Equal resource sharing scheduling for PCS data services

Jeu-Yih Jeng [a,b] and Yi-Bing Lin [a]

[a] *Department of Computer Science and Information Engineering, National Chiao Tung University, Hsinchu, Taiwan, ROC*
[b] *Telecommunication Laboratories, Chunghwa Telecom Co., Ltd., Taiwan, ROC*

For high speed mobile communication applications, the data rate can be increased by using multiple channels (or time slots) instead of one channel. To reduce the high blocking rate of multiple channels assignment, flexible resource allocation strategies have been proposed. This paper proposes the Equal Resource Sharing Allocation scheme (ERSA scheme) for flexible resource allocation. The ERSA scheme dynamically averages the allocated resource to the call requests based on the number of calls in a base station. The scheme accommodates the maximum number of requests while providing acceptable quality to the admitted requests. We developed an analytic model to investigate the performance of ERSA, and conducted simulation experiments to validate the analytic model. We define satisfaction indication SI as the performance measurement of the resource allocation algorithm. The experiment results indicate that the ERSA scheme outperforms other resource allocation algorithms proposed in our previous study.

1. Introduction

Personal Communication Services (PCS) provide telecommunication services to mobile users. They will offer voice, video and data services, all in a single mobile terminal. Various types of applications lead to a number of network requirements including high demand for user bit rates (constant or variable bit rates), flexible transfer mode (connection-oriented or connectionless), and certain degree of quality-of-service (QoS). For high speed data applications, the data rate has been increased by using multiple channels (or time slots) instead of one channel. This approach is adopted in GSM HSCSD defined in ETSI GSM specification phase 2+. The major problem of this approach is that it may cause high blocking (that is, less customers can share the PCS if more radio resources are assigned to individual customers). One solution to this problem is to exercise flexible resource assignment where the customers may request resources with flexibility (i.e., to specify the maximum capacity and the minimum capacity for the request). Based on the current available capacity of a base station, a customer will be assigned any rate between the maximum and the minimum capacities. In [4], we have proposed four channel allocation algorithms for the base station: H1 (always allocates maximum), H2 (always allocates minimum), S1 (allocates maximum unless available resources are not enough), and S2 (allocates resources according to the current blocking statistics of the base station). In this paper we propose an effective scheme, named *Equal Resource Sharing Allocation* scheme (ERSA scheme), for flexible resource allocation. The proposed scheme dynamically "averages" the allocated resource to the call requests based on the number of calls in a base station. The intuition behind this concept is to accommodate the maximum requests while providing acceptable quality to the admitted requests. To investigate the performance of this scheme, an analytic model is proposed, and a simulation model is developed to validate the analytic model.

2. Analytic model

This section proposes an analytic model for the ERSA scheme. We first introduce the traffic model, and then describe the details of the ERSA algorithm. An output measure called the *Satisfaction Indication* SI is derived to investigate the performance for this scheme.

2.1. The traffic model

We assume that the call arrivals to a coverage area form a Poisson process. In figure 1, t_c is the *call holding time* of a portable, which is assumed to be exponentially distributed with the density function

$$f_c(t_c) = \mu e^{-\mu t_c},$$

where the mean call holding time is $E[t_c] = 1/\mu$. The *cell residence time* of a portable at a coverage area i (the intervals that a portable stays in coverage area i) is t_i. In this figure, t_0 is the time that the portable resides at coverage area 0, and t_i (where $i \geqslant 1$) is the residence time at coverage area i. We assume that $t_0, t_1, t_2, \ldots, t_k$ are independent and identically distributed random variables with an arbitrary non-lattice density function $f(\cdot)$ with the mean $1/\eta$. Let $f^*(s)$ be the Laplace transform of the cell residence time distribution. Then

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

Suppose that a call for the portable occurs when the portable is in coverage area 0. Let τ be the interval between when the call arrives and when the portable moves out of coverage area 0. Suppose that a call successfully hands over i times. Let $t_{c,i}$ be the period between the time when the portable moves into coverage i and the time when the call is completed. The period $t_{c,i}$ is called the *excess life* of t_c. When channels are assigned to a new call, the channels are released if the call completes or the portable moves

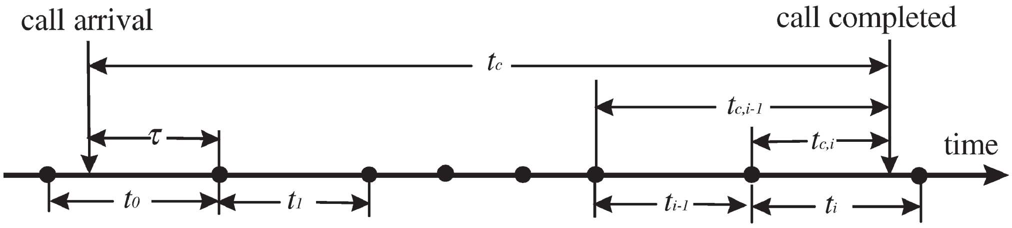


Figure 1. The timing diagram of a PCS portable.

out of the coverage area. Let t_{do} be the *dwell (channel occupation) time* of a new call. Then

$$t_{\text{do}} = \min(t_{\text{c}}, \tau).$$

The expected dwell time of a new call is derived in appendix B.1:

$$E[t_{\text{do}}] = \frac{1}{\mu} - \frac{\eta}{\mu^2}\left[1 - f^*(\mu)\right]. \tag{1}$$

Let t_{dh} be the dwell time of a hand-off call. If a call successfully hands over i times, then at coverage area i,

$$t_{\text{dh}} = \min(t_{\text{c},i}, t_i).$$

The expected dwell time of a hand-off call is derived in appendix B.1:

$$E[t_{\text{dh}}] = \frac{1}{\mu}\left[1 - f^*(\mu)\right]. \tag{2}$$

Let C be the total number of radio resources (e.g., time slots) of a base station (BS). Assume that the minimum capacity required to make a connection call is C_{min} and the maximum capacity is C_{max} for a request. Let λ_{o} and λ_{h} be the *new call arrival rate* and the *hand-off arrival rate* to a coverage area, respectively. Let p_{o} be the probability that a new call is rejected, p_{f} be the probability that a hand-off call is rejected. From appendix B.2, λ_{h} can be expressed as a function of λ_{o}, p_{o} and p_{f}:

$$\lambda_{\text{h}} = \frac{(1 - p_{\text{o}})\eta[1 - f^*(\mu)]\lambda_{\text{o}}}{\mu[1 - (1 - p_{\text{f}})f^*(\mu)]}. \tag{3}$$

The net traffic to a coverage area is

$$\rho = \lambda_{\text{o}}E[t_{\text{do}}] + \lambda_{\text{h}}E[t_{\text{dh}}]. \tag{4}$$

From (1)–(3), (4) can be expressed as

$$\rho = \frac{\lambda_{\text{o}}}{\mu}\left\{1 - \frac{\eta p_{\text{o}}[1 - f^*(\mu)]}{\mu[1 - (1 - p_{\text{f}})f^*(\mu)]}\right\}. \tag{5}$$

Since the ERSA scheme dynamically averages the allocated resource to the call requests based on the number of calls in a BS, most of calls in the BS are allocated the amount C_{min} as the traffic is high. Therefore, the scheme accommodate the maximum requests, and the maximum capacity of a BS is $\lfloor C/\ C_{\text{min}}\rfloor$. Since the blocking probability for an $M/G/c/c$ queue is the same as an $M/M/c/c$ queue with the same arrival process and same capacity [1], from the Erlang's formula,

$$p_{\text{o}} = B\left(\rho, \left\lfloor \frac{C}{C_{\text{min}}}\right\rfloor\right) = \frac{(\rho^{\lfloor C/C_{\text{min}}\rfloor}/\lfloor C/C_{\text{min}}\rfloor!)}{\sum_{i=0}^{\lfloor C/C_{\text{min}}\rfloor}(\rho^i/i!)}, \tag{6}$$

where $B(\rho, \lfloor C/C_{\text{min}}\rfloor)$ is the Erlang loss equation. In this paper, we assume that the hand-off calls and the new calls are indistinguishable and $p_{\text{f}} = p_{\text{o}}$. Therefore, the probability p_{o} can be obtained by assigning an initial value for λ_{h}, and then iterating (3) and (6) until the λ_{h} value converge.

2.2. The equal resource sharing allocation algorithm

In the ERSA scheme, a BS dynamically "averages" the allocated resource to the call requests based on the number of calls in the BS. When a call request arrives at a BS, if the BS has enough resources, then it allocates the maximum number of channels C_{max} to the request. On the other hand, if the total number of calls in the BS exceeds the maximum capacity ($= \lfloor C/C_{\text{min}}\rfloor$), then most of calls in the BS are allocated the amount C_{min}, and the request is blocked or forced terminated. Otherwise, the system allocates the resources evenly to all of the call requests in the BS. The intuition behind this concept is to dynamically adjust the amount of resource allocated to the call requests based on the number of calls in a BS in order to accommodate the maximum requests while providing acceptable quality to the admitted requests. The details of the algorithm are described as follows.

Algorithm ERSA. Let

$$J_i = \left\lfloor \frac{C}{C_{\text{max}} - i}\right\rfloor, \quad \text{where } 0 \leqslant i \leqslant C_{\text{max}} - C_{\text{min}}. \tag{7}$$

When a call request arrives at a BS, let j be the total number of calls served at the BS before the call arrives. One of the three cases is executed:

Case 1. If $1 \leqslant j < J_0$, the system allocates C_{max} channels to the request.

Case 2. For all $i \geqslant 1$ and $J_{i-1} < J_i$, if $J_{i-1} \leqslant j < J_i$, the system allocates $C_{\text{max}} - i + 1$ channels to the first $C - \lfloor C/j\rfloor j$ calls, and allocates $C_{\text{max}} - i$ channels to the other $j(1 + \lfloor C/j\rfloor) - C$ calls.

Case 3. If $j = J_{(C_{\text{max}} - C_{\text{min}})}$, then the system blocks the arrival request.

