Telecommunication Systems 19:2, 125–146, 2002

Call Blocking Performance Study for PCS Networks under More Realistic Mobility Assumptions *

HELEN ZENG helenz@nortelnetworks.com
Nortel Networks, USA

YUGUANG FANG fang@ece.ufl.edu
Department of Electrical and Computer Engineering, University of Florida, 435 Engineering Building, P.O.Box 116130, Gainesville, FL 32611, USA

IMRICH CHLAMTAC chlamtac@utdallas.edu
Erik Jonsson School of Engineering and Computer Science, University of Texas at Dallas, USA

Received 3 July 2000; Revised 13 July 2001

Abstract. In the micro-cell-based PCS networks, due to the high user mobility, handoffs occur more frequently. Hence, the classical assumptions, such as the exponential assumptions for channel holding time and call inter-arrival time, may not be valid. In this paper, we investigate the call blocking performance for PCS networks using a semi-analytic and semi-simulation approach. We first construct a simulation model as the base for our performance study, using which the handoff traffic is studied. Then we present a few possible approximation models from which analytical results for call blocking performance metrics can be obtained and compared with the simulation results. We show that for a certain parameter range, such approximations may provide appropriate results for call blocking performance. Finally, using the simulation model, we investigate how various factors, such as the high moments, the variance of cell residence time, mobility factors and the new call traffic load affect the call blocking performance. Our study shows that all these factors may have a significant impact on call blocking performance metrics such as call blocking probability, call incompletion probability and call dropping probability. This research provides a strong motivation for the necessity of reexamining the validity of analytical results obtained from classical teletraffic theory when dealing with the emerging wireless systems.

Keywords: teletraffic, mobility, handoff traffic, channel holding time, blocking probability, call dropping probability

1. Introduction

In a PCS network [Lin and Chlamtac, 16], the total coverage area is divided into cells, each of which is provisioned with a number of channels. Call arrivals in such a network can be classified as new calls and handoff calls. A new call is the call initiated in the cell

* The work of Yuguang Fang was supported in part by National Science Foundation Faculty Early Career Development Award ANI-0093241. The work of Imrich Chlamtac was supported in part by the U.S. Army Research Office under Contract DAAG55-97-1-0312 and DAAG55-97-1-0382.

by a mobile user whereas a handoff call is an on-going call handed over from another cell. Call blocking occurs when a call arrives at a cell and finds no channel available. Depending on whether the call is a new call or a handoff call, we term the respective call blocking probabilities *the new call blocking probability* and *the handoff call blocking probability* respectively. When a handoff call is blocked, the call is forced to terminate, resulting in a dropped call. In PCS networks, the call blocking performance evaluation is a key issue for system design, resource dimensioning and management [Lin and Chlamtac, 16].

To facilitate our presentation, we first clarify the following three concepts. *Cell residence time* (CRT) is defined to be the time a mobile stays in a cell. *Channel holding time* (CHT) is the time a mobile occupies channel(s) within a cell for the call connection service, i.e., the time that the mobile user utilizes the system resource during its residence in the cell. *Call holding time* is defined as the uninterrupted call duration time (corresponding to the holding time for a wired phone call or the session time in computer system). Their relationships will be clear in the subsequent development.

Two performance metrics, the call blocking probability and the call dropping probability, are key parameters for PCS network design. These parameters are determined by the CHT (corresponding to the service time in queuing networks), cell traffic (the merged traffic of new call and handoff call arrivals), the number of channels from the base station and the channel allocation schemes for call admission control. In earlier studies of wireless cellular systems, the following three assumptions were commonly used:

(1) the call holding time is exponentially distributed,

(2) the channel holding time (CHT) is exponentially distributed, and

(3) the arrival process of the cell traffic follows a Poisson distribution.

Hong and Rappaport [13] proposed a traffic model for cellular mobile radio telephone systems, which has inspired extensive research in teletraffic analysis and network design (see [Fang et al. 9; Tekinay and Jabbari 24] and references therein). Most research works in the current literature use the aforementioned assumptions in order to obtain analytical results. However, simulation studies and field data have shown that some of these assumptions are not valid in PCS networks [Barcelo and Bueno, 1; Barcelo and Jordan, 2; Guerin, 12; Jedrzycki and Leung, 14; Orlik and Rappaport, 19]. It was observed that the CHT is not exponentially distributed [Barcelo and Bueno, 1; Barcelo and Jordan, 2; Guerin, 12; Jedrzycki and Leung, 14] and that the cell traffic is no longer Poissonian [Orlik and Rappaport, 20; Rajaratnam and Takawira, 21; Rajaratnam and Takawira, 22]. Although it is well known from queuing theory [Kelly, 15] that the blocking probability in an $M/G/m/m$ queue is insensitive to service time distribution (corresponding, in our case, to CHT distribution), it is not known whether this is true for a $G/G/m/m$ system. Even though we accept the fact that the cell traffic is Poissonian, as commonly accepted in the current literature, the arrival rate of the cell traffic in the PCS network is in fact affected by the distribution of the CRT (hence indirectly by the distribution of CHT) [Fang and Chlamtac, 8; Fang et al., 11], which signifies the difference between the $M/G/m/m$ queue system we are considering and the ones in the queueing literature. Recently, Orlik and Rappaport [20] and Rajaratnam and Takawira [22] applied two different modeling techniques to investigate the sensitivity property of blocking probabilities to the assumption on handoff traffic and obtained different conclusions: Orlik and Rappaport [20] conclude that the blocking probabilities are insensitive to the assumption on handoff traffic, while Rajaratnam and Takawira [22] claim that blocking probabilities are indeed sensitive to the assumption on handoff traffic and that the Poisson assumption is not appropriate. It is, therefore, imperative to systematically reexamine the validity of the aforementioned assumptions for PCS networks and investigate how the modeling assumptions affect the performance results for the emerging PCS networks.

In a series of papers [Fang et al., 9; Fang et al., 10; Fang et al., 11], the authors demonstrated that the teletraffic performance of a PCS network depends on the users' mobility, which can be characterized by the CRT distribution. It was observed that call holding time and CRT are two independent important time variables, which can completely determine other quantities of interest (such as the CHT and cell traffic). Since the call holding time is totally determined by the calling habits of mobile users, given the group calling patterns of cellular users, the CRT turns out to be the key time variable to characterize wireless network performance. We show that a more general mobility model for the CRT (for mobility) is needed. In the current literature, nontrivial distribution models [Del Re et al., 6; Hong and Rappaport, 13] and generalized gamma distribution model [Zonoozi and Dassanayake, 27] have been used to model the CRT based on the modeling of speed and moving directions of mobiles and the hexagonal cell shape. Lin et al. [17] and Fang et al. [11] propose to model the CRT directly as a random variable and characterized the users' mobility by specifying the distribution of the CRT and show how CRT distribution affects the CHT distribution.

In the current PCS networks, due to the multi-service (voice, data, video), multi-environment (indoor,outdoor), and large variety of user mobility (pedestrians, vehicular mobile users), the mobile communication behavior is more complex. A call may traverse a large number of cells while the mobile user is in the car, or may stay in one cell during the whole call holding time when the mobile user is indoors. In all these cases, a non-exponential CRT tends to be more realistic. As demonstrated in [Fang, 7], the mixed distribution model such as hyper-Erlang distribution is the natural choice for modeling the mixed mobility environments. In order to accurately monitor the system performance, the CHT and cell traffic need to be evaluated based on such non-exponential models. It is important to systematically examine how the cell traffic and call blocking performance is affected by the CRT distribution.

In this paper, we evaluate the handoff traffic and call blocking performance of homogeneous PCS networks under a more general CRT distribution model. In this process, we combine analytical approach with simulation studies. We develop a simulation model in which the CRT is the fundamental random variable we model and parameters, such as handoff arrival rate and CHT, are determined analytically. We demonstrate how the cell traffic and call blocking performance can be affected by the CRT distribution. Our results show that the distribution of CRT can significantly affect the cell traffic, the call blocking probabilities, call incompletion probability and call dropping probability. Specifically, we show that, in addition to the mean of the CRT, the variance is also a major contributing factor, affecting the call blocking performance of PCS networks. This study suggests that the first-moment model should not be used without justification of the underlying assumptions, and that more general mobility models may be necessary in order to provide more appropriate approximations. Based on our previous research [Fang et al., 11], we observe that the CHT is exponentially distributed if and only if the CRT is exponentially distributed, which implies that the only case we can invoke the classical assumption (CHT is exponentially distributed) is when the CRT is exponentially distributed. Therefore, in practice, we can collect data for CRT and test whether the exponential distribution provides a good fit. If so, we can apply traditional results in teletraffic theory. One mobility model, which gives an exponential CRT, is the Markov chain model (each cell is modeled as a node of a queueing network with an exponential server). If the exponential model does not fit the field data, we have to seek other distribution (such as hyper-Erlang distribution model [Fang and Chlamtac, 8]) and apply the technique we develop in this paper to evaluate the system. Many studies [Barcelo and Jordan, 2; Jedrzycki and Leung, 14] show that lognormal distribution provides better fit to field data for CHT. In [Fang, 7; Fang and Chlamtac, 8], we demonstrate that hyper-Erlang distribution also give excellent approximation to the field data for CHT. We expect that hyper-Erlang distribution will be a good choice for CRT as well. Although we use gamma distribution for CRT model due to the explicit interpretation of the two parameters in this distribution and its more general approximation capability [Cox, 5], our approach of analysis can be easily extended to the hyper-Erlang model.

