest Editor for the IEEE TRANSACTIONS ON COMPUTERS Special Issue on Mobile Computing. He is an Associate Editor of IEEE NETWORKS.

Dr. Lin is 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 SIMULATION Magazine, an Area Editor of ACM Mobile Computing and Communication Review, a Columnist of ACM Simulation Digest, a Member of the Editorial Board of the International Journal of Communications Systems, a Member of the Editorial Board of ACM/Baltzer Wireless Networks, a Member of the Editorial Board of Computer Simulation Modeling and Analysis, and Guest Editor for the ACM/Baltzer MONET Special Issue on Personal Communications. He was the 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, and Publicity Chair of ACM Sigmobile.

Ai-Chun Pang received the B.S.C.S.I.E. and M.S.C.S.I.E. degrees from National Chiao Tung University (NCTU), Hsinchu, Taiwan, R.O.C., in 1996 and 1998, respectively. She is currently working toward the Ph.D. degree at NCTU.

Her current research interests include personal communications services, computer telephony integration, and mobile computing.