It is obvious that the probability that a call request is connected is $1 - p_o$. If the number of calls served at the BS is not greater than $\lfloor C/C_{max} \rfloor$, then all of the call requests are allocated C_{max} channels. On the other hand, if $\lfloor C/C_{max} \rfloor <$ the number of calls served at the BS $\leqslant$ the maximum capacity $\lfloor C/C_{min} \rfloor$, then all of the call requests evenly share the total C channels. Let C_a be the *expected number of channels* allocated to a connected call. Then C_a can be expressed as

$$C_a = \frac{1}{1-p_o}\left(\sum_{j=1}^{\lfloor C/C_{max} \rfloor} C_{max}\pi_j + \sum_{j=\lfloor C/C_{max} \rfloor+1}^{\lfloor C/C_{min} \rfloor} \frac{C}{j}\pi_j \right), \tag{8}$$

where π_j is the probability that there are j calls served at the BS, which is expressed as

$$\pi_j = \frac{(\rho^j/j!)}{\sum_{i=0}^{\lfloor C/C_{min} \rfloor}(\rho^i/i!)}. \tag{9}$$

The described ERSA algorithm assumes that the call requests are of the same type with the same C_{max} and C_{min}. In reality, there may be several job types with different C_{max} and C_{min} values. It is apparent that the ERSA described in this paper can be extended for multiple job types. To simplify our discussion, we focus on a single job type.

2.3. The satisfaction indication

In [3] we have defined satisfaction indication SI to measure the performance of the resource allocation. SI indicates the customer's satisfaction about the call connection, which can be used by the service provider as a charge indication. Thus, SI is expressed as

$$\begin{aligned} \mathrm{SI} &= \frac{p_i T_I \sigma + p_c T_C}{E[t_c]}\left(\frac{C_a}{C_{max}}\right) \\ &= \mu(p_i T_I \sigma + p_c T_C)\left(\frac{C_a}{C_{max}}\right), \end{aligned} \tag{10}$$

where p_c is the probability that a call request is completed, p_i is the probability that a call request is connected but is eventually force-terminated, T_I is the *expected effective call holding time* for an incomplete call, T_C is the expected effective call holding time for a complete call, σ ($0 \leqslant \sigma \leqslant 1$) is the *discount factor* for an incomplete call, and $E[t_c] = 1/\mu$ is the expected non-interrupted call holding time. Note that $p_i T_I \sigma + p_c T_C \leqslant 1/\mu$ and $C_a \leqslant C_{max}$, we have $0 \leqslant \mathrm{SI} \leqslant 1$. The intuition behind the discount factor is that if a call is incomplete, it is reasonable to charge the call with discount to increase the customer satisfaction. We derive the SI in appendix B.

In appendix B.3, p_c is derived as

$$p_c = (1-p_o)\left\{1 - \frac{\eta[1-f^*(\mu)]p_f}{\mu[1-(1-p_f)f^*(\mu)]}\right\}. \tag{11}$$

Since $p_o + p_c + p_i = 1$, p_i can be obtained by

$$p_i = 1 - p_o - p_c. \tag{12}$$

In appendix B.3, T_I is derived as

$$\begin{aligned} T_I = {} & \frac{\eta(1-p_o)p_f}{\mu(1-p_c-p_o)[1-f^*(\mu)(1-p_f)]}\left\{\frac{1-f^*(\mu)}{\mu}\right. \\ & \left. + \left[\left.\frac{\mathrm{d}f^*(s)}{\mathrm{d}s}\right|_{s=\mu}\right]\left[\frac{p_f}{1-f^*(\mu)(1-p_f)}\right]\right\} \end{aligned} \tag{13}$$

and T_C is derived as

$$\begin{aligned} T_C = {} & \left(\frac{1-p_o}{p_c}\right)\left\{\frac{1}{\mu} - \left(\frac{\eta}{\mu}\right)\right. \\ & \left. \times \left\{\frac{2[1-f^*(\mu)]}{\mu} + \left.\left[\frac{\mathrm{d}f^*(s)}{\mathrm{d}s}\right]\right|_{s=\mu}\right\}\right\} \\ & + \frac{\eta(1-p_o)[1-f^*(\mu)](1-p_f)}{\mu p_c[1-f^*(\mu)(1-p_f)]} \\ & \times \left\{\frac{2[1-f^*(\mu)]}{\mu} + \left.\frac{\mathrm{d}f^*(s)}{\mathrm{d}s}\right|_{s=\mu}\right. \\ & \left. + \left[\left.\frac{\mathrm{d}f^*(s)}{\mathrm{d}s}\right|_{s=\mu}\right]\left[\frac{p_f}{1-f^*(\mu)(1-p_f)}\right]\right\}. \end{aligned} \tag{14}$$

From (6), (8) and (11)–(14), the SI can be computed.

A simulation model is developed to validate the analytical model. The details of the simulation are described in appendix C. Figure 2 indicates that the analytic results are consistent with the simulation experiments. In all cases, call arrivals are represented by a Poisson process, the call holding times are exponentially distributed, and the cell residence times are Gamma distributed. The cell residence time merits further discussion. In a PCS network, the residence time of a portable at a cell is determined by the radio propagation/interference properties as well as the moving speed/direction of the portable. The residence time data can be directly measured in PCS field trials. To utilize the measured data in our model, the Gamma distribution is selected as a second approximation to the measured data. With Gamma distribution, we can match not only the mean and the standard deviation of the observed data but also match an observed skewness. The Laplace transform of the Gamma distribution is

$$f^*(s) = \left(\frac{\gamma\eta}{s+\gamma\eta}\right)^{\gamma}. \tag{15}$$

Equation (15) is used in (6) and (10)–(14) to compute the desirable output measures.

3. Results and discussions

This section discusses the performance of the ERSA scheme. We assume that every BS has $C = 50$ chan-

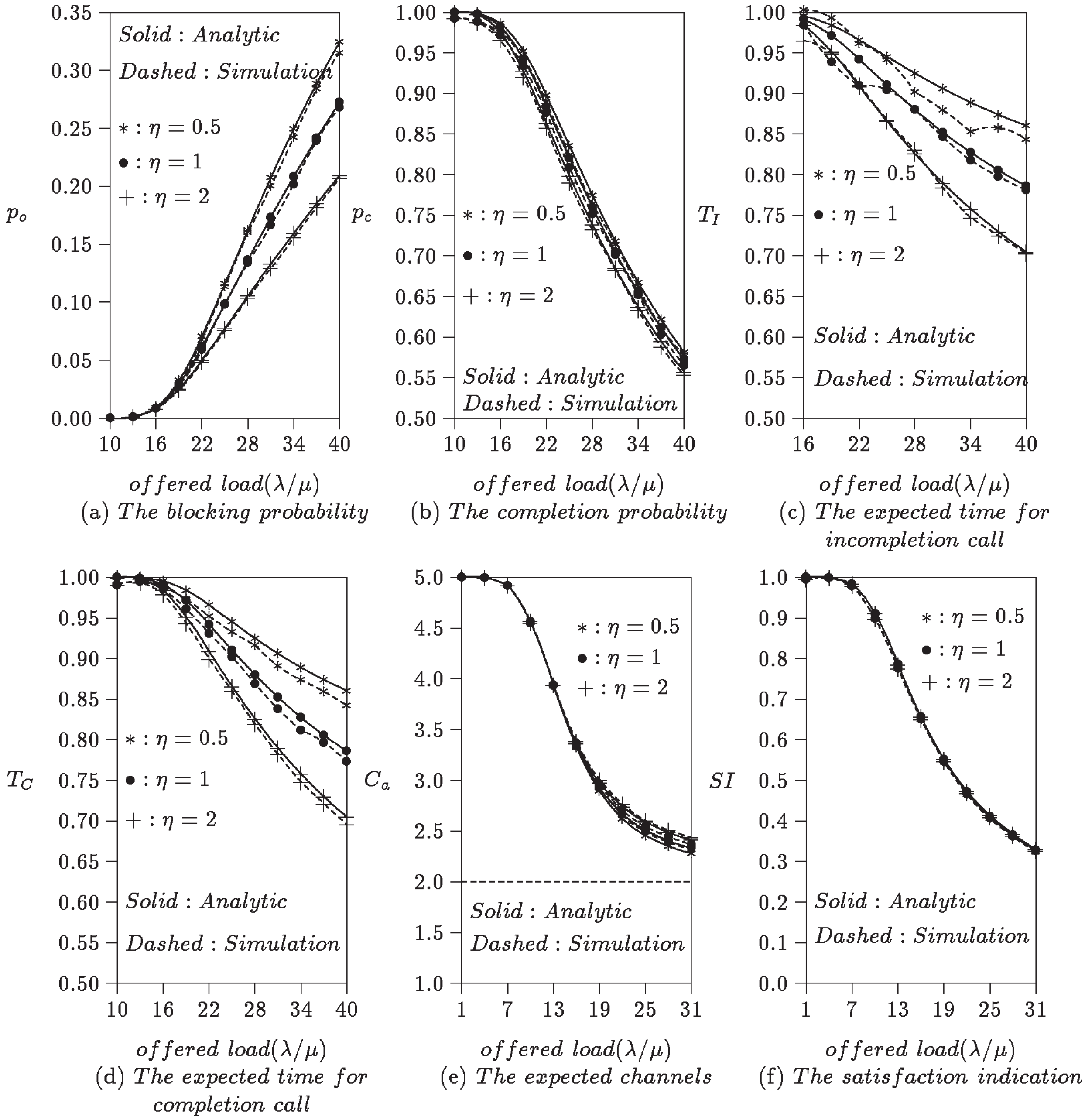


Figure 2. Effects of the mobility η ($\gamma = 1,\ \sigma = 1$).

nels. In most experiments, $C_{\max} = 5$ and $C_{\min} = 2$ are considered.

Effects of the mobility η. Figure 2 illustrates the performance measures effected by mobility η, where $\gamma = 1.0$, $\sigma = 1$ (no discount for incomplete calls), and the offered load ranges from 1 to 40. The figure indicates that p_o, p_c, T_I and T_C decrease as η increases. When η increases, the system experiences large hand-off arrivals with short channel occupancy times. Since a call will experience more hand-offs as η increases, the call is likely to be forced terminated. Thus p_c decreases as η increases (see figure 2(b)). These phenomena are also observed in our previous studies [3,4,8,9]. Figure 2(e) shows that C_a increases as η increases. The reason is that for a high mobility, the BSs will experience shorter channel occupancy times, and then have better chances to allocate more resources to individual customers. The effect becomes significant for heavy offered load. Figure 2(f) shows that the mobility η only has insignificant impact on SI. The SI measure of the ERSA scheme is stable for various mobility values when the discount factor is ignored. The effect of the discount factor will be discussed later.