This paper is organized as follows. In the next section, we present necessary existing analytical results for CHT and some call blocking performance metrics, which will be used in the comparative study. The simulation model is described in section 3. In section 4, we present the comparison results for handoff traffic and call blocking performance. In section 5 we list the conclusions of this study.

2. Performance metrics

The performance metrics we consider in this paper are the call blocking probability, call incompletion probability and call dropping probability. In order to compare some analytical results derived under some exponential assumptions with those obtained under more realistic assumptions via simulations, we first present the computational procedures for the following quantities such as CHT, handoff call arrival rate, call blocking probability, call incompletion probability and call dropping probability, some of which will be used in our comparative study.

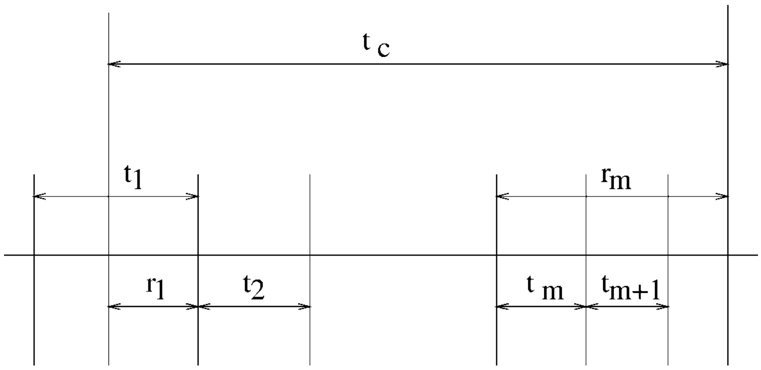


Figure 1. An example of time diagram for call holding time and cell residence time (CRT).

2.1. Channel holding time (CHT)

Channel holding time is the time that a call occupies a channel in a cell. If we model each cell as a queuing system in which the channels assigned to the base station are the servers, while the calls (new calls or handoff calls) form the arrival process, then the CHT is equivalent to the service time. For tractability, many researchers (see [Fang et al. 11] and references therein) have assumed that the CHT is exponential and cell traffic is Poisson, under which the Erlang-B formula can be used to find the call blocking probability. However, as we mentioned in the previous section, such assumptions may not be valid for some PCS networks. In a series of research works, Fang et al. [8, 11] studied CHT under the following two less restrictive assumptions: the new call arrival traffic is Poissonian and the call holding time is exponentially distributed. In this subsection, we present some of the results which will be used in the comparative study.

Figure 1 shows the timing diagram for the call holding time and CRT. Let t_c denote the call holding time for a typical new call, t_m be the CRT at the mth cell a mobile transverses, r_1 be the residual CRT (i.e., the time between the instant the new call is initiated at the first cell and the instant the new call moves out of the cell if the new call is not completed in the cell), and let r_m $(m > 1)$ denote the residual call holding time when the call finishes mth handoff successfully. Let t_{nh} and t_{hh} denote the CHTs for a new call and a handoff call, respectively. Assume that t_c, t_m, r_1, t_{nh} and t_{hh} have density functions $f_c(t)$, $f(t)$, $f_r(t)$, $f_{\mathrm{nh}}(t)$ and $f_{\mathrm{hh}}(t)$ with their corresponding Laplace transforms $f_c^*(s)$, $f^*(s)$, $f_r^*(s)$, $f_{\mathrm{nh}}^*(s)$ and $f_{\mathrm{hh}}^*(s)$, respectively, and with their cumulative distribution functions $F_c(t)$, $F(t)$, $F_r(t)$, $F_{\mathrm{nh}}(t)$, $F_{\mathrm{hh}}(t)$, respectively. Let $1/\mu$ and $1/\eta$ denote the average call holding time and average CRT, respectively.

From figure 1, the CHT for a new call is

$$t_{\mathrm{nh}} = \min\{t_c, r_1\}, \tag{1}$$

and the CHT for a handoff call is

$$t_{\mathrm{hh}} = \min\{r_m, t_m\}. \tag{2}$$

From (1) and (2), we can derive [Fang et al., 11]

$$
\begin{aligned}
f_{\mathrm{nh}}(t) &= f_{\mathrm{c}}(t)\int_t^{\infty} f_r(\tau)\,\mathrm{d}\tau + f_r(t)\int_t^{\infty} f_{\mathrm{c}}(\tau)\,\mathrm{d}\tau,\\
f_{\mathrm{hh}}(t) &= f_{\mathrm{c}}(t)\int_t^{\infty} f(\tau)\,\mathrm{d}\tau + f(t)\int_t^{\infty} f_{\mathrm{c}}(\tau)\,\mathrm{d}\tau,
\end{aligned}
\tag{3}
$$

where $f_r(t) = \eta(1 - F(t))$, the residual life of CRT.

From (3), Fang et al. [11] have shown the CHT for new calls (similarly for handoff calls) is exponentially distributed if and only if the CRT is exponentially distributed. If the CRT is not exponentially distributed, then the CHT is no longer exponentially distributed. For this case, we have shown:

Theorem 1 [Fang et al., 11]. For a PCS network with exponential call holding times and Poisson new call arrivals with arrival rate λ,

(i) the Laplace transform of the probability density function of the new call CHT is given by

$$f^*_{\mathrm{nh}}(s) = \frac{\mu}{s+\mu} + \frac{\eta s}{(s+\mu)^2}\big[1 - f^*(s+\mu)\big], \tag{4}$$

and the expected new call CHT is

$$E[t_{\mathrm{n0}}] = \frac{1}{\mu} - \frac{\eta}{\mu^2}\big[1 - f^*(\mu)\big]; \tag{5}$$

(ii) the Laplace transform of the density function of the handoff call CHT is given by

$$f^*_{\mathrm{hh}}(s) = \frac{\mu}{s+\mu} + \frac{s}{s+\mu} f^*(s+\mu), \tag{6}$$

and the expected handoff call CHT is

$$E[t_{\mathrm{h0}}] = \frac{1}{\mu}\big(1 - f^*(\mu)\big). \tag{7}$$

To illustrate how distribution of CRT affects the CHT, we model the CRT by the following gamma distribution:

$$f(t) = \frac{\beta^{\gamma} t^{\gamma-1}}{\Gamma(\gamma)}\mathrm{e}^{-\beta t}, \qquad f^*(s) = \left(\frac{\beta}{s+\beta}\right)^{\gamma}, \quad \beta = \gamma\eta,$$

where γ is the shape parameter, β is the scale parameter and the $\Gamma(\gamma)$ is the gamma function. In this way, we can study the CHTs for new calls and handoff calls using the above theorem. We notice that the mean and variance of the gamma distribution are $1/\eta$ and $1/(\gamma\eta^2)$, respectively. Figure 2 shows the average CHT for new calls and handoff calls, respectively. We observe that, the average CHT for new calls $E[t_{\mathrm{n0}}]$ and the average CHT for handoff calls $E[t_{\mathrm{h0}}]$ are sometimes significantly different, which are significantly affected by the variance of CRT. This indicates that the CRT distribution

