

## A FINITE POPULATION MODEL AND PRODUCT-FORM APPROXIMATION FOR CELLULAR MOBILE SYSTEMS WITH TRAVELING USERS

Shigeo Shioda

Kenji Ohtsuka  
*Chiba University*

Fumiaki Machihara  
*Tokyo Denki University*

(Received October 30, 2006; Revised July 05, 2007)

*Abstract* We propose a new finite population model for cellular mobile systems with traveling users. In this model, mobile users arrive according to a Poisson process from outside the system, independently travel in the system, and leave the system in due time. A mobile user is either in call (active) or out of call (inactive). We find that if two minor modifications are made to the model, the joint distribution of the number of calls in progress in each cell has the product form. Making the two modifications is referred to as *product-form approximation* in this paper. Under the product-form approximation, the probability that channels are fully occupied in a cell is given by the Erlang-loss formula. We evaluate the accuracy of the product-form approximation through several simulation experiments, and find that the Erlang-loss formula remains applicable to the performance evaluation and channel provisioning of cellular mobile systems.

**Keywords:** Telecommunication, mathematical modeling, cellular mobile systems, Erlang-loss formula, insensitivity

### 1. Introduction

Investigating the performance characteristics of cellular mobile networks has been the focus of considerable research reported in literature. The performance of cellular mobile systems is often described in terms of the probability of call blocking or that of hand-off blocking. Thus, it is crucial to find a simple method to evaluate these performance indices in order to design cellular mobile networks.

In traditional wired telephone networks, the Erlang-loss formula has been used to evaluate the call blocking probability. The Erlang-loss formula holds when new calls are generated according to a Poisson process. In wired telephone networks, calls are generated by users, the number of which is much larger than the number of available channels, and thus the call arrival rate does not depend on the number of calls in progress. Under such infinite population models, the use of the Poisson-arrival assumption is appropriate. However, the infinite population model may not be applicable to cellular mobile systems because the number of mobile users (mobile units) in a small cell cannot be regarded as “infinite” compared with that of available channels. In this sense, the finite population model would suite cellular mobile systems. Note that, due to user mobility, the number of mobile users within a cell should fluctuate over time. Most existing formulas based on finite population models like the Engset formula, however, do not take the fluctuation of the population over time into account.

Recently, Machihara proposed a finite population model for cellular mobile systems [7]. In Machihara’s model, mobile users arrive according to a Poisson process from outside the system, travel independently within the system, and leave the system in due time. He found that, in this model with a certain modification, the call blocking or hand-off blocking probability is given by the Erlang-loss formula. In addition, he showed that an insensitive property holds; the stationary

distribution of the number of calls in progress depends on the dwell-time and the unencumbered-session-time distributions only through their means.

Machihara focused on a single cell, without explicitly considering the cellular networks. The aim of this paper is to extend Machihara's result to cellular networks. Machihara made a modification on his model (a hand-off user who meets with the handoff blocking immediately departs from the system) in order to analyze the system in the product-form framework. The modification made by Machihara, however, does not yield the product-form distribution in cellular networks. In this paper, we find that two minor modifications are required on the model for obtaining the product-form joint distribution of the number of calls in each cell. Making the two modifications is referred to as *product-form approximation* in this paper. Under the product-form approximation, the probability that channels are fully occupied in a cell is given by the Erlang-loss formula. This suggests that call-blocking and handoff-blocking probabilities can also be evaluated by the Erlang-loss formula. Note that the product-form approximation is equivalent to making the state transition of a cell *quasi reversible* [3, 6].

A large number of studies have been made on the performance of cellular mobile systems [4, 5, 11–16]. Hong *et al.* studied prioritized handoff procedures in cellular systems [4]. Rappaport [14] analyzed the performance of a system in which simultaneous multiple handoffs are generated by moving vehicles. Su *et al.* [15] and Takahashi *et al.* [16] focused on the analysis of the effect of the soft handoff in CDMA systems. Ohmikawa *et al.* [11] analyzed the CDMA systems taking into account the dependence of available channels on the channel utilization in neighboring cells. These studies assumed the calls to arrive according to a Poisson process and the dwell time and unencumbered-session time to be exponentially distributed. Orlik *et al.* [12] studied cellular systems in terms of non-exponential dwell-time and session-time distributions. They also considered the analysis based on MMPP for the handoff arrival process [13].

A significant difference between these existing studies cited above and this paper is in the description of the state. A cellular system is often modeled as a large-scale Markov chain. To reduce the size of the Markov chain, all existing studies used certain approximations, like the state aggregation. In contrast, we apply the product-form framework of the queueing theory [3, 6] to cellular mobile systems. This approach allows us to consider the complete state of the system without any state reduction approximations. This framework is also useful in investigating the insensitivity property of the system [10].

This paper is organized as follows. In Section 2, we describe the system under consideration. In Section 3, we show that the number of users of each cell has the product-form distribution. In Section 4, we consider a cellular mobile system with infinite channels. We show that, in infinite-channel cases, the stationary distribution of the number of calls in progress of each cell has the product form. In Section 5, we consider a finite-channel model. We firstly show that, in finite-channel cases, the number of calls in each cell does not have product-form distribution. Next, we show that, if two minor modifications are made on the model, the joint distribution of the number of calls in progress in each cell has the product form. Thus, under the set of the two modifications (the product-form approximation), the probability that channels are fully occupied in a cell is given by the Erlang-loss formula. In Section 6, we evaluate the accuracy of the product-form approximation through several simulation experiments, and find that the Erlang-loss formula remains applicable to the performance evaluation and channel provisioning of cellular mobile systems.

## 2. Network Model

We consider a cellular mobile network, which consists of  $J$  base stations. Each base station has its own zone, which is referred to as *cell*. The  $i$ th cell has  $c_i$  channels. A mobile user makes a call using one of the channels in each cell. We make the following assumptions concerning the mobile user behavior:

- (1) *External arrival.* Mobile users arrive at cell  $i$  from outside the network following a Poisson process with rate  $\lambda_{0i}$ .
- (2) *User mobility.* A mobile user in a cell independently moves to one of the adjacent cells or departs from the network after his stay in the cell. A mobile user in cell  $i$  moves to cell  $j$  with probability  $p_{ij}$  and departs from the network with probability  $p_{i0}$ . The residing time of the mobile users in each cell is subject to arbitrary distribution. In this paper,  $F_i(t)$  and  $h_i$  respectively denote the distribution function and expectation concerning the residing time of the mobile users in cell  $i$ .
- (3) *Active and inactive users.* A mobile user is either in call or out of call respectively. If a mobile user is in (out of) call mode, we say *he is active (inactive)*. The distribution function of the active (inactive) duration,  $G^{(h)}(t)$  ( $G^{(nh)}(t)$ ), is arbitrary. In this paper,  $t_h$  and  $t_{nh}$  respectively denote the average active and inactive durations. We also define  $\rho \stackrel{\text{def}}{=} t_h/(t_h + t_{nh})$ .
- (4) *Call generation.* An inactive user requests a call at the end of the inactive period. If there are available channels at this time, then his request is accepted and he changes into active mode.
- (5) *Call blocking.* A call request of an inactive mobile user is blocked if there is no available channel. When the call request is blocked, the inactive mobile user enters the inactive state again, and after completion of the inactive period, he will try to make a call again.
- (6) *Hand-off blocking.* An active mobile user, who moves into cell  $i$  from one of the adjacent cells, continues calling only when there are available channels in cell  $i$ ; otherwise, the call is interrupted (hand-off blocking). The active mobile user, whose call is interrupted by hand-off blocking, changes his state to inactive, and will try to make a call again after the inactive period has passed.
- (7) *States of new mobile users.* A new mobile user arriving from outside of the network is active with probability  $\rho$  and inactive with probability  $1 - \rho$ . The remaining active time (call holding time) of a new active mobile user has the following distribution:

$$G_e^{(h)}(t) = \frac{1}{t_h} \int_0^t (1 - G^{(h)}(u)) du.$$

Note that  $G_e^{(h)}(t)$  is the equilibrium distribution of  $G^{(h)}(t)$ . Similarly, the remaining inactive time of a new inactive mobile user has the following distribution:

$$G_e^{(nh)}(t) = \frac{1}{t_{nh}} \int_0^t (1 - G^{(nh)}(u)) du,$$

which is the equilibrium distribution of  $G^{(nh)}(t)$ .

## 3. Stationary Distribution of the Number of Mobile Users

In terms of the number of mobile users in each cell, the network model under consideration is equivalent to Kelly (or BCMP) networks where each station has infinite servers [2, 6]. Thus, the joint distribution of the number of mobile users in each cell has the product form. To be more precise, let  $L_j(\geq 0)$  be a random variable expressing the number of mobile users in cell  $j$ , and let

$l_j(\geq 0)$  be an outcome of  $L_j$ . We also let  $\mathbf{L} \stackrel{\text{def}}{=} (L_1, \dots, L_J)$  and  $\mathbf{l} \stackrel{\text{def}}{=} (l_1, \dots, l_J)$ . Then, the stationary distribution of  $\mathbf{L}$ , denoted by  $p_u(\mathbf{l}) \stackrel{\text{def}}{=} P[\mathbf{L} = \mathbf{l}]$ , is given by

$$p_u(\mathbf{l}) = \prod_{i=1}^J \frac{a_i^{l_i}}{l_i!} e^{-a_i} \mathbf{1}(0 \leq l_i), \quad a_i \stackrel{\text{def}}{=} \lambda_i h_i, \tag{3.1}$$

where  $\mathbf{1}(A)$  denotes the indicator function, which is 1 when  $A$  is true and 0 when  $A$  is false, and  $(\lambda_1, \dots, \lambda_J)$  is the solution to the following traffic equation:

$$\lambda_i = \lambda_{0i} + \sum_{j=1; j \neq i}^J \lambda_j p_{ji}. \tag{3.2}$$

Variable  $\lambda_i$  in (3.1) and (3.2) corresponds to the overall arrival rate to cell  $i$ , including both external arrival and internal transitions, and thus  $a_i$  is the average number of mobile users in cell  $i$ . The marginal distribution of the number of mobile users in cell  $i$  is given by  $p_u(l_i) = \frac{a_i^{l_i}}{l_i!} e^{-a_i}$ .

Note that, although the external arrivals to each cell follow a Poisson process, total arrivals to each cell, including internal transitions, do not. This is because there is some correlation between arrivals and the system due to the feedback; a mobile user departing a cell may return to this cell again.

#### 4. The Stationary Distribution of the Number of Calls in Progress: Infinite Channel Case

##### 4.1. Derivation of the stationary distribution

If each cell has an infinite number of channels, the stationary joint distribution of the number of calls in progress in each cell has the product form. To show this, let  $N_j(\geq 0)$  be a random variable expressing the number of calls (number of active users) in cell  $j$ , and let  $n_j(\geq 0)$  be an outcome of  $N_j$ . We also let  $\mathbf{N} \stackrel{\text{def}}{=} (N_1, \dots, N_J)$  and  $\mathbf{n} \stackrel{\text{def}}{=} (n_1, \dots, n_J)$ .