Effects of the cell residence time distribution. Figure 3 plots the performance measures against γ, where $\eta = 1.0\mu$, $\sigma = 1$, and the offered load ranges from 1 to 40.

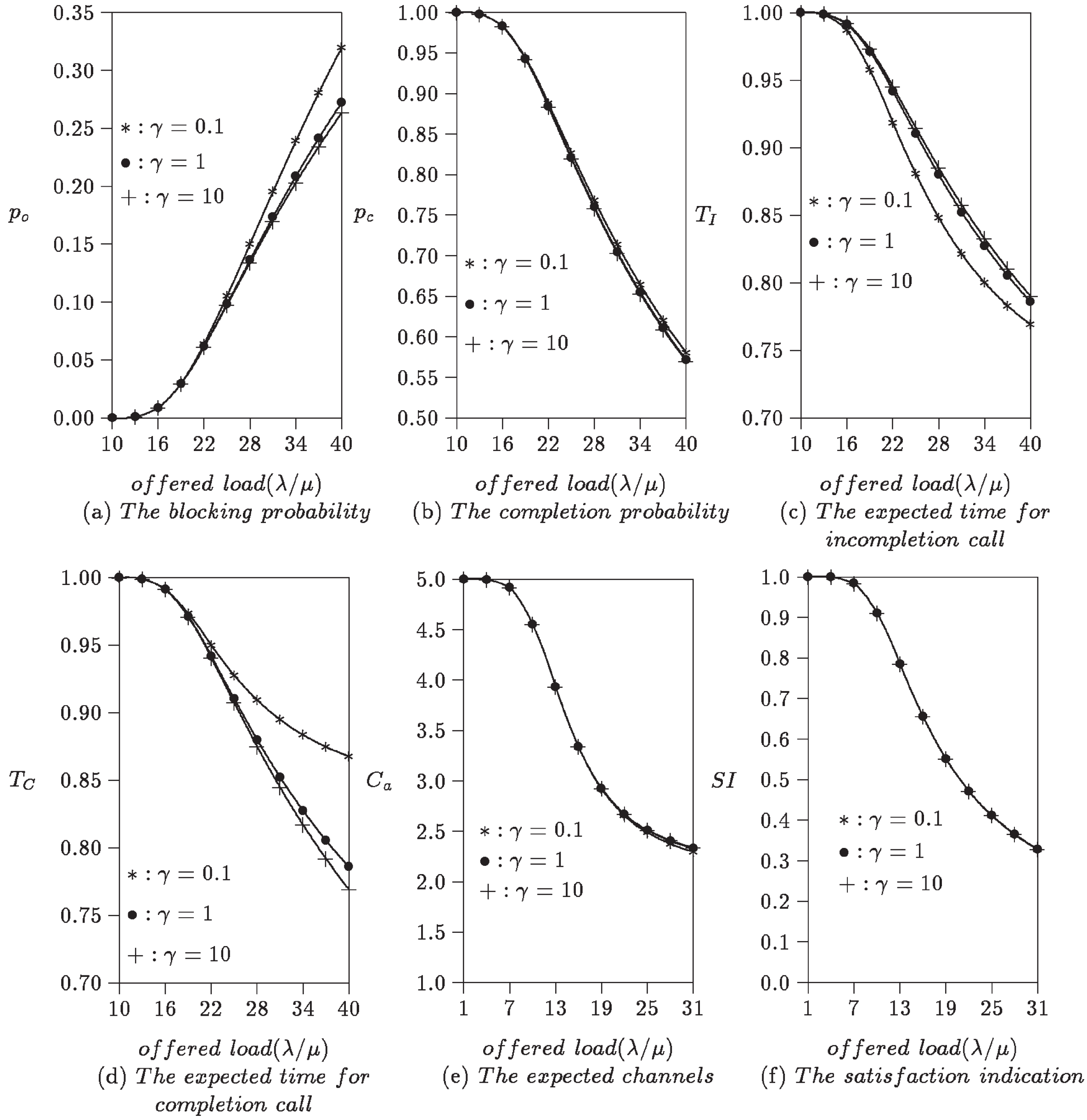


Figure 3. Effects of γ ($\eta = \mu,\ \sigma = 1$).

The figure indicates that p_o, p_c and T_C decrease as γ increases. On the other hand, T_I increases as γ increases. We also find that $T_I = T_C$ as $\gamma = 1$. However, the Gamma parameter only has insignificant impact on SI for all σ values considered. This result implies that the SI measure of the ERSA scheme is insensitive to the cell residence time distributions.

Effects of the discount factor σ. Based on (10), figure 4 shows that the SI decreases as σ decreases ($\gamma = 1.0$, the offered load is 25). For a low mobility value, the discount factor σ only has insignificant impact on SI. On the other hand, σ significantly affects SI when the mobility is high. This phenomenon is due to the fact that p_i ($= 1 - p_o - p_c$) increases as η increases (cf. figures 2(a) and (b)).

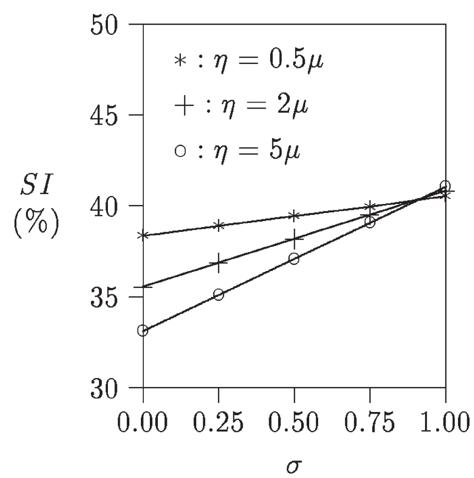


Figure 4. Effect of σ ($\gamma = 1,\ \lambda = 25\mu$).

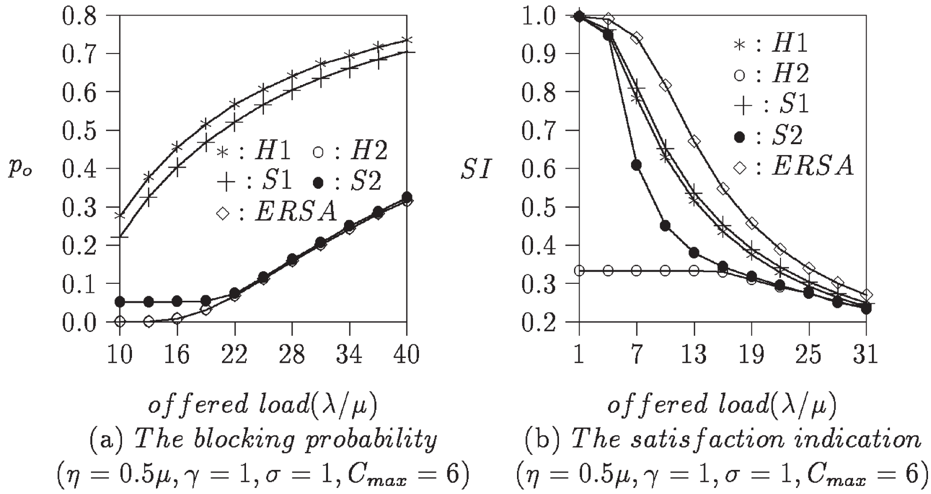


Figure 5. Comparisons among the resource allocation algorithms.

Comparisons with other resource allocation algorithms. Figure 5 compares the performance of the ERSA algorithm with the algorithms H1, H2, S1, and S2 (see section 1 for the descriptions of these algorithms), where $\eta = 0.5\mu$, $\gamma = 1.0$, $\sigma = 1.0$, $C_{\max} = 6$, and the offered load ranges from 1 to 31. The figure indicates that ERSA outperforms the other schemes both on p_o and SI. These phenomena are observed for all σ and η values.

4. Conclusion

Flexible resources allocation strategies can effectively reduce the high blocking rate of PCS data applications utilizing multiple channels (or time slots). In this paper, we propose the Equal Resource Sharing Allocation scheme (ERSA scheme) for flexible resource allocation. The ERSA scheme dynamically averages the allocated resource to the call requests based on the number of calls in a BS. The scheme accommodates the maximum number of requests while providing acceptable quality to the admitted requests. We developed an analytic model to investigate the performance of ERSA, and conducted simulation experiments to validate the analytic model. An output measure called the *satisfaction indication* SI is derived as the performance measurement of the resource allocation. Comparing with the algorithms proposed in our previous study, the ERSA outperforms other algorithms both on p_o and SI. We observed the following results.

- SI is stable for various mobility values if the discount factor is ignored.
- The distribution of the cell residence time has only insignificant impact on SI.
- For a low mobility, the discount factor σ only has insignificant impact on SI. On the other hand, σ significantly affects SI when mobility is high.

Appendix A. Notation

The notation used in this paper is listed below.

- λ_h: the hand-off call arrival rate to a coverage area.
- λ_o: the new call arrival rate to a coverage area.
- $1/\eta = E[t_i]$: the mean portable residence time.
- $1/\mu = E[t_c]$: the mean call holding time.
- τ: suppose that a call arrives when a portable is in coverage area 0; τ is the time between the arrival of the call and when the portable moves out of coverage area 0.
- C: the total number of radio resource (i.e., radio channel) of a base station.
- C_a: the expected number of channels allocated to a connected call.
- $C_{\min}$: the minimum number of channels required to connect a request.
- $C_{\max}$: the maximum number of channels for a request.
- $Cell$: a class used for simulation to record the current status of a base station.
- e: the next event deleted from the event list and processed during the simulation.
- $f_c(t_c)$: the exponential density function of a call holding time t_c.
- $f_{c,i}(t_{c,i})$: the density function of $t_{c,i}$.
- $f(t_i)$: the density function of t_i.
- $f^*(s)$: the Laplace transform of $f(t_i)$.
- $F(t_i)$: the distribution of t_i.
- N: the total number of call arrivals during simulation experiments.
- N_o: the number of blocked originating calls during simulation experiments.
- N_f: the number of forced termination calls during simulation experiments.