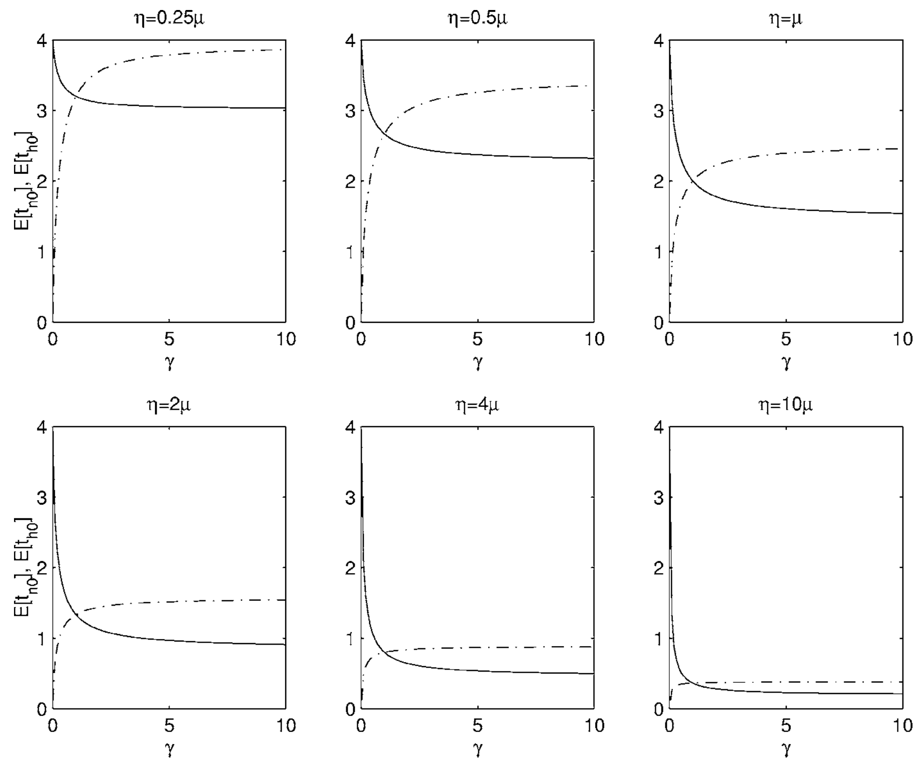


Figure 2. Mean channel holding times, solid: $E[t_{\mathrm{n}0}]$, dashed: $E[t_{\mathrm{h}0}]$. X-axis: shape parameter of gamma distribution, γ.

affects even the traffic situation in the wireless networks. Later, we will show that the call blocking performance is also significantly affected by the distribution of CRT.

2.2. Handoff call arrival rate

If we model each cell as a queuing system, we obtain two traffic streams: the new calls and the handoff calls. We know that the new call arrival rate is λ, we must find the handoff call arrival rate λ_{h}. Let p_0 and p_{f} be the blocking probabilities for new calls and handoff calls, respectively. Applying the results in [Fang et al., 11] to our case, we can obtain

$$\lambda_{\mathrm{h}} = \frac{\eta(1-p_0)[1-f^*(\mu)]\lambda}{\mu[1-(1-p_{\mathrm{f}})f^*(\mu)]}. \tag{8}$$

The derivation of equation (8) can be found in appendix.

2.3. Blocking probabilities

Blocking probabilities of a PCS network are very important parameters for system analysis and design. As we observe, the handoff arrival rate is dependent on the blocking probabilities p_o and p_f, with the two traffic streams (new calls and handoff calls). With appropriate service discipline (channel allocation scheme) for new calls and handoff calls, we can find the blocking probabilities p_o and p_f, which will be dependent on the handoff call arrival rate λ_h. Therefore, we find a set of recursive equations, which can be solved for the blocking probabilities p_o and p_f. For illustration purpose, in this paper, we concentrate on PCS networks using non-prioritized channel allocation scheme, in which case the handoff calls and new calls are not distinguishable. Our procedures here can be easily extended to wireless networks using other prioritized schemes.

In the traditional method [Hong and Rappaport, 13; Tekinay and Jabbari, 24; Yoon and Un, 25], the merged arrival traffic from new calls and handoff calls is assumed to be Poissonian and the CHT is assumed to be exponentially distributed. However, as we mentioned earlier, these assumptions may not be valid for PCS networks. How such assumptions affect the previously known results in the traditional cellular networks is a critical issue which needs to be resolved. We observe that the cell traffic and the CHT are both affected by CRT, by specifying the CRT distribution, we can investigate such effects. We now carry out this study based on the following four cases, each of which represents one approximation potentially used in practice. We assume that the CRT is gamma-distributed for illustration purpose, then we make some assumptions, which may be used by researchers to obtain the estimate for blocking probability. Our purpose is to investigate how much deviation we observe due to the assumption.

Case 1. We develop the simulation model (details are given in the next section) to obtain the call blocking probability p_b. This is the real blocking probability for the PCS network.

Case 2. During the simulation, we can collect the statistics, from which we compute the average call arrival rate to the cell (including new calls and handoff calls) $\lambda + \hat{\lambda}_h$ and the average CHT $\widehat{T}$. Then, we invoke the commonly used assumption: Poisson assumption on the arrival process and we then apply the Erlang-B formula to obtain the blocking probability:

$$\hat{p}_b = \frac{(\hat{\rho}^c/c!)}{\sum_{i=0}^{c}(\hat{\rho}^i/i!)}, \tag{9}$$

where $\hat{\rho} = (\lambda + \hat{\lambda}_h)\widehat{T}$ and c is the number of channels in a cell. In this case, we basically use the mean information for the call arrivals and CHT which can be obtained from experimental data without considering any details of mobility.

Case 3. Another commonly used assumption is to model the CRT as the exponential random variable using the mean information about the CRT collected from experiments, then compute the CHT, from the queuing model $M/G/c/c$ to find the call blocking probability. If this is the case, $f^*(s) = \eta/(s + \eta)$. From theorem 1, we obtain the average CHT

$$E[t_{ch}] = \frac{\lambda}{\lambda + \lambda_h}E[t_{nh}] + \frac{\lambda_h}{\lambda + \lambda_h}E[t_{hh}] = \frac{1}{\eta + \mu}.$$

Let p_1 denote the blocking probability for this case, from the preceding section, we obtain ($p_f = p_1$)

$$\lambda_h = \frac{\eta(1-p_1)(1-\eta/(\eta+\mu))\lambda}{\mu[1-(1-p_1)\eta/(\eta+\mu)]} = \frac{(1-p_1)\lambda}{\sigma+p_1},$$

where $\sigma = \mu/\eta$. Thus, the cell traffic intensity for this case is given by

$$\rho_1 = (\lambda+\lambda_h)E[t_{ch}] = \frac{\sigma+1}{\sigma+p_1}\cdot\frac{\lambda}{\mu+\eta}. \tag{10}$$

From Erlang-B formula, we have the following relationship:

$$p_1 = \frac{(\rho_1^c/c!)}{\sum_{i=0}^{c}(\rho_1^i/i!)}. \tag{11}$$

Solving equations (10) and (11), we can obtain the call blocking probability p_1 under exponential model for the CRT and Poisson cell traffic.

Case 4. In this case, we only assume that the cell traffic is Poisson. Here, we use the gamma distribution to compute the handoff call arrival rate and the average CHT in order to find the traffic intensity. Let p_γ denote the call blocking probability for this case. From theorem 1 and the result in the preceding section, we can obtain the following set of equations:

$$\begin{aligned} f^*(\mu) &= \left(\frac{\gamma\eta}{\mu+\gamma\eta}\right)^{\gamma}, \qquad \lambda_h = \frac{\eta(1-p_\gamma)[1-f^*(\mu)]\lambda}{\mu[1-(1-p_\gamma)f^*(\mu)]}, \\ \rho &= \lambda E[t_{nh}] + \lambda_h E[t_{hh}] = \frac{1}{\mu}\left[\lambda + \frac{\lambda_h\mu-\lambda\eta}{\mu}\big(1-f^*(\mu)\big)\right], \\ p_\gamma &= \frac{(\rho^c/c!)}{\sum_{i=0}^{c}(\rho^i/i!)}. \end{aligned} \tag{12}$$

Solving this set of equations, we can obtain the call blocking probability p_γ.

Case 2 does not assume any detailed information about the mobility, case 3 makes an assumption on the mobility–the exponential model, a commonly used model in current literature, while the last case utilize the full information about the mobility. It is expected that case 4 will yield the same result as case 2 because the arrival rate for the cell traffic and the average CHT should be the same, however, in case 4 we do not need to do any simulation to obtain the necessary parameters for cell traffic and CHT, such parameters can in fact be obtained from the analytical results we developed in our work [Fang et al., 11] and are also given in the second section.

2.4. Call incompletion and call dropping probability

Call incompletion probability and call dropping probability are highly important parameters for customer care. Call incompletion probability is the probability that a call is blocked either at the call initiation or during a handoff. When a call is blocked during a handoff, it results in a call dropping. Since mobile users are more sensitive to call dropping than to call blocking, therefore PCS network service providers have to minimize the call dropping probability for real system design. Let p_{nc} denote the call incompletion probability, p_{d} denote the call dropping probability, then p_{nc} can be derived from the result in [Fang et al., 9]:

$$p_{\mathrm{nc}} = p_0 + p_{\mathrm{d}} \tag{13}$$

$$= p_0 + \frac{\lambda_{\mathrm{h}}}{\lambda} p_{\mathrm{f}} \tag{14}$$

from which we obtain

$$p_{\mathrm{d}} = \frac{\lambda_{\mathrm{h}}}{\lambda} p_{\mathrm{f}}. \tag{15}$$

The call incompletion probability indicates the overall effect of call blocking and call dropping, it can be used to study the tradeoff between the new call blocking and handoff call blocking (i.e., call dropping).