**Theorem 4.1.** *The stationary distribution of  $\mathbf{N}$ , denoted by  $\pi(\mathbf{n}) \stackrel{\text{def}}{=} P[\mathbf{N} = \mathbf{n}]$ , is given by*

$$\pi(\mathbf{n}) = \prod_{i=1}^J \frac{(a_i \rho)^{n_i}}{n_i!} e^{-a_i \rho} \mathbf{1}(0 \leq n_i). \tag{4.1}$$

*Proof.* In the infinite channel case, a mobile user is active with probability  $\rho$  and inactive with probability  $1 - \rho$ . The states (active or inactive) of mobile users are mutually independent from each other. The state of a given mobile user is also independent of the number of users in each cell. Thus, the conditional probability distribution of  $\mathbf{N}$  when the number of users in each cell is  $\mathbf{l}$ , denoted by  $\pi(\mathbf{n}|\mathbf{l}) \stackrel{\text{def}}{=} P[\mathbf{N} = \mathbf{n} | \mathbf{L} = \mathbf{l}]$ , is given by

$$\pi(\mathbf{n}|\mathbf{l}) = \prod_{i=1}^J \binom{l_i}{n_i} \rho^{n_i} (1 - \rho)^{l_i - n_i} \mathbf{1}(0 \leq n_i \leq l_i).$$

Combining (3.1) with the above equation yields

$$\begin{aligned} \pi(\mathbf{n}) &= \sum_{l_1=0}^{\infty} \cdots \sum_{l_J=0}^{\infty} \pi(\mathbf{n}|\mathbf{l}) \prod_{i=1}^J \frac{a_i^{l_i}}{l_i!} e^{-a_i} \\ &= \sum_{l_1=n_1}^{\infty} \cdots \sum_{l_J=n_J}^{\infty} \prod_{i=1}^J \binom{l_i}{n_i} \rho^{n_i} (1 - \rho)^{l_i - n_i} \frac{a_i^{l_i}}{l_i!} e^{-a_i} \mathbf{1}(0 \leq n_i) \end{aligned}$$

$$\begin{aligned}
&= \prod_{i=1}^J \sum_{l_i=n_i}^{\infty} \frac{a_i^{n_i}}{n_i!} \rho^{n_i} \frac{a_i^{l_i-n_i}}{(l_i-n_i)!} (1-\rho)^{l_i-n_i} e^{-a_i} \mathbf{1}(0 \leq n_i) \\
&= \prod_{i=1}^J \frac{(a_i \rho)^{n_i}}{n_i!} e^{-a_i} \sum_{l_i=n_i}^{\infty} \frac{((1-\rho)a_i)^{l_i-n_i}}{(l_i-n_i)!} \mathbf{1}(0 \leq n_i) \\
&= \prod_{i=1}^J \frac{(a_i \rho)^{n_i}}{n_i!} e^{-a_i \rho} \mathbf{1}(0 \leq n_i),
\end{aligned}$$

which completes the proof.  $\square$

Let  $M_j(\geq 0)$  be a random variable denoting the number of inactive users in cell  $j$ , and let  $m_j(\geq 0)$  be an outcome of  $M_j$ . The number of inactive users in each cell,  $\mathbf{M} \stackrel{\text{def}}{=} (M_1, \dots, M_J)$ , also has the product form distribution. In addition to this, the joint distribution of  $N$  and  $\mathbf{M}$  has the product form, as shown in the following theorem:

**Theorem 4.2.** *The joint distribution of  $N$  and  $\mathbf{M}$ , denoted by  $p(\mathbf{n}, \mathbf{m}) \stackrel{\text{def}}{=} P[N = \mathbf{n}, \mathbf{M} = \mathbf{m}]$ , is given by*

$$p(\mathbf{n}, \mathbf{m}) = \prod_{i=1}^J e^{-a_i} \frac{(\rho a_i)^{n_i}}{n_i!} \frac{((1-\rho)a_i)^{m_i}}{m_i!} \mathbf{1}(0 \leq n_i) \mathbf{1}(0 \leq m_i). \quad (4.2)$$

*Proof.* Observe that

$$\begin{aligned}
p(\mathbf{n}, \mathbf{m}) &= P[N = \mathbf{n}, \mathbf{M} = \mathbf{m}] \\
&= P[N = \mathbf{n}, \mathbf{L} = \mathbf{m} + \mathbf{n}] \\
&= P[N = \mathbf{n} | \mathbf{L} = \mathbf{m} + \mathbf{n}] P[\mathbf{L} = \mathbf{m} + \mathbf{n}] \\
&= \pi(\mathbf{n} | \mathbf{n} + \mathbf{m}) p_u(\mathbf{n} + \mathbf{m}) \\
&= \prod_{i=1}^J \frac{a_i^{(n_i+m_i)}}{(n_i+m_i)!} e^{-a_i} \binom{n_i+m_i}{n_i} \rho^{n_i} (1-\rho)^{m_i} \mathbf{1}(0 \leq n_i \leq n_i+m_i) \mathbf{1}(0 \leq n_i+m_i) \\
&= \prod_{i=1}^J e^{-a_i} \frac{(\rho a_i)^{n_i}}{n_i!} \frac{((1-\rho)a_i)^{m_i}}{m_i!} \mathbf{1}(0 \leq n_i) \mathbf{1}(0 \leq m_i),
\end{aligned}$$

which completes the proof.  $\square$

The above theorem implies that the number of active users and that of inactive users are mutually independent in infinite channel cases.

#### 4.2. Balance equation

If the residing time, the active duration, and the inactive duration are all exponentially distributed, the transition of  $\{N, \mathbf{M}\}$  is governed by a Markov chain. For comparison with the finite channel case, we show the balance equation of the Markov chain. In the following, we let  $\mu_i \stackrel{\text{def}}{=} 1/h_i$ ,  $\mu^{(h)} \stackrel{\text{def}}{=} 1/t_h$ , and  $\mu^{(nh)} \stackrel{\text{def}}{=} 1/t_{nh}$ . We also let  $q((\mathbf{n}, \mathbf{m}), (\mathbf{n}', \mathbf{m}'))$  be the transition rate from state  $(\mathbf{n}, \mathbf{m}) \stackrel{\text{def}}{=} \{N = \mathbf{n}, \mathbf{M} = \mathbf{m}\}$  to state  $(\mathbf{n}', \mathbf{m}') \stackrel{\text{def}}{=} \{N = \mathbf{n}', \mathbf{M} = \mathbf{m}'\}$ .

In the network under consideration, there are eight driving processes that cause the state transitions: 1) new active user arrivals, 2) new inactive user arrivals, 3) generation of new calls, 4) completion of calls, 5) arrival of hand-offs (arrival of active users from neighboring cells), 6) arrival of inactive users from neighboring cells, 7) active user departures, and 8) inactive user departures. The transition rates of these driving processes from state  $(\mathbf{n}, \mathbf{m})$ , where  $\mathbf{0} \leq \mathbf{n}, \mathbf{0} \leq \mathbf{m}^1$ ,

<sup>1</sup>The inequality between vectors means component-wise inequality;  $\mathbf{0}$  denotes a vector whose elements are all equal to 0.

are given below:

- 1)  $q((\mathbf{n}, \mathbf{m}), (\mathbf{n} + \mathbf{e}_i, \mathbf{m})) = \rho\lambda_{0i}$ ,
- 2)  $q((\mathbf{n}, \mathbf{m}), (\mathbf{n}, \mathbf{m} + \mathbf{e}_i)) = (1 - \rho)\lambda_{0i}$ ,
- 3)  $q((\mathbf{n}, \mathbf{m}), (\mathbf{n} + \mathbf{e}_i, \mathbf{m} - \mathbf{e}_i)) = \mu^{(nh)}m_i$ ,
- 4)  $q((\mathbf{n}, \mathbf{m}), (\mathbf{n} - \mathbf{e}_i, \mathbf{m} + \mathbf{e}_i)) = \mu^{(h)}n_i$ ,
- 5)  $q((\mathbf{n}, \mathbf{m}), (\mathbf{n} - \mathbf{e}_i + \mathbf{e}_j, \mathbf{m})) = \mu_in_ip_{ij}$ ,
- 6)  $q((\mathbf{n}, \mathbf{m}), (\mathbf{n}, \mathbf{m} - \mathbf{e}_i + \mathbf{e}_j)) = \mu_im_ip_{ij}$ ,
- 7)  $q((\mathbf{n}, \mathbf{m}), (\mathbf{n} - \mathbf{e}_i, \mathbf{m})) = \mu_in_ip_{i0}$ ,
- 8)  $q((\mathbf{n}, \mathbf{m}), (\mathbf{n}, \mathbf{m} - \mathbf{e}_i)) = \mu_im_ip_{i0}$ ,

where  $\mathbf{e}_i$  denote the vector whose  $i$ th element is 1 and other elements are 0. The balance equation of the system, which equates the probability flow out of state  $(\mathbf{n}, \mathbf{m})$  with the probability flow into the same state, is given by

$$\begin{aligned}
 & p(\mathbf{n}, \mathbf{m}) \sum_{i=1}^J [\lambda_{0i} + \mu_i(n_i + m_i) + \mu^{(nh)}m_i + \mu^{(h)}n_i] \\
 &= \sum_{i=1}^J [p(\mathbf{n} - \mathbf{e}_i, \mathbf{m})\rho\lambda_{0i}\mathbf{1}(n_i > 0) + p(\mathbf{n}, \mathbf{m} - \mathbf{e}_i)(1 - \rho)\lambda_{0i}\mathbf{1}(m_i > 0)] \\
 &+ \sum_{i=1}^J [p(\mathbf{n} + \mathbf{e}_i, \mathbf{m})\mu_i(n_i + 1)p_{i0} + p(\mathbf{n}, \mathbf{m} + \mathbf{e}_i)\mu_i(m_i + 1)p_{i0}] \\
 &+ \sum_{i=1}^J [p(\mathbf{n} + \mathbf{e}_i, \mathbf{m} - \mathbf{e}_i)\mu^{(h)}(n_i + 1)\mathbf{1}(m_i > 0) + p(\mathbf{n} - \mathbf{e}_i, \mathbf{m} + \mathbf{e}_i)\mu^{(nh)}(m_i + 1)\mathbf{1}(n_i > 0)] \\
 &+ \sum_{i=1}^J \sum_{j \neq i}^J [p(\mathbf{n} + \mathbf{e}_i - \mathbf{e}_j, \mathbf{m})\mu_i(n_i + 1)p_{ij}\mathbf{1}(n_j > 0) + p(\mathbf{n}, \mathbf{m} + \mathbf{e}_i - \mathbf{e}_j)\mu_i(m_i + 1)p_{ij}\mathbf{1}(m_j > 0)].
 \end{aligned} \tag{4.3}$$

Now we show (4.2) satisfies the balance equation (4.3). For this purpose, we divide (4.3) into the following three partial balance equations:

$$\begin{aligned}
 & p(\mathbf{n}, \mathbf{m}) \sum_{i=1}^J [\mu^{(nh)}m_i + \mu^{(h)}n_i] \\
 &= \sum_{i=1}^J [p(\mathbf{n} + \mathbf{e}_i, \mathbf{m} - \mathbf{e}_i)(n_i + 1)\mu^{(h)}\mathbf{1}(m_i > 0) + p(\mathbf{n} - \mathbf{e}_i, \mathbf{m} + \mathbf{e}_i)\mu^{(nh)}(m_i + 1)\mathbf{1}(n_i > 0)],
 \end{aligned} \tag{4.4}$$

$$p(\mathbf{n}, \mathbf{m}) \sum_{i=1}^J \lambda_{0i} = \sum_{i=1}^J [p(\mathbf{n} + \mathbf{e}_i, \mathbf{m})\mu_i(n_i + 1)p_{i0} + p(\mathbf{n}, \mathbf{m} + \mathbf{e}_i)\mu_i(m_i + 1)p_{i0}], \tag{4.5}$$

$$\begin{aligned}
 & p(\mathbf{n}, \mathbf{m}) \sum_{i=1}^J \mu_i(n_i + m_i) = \sum_{i=1}^J [p(\mathbf{n} - \mathbf{e}_i, \mathbf{m})\rho\lambda_{0i}\mathbf{1}(n_i > 0) + p(\mathbf{n}, \mathbf{m} - \mathbf{e}_i)(1 - \rho)\lambda_{0i}\mathbf{1}(m_i > 0)] \\
 &+ \sum_{i=1}^J \sum_{j \neq i}^J [p(\mathbf{n} + \mathbf{e}_i - \mathbf{e}_j, \mathbf{m})\mu_i(n_i + 1)p_{ij}\mathbf{1}(n_j > 0) + p(\mathbf{n}, \mathbf{m} + \mathbf{e}_i - \mathbf{e}_j)\mu_i(m_i + 1)p_{ij}\mathbf{1}(m_j > 0)].
 \end{aligned} \tag{4.6}$$