- N_c: the number of completion calls during simulation experiments.
- p_f: the probability that a handoff request is rejected at a cell.
- p_i: the probability that a call request is connected but is eventually force-terminated.
- p_c: the probability that a call request is completed.
- p_o: the probability that a new request is rejected at a cell.
- $R(\tau)$: the distribution of τ.
- $r(\tau)$: the density function of τ.
- $r^*(s)$: the Laplace transform of $r(\tau)$.
- SI: the satisfaction indication for a connected call (either complete or incomplete).
- t_dh: the dwell (channel occupation) time of a hand-off call.
- t_do: the dwell (channel occupation) time of a new call.
- t_c: the call holding time of a portable.
- t_i: the residence time of a portable at a coverage area i.
- $t_{\mathrm{c},i}$: Suppose that a call successfully hands over i coverage areas. Then $t_{\mathrm{c},i}$ is the period between the time when the portable moves into coverage area i. $t_{\mathrm{c},i}$ is called the *excess life* of t_c.
- T_I: the expected effective call holding time for an incomplete call.
- T_C: the expected effective call holding time for a complete call.
- $t_\mathrm{inc}[1 : N_\mathrm{f}]$: a data array used for simulation to record the call holding times for each incomplete calls.
- $t_\mathrm{com}[1 : N_\mathrm{c}]$: a data array used for simulation to record the call holding times for each completion calls.
- π_k: the probability that there are k calls served at a BS.
- σ: the discount factor for an incomplete call.
- ρ: the net traffic to a BS.

Appendix B. Analytic model

This appendix derives the hand-off traffic, the expected dwell times, and the effective call holding time. Consider figure 1. Assume that the random variable τ has a distribution function $R(\tau)$, the density function $r(\tau)$, and the Laplace transform $r^*(s)$. The function $r(t)$ can be derived using $f(t)$. Suppose that $f(t)$ is *non-lattice*. Since the call arrivals to a portable form a Poisson process, a call arrival is a random observer of the time interval t_0. From [10] we have

$$r(t) = \eta \int_{\tau=t}^{\infty} f(t)\,\mathrm{d}t = \eta\big[1 - F(t)\big] \tag{B.1}$$

and the Laplace transform of $r(t)$ is

$$\begin{aligned} r^*(s) &= \int_{t=0}^{\infty} \eta\big[1 - F(t)\big]\mathrm{e}^{-st}\,\mathrm{d}t \\ &= \eta\left[\int_{t=0}^{\infty} \mathrm{e}^{-st}\,\mathrm{d}t - \int_{t=0}^{\infty} F(t)\mathrm{e}^{-st}\,\mathrm{d}t\right] \\ &= \frac{\eta}{s}\big[1 - f^*(s)\big]. \end{aligned} \tag{B.2}$$

Equation (B.2) will be used in appendices B.1–B.3.

B.1. The expected dwell time

The expected values $E[t_\mathrm{do}]$ and $E[t_\mathrm{dh}]$ are derived as follows:

$$\begin{aligned} E[t_\mathrm{do}] &= E\big[\min(t_\mathrm{c}, \tau)\big] \\ &= \int_{t_\mathrm{c}=0}^{\infty}\int_{\tau=0}^{t_\mathrm{c}} \tau r(\tau) f_\mathrm{c}(t_\mathrm{c})\,\mathrm{d}\tau\,\mathrm{d}t_\mathrm{c} \\ &\quad + \int_{t_\mathrm{c}=0}^{\infty}\int_{\tau=t_\mathrm{c}}^{\infty} t_\mathrm{c} r(\tau) f_\mathrm{c}(t_\mathrm{c})\,\mathrm{d}\tau\,\mathrm{d}t_\mathrm{c} \\ &= A + B, \end{aligned} \tag{B.3}$$

where the first item of the right hand side of (B.3) is

$$\begin{aligned} A &= \int_{t_\mathrm{c}=0}^{\infty}\int_{\tau=0}^{t_\mathrm{c}} \tau r(\tau)\mu\mathrm{e}^{-\mu t_\mathrm{c}}\,\mathrm{d}\tau\,\mathrm{d}t_\mathrm{c} \\ &= \int_{t_\mathrm{c}=0}^{\infty}\left[\tau R(\tau)\Big|_{\tau=0}^{t_\mathrm{c}} - \int_{\tau=0}^{t_\mathrm{c}} R(\tau)\,\mathrm{d}\tau\right]\mu\mathrm{e}^{-\mu t_\mathrm{c}}\,\mathrm{d}t_\mathrm{c}. \end{aligned} \tag{B.4}$$

Let $R^{(2)}(t_\mathrm{c}) = \int_{\tau=0}^{t_\mathrm{c}} R(\tau)\,\mathrm{d}\tau$ then (B.4) is re-written as

$$\begin{aligned} A &= \int_{t_\mathrm{c}=0}^{\infty} t_\mathrm{c} R(t_\mathrm{c})\mu\mathrm{e}^{-\mu t_\mathrm{c}}\,\mathrm{d}t_\mathrm{c} \\ &\quad - \int_{t_\mathrm{c}=0}^{\infty} R^{(2)}(t_\mathrm{c})\mu\mathrm{e}^{-\mu t_\mathrm{c}}\,\mathrm{d}t_\mathrm{c} \qquad \text{(B.5)} \\ &= \mu\left\{(-1)\frac{\mathrm{d}}{\mathrm{d}s}\left[\frac{r^*(s)}{s}\right]\Bigg|_{s=\mu} - \frac{r^*(s)}{s^2}\Bigg|_{s=\mu}\right\}, \qquad \text{(B.6)} \end{aligned}$$

where (B.6) is derived from (B.5) and the following identity [12]:

$$\int_{t=0}^{\infty} t f(t)\mathrm{e}^{-st}\,\mathrm{d}t = (-1)\frac{\mathrm{d}f^*(s)}{\mathrm{d}s}. \tag{B.7}$$

The second item of the right hand side of (B.3) is

$$\begin{aligned} B &= \int_{t_\mathrm{c}=0}^{\infty}\int_{\tau=t_\mathrm{c}}^{\infty} t_\mathrm{c} r(\tau)\mu\mathrm{e}^{-\mu t_\mathrm{c}}\,\mathrm{d}\tau\,\mathrm{d}t_\mathrm{c} \\ &= \frac{1}{\mu} - \mu\left\{(-1)\frac{\mathrm{d}}{\mathrm{d}s}\left[\frac{r^*(s)}{s}\right]\Bigg|_{s=\mu}\right\}. \end{aligned} \tag{B.8}$$

From (B.6) and (B.8), (B.3) is re-written as

$$E[t_\mathrm{do}] = \frac{1}{\mu} - \mu\left[\frac{r^*(\mu)}{\mu^2}\right] = \frac{1}{\mu} - \frac{\eta}{\mu^2}\big[1 - f^*(\mu)\big].$$

Due to the memoryless property, the excess life of a call has the same (exponential) distribution as the original call holding time. In other words, for $i \geqslant 1$,

$$f_{c,i}(t) = f_c(t) = \mu e^{-\mu t_{c,i}}. \tag{B.9}$$

We have

$$\begin{aligned} E[t_{dh}] &= E\big[\min(t_{c,i}, t_i)\big] \\ &= \int_{t_{c,i}=0}^{\infty} \int_{t_i=0}^{t_{c,i}} t_i f(t_i) \mu e^{-\mu t_{c,i}} \, dt_i \, dt_{c,i} \\ &\quad + \int_{t_{c,i}=0}^{\infty} \int_{t_i=t_{c,i}}^{\infty} t_{c,i} f(t_i) \mu e^{-\mu t_{c,i}} \, dt_i \, dt_{c,i} \\ &= \frac{1}{\mu}\big[1 - f^*(\mu)\big]. \end{aligned}$$

B.2. The hand-off traffic

This subsection derives the hand-off call arrival rate λ_h at a coverage. Suppose that the new call arrival rate to the coverage area is λ_o. Let p_o be the new call blocking probability at the coverage area and p_f be the forced termination probability. Let $\Pr[t_c > \tau]$ ($\Pr[t_{c,i} > t_i]$) be the probability that a new (hand-off) call at the coverage area is not completed before the portable moves out of the coverage area. Then

$$\begin{aligned} \lambda_h = \lambda_h(1 - p_f) \Pr[t_{c,i} > t_{m,i}] \\ + \lambda_o(1 - p_o) \Pr[t_c > \tau_{m,0}]. \end{aligned} \tag{B.10}$$

The probabilities $\Pr[t_c > \tau_{m,0}]$ and $\Pr[t_{c,i} > t_{m,i}]$ are derived as follows:

$$\begin{aligned} \Pr[t_c > \tau] &= \int_{t_c=0}^{\infty} \int_{\tau=0}^{t_c} r(\tau) \mu e^{-\mu t_c} \, d\tau \, dt_c \\ &= \mu \int_{t_c=0}^{\infty} R(t_c) e^{-\mu t_c} \, dt_c && \text{(B.11)} \\ &= \mu \left[\frac{r^*(s)}{s}\right]\bigg|_{s=\mu} && \text{(B.12)} \\ &= r^*(\mu) && \text{(B.13)} \\ &= \frac{\eta}{\mu}\big[1 - f^*(\mu)\big]. && \text{(B.14)} \end{aligned}$$

Note that (B.12) is derived from (B.11), and (B.14) is derived from (B.13) and (B.2). From (B.9), $f_{c,i}(t) = \mu e^{-\mu t}$, and

$$\begin{aligned} \Pr[t_{c,i} > t_i] &= \int_{t_{c,i}=0}^{\infty} \int_{t_i=0}^{t_{c,i}} f(t_i) \mu e^{-\mu t_{c,i}} \, dt_i \, dt_{c,i} \\ &= \mu \left[\frac{f^*(s)}{s}\right]\bigg|_{s=\mu} \\ &= f^*(\mu). \end{aligned} \tag{B.15}$$

From (B.14) and (B.15), (B.10) is re-written as

$$\begin{aligned} \lambda_h &= \frac{(1 - p_o) \Pr[t_c > \tau] \lambda_o}{1 - (1 - p_f) \Pr[t_{c,i} > t_i]} \\ &= \frac{\eta(1 - p_o)[1 - f^*(\mu)]\lambda_o}{\mu[1 - (1 - p_f) f^*(\mu)]}. \end{aligned}$$

B.3. The effective call holding time

This subsection derives the call completion probability, the call incompletion probability, the effective call holding time for an incomplete call, and the effective call holding time for a complete call.