3. Simulation model

The critical issue for the performance evaluation of the PCS networks is the handoff traffic (cell traffic) characterization. Since analytical model for handoff traffic is not available in the literature and we do not think there will be an appropriate analytical model to fully characterize the handoff traffic accurately in the near future, we propose the following approach to carry out the performance evaluation. We assume that the whole geographical area is divided into hexagonal cells. In order to eliminate the edge effect, we use a wrap around model, which can guarantee that each cell has six neighbors so that the handoff departure from the edge cells will not be ignored. For example, in figure 3, the edge cells and their neighbors are: 32{33, 16, 31, 23, 35, 29}; 31{32, 16, 15, 30, 36, 24}; 30{31, 15, 14, 28, 29, 37, 25}; 29{30, 14, 28, 32, 20, 26}; 28{14, 13, 27, 33, 21, 29}; 27{13, 12, 26, 34, 22, 28}; 26{12, 25, 29, 35, 23, 27}. We use the general distribution model for the CRT (for users' mobility), then apply our analytical results to model the CHTs for new calls and handoff calls, respectively. In each cell, an incomplete call will be routed to one of the neighboring cells by a randomization procedure. In this approach, we do not make any assumptions on the handoff traffic, while the overall network dynamics will mimic that of a homogeneous PCS networks. This semi-simulation study seems to work well for our study. The details are described as follows.

This simulation model can capture the essence of dynamics of the homogeneous networks, which share the following common characteristics: each cell has the same number of channels, the same new call arrival rate, the same call holding time distribution, the same CRT distribution and the same user moving pattern. New calls are generated independently in each cell according to Poisson distribution, each new call is assigned a call holding time with exponential probability distribution. Handoff calls are generated based on the CHT according to (1) and (2). An unfinished call of a mobile moving out of a cell will be handed over with equal probability $1/6$ to any one of

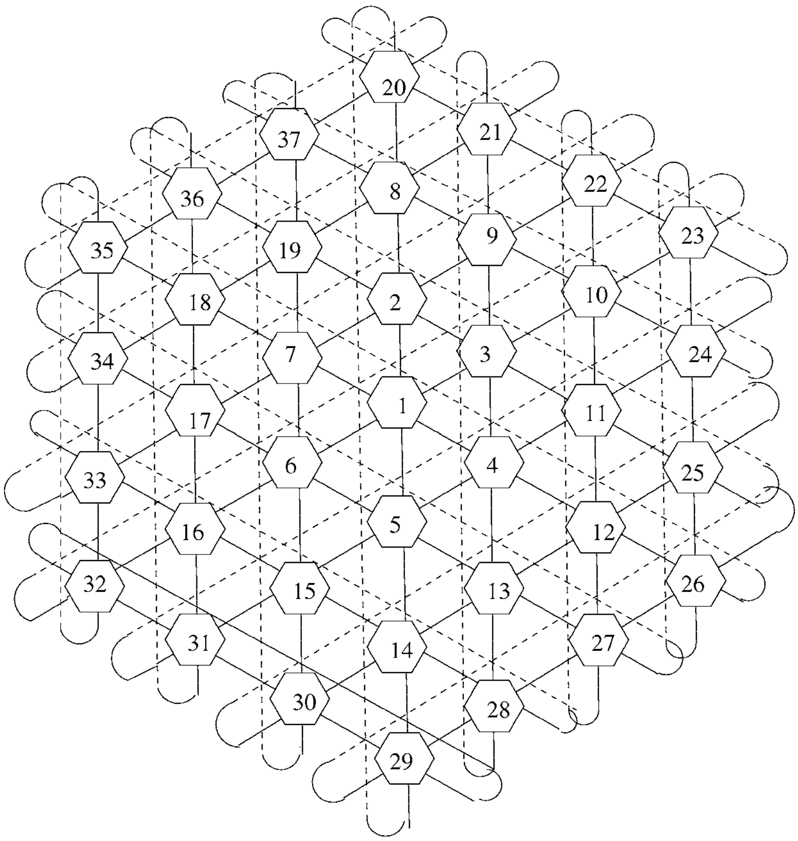


Figure 3. The wrap-around simulation model.

its neighboring cells in a hexagonal layout (this is called the randomization procedure). The input and output parameters for this simulation study are described below.

Input parameters: the new call arrival rate λ, the average new call holding time $1/\mu$, the number c of channels in one cell, and CRT distribution, in particular, the gamma distribution for illustration purpose.

Output parameters: the handoff arrival rate H_a, the handoff departure rate H_d, the inter-arrival time of handoff call arrival process, the CHT distribution for new calls and handoff calls, call blocking probability p_0 (for new calls), p_f (for handoff calls), p_b (for total calls), call incompletion probability p_{nc}, call dropping probability p_d, call blocking probability obtained under the three classical assumptions.

4. Results and discussions

In this section, we present our comparative study for handoff traffic, CHT, call blocking probability, call incompletion probability and call dropping probability.

4.1. Handoff traffic

We first study the handoff traffic and observe whether the Poisson model is valid. For this purpose, we use the following parameters: there are $c = 12$ channels in each cell, we use 4-minute mean call holding time ($\mu = 0.25$), different mean CRT ($\eta = 0.25$, 0.5, 1.0, 2.5), and different new call arrival rate ($\lambda = 0.5, 1.0, 1.5, 2.0, 2.5$). For each set of parameters (η, λ), we vary the variance of the CRT, say, $1/(\gamma\eta^2)$ (i.e., the change of γ), then observe the effect of variance of CRT on the handoff traffic distribution.

As we mentioned before, the cell traffic consists of two streams, the new calls with arrival rate λ and the handoff calls with arrival rate λ_h. If we let $f_{ch}(t)$ denote the probability density function of CHT for the cell traffic, then we have

$$f_{ch}(t) = \frac{\lambda}{\lambda + \lambda_h} f_{no}(t) + \frac{\lambda_h}{\lambda + \lambda_h} f_{hh}(t).$$

Figure 4 illustrates the relationship between the CRT and CHT. The CHT is exponential if and only if the CRT is exponential, which is consistent with the analytical results we obtained in [Fang et al., 11].

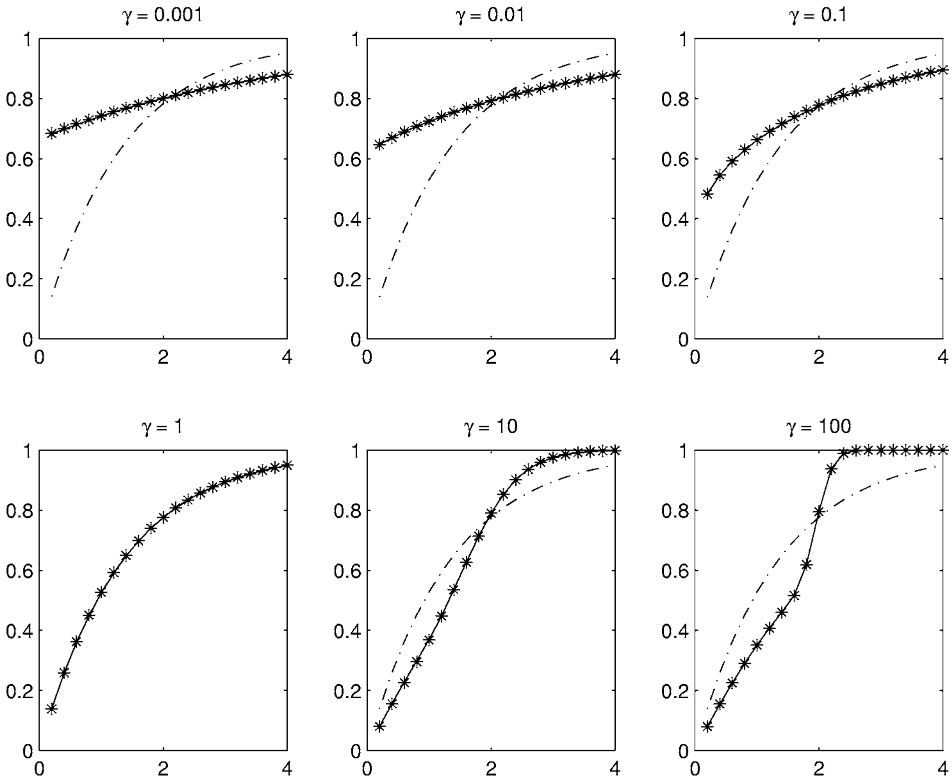


Figure 4. Channel holding time CDF: $\lambda = 0.5$, $\eta = 0.5$, $\mu = 0.25$, $c = 12$, solid curve: simulation results, $*$: analytical results, dashed curve: exponential fitting.