If (4.2) satisfies all the partial balance equations, then (4.2) satisfies the balance equation. It follows from  $\mu^{(h)}/\mu^{(nh)} = (1 - \rho)/\rho$  that (4.2) satisfies (4.4). Next, observe that (4.2) yields

$$p(\mathbf{n} - \mathbf{e}_i, \mathbf{m})\rho\lambda_{0i}\mathbf{1}(n_i > 0) + p(\mathbf{n}, \mathbf{m} - \mathbf{e}_i)(1 - \rho)\lambda_{0i}\mathbf{1}(m_i > 0) = p(\mathbf{n}, \mathbf{m})\frac{(n_i + m_i)\lambda_{0i}}{a_i},$$

$$\begin{aligned}
& \sum_{i \neq j}^J [p(\mathbf{n} + \mathbf{e}_i - \mathbf{e}_j, \mathbf{m}) \mu_i (n_i + 1) p_{ij} \mathbf{1}(n_j > 0) + p(\mathbf{n}, \mathbf{m} + \mathbf{e}_i - \mathbf{e}_j) \mu_i (m_i + 1) p_{ij} \mathbf{1}(m_j > 0)] \\
&= p(\mathbf{n}, \mathbf{m}) \sum_{i \neq j}^J \frac{\lambda_i}{\lambda_j} \mu_j p_{ij} (n_j + m_j) \\
&= p(\mathbf{n}, \mathbf{m}) \left(1 - \frac{\lambda_{0j}}{\lambda_j}\right) \mu_j (n_j + m_j) \\
&= p(\mathbf{n}, \mathbf{m}) \left(\mu_j (n_j + m_j) - \frac{(n_j + m_j) \lambda_{0j}}{a_j}\right),
\end{aligned}$$

from which (4.6) follows. Also observe that (4.2) yields

$$p(\mathbf{n} + \mathbf{e}_i, \mathbf{m}) \mu_i (n_i + 1) p_{i0} + p(\mathbf{n}, \mathbf{m} + \mathbf{e}_i) \mu_i (m_i + 1) p_{i0} = p(\mathbf{n}, \mathbf{m}) a_i \mu_i p_{i0}.$$

Thus, (4.5) also follows if

$$\sum_{i=1}^J \lambda_{0i} = \sum_{i=1}^J a_i \mu_i p_{i0},$$

which can be derived from the traffic equation (3.2). From these considerations, we know that (4.2) certainly satisfies the balance equation and thus (4.2) is indeed the stationary distribution.

**Remark 4.3.** Strictly speaking, the fact that (4.2) is a probability solution satisfying the balance equation does not always assure that (4.2) is the stationary distribution of the continuous-time Markov chain. If the traffic equation (3.2) has the unique solution, however, (4.2) is indeed the unique stationary distribution. To see this, first observe that the considered Markov chain is irreducible because all states are accessible from and to state  $(\mathbf{0}, \mathbf{0})$ . Next, observe that

$$\begin{aligned}
& \sum_{\mathbf{n}} \sum_{\mathbf{m}} p(\mathbf{n}, \mathbf{m}) \sum_{i=1}^J [\lambda_{0i} + \mu_i (n_i + m_i) + \mu^{(nh)} m_i + \mu^{(h)} n_i] \\
&= \sum_{i=1}^J [\lambda_{0i} + a_i \mu_i + \rho a_i \mu^{(h)} + (1 - \rho) a_i \mu^{(nh)}] < \infty.
\end{aligned} \tag{4.7}$$

It thus follows from Corollary 4.4 in [1] (page 53) that the considered Markov chain is ergodic and thus (4.2) is the unique stationary distribution.

## 5. The Stationary Distribution of the Number of Calls in Progress: Finite Channel Case

Next, we consider the finite channel case; that is, where each cell has a finite number of channels. In the finite channel case, call or hand-off blockings occur, which yields the correlation between the behaviors of mobile users. Due to this correlation, the distribution of the number of calls in progress is no longer given by the product form. In this section, we discover that, by making two modifications to the model, we can analyze the system in the product-form framework.

### 5.1. Product-form approximation: single cell

Firstly, we show that, in the finite-channel case, the model under consideration does not have the product-form distribution such as (4.1) or (4.2). To this end, we consider a single cell where the active and inactive durations are all exponentially distributed. In this case, the state of the system is described by  $\{N_1, M_1\}$ , the transition of which is governed by a Markov chain. The transition rates from state  $(n_1, m_1) \stackrel{\text{def}}{=} \{N_1 = n_1, M_1 = m_1\}$ , where  $0 \leq n_1 \leq c_1, 0 \leq m_1$ , are given below:

$$q((n_1, m_1), (n_1 + 1, m_1)) = \rho \lambda_{01} \mathbf{1}(n_1 < c_1),$$

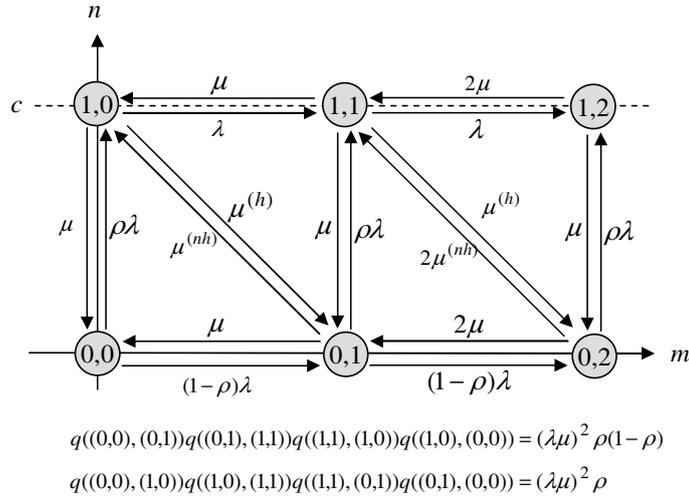


Figure 1: State transition in the finite-channel case: without modification

$$\begin{aligned}
 q((n_1, m_1), (n_1, m_1 + 1)) &= \begin{cases} (1-\rho)\lambda_{01} & n_1 < c_1, \\ \lambda_{01} & n_1 = c_1, \end{cases} \\
 q((n_1, m_1), (n_1 + 1, m_1 - 1)) &= \mu^{(nh)} m_1 \mathbf{1}(n_1 < c_1), \\
 q((n_1, m_1), (n_1 - 1, m_1 + 1)) &= \mu^{(h)} n_1, \\
 q((n_1, m_1), (n_1 - 1, m_1)) &= \mu_1 n_1, \\
 q((n_1, m_1), (n_1, m_1 - 1)) &= \mu_1 m_1.
 \end{aligned}$$

Now, we assume that  $(N_1, M_1)$  has the following stationary distribution:

$$\begin{aligned}
 p(n_1, m_1) &\stackrel{\text{def}}{=} P[N_1 = n_1, M_1 = m_1] \\
 &= g \frac{(\rho a_1)^{n_1}}{n_1!} e^{-(1-\rho)a_1} \frac{((1-\rho)a_1)^{m_1}}{m_1!} \mathbf{1}(0 \leq n_1 \leq c_1) \mathbf{1}(0 \leq m_1), \\
 g &\stackrel{\text{def}}{=} 1 / \sum_{k=0}^{c_1} \frac{(\rho a_1)^k}{k!}.
 \end{aligned} \tag{5.1}$$

Let us consider the time-reversal Markov Chain [3], the transition rate of which is given by

$$\tilde{q}((n, m), (n', m')) = q((n', m'), (n, m)) \frac{p(n', m')}{p(n, m)}.$$

If the stationary distribution of the original Markov Chain is given by (5.1), the transition rates in the time-reversed network are

$$\begin{aligned}
 \tilde{q}((n_1, m_1), (n_1 + 1, m_1)) &= \rho \lambda_{01} \mathbf{1}(n_1 < c_1), \\
 \tilde{q}((n_1, m_1), (n_1, m_1 + 1)) &= (1-\rho)\lambda_{01}, \\
 \tilde{q}((n_1, m_1), (n_1 + 1, m_1 - 1)) &= \mu^{(nh)} m_1 \mathbf{1}(n_1 < c_1), \\
 \tilde{q}((n_1, m_1), (n_1 - 1, m_1 + 1)) &= \mu^{(h)} n_1, \\
 \tilde{q}((n_1, m_1), (n_1 - 1, m_1)) &= \mu_1 n_1, \\
 \tilde{q}((n_1, m_1), (n_1, m_1 - 1)) &= \mu_1 m_1,
 \end{aligned}$$

where we use the fact that  $a_1 \mu_1 = \lambda_1$ . As you see, the transition rates in the original Markov chain and those in the time-reversed Markov chain differ. Because of this, the original Markov chain is not quasi reversible [3, 6] and (5.1) does not satisfy the balance equation.

This fact is also confirmed by the following observation: in the finite-channel case, the transition rates  $q((n_1, m_1), (n_1, m_1 + 1))$  with  $n_1 < c_1$  and  $q((c_1, m_1), (c_1, m_1 + 1))$  are different. Thus, two paths in the state space, which start at  $(0, 0)$ , go to  $(1, 1)$ , and return to  $(0, 0)$  again, have different probabilities as shown in Figure 1. This means that Kolmogorov's criterion [6] is not satisfied and thus the Markov chain under consideration is not reversible.

Now, we make the following two modifications to the model:

**[Modification 1]** An active mobile user, who enters cell  $i$  but encounters hand-off blocking, changes his state into the *call-suspending* state. The call-suspending user releases the channel and does not change the state until leaving cell  $i$ .

**[Modification 2]** Call-suspending users in cell  $i$  depart from cell  $i$  according to a Poisson process with rate  $\rho\lambda_i$  only when the channels in cell  $i$  are all occupied. A call-suspending user, who departs from cell  $i$ , goes to cell  $j$  with probability  $p_{ij}$  or departs from the network with probability  $p_{i0}$ . If he moves to one of the neighboring cells, he tries to hold a channel in order to restart the call as soon as he enters there.

With modifications 1 and 2, the transition rate  $q((n_1, m_1), (n_1, m_1 + 1))$  of the Markov chain is always equal to  $(1 - \rho)\lambda_{01}$  despite the value of  $n_1$ .

As shown in Figure 2, under the above mentioned modifications, Kolmogorov's criterion [6] is satisfied, and thus the Markov chain governing the transition of  $(n_1, m_1)$  is reversible. It is also easy to see that (5.1) is the stationary distribution of  $(n_1, m_1)$  of the original Markov chain under the two modifications. Making modifications 1 and 2 on the model is referred to as the *product-form approximation* in the paper.