Suppose that the portable resides in coverage area 0 when a call arrived and the call will be is forced terminated at coverage area $k + 1$ (see figure 6(a)). Let $t = \tau + t_1 + t_2 + \cdots + t_k$. The density function $f_k(t)$ of t is expressed as

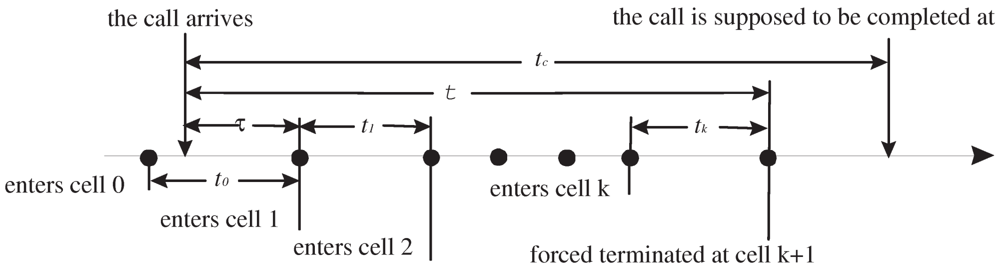
$$f_k(t) = \int_{\tau=0}^{t} \int_{t_1=0}^{t} \cdots \int_{t_{k-1}=0}^{t-t_1-\cdots-t_{k-2}} r(\tau) f(t_1) \cdots f(t_{k-1}) f(t - t_1 - \cdots - t_{k-1}) \, dt_1 \, dt_2 \, \ldots \, dt_{k-1}. \tag{B.16}$$

From the convolution property, the Laplace transform of $f_k(t)$ can be expressed as

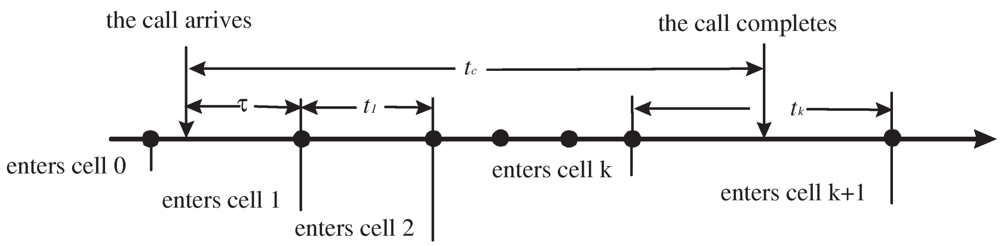
$$\begin{aligned} f_k^*(s) &= r^*(s)\big[f^*(s)\big]^k && \text{(B.17)} \\ &= \frac{\eta}{s}\big[1 - f^*(s)\big]\big[f^*(s)\big]^k. && \text{(B.18)} \end{aligned}$$

Let p_o be the *new call blocking probability* and p_f be the *forced termination probability*, or the probability that no radio channel is available when a handoff call arrives. Let p_c be the *call completion* probability or the probability that a call is completed. Let p_i be the *call incompletion* probability or the probability that a call is connected but is eventually forced terminated. Then it is apparent that

$$p_i + p_o + p_c = 1.$$



(a) The timing diagram for the case when a call is forced terminated



(b) The timing diagram for the case when a call is completed

Figure 6. The timing diagram for the effective call holding times.

Let $g_\mathrm{i}(t)$ be the density function for the effective call holding time t of an incomplete call that is forced terminated at the crossing of the $(k+1)$st cell. Since the non-interrupted call holding time t_c has the exponential distribution with mean $1/\mu$, from the definitions and (B.16) we have

$$g_\mathrm{i}(t) = \left(\frac{1}{p_\mathrm{i}}\right)\left[\sum_{k=0}^{\infty} f_k(t)(1-p_\mathrm{o})(1-p_\mathrm{f})^k p_\mathrm{f} \int_{t_\mathrm{c}=t}^{\infty} \mu \mathrm{e}^{-\mu t_\mathrm{c}}\,\mathrm{d}t_\mathrm{c}\right] \tag{B.19}$$

$$= \left(\frac{1-p_\mathrm{o}}{1-p_\mathrm{c}-p_\mathrm{o}}\right)\left[\sum_{k=0}^{\infty} f_k(t)(1-p_\mathrm{f})^k p_\mathrm{f} \mathrm{e}^{-\mu t}\right]. \tag{B.20}$$

From (B.20) and because

$$\int_{t=0}^{\infty} g_\mathrm{i}(t)\,\mathrm{d}t = 1$$

we have

$$(1-p_\mathrm{o})\int_{t=0}^{\infty}\sum_{k=0}^{\infty} f_k(t)(1-p_\mathrm{f})^k p_\mathrm{f}\mathrm{e}^{-\mu t}\,\mathrm{d}t = 1-p_\mathrm{c}-p_\mathrm{o}$$

$$\Rightarrow (1-p_\mathrm{o})\sum_{k=0}^{\infty}(1-p_\mathrm{f})^k p_\mathrm{f} r^*(\mu)\left[f^*(\mu)\right]^k = 1-p_\mathrm{c}-p_\mathrm{o}$$

$$\Rightarrow (1-p_\mathrm{o})r^*(\mu)p_\mathrm{f}\sum_{k=0}^{\infty}\left[(1-p_\mathrm{f})f^*(\mu)\right]^k = 1-p_\mathrm{c}-p_\mathrm{o}$$

$$\Rightarrow \frac{(1-p_\mathrm{o})r^*(\mu)p_\mathrm{f}}{1-(1-p_\mathrm{f})f^*(\mu)} = 1-p_\mathrm{c}-p_\mathrm{o}$$

$$\Rightarrow p_\mathrm{c} = (1-p_\mathrm{o})\left[1-\frac{r^*(\mu)p_\mathrm{f}}{1-(1-p_\mathrm{f})f^*(\mu)}\right]. \tag{B.21}$$

From (B.2) and (B.21), we have

$$p_\mathrm{c} = (1-p_\mathrm{o})\left\{1-\frac{\eta[1-f^*(\mu)]p_\mathrm{f}}{\mu[1-(1-p_\mathrm{f})f^*(\mu)]}\right\}. \tag{B.22}$$

The effective call holding time T_{I} for an incomplete call is

$$\begin{aligned}
T_{\mathrm{I}} &= \int_{t=0}^{\infty} t g_{\mathrm{i}}(t)\,\mathrm{d}t \\
&= \int_{t=0}^{\infty} t\left(\frac{1-p_{\mathrm{o}}}{1-p_{\mathrm{c}}-p_{\mathrm{o}}}\right)\left[\sum_{k=0}^{\infty} f_k(t)(1-p_{\mathrm{f}})^k p_{\mathrm{f}}\mathrm{e}^{-\mu t}\right]\mathrm{d}t \\
&= \left(\frac{1-p_{\mathrm{o}}}{1-p_{\mathrm{c}}-p_{\mathrm{o}}}\right)\left[\sum_{k=0}^{\infty}(1-p_{\mathrm{f}})^k p_{\mathrm{f}}\int_{t=0}^{\infty} t f_k(t)\mathrm{e}^{-\mu t}\,\mathrm{d}t\right] \\
&= \left[\frac{(1-p_{\mathrm{o}})p_{\mathrm{f}}}{1-p_{\mathrm{c}}-p_{\mathrm{o}}}\right]\left\{\sum_{k=0}^{\infty}(1-p_{\mathrm{f}})^k\left\{-\frac{\mathrm{d}}{\mathrm{d}s}\left\{r^*(s)\left[f^*(s)\right]^k\right\}\bigg|_{s=\mu}\right\}\right\}.
\end{aligned} \tag{B.23}$$

The derivative of the right hand side of (B.23) is re-written as

$$\begin{aligned}
&-\frac{\mathrm{d}}{\mathrm{d}s}\left\{r^*(s)\left[f^*(s)\right]^k\right\} \\
&= \frac{\eta}{s^2}\left[f^*(s)\right]^k - \frac{\eta}{s^2}\left[f^*(s)\right]^{k+1} - \begin{cases} -\left(\dfrac{\eta}{s}\right)\left[\dfrac{\mathrm{d}f^*(s)}{\mathrm{d}s}\right], & k=0, \\ \dfrac{\eta}{s}k\left[\dfrac{\mathrm{d}f^*(s)}{\mathrm{d}s}\right]\left[f^*(s)\right]^{k-1} - \left(\dfrac{\eta}{s}\right)(k+1)\left[\dfrac{\mathrm{d}f^*(s)}{\mathrm{d}s}\right]\left[f^*(s)\right]^k, & k>0 \end{cases} \\
&= A+B+C,
\end{aligned}$$

where

$$\begin{aligned}
A &= \frac{\eta}{s^2}\left[1-f^*(s)\right]\left[f^*(s)\right]^k, \\
B &= \frac{\eta}{s}k\left[\frac{\mathrm{d}f^*(s)}{\mathrm{d}s}\right]\left[f^*(s)\right]^{k-1}, \\
C &= -\left(\frac{\eta}{s}\right)(k+1)\left[\frac{\mathrm{d}f^*(s)}{\mathrm{d}s}\right]\left[f^*(s)\right]^k.
\end{aligned} \tag{B.24}$$