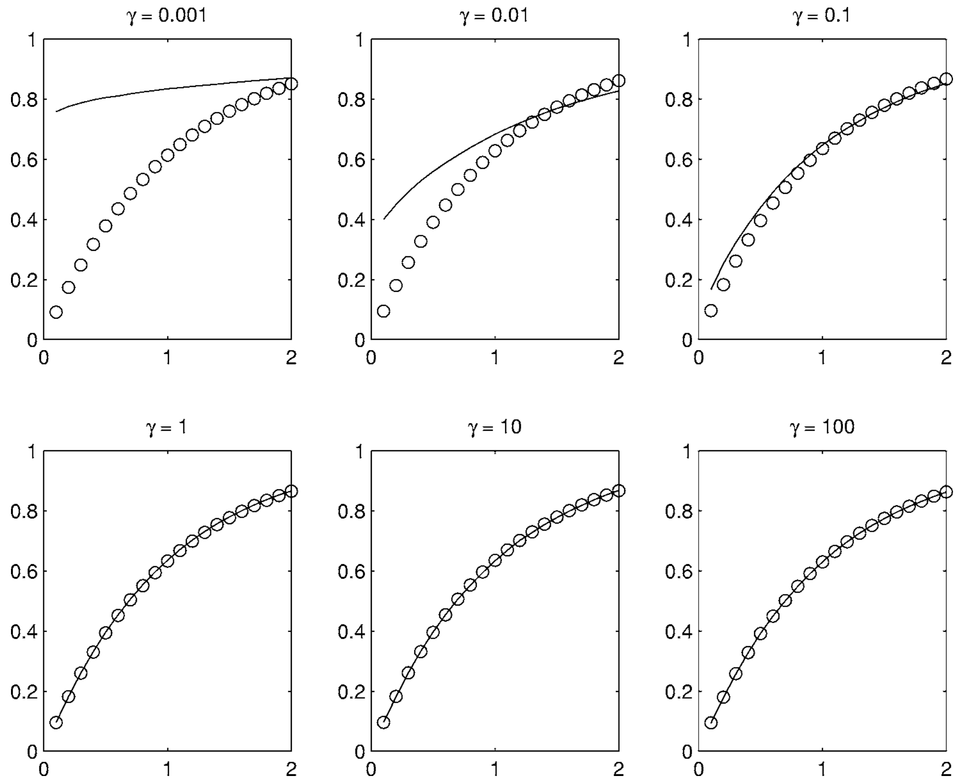


Figure 5. Handoff call inter-arrival time CDF: $\lambda = 0.5$, $\eta = 0.5$, $\mu = 0.25$, $c = 12$, solid curve: simulation results, o: exponential fitting.

We also observe that the CHT distribution obtained from the simulation coincides with that obtained from the analytical results (in theorem 1).

Figure 5 shows the characterization of handoff traffic. We observe that the handoff traffic is not Poisson when the variance of CRT is large (i.e., γ is very small). When the variance of CRT increases, the mismatch between the handoff traffic distribution and the Poisson fitting increases.

In figure 4 we show that, when variance of CRT is large (i.e., $\gamma < 1$), the CHT is not exponential, while in figure 5, when variance of CRT is small, the handoff call arrival process is Poisson, that means when CHT is not exponential, the handoff arrival traffic can still be appropriately modeled by Poissonian process. This can be explained by the fact that the handoff traffic depends on CHT. This dependency needs to be taken into account in teletraffic analysis. Another observation is that when the CRT is exponentially distributed ($\gamma = 1$), the handoff traffic is Poisson, however, the reverse is not true, i.e., although the handoff traffic is Poisson, the CRT may not be exponential. This corrects the wrong claim we made in [Fang et al., 11] regarding the handoff traffic.

When the new call arrival rate λ is small, we call this environment low blocking environment, such as in figure 5, $\lambda = 0.5$. When the new call arrival rate λ increases,

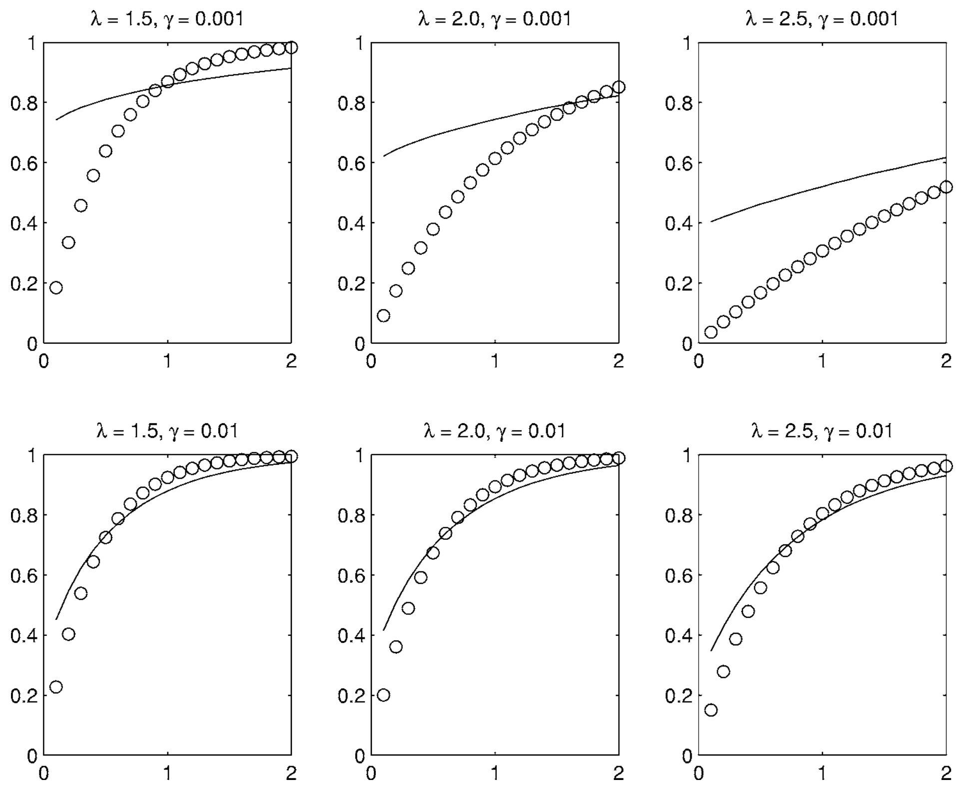


Figure 6. Handoff call inter-arrival time CDF under different new call arrival rate: $\eta = 0.5$, $\mu = 0.25$, $c = 12$, solid curve: simulation results, o: Poisson fitting.

the blocking probability increases, and the low blocking environment changes to high blocking environment. Figure 6 shows the effect of the new call arrival rate on the handoff traffic. Comparing with the handoff traffic in figure 5, in low blocking environment, ($\lambda = 0.5$–1.5), when the new call arrival rate increases, the mismatch between the handoff traffic and its Poisson fitting decreases; however, in a high blocking environment ($\lambda = 1.5$–2.5), the higher the new call arrival rate, the higher the mismatch between the handoff traffic and its Poisson fitting. From figure 8, we can see, in high blocking environment, when the new call arrival rate increases, the handoff arrival rate decreases, which is opposite to the low blocking environment.

Next, we show that not only the variance of CRT, but also its mean affects the handoff call inter-arrival distribution. Let the value of η/μ to characterize the user's mobility. It is obvious that large η/μ implies high mobility, small η/μ implies low mobility, this is consistent with our intuition. We compare the handoff traffic under different user mobility: $\eta = 2\mu$, $\eta = 4\mu$, and $\eta = 10\mu$. Figure 7 shows that, higher user mobility will cause less mismatch between handoff traffic and Poisson fitting, and less significant effect of variance of CRT on handoff traffic. In [Chlamtac et al., 3], we found that, for higher user mobility, the handoff call blocking probability is lower and