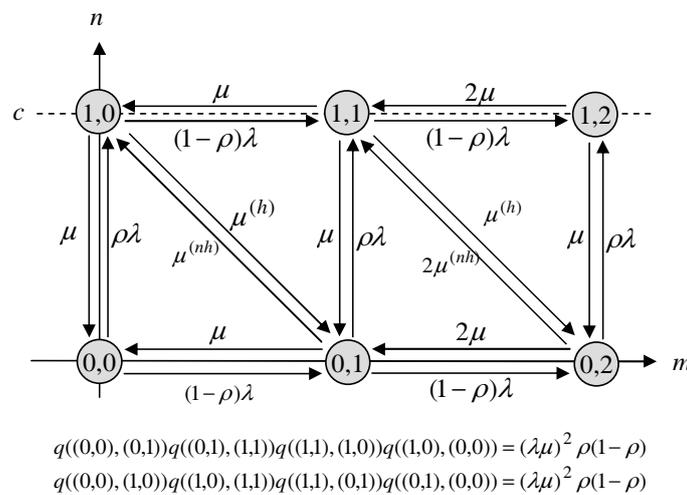


Figure 2: State transition in the finite-channel case: under the product-form condition

The probability that channels are all occupied is given by the Erlang-loss formula

$$b = \frac{(\rho a_1)^{c_1}}{c_1!} / \sum_{k=0}^{c_1} \frac{(\rho a_1)^k}{k!}. \tag{5.2}$$

Equation (5.2) gives the probability that channels are fully occupied at an arbitrary epoch. We call it the *channel-full-occupancy probability*. Note that call-blocking probability is defined by the one that channels are fully occupied at the call generation epoch of an inactive user, and the handoff-blocking probability is the one that channels are fully occupied at the arrival epoch of

an active user. In spite of the difference in their definition, we could conjecture that the call-blocking, the handoff-blocking, and the channel-full-occupancy probabilities are all the same. For a single cell case, we can easily confirm this conjecture. To see this, first note that active users arrive according to a Poisson process from outside the network<sup>2</sup>. Since Poisson arrivals see the time-average behavior of the cell (PASTA property), the handoff-blocking probability is equal to the channel-full-occupancy probability. Next observe that, when the inactive duration is exponentially distributed, inactive users generate calls according to a Poisson process and call generation ratio depends only on the number of inactive users, which is statistically independent of the number of occupied channels as shown in the product form of (5.1). Thus, it also follows from the PASTA property that the call-blocking probability is equal to the channel-full-occupancy probability. Concerning the case where the inactive duration is not exponential, please see Section 5.4.

**Remark 5.1.** Although modification 2 is not required in a single-cell model, it is essential in cellular networks. Both modifications are necessary for making the transition of the state *quasi reversible* [3, 6].

**Remark 5.2.** The departure rate of call-suspending users from cell  $i$  is always fixed at  $\rho\lambda_1$  regardless of how many users have encountered hand-off blocking and thus have become call-suspending users so far. This means that the number of call-suspending users does not affect the state transition of the Markov Chain considered in this section. The number of call-suspending users is not a variable describing the Markov Chain. (The number of mobile users in the call-suspending state might become negative.) Assuming modifications 1 and 2 is equivalent to that users encountering the hand-off blocking once disappear from the system but appear again in the same cell with a constant rate and immediately move to one of adjacent cells.

**Remark 5.3.** Machihara made the following modification in [7]: active mobile users, whose calls are interrupted by hand-off blocking, change their state to inactive and no longer make calls. This modification is equivalent to modification 1 in a single cell model. It is not, however, sufficient for cellular networks and thus modification 2 is required.

### 5.2. Insensitivity

In Sec. 5.1, we derive the stationary distribution of  $\{N_1, M_1\}$  when the residing time, active duration, and inactive duration are all exponentially distributed. Machihara [7] showed an insensitivity property of (5.1) in the sense that (5.1) depends on the distributions of the residing time and the active duration only through their means. Here we show that a stronger insensitivity property holds; that is, (5.1) depends on the distributions of the residing time, active duration, and inactive duration only through their means. To do this, we let  $\mathbf{R} \stackrel{\text{def}}{=} (R_1, \dots, R_{n_1+m_1})$ ,  $\mathbf{R}^{(h)} \stackrel{\text{def}}{=} (R_1^{(h)}, \dots, R_{n_1}^{(h)})$ ,  $\mathbf{R}^{(nh)} \stackrel{\text{def}}{=} (R_1^{(nh)}, \dots, R_{m_1}^{(nh)})$ , where  $R_i$  is the residual residing time,  $R_i^{(h)}$  is the residual active time, and  $R_i^{(nh)}$  is the residual inactive time of mobile user  $i$  when there are  $n_1$  active users and  $m_1$  inactive users in the cell.

#### Theorem 5.4.

$$\begin{aligned}
 &P[(N_1, M_1) = (n_1, m_1), \mathbf{R} \leq \mathbf{t}, \mathbf{R}^{(h)} \leq \mathbf{t}^{(h)}, \mathbf{R}^{(nh)} \leq \mathbf{t}^{(nh)}] \\
 &= g \frac{(\rho a_1)^{n_1}}{n_1!} e^{-(1-\rho)a_1} \frac{((1-\rho)a_1)^{m_1}}{m_1!} \\
 &\quad \times \prod_{i=1}^{n_1+m_1} F_e(t_i) \prod_{i=1}^{n_1} G_e^{(h)}(t_i^{(h)}) \prod_{i=1}^{m_1} G_e^{(nh)}(t_i^{(nh)}) \mathbf{1}(0 \leq n_1 \leq c_1) \mathbf{1}(0 \leq m_1) \tag{5.3}
 \end{aligned}$$

<sup>2</sup>It is not the case for a multi-cell network.

where,  $\mathbf{t} \stackrel{\text{def}}{=} (t_1, \dots, t_{n_1+m_1})$ ,  $\mathbf{t}^{(h)} \stackrel{\text{def}}{=} (t_1^{(h)}, \dots, t_{n_1}^{(h)})$ ,  $\mathbf{t}^{(nh)} \stackrel{\text{def}}{=} (t_1^{(nh)}, \dots, t_{m_1}^{(nh)})$ .  $F_e(\mathbf{t})$  is the equilibrium distribution of  $F(\mathbf{t})$ .

*Proof.* Please see Appendix. □

The insensitivity property of (5.1) is a direct consequence of Theorem 5.4.

### 5.3. Product-form approximation: cellular networks

The product-form approximation introduced in Sec. 5.1 makes the state-transition of the cell quasi-reversible. Thus, we can expect that, under the product-form approximation, the number of calls in progress in each cell has the product form. This conjecture is confirmed by the following theorem:

**Theorem 5.5.** *If the residing time, active time, and inactive time are all exponentially distributed, then the joint distribution of  $\mathbf{N}$  and  $\mathbf{M}$  has the following product form under the product-form approximation:*

$$p(\mathbf{n}, \mathbf{m}) = g \prod_{i=1}^J e^{-(1-\rho)a_i} \frac{(\rho a_i)^{n_i}}{n_i!} \frac{((1-\rho)a_i)^{m_i}}{m_i!} \mathbf{1}(0 \leq n_i \leq c_i) \mathbf{1}(0 \leq m_i),$$

$$g \stackrel{\text{def}}{=} 1 / \prod_{i=1}^J \sum_{k=0}^{c_i} \frac{(\rho a_i)^k}{k!}. \quad (5.4)$$

*Proof.* In the cellular network model under the product-form approximation, there are nine driving processes that cause the state transitions: 1) new active user arrivals, 2) new inactive user arrivals, 3) generation of new calls, 4) completion of calls, 5) arrival of hand-offs (arrival of active users from neighboring cells), 6) arrival of inactive users from neighboring cells, 7) arrival of suspending users from neighboring cells, 8) active user departures, and 9) inactive user departures. The transition rates by these driving processes from state  $(\mathbf{n}, \mathbf{m})$ , where  $\mathbf{0} \leq \mathbf{n} \leq \mathbf{c}$ ,  $\mathbf{0} \leq \mathbf{m}$ , are given below:

- 1)  $q((\mathbf{n}, \mathbf{m}), (\mathbf{n} + \mathbf{e}_i, \mathbf{m})) = \lambda_{0i} \rho \mathbf{1}(n_i < c_i)$ ,
- 2)  $q((\mathbf{n}, \mathbf{m}), (\mathbf{n}, \mathbf{m} + \mathbf{e}_i)) = (1 - \rho) \lambda_{0i}$ ,
- 3)  $q((\mathbf{n}, \mathbf{m}), (\mathbf{n} + \mathbf{e}_i, \mathbf{m} - \mathbf{e}_i)) = \mu^{(nh)} m_i \mathbf{1}(n_i < c_i)$ ,
- 4)  $q((\mathbf{n}, \mathbf{m}), (\mathbf{n} - \mathbf{e}_i, \mathbf{m} + \mathbf{e}_i)) = \mu^{(h)} n_i$ ,
- 5)  $q((\mathbf{n}, \mathbf{m}), (\mathbf{n} - \mathbf{e}_i + \mathbf{e}_j, \mathbf{m})) = \mu_i n_i p_{ij} \mathbf{1}(n_j < c_j)$ ,
- 6)  $q((\mathbf{n}, \mathbf{m}), (\mathbf{n}, \mathbf{m} - \mathbf{e}_i + \mathbf{e}_j)) = \mu_i m_i p_{ij}$ ,
- 7)  $q((\mathbf{n}, \mathbf{m}), (\mathbf{n} + \mathbf{e}_i, \mathbf{m})) = \sum_{j \neq i} \rho \lambda_j p_{ji} \mathbf{1}(n_j = c_j) \mathbf{1}(n_i < c_i)$ ,
- 8)  $q((\mathbf{n}, \mathbf{m}), (\mathbf{n} - \mathbf{e}_i, \mathbf{m})) = \mu_i n_i p_{i0} + \mu_i n_i \sum_{j \neq i} p_{ij} \mathbf{1}(n_j = c_j)$ ,
- 9)  $q((\mathbf{n}, \mathbf{m}), (\mathbf{n}, \mathbf{m} - \mathbf{e}_i)) = \mu_i m_i p_{i0}$ .