From (B.25), the summation term in the right hand side of (B.23) is re-written as

$$\begin{aligned}
\sum_{k=0}^{\infty} A(1-p_{\mathrm{f}})^k &= \sum_{k=0}^{\infty}\left(\frac{\eta}{s^2}\right)\left[1-f^*(s)\right]\left[f^*(s)(1-p_{\mathrm{f}})\right]^k \\
&= \frac{\eta[1-f^*(s)]}{s^2[1-f^*(s)(1-p_{\mathrm{f}})]},
\end{aligned} \tag{B.25}$$

$$\begin{aligned}
\sum_{k=0}^{\infty} B(1-p_{\mathrm{f}})^k &= \sum_{k=0}^{\infty}\left(-\frac{\eta}{s}\right)k\left[\frac{\mathrm{d}f^*(s)}{\mathrm{d}s}\right]\left[f^*(s)\right]^{k-1}(1-p_{\mathrm{f}})^k \\
&= \left[-\frac{\eta}{sf^*(s)}\right]\left[\frac{\mathrm{d}f^*(s)}{\mathrm{d}s}\right]\left\{\sum_{k=0}^{\infty} k\left[f^*(s)(1-p_{\mathrm{f}})\right]^k\right\} \\
&= -\left(\frac{\eta}{s}\right)\left[\frac{\mathrm{d}f^*(s)}{\mathrm{d}s}\right]\left\{\frac{1-p_{\mathrm{f}}}{[1-f^*(s)(1-p_{\mathrm{f}})]^2}\right\},
\end{aligned} \tag{B.26}$$

$$\begin{aligned}
\sum_{k=0}^{\infty} C(1-p_{\mathrm{f}})^k &= -\sum_{k=0}^{\infty}\left(\frac{\eta}{s}\right)(k+1)\left[\frac{\mathrm{d}f^*(s)}{\mathrm{d}s}\right]\left[f^*(s)(1-p_{\mathrm{f}})\right]^k \\
&= -\left(\frac{\eta}{s}\right)\left[\frac{\mathrm{d}f^*(s)}{\mathrm{d}s}\right]\frac{1}{f^*(s)(1-p_{\mathrm{f}})}\left\{\sum_{k=0}^{\infty} k\left[f^*(s)(1-p_{\mathrm{f}})\right]^k\right\} \\
&= -\left(\frac{\eta}{s}\right)\left[\frac{\mathrm{d}f^*(s)}{\mathrm{d}s}\right]\frac{1}{[1-f^*(s)(1-p_{\mathrm{f}})]^2}.
\end{aligned} \tag{B.27}$$

From (B.25)–(B.27), we have

$$\sum_{k=0}^{\infty}(A+B+C)(1-p_{\mathrm{f}})^k = \frac{\eta}{s[1-f^*(s)(1-p_{\mathrm{f}})]}\left\{\frac{1-f^*(s)}{s} - \left[\frac{\mathrm{d}f^*(s)}{\mathrm{d}s}\right]\left[\frac{p_{\mathrm{f}}}{1-f^*(s)(1-p_{\mathrm{f}})}\right]\right\}. \tag{B.28}$$

From (B.28), (B.23) is written as

$$T_{\mathrm{I}} = \frac{\eta(1-p_{\mathrm{o}})p_{\mathrm{f}}}{\mu(1-p_{\mathrm{c}}-p_{\mathrm{o}})[1-f^*(\mu)(1-p_{\mathrm{f}})]}\left\{\frac{1-f^*(\mu)}{\mu} + \left[\frac{\mathrm{d}f^*(s)}{\mathrm{d}s}\bigg|_{s=\mu}\right]\left[\frac{p_{\mathrm{f}}}{1-f^*(\mu)(1-p_{\mathrm{f}})}\right]\right\}. \tag{B.29}$$

Let $g_{\mathrm{c}}(t_{\mathrm{c}})$ be the density function for the effective call holding time t_{c} of a complete call. Figure 6(b) illustrates that a call is completed when the portable is in coverage area k. Then

$$\text{for } k = 0, \quad 0 \leqslant t_{\mathrm{c}} \leqslant \tau, \tag{B.30}$$

$$\text{for } k \geqslant 1, \quad \tau + t_1 + \cdots + t_k \leqslant t_{\mathrm{c}} \leqslant \tau + t_1 + \cdots + t_{k+1}. \tag{B.31}$$

From (B.30), (B.31) and by following an argument similar to that for the derivation of (B.20), we have

$$g_{\mathrm{c}}(t_{\mathrm{c}}) = U(t_{\mathrm{c}}) + W(t_{\mathrm{c}}), \tag{B.32}$$

where $U(t_{\mathrm{c}})$ corresponds to (B.30) and is expressed as

$$U(t_{\mathrm{c}}) = \left(\frac{1-p_{\mathrm{o}}}{p_{\mathrm{c}}}\right)\left[\int_{\tau=t_{\mathrm{c}}}^{\infty} r(\tau)\mu \mathrm{e}^{-\mu t_{\mathrm{c}}}\,\mathrm{d}\tau\right] \tag{B.33}$$

and $W(t_{\mathrm{c}})$ corresponds to (B.31) and is expressed as

$$W(t_{\mathrm{c}}) = \left(\frac{1-p_{\mathrm{o}}}{p_{\mathrm{c}}}\right)\left[\sum_{k=1}^{\infty}\int_{t=0}^{t_{\mathrm{c}}}\int_{\tau=t_{\mathrm{c}}-t}^{\infty} f_k(t)(1-p_{\mathrm{f}})^k r(\tau)\mu \mathrm{e}^{-\mu t_{\mathrm{c}}}\,\mathrm{d}\tau\,\mathrm{d}t\right]. \tag{B.34}$$

The expected effective call holding T_{C} for a complete call is

$$T_{\mathrm{C}} = \int_{t_{\mathrm{c}}=0}^{\infty} t_{\mathrm{c}} g_{\mathrm{c}}(t_{\mathrm{c}})\,\mathrm{d}t_{\mathrm{c}} = \int_{t_{\mathrm{c}}=0}^{\infty} t_{\mathrm{c}} U(t_{\mathrm{c}})\,\mathrm{d}t_{\mathrm{c}} + \int_{t_{\mathrm{c}}=0}^{\infty} t_{\mathrm{c}} W(t_{\mathrm{c}})\,\mathrm{d}t_{\mathrm{c}}. \tag{B.35}$$

From (B.33) we have

$$\begin{aligned}
\int_{t_{\mathrm{c}}=0}^{\infty} t_{\mathrm{c}} U(t_{\mathrm{c}})\,\mathrm{d}t_{\mathrm{c}} &= \int_{t_{\mathrm{c}}=0}^{\infty} t_{\mathrm{c}}\left(\frac{1-p_{\mathrm{o}}}{p_{\mathrm{c}}}\right)\left[\int_{\tau=t_{\mathrm{c}}}^{\infty} r(\tau)\mu \mathrm{e}^{-\mu t_{\mathrm{c}}}\mathrm{d}\tau\right]\mathrm{d}t_{\mathrm{c}} \\
&= \left(\frac{1-p_{\mathrm{o}}}{p_{\mathrm{c}}}\right)\left\{\frac{1}{\mu} - \mu\int_{t_{\mathrm{c}}=0}^{\infty}\left[\int_{\tau=0}^{t_{\mathrm{c}}} t_{\mathrm{c}} r(\tau)\,\mathrm{d}\tau\right]\mathrm{e}^{-\mu t_{\mathrm{c}}}\,\mathrm{d}t_{\mathrm{c}}\right\} \\
&= \left(\frac{1-p_{\mathrm{o}}}{p_{\mathrm{c}}}\right)\left\{\frac{1}{\mu} + \frac{\mathrm{d}}{\mathrm{d}s}\left\{\frac{\eta\mu[1-f^*(s)]}{s^2}\right\}\bigg|_{s=\mu}\right\} \\
&= \left(\frac{1-p_{\mathrm{o}}}{p_{\mathrm{c}}}\right)\left\{\frac{1}{\mu} + \left\{(-2)\frac{\eta\mu[1-f^*(s)]}{s^3} - \left(\frac{\eta\mu}{s^2}\right)\left[\frac{\mathrm{d}f^*(s)}{\mathrm{d}s}\right]\right\}\bigg|_{s=\mu}\right\} \\
&= \left(\frac{1-p_{\mathrm{o}}}{p_{\mathrm{c}}}\right)\left\{\frac{1}{\mu} - \left(\frac{\eta}{\mu}\right)\left\{\frac{2[1-f^*(\mu)]}{\mu} + \left[\frac{\mathrm{d}f^*(s)}{\mathrm{d}s}\right]\bigg|_{s=\mu}\right\}\right\}.
\end{aligned} \tag{B.36}$$

From (B.34) we have

$$\begin{aligned}
\int_{t_{\mathrm{c}}=0}^{\infty} t_{\mathrm{c}} W(t_{\mathrm{c}})\,\mathrm{d}t_{\mathrm{c}} &= \int_{t_{\mathrm{c}}=0}^{\infty}\left(\frac{1-p_{\mathrm{o}}}{p_{\mathrm{c}}}\right)\sum_{k=1}^{\infty}\int_{t=0}^{t_{\mathrm{c}}}\int_{\tau=t_{\mathrm{c}}-t}^{\infty} f_k(t)(1-p_{\mathrm{f}})^k r(\tau) t_{\mathrm{c}}\mu \mathrm{e}^{-\mu t_{\mathrm{c}}}\,\mathrm{d}\tau\,\mathrm{d}t\,\mathrm{d}t_{\mathrm{c}} \\
&= \left(\frac{1-p_{\mathrm{o}}}{p_{\mathrm{c}}}\right)\sum_{k=1}^{\infty}\int_{t=0}^{\infty}\int_{\tau=0}^{\infty}\int_{t_{\mathrm{c}}=0}^{t+\tau} f_k(t)(1-p_{\mathrm{f}})^k r(\tau) t_{\mathrm{c}}\mu \mathrm{e}^{-\mu t_{\mathrm{c}}}\,\mathrm{d}t_{\mathrm{c}}\,\mathrm{d}\tau\,\mathrm{d}t \\
&= \left(\frac{1-p_{\mathrm{o}}}{p_{\mathrm{c}}}\right)\sum_{k=1}^{\infty}\int_{t=0}^{\infty}\int_{\tau=0}^{\infty} f_k(t)(1-p_{\mathrm{f}})^k r(\tau) \\
&\quad\times\left[\left(t\mathrm{e}^{-\mu t} + \frac{\mathrm{e}^{-\mu t}}{\mu}\right) - \left(t\mathrm{e}^{-\mu t} + \frac{\mathrm{e}^{-\mu t}}{\mu}\right)\mathrm{e}^{-\mu\tau} - \mathrm{e}^{-\mu t}\tau \mathrm{e}^{-\mu\tau}\right]\mathrm{d}\tau\,\mathrm{d}t \\
&= \left(\frac{1-p_{\mathrm{o}}}{p_{\mathrm{c}}}\right)\sum_{k=1}^{\infty}\int_{t=0}^{\infty} f_k(t)(1-p_{\mathrm{f}})^k
\end{aligned}$$