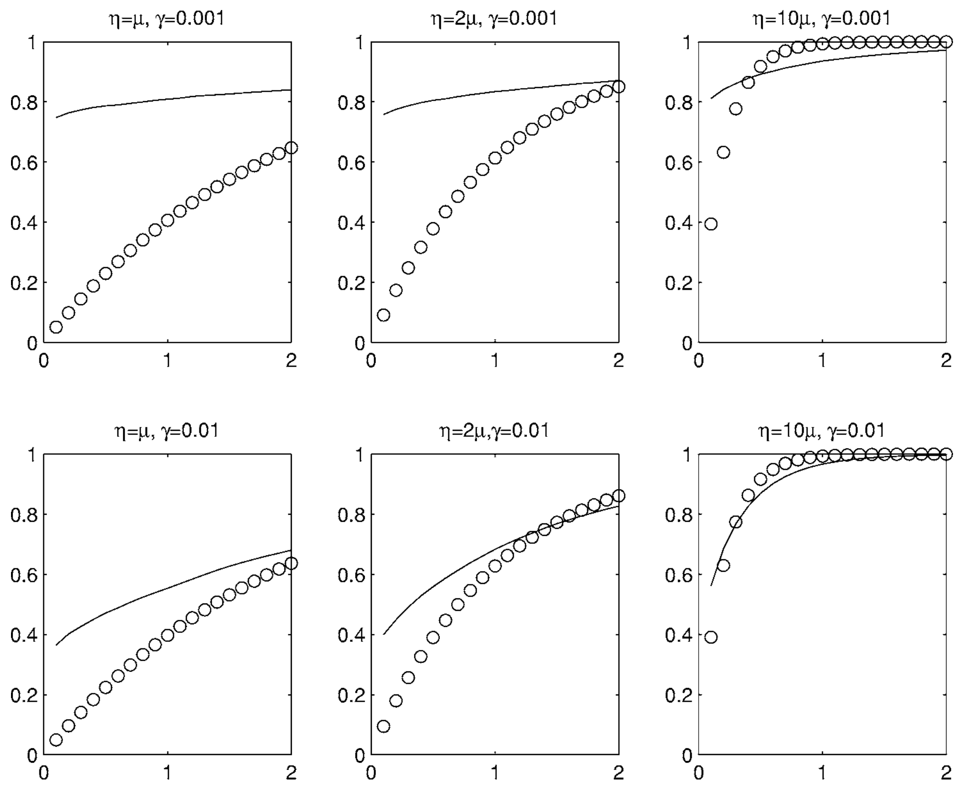


Figure 7. Handoff call inter-arrival CDF under different users' mobility: $\lambda = 1.5$, $\mu = 0.25$, $c = 12$, solid curve: simulation results, o: Poisson process fitting

less affected by CRT distribution. Those results are consistent with the results here. Since in higher user mobility environment, the probability for a call to be handed off to other cells is higher, the handoff traffic in cells have more interaction on each other, but the homogeneous property makes them similar enough, thus the handoff traffic is less sensitive to the CRT distribution.

From above we conclude that the handoff traffic is affected by the new call traffic load, the high moments (e.g., the variance) of CRT, and the users' mobility. The handoff arrival rate λ_h is also affected by CRT distribution. When the variance of CRT decreases, λ_h increases (figure 8). The handoff arrival rate obtained from simulation and that obtained from analytical result (i.e., equation (8)) are identical, which in turn confirms the validity of our analytical results.

4.2. *Call blocking performance*

In this subsection, we study the call blocking probability, the call incompletion probability and the call dropping probability in the following cases:

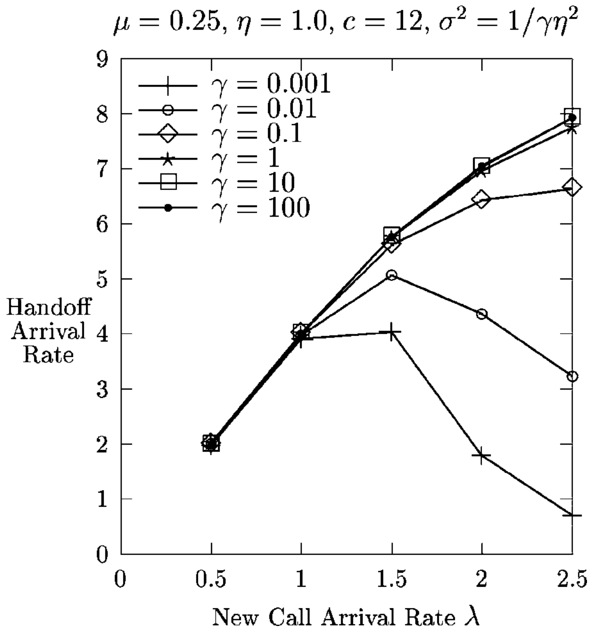
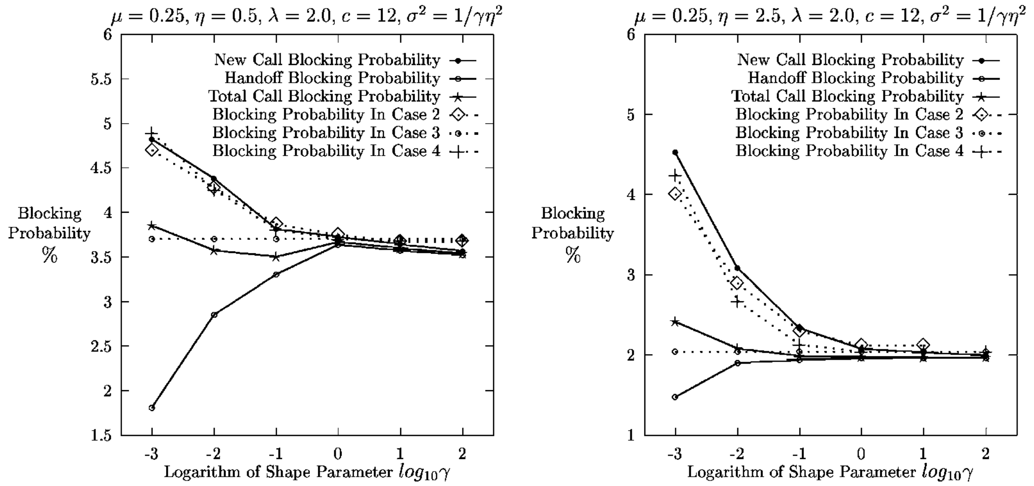


Figure 8. Effect of CRT distribution on handoff call arrival rate: $\mu = 0.25$, $\eta/\mu = 4$, $c = 12$.

1. Keeping the same new call arrival rate, and the same mean CRT, changing the shape parameter γ (i.e., changing the variance), and then observing the effect on the call blocking probability, comparing the result with the classical result in cases 2–4 (as described in the previous section), and comparing p_0, p_f, p_b (figures 9 and 10).
2. Similar to 1, also changing the variance of CRT under different mean value (i.e., different user mobility), observing the effect on the call blocking probability, p_0, p_f, p_b and the results obtained in cases 2–4 (figure 10).
3. Keeping the same new call arrival rate, changing the variance of CRT under different mean value (i.e., different user mobility), then observing its effect on the new call blocking probability and the handoff call blocking probability (figures 10 and 11).
4. Keeping the same mean CRT, changing the variance of CRT under different new call arrival rate, then observing the effect on the call blocking probabilities (figure 10).
5. Keeping the same mean CRT, changing the variance of CRT under different new call arrival rate, and observing the effect on the call incompletion probability and the call dropping probability (figure 12).

In figures 9 and 10, we observe that:

1. When the variance of CRT is low, all four cases (in figure 9) give almost the same results for blocking probabilities, which implies that all four approximation model can be used to estimate the blocking probabilities. This is not surprising, because we observe that when the variance of CRT is low, the handoff traffic can be approximated by the Poisson process, so is the cell traffic, which validates the assumptions



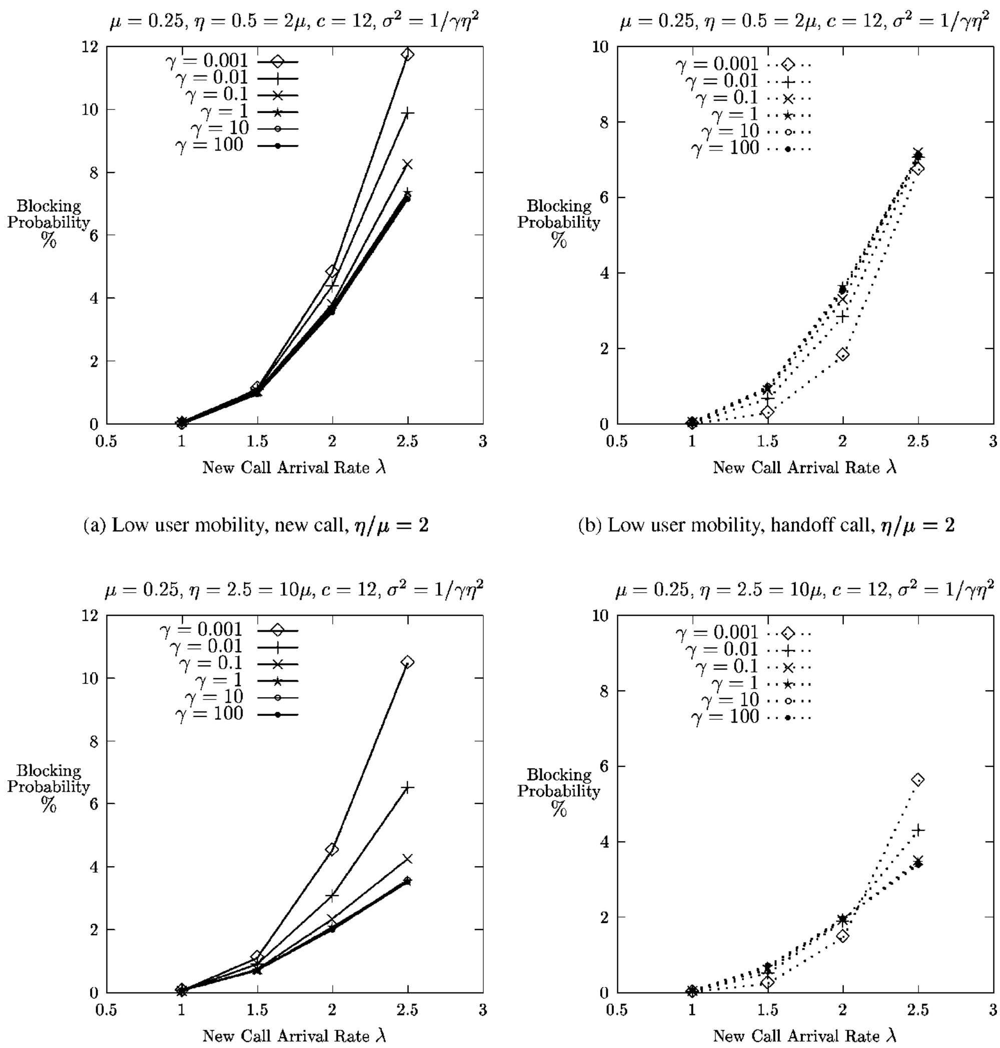
(a) $\lambda = 2.0, \eta/\mu = 2$ (b) $\lambda = 2.0, \eta/\mu = 10$