The balance equation is given by

$$p(\mathbf{n}, \mathbf{m}) \sum_{i=1}^J [\lambda_{0i} ((1 - \rho) + \rho \mathbf{1}(n_i < c_i)) + \mu_i (n_i + m_i) + \mu^{(nh)} m_i \mathbf{1}(n_i < c_i) + \mu^{(h)} n_i$$

$$+ \sum_{j \neq i} \rho \lambda_j p_{ij} \mathbf{1}(n_i = c_i) \mathbf{1}(n_j < c_j)]$$

$$= \sum_{i=1}^J [p(\mathbf{n} - \mathbf{e}_i, \mathbf{m}) \rho \lambda_{0i} \mathbf{1}(n_i > 0) + p(\mathbf{n}, \mathbf{m} - \mathbf{e}_i) (1 - \rho) \lambda_{0i} \mathbf{1}(m_i > 0)]$$

$$\begin{aligned}
 & + \sum_{i=1}^J [p(\mathbf{n} + \mathbf{e}_i, \mathbf{m})\mu_i(n_i + 1)p_{i0}\mathbf{1}(n_i < c_i) + p(\mathbf{n}, \mathbf{m} + \mathbf{e}_i)\mu_i(m_i + 1)p_{i0}] \\
 & + \sum_{i=1}^J [p(\mathbf{n} + \mathbf{e}_i, \mathbf{m} - \mathbf{e}_i)\mu^{(h)}(n_i + 1)\mathbf{1}(n_i < c_i)\mathbf{1}(m_i > 0) \\
 & \quad + p(\mathbf{n} - \mathbf{e}_i, \mathbf{m} + \mathbf{e}_i)(m_i + 1)\mu^{(nh)}\mathbf{1}(n_i > 0)] \\
 & + \sum_{i=1}^J \sum_{j \neq i}^J [p(\mathbf{n} + \mathbf{e}_i - \mathbf{e}_j, \mathbf{m})\mu_i(n_i + 1)p_{ij}\mathbf{1}(n_i < c_i)\mathbf{1}(n_j > 0) \\
 & \quad + p(\mathbf{n} - \mathbf{e}_j, \mathbf{m})\rho\lambda_i p_{ij}\mathbf{1}(n_i = c_i)\mathbf{1}(n_j > 0) \\
 & \quad + p(\mathbf{n} + \mathbf{e}_i, \mathbf{m})\mu_i(n_i + 1)p_{ij}\mathbf{1}(n_i < c_i)\mathbf{1}(n_j = c_j) \\
 & \quad + p(\mathbf{n}, \mathbf{m} + \mathbf{e}_i - \mathbf{e}_j)\mu_i(m_i + 1)p_{ij}\mathbf{1}(m_j > 0)]. \tag{5.5}
 \end{aligned}$$

To prove the theorem, we first show that (5.4) satisfies the balance equation. For this purpose, we divide the balance equation into the following partial balance equations:

$$\begin{aligned}
 & p(\mathbf{n}, \mathbf{m}) \sum_{i=1}^J [\mu^{(nh)}m_i\mathbf{1}(n_i < c_i) + \mu^{(h)}n_i] \\
 & = \sum_{i=1}^J [p(\mathbf{n} + \mathbf{e}_i, \mathbf{m} - \mathbf{e}_i)\mu^{(h)}(n_i + 1)\mathbf{1}(n_i < c_i)\mathbf{1}(m_i > 0) \\
 & \quad + p(\mathbf{n} - \mathbf{e}_i, \mathbf{m} + \mathbf{e}_i)\mu^{(nh)}(m_i + 1)\mathbf{1}(n_i > 0)], \tag{5.6}
 \end{aligned}$$

$$\begin{aligned}
 & p(\mathbf{n}, \mathbf{m}) \sum_{i=1}^J ((1 - \rho) + \rho\mathbf{1}(n_i < c_i))\lambda_{0i} \\
 & = \sum_{i=1}^J [p(\mathbf{n} + \mathbf{e}_i, \mathbf{m})\mu_i(n_i + 1)p_{i0}\mathbf{1}(n_i < c_i) + p(\mathbf{n}, \mathbf{m} + \mathbf{e}_i)\mu_i(m_i + 1)p_{i0}], \tag{5.7}
 \end{aligned}$$

$$\begin{aligned}
 & p(\mathbf{n}, \mathbf{m}) \sum_{i=1}^J \mu_i(n_i + m_i) \\
 & = \sum_{i=1}^J [p(\mathbf{n} - \mathbf{e}_i, \mathbf{m})\rho\lambda_{0i}\mathbf{1}(n_i > 0) + p(\mathbf{n}, \mathbf{m} - \mathbf{e}_i)(1 - \rho)\lambda_{0i}\mathbf{1}(m_i > 0)] \\
 & \quad + \sum_{i=1}^J \sum_{j \neq i}^J [p(\mathbf{n} + \mathbf{e}_i - \mathbf{e}_j, \mathbf{m})\mu_i(n_i + 1)p_{ij}\mathbf{1}(n_i < c_i)\mathbf{1}(n_j > 0) \\
 & \quad \quad + p(\mathbf{n} - \mathbf{e}_j, \mathbf{m})\rho\lambda_i p_{ij}\mathbf{1}(n_i = c_i)\mathbf{1}(n_j > 0) \\
 & \quad \quad + p(\mathbf{n}, \mathbf{m} + \mathbf{e}_i - \mathbf{e}_j)\mu_i(m_i + 1)p_{ij}\mathbf{1}(m_j > 0)], \tag{5.8}
 \end{aligned}$$

$$\begin{aligned}
 & p(\mathbf{n}, \mathbf{m}) \sum_{i=1}^J \sum_{j \neq i}^J \rho\lambda_i p_{ij}\mathbf{1}(n_i = c_i)\mathbf{1}(n_j < c_j) \\
 & = \sum_{i=1}^J \sum_{j \neq i}^J p(\mathbf{n} + \mathbf{e}_i, \mathbf{m})\mu_i(n_i + 1)p_{ij}\mathbf{1}(n_i < c_i)\mathbf{1}(n_j = c_j). \tag{5.9}
 \end{aligned}$$

First observe that (5.4) yields

$$\begin{aligned}
 & p(\mathbf{n} + \mathbf{e}_i, \mathbf{m} - \mathbf{e}_i)\mu^{(h)}(n_i + 1)\mathbf{1}(n_i < c_i)\mathbf{1}(m_i > 0) + p(\mathbf{n} - \mathbf{e}_i, \mathbf{m} + \mathbf{e}_i)\mu^{(nh)}(m_i + 1)\mathbf{1}(n_i > 0) \\
 & = p(\mathbf{n}, \mathbf{m})(\mu^{(nh)}m_i\mathbf{1}(n_i < c_i) + \mu^{(h)}n_i)
 \end{aligned}$$

from which (5.6) follows. Next, it follows from (5.4) that

$$p(\mathbf{n} - \mathbf{e}_i, \mathbf{m})\rho\lambda_{0i}\mathbf{1}(n_i > 0) + p(\mathbf{n}, \mathbf{m} - \mathbf{e}_i)(1 - \rho)\lambda_{0i}\mathbf{1}(m_i > 0) = p(\mathbf{n}, \mathbf{m})\frac{(n_i + m_i)\lambda_{0i}}{a_i},$$

$$\begin{aligned}
 & \sum_{i \neq j}^J [p(\mathbf{n} + \mathbf{e}_i - \mathbf{e}_j, \mathbf{m}) \mu_i(n_i + 1) p_{ij} \mathbf{1}(n_i < c_i) \mathbf{1}(n_j > 0) \\
 & \quad + p(\mathbf{n} - \mathbf{e}_j, \mathbf{m}) \rho \lambda_i p_{ij} \mathbf{1}(n_i = c_i) \mathbf{1}(n_j > 0) + p(\mathbf{n}, \mathbf{m} + \mathbf{e}_i - \mathbf{e}_j) \mu_i(m_i + 1) p_{ij} \mathbf{1}(n_j > 0)] \\
 & = p(\mathbf{n}, \mathbf{m}) \sum_{i \neq j}^J \frac{\lambda_i}{\lambda_j} (n_j \mathbf{1}(n_i \leq c_i) + m_j) \mu_j p_{ij} \\
 & = p(\mathbf{n}, \mathbf{m}) \left( 1 - \frac{\lambda_{0j}}{\lambda_j} \right) (n_j + m_j) \mu_j \\
 & = p(\mathbf{n}, \mathbf{m}) \left( \mu_j (n_j + m_j) - \frac{(n_j + m_j) \lambda_{0j}}{a_j} \right)
 \end{aligned}$$

from which (5.8) follows. Similarly, it follows from (5.4) that

$$p(\mathbf{n} + \mathbf{e}_i, \mathbf{m}) \mu_i(n_i + 1) p_{ij} \mathbf{1}(n_i < c_i) \mathbf{1}(n_j = c_j) = p(\mathbf{n}, \mathbf{m}) \rho \lambda_i p_{ij} \mathbf{1}(n_i < c_i) \mathbf{1}(n_j = c_j),$$

from which (5.9) follows. Finally, (5.7) follows if

$$\begin{aligned}
 & p(\mathbf{n} + \mathbf{e}_i, \mathbf{m}) \mu_i(n_i + 1) p_{i0} \mathbf{1}(n_i < c_i) + p(\mathbf{n}, \mathbf{m} + \mathbf{e}_i) \mu_i(m_i + 1) p_{i0} \\
 & \quad = p(\mathbf{n}, \mathbf{m}) a_i \mu_i p_{i0} (\rho \mathbf{1}(n_i < c_i) + (1 - \rho)), \\
 & \sum_{i=1}^J \lambda_{0i} = \sum_{i=1}^J a_i \mu_i p_{i0}.
 \end{aligned}$$

The first equality follows from (5.4). The second equality follows from the traffic equation (3.2). Thus, (5.4) certainly satisfies the balance equation.

To finish the proof, we should show that the considered Markov chain is ergodic (also see Remark 4.3: here we assume that (3.2) has the unique solution). To this end, observe that

$$\begin{aligned}
 & \sum_{\mathbf{n}} \sum_{\mathbf{m}} p(\mathbf{n}, \mathbf{m}) \sum_{i=1}^J \left[ \lambda_{0i} ((1 - \rho) + \rho \mathbf{1}(n_i < c_i)) + \mu_i(n_i + m_i) + \mu^{(nh)} m_i \mathbf{1}(n_i < c_i) + \mu^{(h)} n_i \right. \\
 & \quad \left. + \sum_{j \neq i} \rho \lambda_j p_{ij} \mathbf{1}(n_i = c_i) \mathbf{1}(n_j < c_j) \right] \\
 & < \sum_{\mathbf{n}} \sum_{\mathbf{m}} p(\mathbf{n}, \mathbf{m}) \sum_{i=1}^J \left[ \lambda_{0i} + \mu_i(n_i + m_i) + \mu^{(nh)} m_i + \mu^{(h)} n_i + \rho \lambda_i \right] \\
 & = \sum_{i=1}^J \left[ \lambda_{0i} + a_i \mu_i + \rho a_i \mu^{(h)} + (1 - \rho) a_i \mu^{(nh)} + \rho \lambda_i \right] < \infty.
 \end{aligned} \tag{5.10}$$

Thus the ergodicity follows from Corollary 4.4 in [1] (page 53), which finishes the proof. □

**Remark 5.6.** Under the product-form approximation, the insensitivity on the product-form distribution (5.4) also follows; (5.4) depends on the residing-time, the active-duration, and the inactive-duration distributions only through their means. To prove the insensitivity, we describe the system by the generalized semi-Markov process and apply Theorem 4.1 in Miyazawa [10].

**5.4. Remark on the call-blocking and handoff-blocking probabilities**

It is a direct consequence of Theorem 5.4 that the channel-full-occupancy probability of cell  $i$ ,  $P[N_i = c_i]$ , is given by the Erlang-loss formula. This does not, however, always conclude that call-blocking and handoff-blocking probabilities are also given by the Erlang-loss formula because they are differently defined as explained in Section 5.1.