$$\times \left\{ \left[1 - f^*(\mu)\right] t\mathrm{e}^{-\mu t} + \left[\frac{1}{\mu} - \frac{f^*(\mu)}{\mu} + \frac{\mathrm{d}f^*(s)}{\mathrm{d}s}\bigg|_{s=\mu}\right] \mathrm{e}^{-\mu t} \right\} \mathrm{d}t$$

$$= \left(\frac{1-p_\mathrm{o}}{p_\mathrm{c}}\right) \sum_{k=1}^{\infty} (1-p_\mathrm{f})^k (X+Y), \tag{B.37}$$

where

$$X = \left[\frac{1}{\mu} - \frac{f^*(\mu)}{\mu} + \frac{\mathrm{d}f^*(s)}{\mathrm{d}s}\bigg|_{s=\mu}\right] f_k^*(\mu), \tag{B.38}$$

$$Y = \left[1 - f^*(\mu)\right] \left[-\frac{\mathrm{d}f_k^*(s)}{\mathrm{d}s}\bigg|_{s=\mu}\right]. \tag{B.39}$$

From (B.38) and (B.39), we have

$$\sum_{k=1}^{\infty} (1-p_\mathrm{f})^k X = \frac{\eta[1-f^*(\mu)](1-p_\mathrm{f})}{\mu[1-f^*(\mu)(1-p_\mathrm{f})]} \left[\frac{1}{\mu} - \frac{f^*(\mu)}{\mu} + \frac{\mathrm{d}f^*(s)}{\mathrm{d}s}\bigg|_{s=\mu}\right], \tag{B.40}$$

$$\sum_{k=1}^{\infty} (1-p_\mathrm{f})^k Y = \frac{\eta[1-f^*(\mu)](1-p_\mathrm{f})}{\mu[1-f^*(\mu)(1-p_\mathrm{f})]} \left\{ \left[\frac{\mathrm{d}f^*(s)}{\mathrm{d}s}\bigg|_{s=\mu}\right] \left[\frac{p_\mathrm{f}}{1-f^*(\mu)(1-p_\mathrm{f})}\right] + \frac{1-f^*(\mu)}{\mu} \right\}. \tag{B.41}$$

From (B.36), (B.40) and (B.41), (B.35) is expressed as

$$T_\mathrm{C} = \left(\frac{1-p_\mathrm{o}}{p_\mathrm{c}}\right) \left\{ \frac{1}{\mu} - \left(\frac{\eta}{\mu}\right) \left\{ \frac{2[1-f^*(\mu)]}{\mu} + \left[\frac{\mathrm{d}f^*(s)}{\mathrm{d}s}\right]\bigg|_{s=\mu} \right\} \right\}$$
$$+ \frac{\eta(1-p_\mathrm{o})[1-f^*(\mu)](1-p_\mathrm{f})}{\mu p_\mathrm{c}[1-f^*(\mu)(1-p_\mathrm{f})]} \left\{ \frac{2[1-f^*(\mu)]}{\mu} + \frac{\mathrm{d}f^*(s)}{\mathrm{d}s}\bigg|_{s=\mu} + \left[\frac{\mathrm{d}f^*(s)}{\mathrm{d}s}\bigg|_{s=\mu}\right] \left[\frac{p_\mathrm{f}}{1-f^*(\mu)(1-p_\mathrm{f})}\right] \right\}.$$

Appendix C. Simulation model

To simulate a very large PCS network, we use a wrap-around topology for simulation [7]. This approach eliminates the boundary effect occurs in an unwrapped topology. An 8×8 wrapped mesh topology is considered in this paper. The mobility behavior of users in the simulation is described by a two-dimensional random walk proposed in [5]. In this model, the call arrivals to each BS form a Poisson process with arrival rate λ. The call holding times are exponentially distributed with mean $1/\mu$. A mobile user stays in the coverage areas of a period of time that has a Gamma distribution with mean $1/\eta$. Then, the user moves to one of the four neighboring cells with the same routing probabilities 0.25. If the user moves from a cell to another cell before the call is completed, a handover is performed. The user first requests resources at new BS, and releases the resources of the old BS. The number of resources assigned to the request may be the same or different after the handover process.

We develop a discrete event simulation model for the ERSA scheme. Each event contains the following attributes:

Type attribute indicates the type of event. There are three types of events in the simulation: the ARRIVAL events represent call arrivals, the COMPLETION events represent call completions, and the HANDOFF events represent the calls which hand over from a BS to another BS.

Timestamp attribute indicates the timestamp of the event.

Residual_time is used in a HANDOFF event to indicate the residual call holding time of the call represented by the event.

Occupied_time is used in a COMPLETION or a HANDOFF event to indicate the total call connection time of the call (represented by the event) before the event been processed, where the occupied_time plus the residual_time is equal to the call holding time of the call.

New_cell specifies the BS where the event occurs.

Old_cell is used in a HANDOFF event to specify the BS of the call (represented by the event) before hand-off occurs.

The events are inserted into an event list, and are deleted (and processed) from the event list in the non-decreasing timestamp order. A simulation clock is maintained to indicate the progress of the simulation. In other words, the clock value is the timestamp of the event being processed. The output measures of the simulation are the total number of call arrivals (N), the number of blocked originating call (N_o), the number of forced termination (N_f), and the number of completion calls (N_c). In every experiment, 100,000 incoming calls are simulated to ensure that the simulation results are stable. The above output measures are used to compute p_o, p_c, t_i, t_c, C_a and SI.

To simulate the behavior of each BS, a data structure or class $cell$ is introduced to record the current status of the BS. Each $cell$ consists of the following attributes:

N_{call} indicates the number of outstanding calls (the number of calls in progress) in the base station. Thus, $0 \leqslant N_{\text{call}} \leqslant \lfloor C/C_{\min} \rfloor$.

L_t indicates the time that last event (either ARRIVAL, COMPLETION, or HANDOFF) occurred in the BS.

$X[k]$ ($0 \leqslant k \leqslant \lfloor C/C_{\min} \rfloor$) indicates the cumulative occupation time that there are k calls in progress in the BS.

Two data arrays $t_{\text{inc}}[i]$ and $t_{\text{com}}[j]$ are used to record the call holding times for each incomplete call and complete call, respectively, where i ($1 \leqslant i \leqslant N_{\text{f}}$) indicates the ith incomplete call and j ($1 \leqslant j \leqslant N_{\text{c}}$) indicates the jth complete call. The details of the simulation are given in figure 7.

Steps (1) and (2) initialize the simulation and generate an ARRIVAL event for each BS. The event generation is completed in the following steps:

(1) Allocate the storage for the event.

(2) Determine the type (either ARRIVAL, COMPLETION, or HANDOFF) of the event.

(3) Determine the timestamp of the event. If the event type is ARRIVAL, then the timestamp is the clock value plus the inter-arrival time (drawn from an exponential distribution with mean $1/\lambda$). If the event is COMPLETION, then the timestamp is the clock value plus the channels occupancy time (drawn from exponential distribution with mean $1/\mu$). If the event is HANDOFF, then the timestamp is the clock value plus the cell residence time (drawn from Gamma distribution with mean $1/\eta$).

(4) Determine other attributes according to the type of event been processed below.

(5) Insert the event into the event list.

The next event e is deleted from the event list, and is processed based on its type (cf. steps (3) and (4)). For an ARRIVAL event, if $N = 100,000$ (cf. step (5)) then the simulation terminates, and the performance measures are calculated (cf. step (6)) as follows:

- $p_{\text{o}} = N_{\text{o}}/N$,
- $p_{\text{f}} = N_{\text{f}}/N_{\text{h}}$,
- $p_{\text{c}} = N_{\text{c}}/N$,
- $p_{\text{i}} = 1 - p_{\text{c}} - p_{\text{o}}$,
- $T_{\text{I}} = \sum_{k=1}^{N_{\text{f}}} t_{\text{inc}}[k]/N_{\text{f}}$,
- $T_{\text{C}} = \sum_{k=1}^{N_{\text{c}}} t_{\text{com}}[k]/N_{\text{c}}$,
-

$$C_{\text{a}} = \frac{1}{1-p_{\text{o}}}\left(\sum_{k=1}^{\lfloor C/C_{\max}\rfloor} \pi_k * C_{\max} + \sum_{k=C_{\max}+1}^{\lfloor C/C_{\min}\rfloor} \pi_k * \left(\frac{C}{k}\right)\right),$$

where π_k is the probability that there are k calls served at a BS, and can be obtained by

$$\pi_k = \sum_{i=1}^{8}\sum_{j=1}^{8} \mathrm{X}_{i,j}[k] \bigg/ \left(\sum_{i=1}^{8}\sum_{j=1}^{8}\sum_{l=0}^{\lfloor C/C_{\min}\rfloor} \mathrm{X}_{i,j}[l]\right),$$

- $\mathrm{SI} = \mu(p_{\text{i}} * T_{\text{I}} * \sigma + p_{\text{c}} * T_{\text{C}})(C_{\text{a}}/C_{\max})$.

If $N < 100,000$ at step (5), then the next ARRIVAL event is generated in the BS specified by $e->new\_cell$ (step (6)). If the number of calls N_{call} in the BS is less than the maximum capacity $\lfloor C/C_{\min} \rfloor$ (cf. step (7)), then the call arrival is connected. The cumulative occupation time $X[N_{\text{call}}]$ of the BS is incremented by the difference of the clock value and L_t, L_t is updated to the clock value, and N_{call} is incremented by 1 (cf. step (8)). Steps (9) and (10) decide whether the call arrival is completed in this BS or handoff to another BS. If the call holding time t_{c} (generated by an exponential distribution random number generator) is smaller than the cell residence time t_{m} (generated by a Gamma distribution random number generator), then the call will be completed in this BS, and a COMPLETION event is generated. The occupied_time of the COMPLETION event is set to t_{c}. Otherwise, the call will handoff to another BS which are determined by a uniform distribution random number generator (cf. step (13)), and a HANDOFF event is generated. The occupied_time of the HANDOFF event is set to t_{m} and the residual_time is set to $t_{\text{c}} - t_{\text{m}}$. If N_{call} of the BS $\geqslant \lfloor C/C_{\min} \rfloor$ at step (7), then the call arrival is blocking and N_{o} is incremented by 1 (cf. step (14)).