Figure 9. Effect of CRT distribution on the call blocking probability under different assumptions.

we used in the four cases. However, when the variance is high (i.e., $\gamma < 1$), the results from four cases and simulations are different, cases 2 and 4 results are closer to simulation result of new call blocking probability p_0 than the case 3 result. Obviously, we observe that when the variance of CRT is high, the handoff (cell) traffic is no longer Poissonian.

2. The new call blocking probability p_0 is different from the handoff call blocking probability p_f. The new call blocking probability p_0 is higher than handoff blocking probability p_f. In [Hong and Rappaport, 13], when no priority is given to handoff calls over new calls, no difference exists between their blocking probabilities. However, we observe here that even under the non-prioritized channel allocation scheme, the new call arrivals and handoff call arrivals are significantly different. From the previous section, we know that the new calls and handoff calls have different CHT distribution, the handoff arrival process is not Poissonian. In [23], Sidi and Starobinski found that there exists fundamental difference between the new call blocking probability and the handoff blocking probability based on two assumptions: (1) the calls exiting a group of cells are ignored; (2) the handoff calls entering the group from outside is Poissonian. We use a wrap-around model as in [Chlamtac et al., 3; Zeng and Chlamtac, 26], in which the handoff traffic going out of any cell has not been ignored, the Poisson assumption for the handoff traffic into any cell is not used. Our model thus provides more realistic approximation to the new call blocking probability and the handoff call blocking probability.

3. The variance of CRT affects the blocking probability. We observe that the higher the variance of CRT, the higher the new call blocking probability, and the greater the



(a) Low user mobility, new call, $\eta/\mu = 2$

(b) Low user mobility, handoff call, $\eta/\mu = 2$

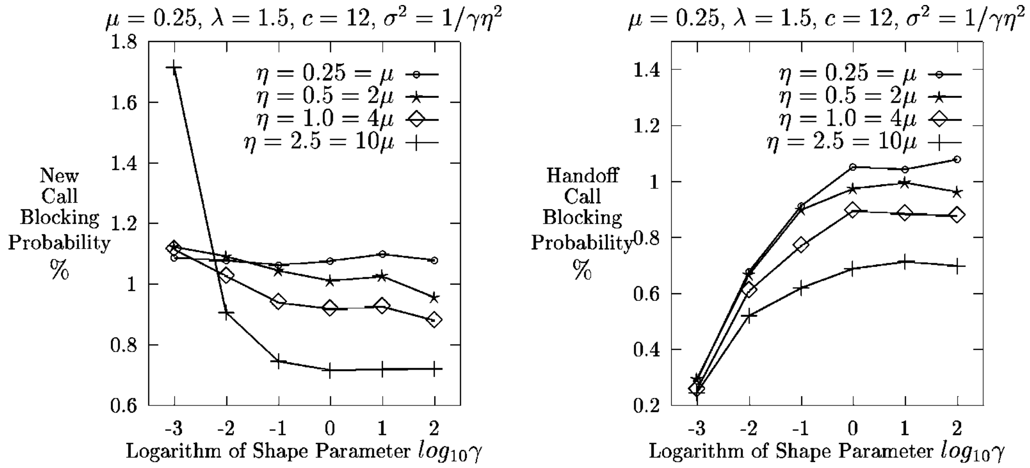
(c) High user mobility, new call, $\eta/\mu = 10$

(d) High user mobility, handoff call, $\eta/\mu = 10$

Figure 10. Comparison of the effect of CRT on new call and handoff call blocking probability under different user mobility.

difference between the new call blocking probability and the handoff call blocking probability.

4. The new call traffic load also affects the blocking probabilities. When the new call arrival rate is low, no matter whether the user mobility is low or high ($\eta/\mu = 2$ or $\eta/\mu = 10$), the new call blocking probability is not very sensitive to the variance of CRT. However, when the new call arrival rate is high, the new call blocking probability is very sensitive to the variance of CRT.



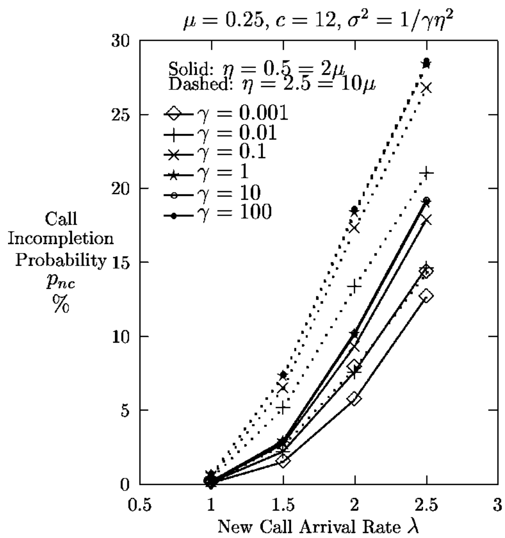
(a) New call (b) Handoff call

Figure 11. Effect of user mobility on call blocking probability, $\lambda = 1.5$.

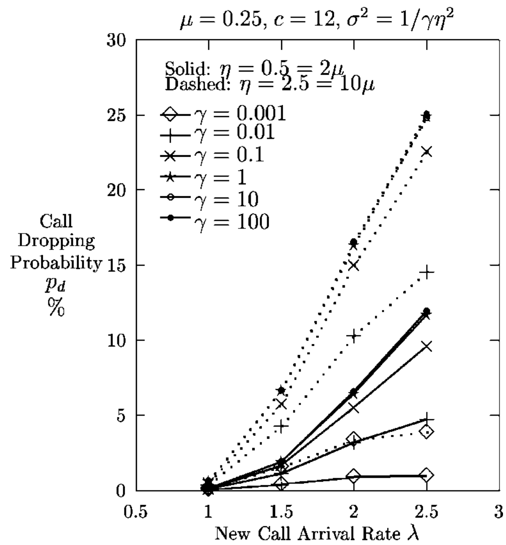
5. The user mobility affects the blocking probabilities. When the user mobility is low, $\eta/\mu = 2$, the new call blocking probability obtained from simulation and those from other assumptions (cases 2 and 4) are almost the same, the blocking probability obtained from case 3 is constant. When the user mobility is high, $\eta/\mu = 10$, the new call blocking probability is underestimated in cases 2 and 4, while the blocking probability obtained from case 3 is still constant.

From figure 11, we observe that the user mobility has opposite effects on the new call blocking probability and the handoff call blocking probability. When the user mobility is high (η/μ is large), the variance of CRT has more significant effect on the new call blocking probability, while when the user mobility is low (η/μ is small), the variance of CRT has more significant effect on the handoff call blocking probability. Thus, when user mobility is high, the difference between the new call blocking probability and the handoff blocking probability is significantly large.