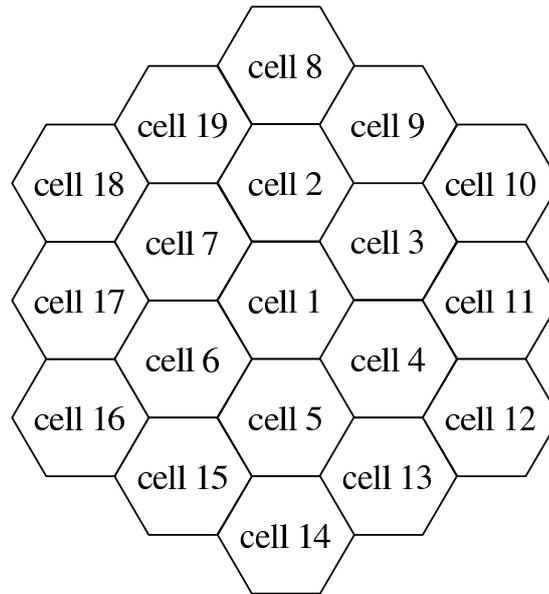


Figure 3: Cell network used in the simulation

Fortunately, it can be easily seen that the handoff-blocking probability is equal to the channel-full-occupancy probability. This comes from the fact that, in the network of quasi-reversible queues, the arrival to any node observes time average, with the arrived job itself exclude (Theorem 4.14 in [3]). That is, active users see the time-average behavior of a cell at their arrival epochs to the cell, and thus the handoff-blocking probability is equal to the channel-full-occupancy probability. In addition to this, if the inactive duration is exponentially distributed, then the call blocking probability is also equal to the channel-full-occupancy probability as explained in Section 5.1. Thus, in usual cases, the call-blocking and the handoff-blocking probabilities are also given by the Erlang-loss formula.

For the case when the inactive duration is not exponentially distributed, we need subtle discussion because the call generation process by inactive users is no longer Poisson process. To discuss non-exponential-inactive-duration cases, we note that the cellular mobile system discussed in this paper can be modeled by a Generalized Semi-Markov Process (GSMP) (see Appendix). A GSMP having the stationary distribution like (5.3) is called *product-form decomposable*. It is known that a GSMP is product-form decomposable, then the state transition epochs of the GSMP sees the stationary distribution of the system [10] (please also see Remark A.3). This observation suggests the following conjecture; the call-blocking probability is equal to the channel-full-occupancy probability even if the inactive duration is not exponentially distributed. The rigorous proof of this conjecture remains a future study.

## 6. Numerical Examples

In the previous section, we show that, under the product-form approximation, the call-blocking probability is given by the Erlang-loss formula. In this section, we evaluate the accuracy of the product-form approximation using the simulation experiments.

### 6.1. Network model

To see the accuracy of the product-form approximation, we conducted simulation experiments using the network model depicted in Figure 3, which was made of 19 cells. In the simulation, we let  $h_i = 120$  min and  $t_h = 3$  min. Concerning the average inactive duration, we considered two

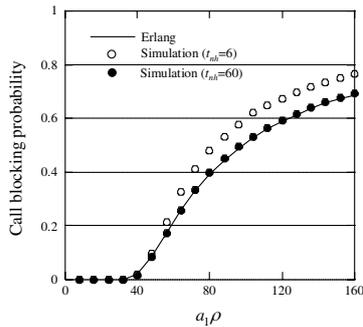


Figure 4: Call-blocking probability in cell 1

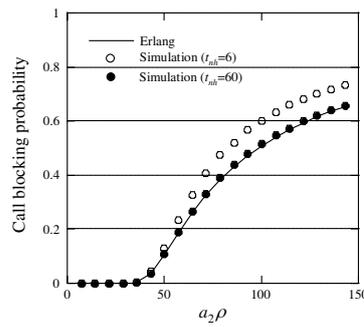


Figure 5: Call-blocking probability in cell 2

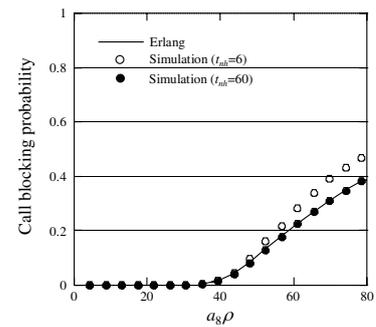


Figure 6: Call-blocking probability in cell 8

different cases:  $t_{nh} = 60$  min or  $t_{nh} = 6$  min. We assumed that each cell has the same external traffic and the same number of channels ( $\lambda_{01} = \lambda_{02} = \dots = \lambda_{019}$  and  $c_1 = c_2 = \dots = c_{19}$ ). Note that the external arrival of mobile users occurred even in inner cells (for example, cell 1); the external arrival of a mobile user corresponds to the event whereby he switches on his mobile phone.

## 6.2. Accuracy of the product-form approximation

First, we compared the call-blocking probabilities in the simulation with those evaluated by the Erlang-loss formula (5.2). The number of channels of each cell was set at 50 and the probability  $p_{ij}$  ( $i, j \neq 0$ ) and  $p_{i0}$  ( $i \neq 0$ ) were given by the following:

$$p_{ij} = \begin{cases} 1/7 & \text{if cell } i \text{ is neighbor to cell } j, \\ 0 & \text{otherwise,} \end{cases} \quad (6.1)$$

$$p_{i0} = \begin{cases} 1/7 & i = 1, 2, 3, 4, 5, 6, \text{ or } 7, \\ 3/7 & i = 9, 11, 13, 15, 17, \text{ or } 19, \\ 4/7 & i = 8, 10, 12, 14, 16, \text{ or } 18. \end{cases} \quad (6.2)$$

In the simulation, we evaluated the call blocking probability by changing  $a_i\rho$  from 0 to 160. The results are shown in Figure 4 (cell 1), Figure 5 (cell 2), and Figure 6 (cell 8). Note that, under the product-form approximation,  $a_i\rho$  corresponds to the offered traffic of cell  $i$ . Thus, we refer  $a_i\rho$  to as the *model traffic* in this section.

The figures show that, when  $t_{nh} = 60$  min, the call blocking probability in the simulation and that evaluated by (5.2) were very close. This suggests that the error caused by the product-form approximation is almost negligible when the average inactive duration is much longer than the average active duration. However, when  $t_{nh} = 6$  min, we see some gap between the call blocking probability in the simulation and that by the Erlang-loss formula. This gap was caused by the call requests by hand-off users who meet with hand-off blocking; the product-form approximation neglects this effect.

The WHITE PAPER of Information and Communications in Japan published by Ministry of Internal Affairs and Communications [8] reported that the average total holding time per contract in a day is 3 min and 16 sec. This report indicates that the case  $t_{nh} = 6$  min (and  $t_h = 3$  min) is very unusual, and thus the use of Erlang-loss formula does not yield large error in the estimation of the call blocking probability in usual cases.

Under the setting of transition probability (6.1) and (6.2), the offered traffic in inner cells (for example, cell 1) is much larger than the offered traffic in outer cells (for example, cell 8). So, we

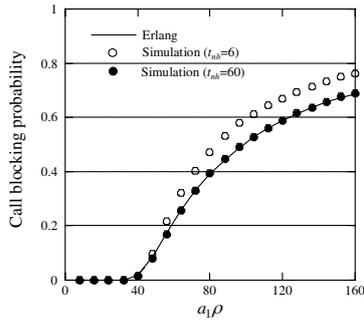


Figure 7: Call-blocking probability in cell 1

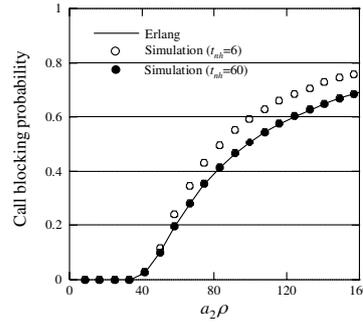


Figure 8: Call-blocking probability in cell 2

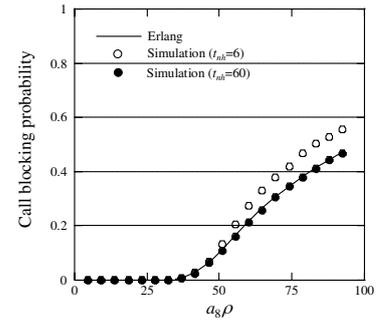


Figure 9: Call-blocking probability in cell 8

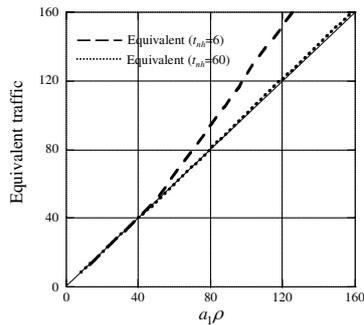


Figure 10: Model traffic v.s. equivalent traffic in cell 1

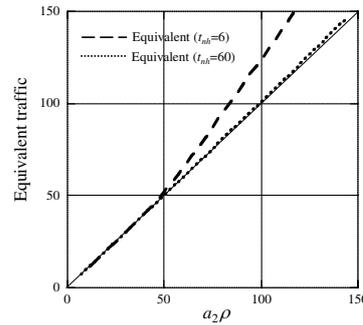


Figure 11: Model traffic v.s. equivalent traffic in cell 2

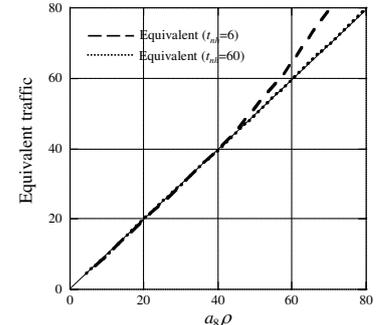


Figure 12: Model traffic v.s. equivalent traffic in cell 8

also conducted another simulation experiment by letting

$$\begin{aligned}
 &\text{for } i = 1, 2, 3, 4, 5, 6, \text{ or } 7 & p_{ij} &= \begin{cases} 1/7 & \text{cell } i \text{ is neighbor to cell } j, \\ 0 & \text{otherwise,} \end{cases} \\
 &\text{for } i = 9, 11, 13, 15, 17, \text{ or } 19 & p_{ij} &= \begin{cases} 3/14 & \text{cell } i \text{ is neighbor to cell } j, \\ 0 & \text{otherwise,} \end{cases} \\
 &\text{for } i = 8, 10, 12, 14, 16, \text{ or } 18 & p_{ij} &= \begin{cases} 2/7 & \text{cell } i \text{ is neighbor to cell } j, \\ 0 & \text{otherwise,} \end{cases}
 \end{aligned} \tag{6.3}$$

and

$$\text{for } i = 1, \dots, 19 \quad p_{i0} = 1/7. \tag{6.4}$$

The results are shown in Figure 7 (cell 1), Figure 8 (cell 2), and Figure 9 (cell 8), where we see the results similar with those in Figures 4, 5, and 6.

### 6.3. Equivalent traffic

The model traffic  $a_i\rho$  cannot be directly observed because model parameters  $\rho$  and  $h_i$  are generally hard to observe. In actual situations, the following value seems to be used as the load of cell  $i$ :

$$\alpha_i = \frac{\langle n_i \rangle}{1 - b_i}, \tag{6.5}$$

where  $\langle n_i \rangle$  is the average number of calls in cell  $i$  and  $b_i$  is the call blocking probability of cell  $i$ . We refer  $\alpha_i$  to as the ‘‘equivalent traffic’’. Note that  $\alpha_i$  differs to  $a_i\rho$ .