If e is a COMPLETION event at step (4), then the call represented by e is completed. For the BS specified by $e->new\_cell$, the cumulative occupation time $X[N_{\text{call}}]$ is incremented by the difference of the clock value and L_t, L_t is updated to the clock value, and N_{call} is decremented by 1 (cf. step (15)). The number of complete calls N_{c} is then incremented by 1, and a new element is appended to t_{com} array (cf. step (16)), where

$$t_{\text{com}}[N_{\text{c}}] = e-> occupied\_time.$$

If e is a HANDOFF event at step (4), then the number of handoff calls N_{h} is incremented by 1 (cf. step (17)). The behavior of the original BS specified by $e->old\_cell$ is the same as that for a COMPLETION event given at step (15) (cf. step (18)). For the handoff BS specified by $e->new\_cell$, if $N_{\text{call}} < \lfloor C/C_{\min} \rfloor$ (cf. step (20)), then the handoff call represented by e is connected. The behavior of the BS is the same as that for a ARRIVAL event given at step (8) (cf. step (21)). The cell residence time t_{m} for the handoff call is generated by a Gamma distribution random number generator (cf. step (22)). If $e->residual\_time$ is less than t_{m} (cf. step (23)), then the call will be completed in this BS and a COMPLETION event is generated (cf. step (24)). The occupied_time of the COMPLETION event is set to $e->residual\_time + e->occupied\_time$. Otherwise, the call will handoff to another BS (cf. steps (25)

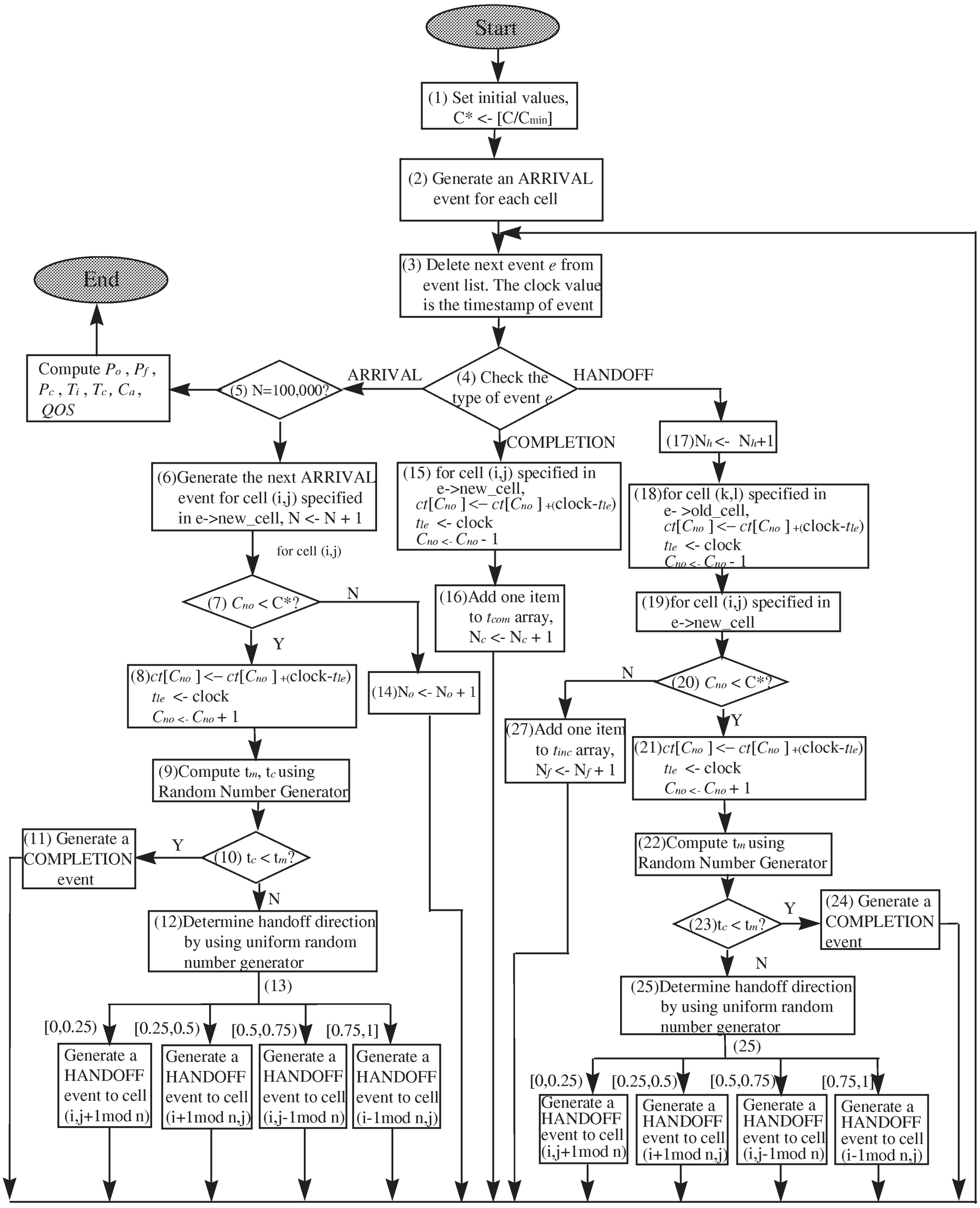


Figure 7. The simulation flow chart.

and (26)), and a new HANDOFF event is generated. The occupied_time of the new HANDOFF event is set to $t_\mathrm{m} + e- > occupied\_time$ and the residual_time is set to $e- > residual\_time - t_\mathrm{m}$. If N_call of the BS $\geqslant \lfloor C/C_\mathrm{min} \rfloor$ at step (20), then the call represented by e is force terminated. The number of incomplete calls N_f is then incremented by 1, and a new element is appended to t_inc array (cf. step (27)), where

$$t_\mathrm{inc}[N_\mathrm{f}] = e- > occupied\_time.$$

References

[1] D. Gross and C.M. Harris, *Fundamentals of Queueing Theory* (Wiley, New York, 1985).

[2] D. Hong and S.S. Rappaport, Traffic model and performance analysis for cellular mobile telephone systems with prioritized and no-protection handoff procedure, IEEE Trans. Vehicular Technol. 35(3) (August 1986).

[3] J.-Y. Jeng, Y.-B. Lin and W. Chen, Dynamic channel assignment strategies for wireless ATM, Information Sciences 109 (1998) 211–225.

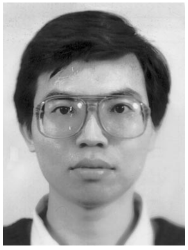
[4] J.-Y. Jeng, C.-W. Lin and Y.-B. Lin, Dynamic resource scheduling for GSM data services, IEICE Trans. Commun. 80-B(2) (February 1997).

[5] Y.-B. Lin, Reducing location update cost in a PCS network, IEEE/ACM Trans. Networking 5(1) (1997) 25–33.

[6] Y.-B. Lin and I. Chlamtac, Effective call holding time for a PCS network, in: *IEEE PIMRC '96*.

[7] Y.-B. Lin and V.K. Mak, Eliminating the boundary effect of a large scale personal communication service network simulation, ACM Trans. Modeling Comput. Simulation 4(2) (1994).

[8] Y.-B. Lin, S. Mohan and A. Noerpel, Queueing priority channel assignment strategies for handoff and initial access for a PCS network, IEEE Trans. Vehicular Technol. 43(3) (1994) 704–712.
[9] Y.-B. Lin, A. Noerpel and D. Harasty, The sub-rating channel assignment strategy for PCS handoffs, IEEE Trans. Vehicular Technol. 45(1) (February 1996).
[10] S.M. Ross, *Stochastic Process* (Wiley, New York, 1983).
[11] S. Tekinary and B. Jabbari, A measurement based prioritized scheme for handovers in cellular and microcellular networks, IEEE J. Selected Areas Commun. (October 1992) 1343–1350.
[12] E.J. Watson, *Laplace Transforms and Applications* (Birkhäuser, Basel, 1981).



Jeu-Yih Jeng received his B.S. degree in mathematics from Fu-Jen University, his M.S. degree in applied mathematics from National Chiao Tung University and his Ph.D. degree in computer science and information engineering, National Chiao Tung University, in 1983, 1985 and 1998. Since 1985, he has been with the Information Technology Laboratory of Telecommunication Laboratories, Chunghwa Telecom Co., Ltd, where he is currently a senior researcher. His research interests include design and analysis of personal communications services network, and performance modeling.

Yi-Bing Lin received his BSEE degree from National Cheng Kung University in 1983, and his Ph.D. degree in computer science from the University of Washington in 1990. From 1990 to 1995, he was with the Applied Research Area at Bell Communications Research (Bellcore), Morristown, NJ. In 1995, he was appointed as a Professor in the Department of Computer Science and Information Engineering (CSIE), National Chiao Tung University (NCTU). In 1996, he was appointed as Deputy Director of Microelectronics and Information Systems Research Center, NCTU. Since 1997, he has been Chairman of CSIE, NCTU. His current research interests include design and analysis of personal communications services network, mobile computing, distributed simulation, and performance modeling. Dr. Lin is a subject area editor of the Journal of Parallel and Distributed Computing, an associate editor of the International Journal on Computer Simulation, an associate editor of IEEE Network, an associate editor of the 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 International Journal of Communications Systems, an editor of Journal of Information Science and Engineering, a member of the editorial board of ACM/URSI/Baltzer Wireless Networks, 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 MONET special issue on Personal Communications Services, and Guest Editor for IEEE Transactions on Computers special issue on Mobile Computing.
E-mail: liny@scie.nctu.edu.tw