Recall in figure 2, the average CHTs for new calls and handoff calls ($E[t_{n0}]$, $E[t_{h0}]$) are generally different. When the variance of CRT is high ($\gamma < 1$), $E[t_{n0}] > E[t_{h0}]$, the higher the variance, the greater the difference between $E[t_{n0}]$ and $E[t_{h0}]$. Thus, when new calls and handoff calls compete for channels, more handoff calls will be accommodated, the higher the variance of CRT, the greater the difference between p_0 and p_f. When the variance of CRT is low ($\gamma > 1$), $E[t_{n0}] < E[t_{h0}]$, the difference between $E[t_{n0}]$ and $E[t_{h0}]$ is less, thus the difference between p_0 and p_f is less. When CRT is exponential, $E[t_{n0}] = E[t_{h0}]$, so p_0 and p_f are almost equal. When the user mobility is high, both $E[t_{n0}]$ and $E[t_{h0}]$ are small, thus p_0 and p_f are small, but the relative ratio of $E[t_{h0}]/E[t_{n0}]$ is large, thus the difference between p_0 and p_f is significant. We can think an extreme case when the user mobility is extremely high (i.e., $\eta/\mu = \infty$), a handoff call



(a) Call incompletion probability



(b) Call dropping probability

Figure 12. Effect of CRT distribution on call incompletion probability and call dropping probability under different user mobility.

stays in a cell for 0 unit of time, the handoff call blocking probability is 0, but the new call blocking probability is not 0, so the difference between p_0 and p_{f} is very significant.

In figure 12, we show that when the variance of CRT decreases, the call incompletion probability (p_{nc}) increases. When the user mobility is higher, the effect of variance of CRT on call incompletion probability is more significant. Moreover, when the variance of CRT decreases, the call dropping probability (p_{d}) increases. When the user mobility is higher, the effect of variance of CRT on call incompletion probability is more significant.

5. Conclusions

In this paper we carry out a systematic performance study of handoff traffic, the call blocking probability, the call incompletion probability and call dropping probability in PCS networks under generalized CRT distributions. This paper has demonstrated the utility of building a new approach (semi-analytic and semi-simulation) for performance modeling and analysis of wireless mobile networks. Results for some key performance metrics in the high-blocking and the low-blocking environments have been presented and it has been shown that the distribution of CRT, the users' mobility and the new call traffic load all make significant contributions to call blocking performance metrics such as call blocking probabilities, call incompletion probability and call dropping probability. Our research in this paper shows that we should model the CRT with more general model (such as gamma distribution or hyper-Erlang distribution), then apply our semi-analytic and semi-simulation approach to obtain the appropriate approximation for the performance metrics of interest. The results presented can provide guidelines for PCS network design.

Appendix. Handoff call arrival rate

We first define the *handoff probability*. The handoff probability is the probability for a call to generate further handoff(s). The handoff probability for a new call is calculated as follows:

$$P_r(t_c > r_1) = \int_{t_c=0}^{\infty} \int_{r_1=0}^{t_c} f_r(r_1)\mu e^{-\mu t_c}\, dr_1\, dt_c = \mu \int_{t_c=0}^{\infty} F_r(t_c) e^{-\mu t_c}\, dt_c$$
$$= \mu \frac{f_r^*(s)}{s}\bigg|_{s=\mu} = f_r^*(\mu) = \frac{\eta}{\mu}\big[1 - f * (\mu)\big]. \quad (16)$$

The handoff probability for a handoff call is calculated as follows:

$$P_r(r_m > t_m) = \int_{r_m}^{\infty} \int_{t_m=0}^{r_m} f(t_m)\mu e^{-\mu r_m}\, dt_m\, dr_m = \mu \frac{f^*(s)}{s}\bigg|_{s=\mu} = f^*(\mu). \quad (17)$$

A simple way to obtain the handoff arrival rate is as follows: suppose the network is homogeneous, then, the handoff arrival rate is equal to the handoff departure rate:

$$\lambda_h = \lambda_h(1 - p_f)P_r(r_m > t_m) + \lambda(1 - p_0)P_r(t_c > r_1) \quad (18)$$

where p_0 and p_f are the new call blocking probability and the handoff blocking probability, respectively. After plugging in the equation (16) and (17), we obtain

$$\lambda_h = \frac{\eta(1 - p_0)[1 - f^*(\mu)]\lambda}{\mu[1 - (1 - p_f)f^*(\mu)]}. \quad (19)$$

For a complete alternative derivation of handoff arrival rate, please refer to [Fang et al., 11].

References

[1] F. Barcelo and S. Bueno, Idle and inter-arrival time statistics in public access mobile radio (PAMR) systems, in: *Proc. of IEEE Globecom'97*, Phoenix, AZ, November 1997.

[2] F. Barcelo and J. Jordan, Channel holding time distribution in cellular telephony, in: *The 9th Internat. Conf. on Wireless Communications (Wireless'97)*, Vol. 1, Alberta, Canada, 9–11 July 1997, pp. 125–134.

[3] I. Chlamtac, Y. Fang and H. Zeng, Call blocking analysis for PCS networks under general cell residence time, in: *WCNC'99*, New Orleans, September 1999.

[4] E. Chlebus and W. Ludwin, Is handoff traffic really Poissonian? in: *IEEE ICUPC'95*, Tokyo, Japan, 6–10 November 1995, pp. 348–353.

[5] D.R. Cox, *Renewal Theory* (Wiley, New York, 1962).

[6] E. Del Re, R. Fantacci and G. Giambene, Handover and dynamic channel allocation techniques in mobile cellular networks, IEEE Transactions on Vehicular Technology 44 (May 1995) 229–237.
[7] Y. Fang, Hyper-Erlang distribution and its applications in wireless and mobile networks, Wireless Networks (WINET) 7(3) (May 2001) 211–219.
[8] Y. Fang and I. Chlamtac, Teletraffic analysis and mobility modeling for PCS networks, IEEE Transactions on Communications 47(7) (July 1999) 1062–1072.
[9] Y. Fang, I. Chlamtac and Y.B. Lin, Call performance for a PCS networks, IEEE Journal on Selected Areas in Communications 15(8) (October 1997) 1568–1581.
[10] Y. Fang, I. Chlamtac and Y.B. Lin, Modeling PCS networks under general call holding times and cell residence time distributions, IEEE Transactions on Networking 5(6) (December 1997) 893–906.
[11] Y. Fang, I. Chlamtac and Y.B. Lin, Channel occupancy times and handoff rate for PCS networks, IEEE Transactions on Computers 47(6) (1998) 679–692.
[12] R.A. Guerin, Channel occupancy time distribution in a cellular radio system, IEEE Transactions on Vehicular Technology 35(3) (1997) 89–99.
[13] D. Hong and S.S. Rappaport, Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and non-prioritized handoff procedures, IEEE Transactions on Vehicular Technology 35(3) (1986) 77–92.
[14] C. Jedrzycki and V.C.M. Leung, Probability distribution of channel holding time in cellular telephony systems, in: *IEEE Vehicular Technology Conference (VTC'96)*, 1996, pp. 247–251.
[15] F.P. Kelly, Loss networks, The Annals of Applied Probability 1(3) (1991) 319–378.
[16] Y.B. Lin and I. Chlamtac, *Wireless and Mobile Network Architectures* (Wiley, New York).
[17] Y.B. Lin, S. Mohan and A. Noerpel, Queueing priority channel assignment strategies for handoff and initial access for a PCS network, IEEE Transactions on Vehicular Technology 43(3) (1994) 704–712.
[18] I. Mitrani, *Modeling of Computer and Communication Systems* (Cambridge Univ. Press, Cambridge, 1987).
[19] P. Orlik and S.S. Rappaport, A model for teletraffic performance and channel holding time characterization in wireless cellular communication with general session and dwell time distributions, IEEE Journal on Selected Areas in Communications 16(5) (1998) 788–803.
[20] P. Orlik and S.S. Rappaport, On handoff arrival process in cellular communications, in: *Proc. of IEEE WCNC'99*, New Orleans, September 1999.
[21] M. Rajaratnam and F. Takawira, Hand-off traffic modeling in cellular networks, in: *Proc. of Globecom'97*, Phoenix, AZ, November 1997.
[22] M. Rajaratnam and F. Takawira, Nonclassical traffic modleing and performance analysis of cellular mobile networks with and without channel reservation, IEEE Transactions on Vehicular Technology 49(3) (May 2000) 817–834.
[23] M. Sidi and D. Starobinski, New call blocking versus handoff blocking in cellular networks, Wireless Networks 3 (1997) 15–27.
[24] S. Tekinay and B. Jabbari, A measurement-based prioritization scheme for handovers in mobile cellular networks, IEEE Journal on Selected Areas in Communications 10(8) (1992) 1343–1350.
[25] C.H. Yoon and C.K. Un, Performance of personal portable radio telephone systems with and without guard channels, IEEE Journal on Selected Areas in Communications 11(6) (1993) 911–917.
[26] H. Zeng and I. Chlamtac, Handoff traffic distribution in cellular networks, in: *WCNC'99*, New Orleans, September 1999.
[27] M.M. Zonoozi and P. Dassanayake, User mobility modeling and characterization of mobility patterns, IEEE Journal on Selected Areas in Communications 15(7) (September 1997).