Figures 10, 11, and 12 respectively compare the model traffic with the equivalent traffic of cells 1, 2, and 8. All results were obtained when the transition probabilities were given by (6.1)

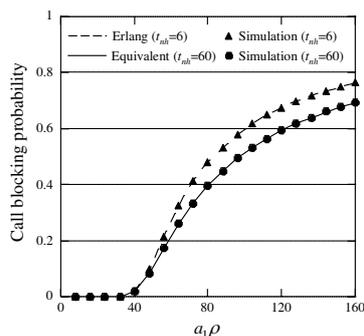


Figure 13: Blocking prob. by equivalent traffic in cell 1

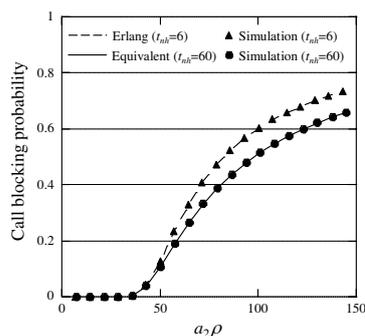


Figure 14: Blocking prob. by equivalent traffic in cell 2

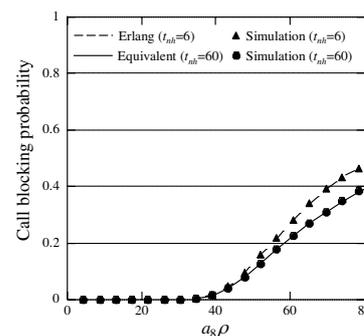


Figure 15: Blocking prob. by equivalent traffic in cell 8

and (6.2). As expected, when  $t_{nh} = 6$  min, there is some discrepancy between the model traffic and the equivalent traffic. This is because the equivalent traffic naturally takes into account the call requests by hand-off users who encounter hand-off blockings, while the model traffic does not. Thus, we can expect the use of the equivalent traffic to enable us to evaluate the call-blocking probability or the necessary number of channels with high accuracy, even for unusual cases like  $t_{nh} = 6$  min.

To confirm this expectation, in Figures 13, 14 and 15, we respectively compared the call-blocking probabilities in cells 1, 2, and 8 obtained by the simulation with those evaluated by the Erlang-loss formula with the equivalent traffic. Even for the case  $t_{nh} = 6$ , both call blocking probabilities are very close. In other words, we can accurately evaluate the call blocking probability based on the Erlang-loss formula if we use the equivalent traffic.

#### 6.4. Channel provisioning based on equivalent traffic

Finally, we investigated the possibility of channel provisioning (evaluation of the necessary number of channels) based on the Erlang-loss formula using the equivalent traffic. To this end, we initially conducted the simulation with various numbers of initial channels and evaluated the equivalent traffic (6.5) in cell 1. The model traffic was fixed at 100. Subsequently, based on the Erlang-loss formula with the equivalent traffic, we calculated the number of *designed channels*, which are necessary and sufficient to ensure a call blocking probability of below 0.01. The equivalent traffic and the number of designed channels are summarized in Tables 1 and 2. Note that we confirmed through the simulation that the exact number of channels necessary and sufficient for ensuring call-blocking probabilities of below 0.01 in cell 1 was 117 when  $t_{nh} = 6$  min and 118 when  $t_{nh} = 60$  min. The transition probabilities were given by (6.1) and (6.2).

The tables indicate that the channel provisioning based on the equivalent traffic was accurate besides unusual cases like  $t_{nh} = 6$  min and the number of initial channels was much smaller than that of those properly designed (118). In the case where  $t_{nh} = 6$  min and the number of initial channels was much smaller than 118, the equivalent traffic and the number of designed channels were overestimated. However, there was no case where the number of designed channels was underestimated.

In Figure 16, we show the exact call blocking probability observed in the simulation when the channels were provisioned according to the number of designed channels estimated from the equivalent traffic. As explained above, when  $t_{nh} = 6$  min and the number of initial channels was much smaller than the proper number of channels (117), the channels were overprovisioned and thus the resultant call blocking probability was much lower than 0.01. Except for these unusual cases, the call blocking probability was close to 0.01. These results confirm that, by using the

Table 1: Equivalent traffic and the necessary number of channels ( $t_{nh} = 6$  min)

Number of initial channels	50	60	70	80	90	100
Equivalent traffic	128.65	118.81	114.61	109.83	105.65	103.14
Number of designed channels	147	137	133	128	123	121
Number of initial channels	110	120	130	140	150	160
Equivalent traffic	100.67	100.35	100.20	100.23	100.12	100.25
Number of designed channels	118	118	118	118	118	118

Table 2: Equivalent traffic and the necessary number of channels ( $t_{nh} = 60$  min)

Number of initial channels	50	60	70	80	90	100
Equivalent traffic	101.63	101.30	100.91	100.52	100.28	100.12
Number of designed channels	119	119	118	118	118	118
Number of initial channels	110	120	130	140	150	160
Equivalent traffic	99.96	99.88	99.85	99.87	99.86	99.86
Number of designed channels	117	117	117	117	117	117

notion of equivalent traffic, channel provisioning based on the Erlang-loss formula is possible, even in cellular mobile networks.

## 7. Conclusion

In this paper, we propose a network model for cellular mobile systems and show the analysis in the product-form framework. The model explained in this paper does not consider either the channel reservation for hand-off users [4] or the soft-handover used in CDMA systems [15, 16]. The proposal of a network model that can take these detailed system features into account remains an area for future study.

### A. Proof of Theorem 5.4

Here, we prove Theorem 5.4 by describing the transition of the state  $(n_1, m_1)$  by a GSMP (Generalized Semi-Markov Process) [10].

The GSMP is a probability model in which many clocks fixed at active sites are finitely running under a given global state called a macro state. When one clock completes its lifetime, which is called the expiry of a site, a transition of the macro state occurs and new sites may be activated with new lifetimes while other clocks retain their residual lifetimes.

Our model has four types of sites. The first has a clock which counts the inter-arrival time between users. We call it a *type-0* site. The second has a clock which counts the remaining residual time of a user at a cell. We call this type of site *type 1*. The third has a clock which counts the remaining active duration of an active user, which we call a *type 2* site. The last has a clock for counting the remaining inactive duration of an inactive user, which we call a *type 3* site. We define the following:

$$\begin{aligned} X(t) &\stackrel{\text{def}}{=} (n_1(t), m_1(t)) \\ Y(t) &\stackrel{\text{def}}{=} (\tau(t), R_1(t), \dots, R_{n_1+m_1}(t), R_1^{(h)}(t), \dots, R_{n_1}^{(h)}(t), R_1^{(nh)}(t), \dots, R_{m_1}^{(nh)}(t)), \end{aligned} \quad (\text{A.1})$$

where,  $\tau(t)$  denotes the interval between the current time  $t$  and the next user arrival instant. In the context of the GSMP,  $X(t)$  corresponds to the macro state and each element of  $Y(t)$  denotes the

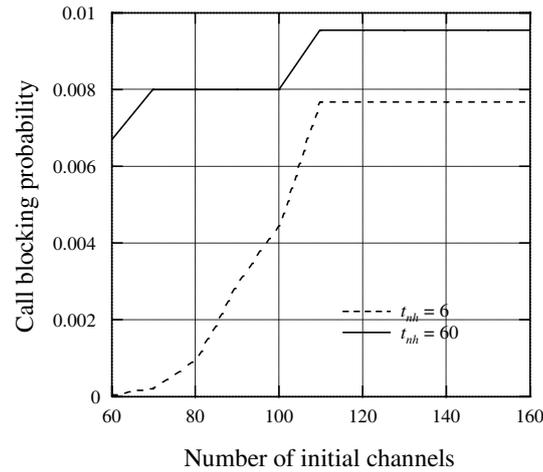


Figure 16: Call blocking probability when channels were provisioned based on the equivalent traffic

remaining lifetime of each active site.

When  $\mathbf{X}(t) = (n_1(t), m_1(t))$ , there is one type-0 site,  $(n_1 + m_1)$ -type-1 sites,  $n_1$ -type-2 sites, and  $m_1$ -type-3 sites. An arriving active user, finding the macro state was in  $(n_1(t), m_1(t))$ , gets the  $s_1$ th type-1 site with probability  $1/(n_1 + m_1 + 1)$  for  $s_1 = 1, \dots, n_1 + m_1 + 1$ , and gets the  $s_2$ th type-2 site with probability  $1/(n_1 + 1)$  for  $s_2 = 1, \dots, n_1 + 1$ . In this case, a type-1 site in position  $l_1 (\neq s_1)$  moves to  $\gamma_1(l_1)$  where  $\gamma_1$  is a permutation of the indices  $1, \dots, n_1 + m_1 + 1$  except  $s_1$ . Similarly, a type-2 site in position  $l_2 (\neq s_2)$  moves to  $\gamma_2(l_2)$  where  $\gamma_2$  is a permutation of the indices  $1, \dots, n_1 + 1$  except  $s_2$ . When a user who possesses the  $s_1$ th-type-1 site and the  $s_2$ th-type-2 site departs the cell, leaving the macro state in  $(n_1(t), m_1(t))$ , a type-1 sites in position  $l_1 (\neq s_1)$  moves to  $\beta_1(l_1)$  and a type-2 site in position  $l_2 (\neq s_2)$  moves to  $\beta_2(l_2)$ , where  $\beta_1$  is a permutation of the indices  $1, \dots, n_1 + m_1$  and  $\beta_2$  is a permutation of the indices  $1, \dots, n_1$ .

As explained above, in our model, two new sites can be activated at once. More precisely, at the arrival of a new user, a type-1 site and a type-2 (or type-3) site are simultaneously activated. Similarly, at the departure of a user, a type-1 site and a type-2 (or type-3) site expire simultaneously. Note that, the expiration of the type-2 (or type-3) site at the departure of a user corresponds to the so-called *interruption*. Here, the interruption means that a site is ended although its remaining lifetime does not reach zero. Some GSMPs allow such site interruptions (see [10]). In our model, each element of  $\mathbf{Y}(t)$  (remaining lifetime of each active site) decreases at rate one.

We use the following notation:

- $\Gamma(\mathbf{x})$ : a set of sites that are active when the macro state is  $\mathbf{x}$ .
- $\Gamma(\mathbf{x}, \mathbf{x}'; s)$ : a set of sites that are active before and after the transition  $\mathbf{x} \rightarrow \mathbf{x}'$  caused by the expiration of site  $s$ , and that their remaining lifetimes are not changed at the transition.
- $p(\mathbf{x}, \mathbf{x}'; s)$ : the probability that the macro state moves to  $\mathbf{x}'$  by the transition from  $\mathbf{x}$  caused by the expiration of site  $s$ .
- $P[\mathbf{x}]$ : probability that  $\mathbf{X} = \mathbf{x}$ .
- $\mu_s$ : the inverse of the mean lifetime of site  $s$ .
- $S$ : a set of sites.
- $S_i$ : a set of type- $i$  sites. ( $i = 0, 1, 2, 3$ )

In the following,  $N_s$  denotes a point process composed of expiration instants of site  $s$ , and  $\lambda_s$  denotes the intensity of  $N_s$ . Similarly, when  $U$  is a set of sites, we let  $N_U$  denote a point process composed of expiration instants of sites in  $U$  and let  $\lambda_{S_U}$  denote its intensity. We also let  $P_s$

$(P_U)$  denote the Palm measure with respect to the point process  $N_s$  ( $N_U$ ). We further define the following:

$$P[\mathbf{x}, \boldsymbol{\theta}] \stackrel{\text{def}}{=} E[e^{-\boldsymbol{\theta} \cdot \mathbf{Y}} \mathbf{1}(X = \mathbf{x})], \quad P_s[\mathbf{x}, \boldsymbol{\theta}] \stackrel{\text{def}}{=} E_s[e^{-\boldsymbol{\theta} \cdot \mathbf{Y}} \mathbf{1}(X = \mathbf{x})], \quad (\text{A.2})$$

where  $\boldsymbol{\theta} = \{\theta_i\}$  is a vector and  $E_s[\cdot]$  is the expectation with respect to the Palm measure  $P_s$ . In some cases, instead of  $\boldsymbol{\theta}$ , we use a vector  $\boldsymbol{\theta}(s)$  which is defined by

$$\boldsymbol{\theta}(s) \stackrel{\text{def}}{=} \{\theta_i(s)\}, \quad \theta_i(s) \stackrel{\text{def}}{=} \begin{cases} \theta_i & i \neq s, \\ 0 & i = s. \end{cases}$$

The discussion of this section follows the paper by Miyazawa [10]. However, there is a small difference between our model and that in [10]. Our model allows simultaneous activations of multiple sites, as explained in the above, while the model in [10] does not. In addition, at the arrival instant of a user, a lifetime of the type-2 (type-3) site is sampled from the distribution  $G_e^{(h)}(t)$  (or  $G_e^{(nh)}(t)$ ), while, at an instant of call generation (completion), a lifetime of the type-2 (type-3) site is sampled from the distribution  $G^{(h)}(t)$  (or  $G^{(nh)}(t)$ ). To consider the difference mentioned above, we slightly modify the results in [10]. However, this modification is not essential, so here we show the outline of the proof. For the detailed discussion, please see [10].

**Lemma A.1** (Miyazawa [10] Lemma 4.1). *If the joint distribution of  $(X, Y)$  has the form such that*

$$P[\mathbf{x}, \mathbf{t}] \stackrel{\text{def}}{=} P[X = \mathbf{x}, Y \leq \mathbf{t}] = P[\mathbf{x}] \prod_s H_s(t_s), \quad (\text{A.3})$$

where  $H_s(t)$  is a distribution depending only on the index  $s \in S$ , then

$$\lambda_s P_s[\mathbf{x}, \boldsymbol{\theta}(s)] = H'_s(0) P[\mathbf{x}] \mathbf{1}(s \in \Gamma(\mathbf{x})) \prod_{l \neq s} \hat{H}_l(\theta_l),$$

where  $\hat{H}_s(\theta)$  is the Laplace transform of  $H_s(t)$ ,  $H'_s(0)$  is a right-hand derivative of  $H_s(t)$  at  $t = 0$ . Moreover, if (A.3) holds, we have

$$\sum_{\mathbf{x}} \sum_s P[\mathbf{x}] \mathbf{1}(s \in \Gamma(\mathbf{x})) H'_s(0) = \sum_{\mathbf{x}} \sum_{\mathbf{x}'} \sum_s P[\mathbf{x}'] \mathbf{1}(s \in \Gamma(\mathbf{x}')) H'_s(0) p(\mathbf{x}', \mathbf{x}; s). \quad (\text{A.4})$$

**Remark A.2.** If the joint distribution of  $(X, Y)$  has the form such as (A.3), then the GSMP is called product-form decomposable. In our model,  $H_s(t)$  should depend only on the type of site to which  $s$  belongs, that is

$$H_s(t) = \begin{cases} H^{(r)}(t) & s \in S_1, \\ H^{(h)}(t) & s \in S_2, \\ H^{(nh)}(t) & s \in S_3. \end{cases}$$

Thus, the stationary distribution (A.3) can be written as

$$\begin{aligned} & P[(N_1, M_1) = (n_1, m_1), \mathbf{R} \leq \mathbf{t}, \mathbf{R}^{(h)} \leq \mathbf{t}^{(h)}, \mathbf{R}^{(nh)} \leq \mathbf{t}^{(nh)}] \\ &= g \frac{(\rho a_1)^{n_1}}{n_1!} e^{-(1-\rho)a_1} \frac{((1-\rho)a_1)^{m_1}}{m_1!} \prod_{i=1}^{n_1+m_1} H^{(r)}(t_i) \prod_{i=1}^{n_1} H^{(h)}(t_i^{(h)}) \prod_{i=1}^{m_1} H^{(nh)}(t_i^{(nh)}). \quad (\text{A.5}) \end{aligned}$$

**Remark A.3.** Lemma A.1 implies that, if GSMP is product-form decomposable, then the point process  $N_s$  sees the stationary distribution of the macro state  $\mathbf{X}$  and the remaining lifetime of the macro state  $\mathbf{Y}$  besides site  $s$ .

**Lemma A.4.** *If the joint distribution of  $X$  and  $Y$  is product-form decomposable, then*

$$H_s(t) = \begin{cases} F_e(t) & s \in S_1, \\ G_e^{(h)}(t) & s \in S_2, \\ G_e^{(nh)}(t) & s \in S_3. \end{cases}$$

*Proof.* Since type-1 sites are not interrupted, the straightforward application of Lemma 4.2 in Miyazawa [10] proves that  $H_s(t) = F_e(t)$  for  $s \in S_1$ . To get the result for  $s \in S_2$ , we define

$$Z(t) = e^{-\theta R_s^{(h)}(t)} \quad \text{for } s \in S_2.$$

Applying the rate conservation law of Miyazawa [9] to  $Z(t)$  yields

$$\theta E[e^{-R_s^{(h)}(0)\theta}] = \lambda_s E_S[e^{-R_s^{(h)}(0-)\theta}] - \lambda_s E_S[e^{-R_s^{(h)}(0+)\theta}], \tag{A.6}$$

where from the product-form decomposability

$$\theta E[e^{-R_s^{(h)}(0)\theta}] = \theta \hat{H}_s(\theta) = \theta \hat{H}^{(h)}(\theta). \tag{A.7}$$

On the other hand, it follows from Lemma A.1 that

$$\begin{aligned} & \lambda_s E_S[e^{-R_s^{(h)}(0-)\theta}] - \lambda_s E_S[e^{-R_s^{(h)}(0+)\theta}] \\ &= \lambda_{S_1} (E_{S_1}[e^{-R_s^{(h)}(0-)\theta}] - E_{S_1}[e^{-R_s^{(h)}(0+)\theta}]) + \lambda_{S_2} (E_{S_2}[e^{-R_s^{(h)}(0-)\theta}] - E_{S_2}[e^{-R_s^{(h)}(0+)\theta}]) \\ & \quad + \lambda_{S_3} (E_{S_3}[e^{-R_s^{(h)}(0-)\theta}] - E_{S_3}[e^{-R_s^{(h)}(0+)\theta}]) \\ &= a_s(1 - \hat{G}_e^{(h)}(\theta)) + b_s(1 - \hat{G}^{(h)}(\theta)) + c_s(\hat{H}^{(h)}(\theta) - 1), \end{aligned} \tag{A.8}$$

where

$$\begin{aligned} a_s &\stackrel{\text{def}}{=} \lambda_{01} \sum_{\mathbf{x}} \sum_{\mathbf{x}'} P[\mathbf{x}'] p(\mathbf{x}', \mathbf{x}; 0) \mathbf{1}(s \notin \Gamma(\mathbf{x}', \mathbf{x}; 0)), \\ b_s &\stackrel{\text{def}}{=} \sum_{\mathbf{x}} \sum_{\mathbf{x}'} \sum_{d \in S_3} P[\mathbf{x}'] \mathbf{1}(s \in \Gamma(\mathbf{x})) H'_d(0) p(\mathbf{x}', \mathbf{x}; d) \mathbf{1}(s \notin \Gamma(\mathbf{x}', \mathbf{x}; d)), \\ c_s &\stackrel{\text{def}}{=} \sum_{\mathbf{x}} \sum_{\mathbf{x}'} \sum_{d \in S_1} P[\mathbf{x}'] \mathbf{1}(s \in \Gamma(\mathbf{x})) H'_d(0) p(\mathbf{x}, \mathbf{x}'; d) \mathbf{1}(s \notin \Gamma(\mathbf{x}, \mathbf{x}'; d)). \end{aligned}$$

Note that  $a_s$  denotes the creation ratio of site  $s$  caused by the arrival of active users,  $b_s$  denotes the creation ratio of site  $s$  caused by the call requests by inactive users, and  $c_s$  denotes the termination ratio of site  $s$  caused by the departures of users.

Substituting (A.7) and (A.8) into (A.6) yields

$$\theta \hat{H}^{(h)}(\theta) = a_s(1 - \hat{G}_e^{(h)}(\theta)) + b_s(1 - \hat{G}^{(h)}(\theta)) + c_s(\hat{H}^{(h)}(\theta) - 1),$$

or likewise

$$\frac{d}{dt} H^{(h)}(t) = a_s(1 - G_e^{(h)}(t)) + b_s(1 - G^{(h)}(t)) + c_s(H^{(h)}(t) - 1). \tag{A.9}$$

Now we show that (A.9) has the solution  $H^{(h)}(t) = G_e^{(h)}(t)$ . First observe that the symmetry of arrival and departure leads to  $a_s = c_s$ . Thus,

$$\frac{d}{dt} H^{(h)}(t) = a_s H^{(h)}(t) - a_s G_e^{(h)}(t) + b_s(1 - G^{(h)}(t)).$$

By solving the above equation, we get

$$\begin{aligned} H^{(h)}(t) &= e^{ast} \int_0^t e^{-asu} \{b_s(1 - G^{(h)}(u)) - a_s G_e^{(h)}(u)\} du \\ &= G_e^{(h)}(t) + e^{ast} \int_0^t e^{-asu} (b_s - \frac{1}{t_h})(1 - G^{(h)}(u)) du. \end{aligned}$$

Since  $H^{(h)}(\infty) = 1$ ,

$$\lim_{t \rightarrow \infty} e^{ast} \int_0^t e^{-asu} (b_s - \frac{1}{t_h})(1 - G^{(h)}(u)) du = 0,$$

and thus  $b_s$  has to be equal to  $1/t_h$ . Hence, we obtain  $H^{(h)}(t) = G_e^{(h)}(t)$ . The similar argument yields  $H_s(t) = G_e^{(nh)}(t)$  when  $s \in S_3$ . □

**Lemma A.5.** *If the joint distribution of  $X$  and  $Y$  is the product-form decomposable, then*

$$\mathbf{1}(s \in \Gamma(\mathbf{x}))\mu_s P[\mathbf{x}] = \sum_{\mathbf{x}'} \sum_d \mu_d P[\mathbf{x}'] p(\mathbf{x}', \mathbf{x}; d) \mathbf{1}(s \notin \Gamma(\mathbf{x}', \mathbf{x}; d)). \tag{A.10}$$

*In addition, if  $s \in S_2$  ( $s \in S_3$ ), then the summation on  $d$  in the right hand side of (A.10) is to be taken for sites in  $S_3$  ( $S_2$ ). On the other hand, if (A.10) and (A.4) holds, then the joint distribution of  $X$  and  $Y$  is the product-form decomposable such as (5.3).*

*Proof.* We first show that the product-form decomposability leads to (A.10). It follows from (4.16) in [10] that for  $s \in S_1$

$$\theta \mathbf{1}(s \in \Gamma(\mathbf{x})) P[\mathbf{x}] \hat{H}_s(\theta) = K_1(\mathbf{x}, s)(1 - \hat{F}(\theta)) - K_2(\mathbf{x}, s)(1 - \hat{H}_s(\theta)),$$

where

$$\begin{aligned} K_1(\mathbf{x}, s) &= \sum_{\mathbf{x}'} \sum_d P[\mathbf{x}'] H'_d(0) p(\mathbf{x}', \mathbf{x}; d) \mathbf{1}(s \notin \Gamma(\mathbf{x}', \mathbf{x}; d)), \\ K_2(\mathbf{x}, s) &= \sum_{d \neq s} P[\mathbf{x}] \mathbf{1}(d \in \Gamma(\mathbf{x})) H'_s(0) - \sum_{\mathbf{x}'} \sum_d P[\mathbf{x}'] H'_d(0) p(\mathbf{x}', \mathbf{x}; d) \mathbf{1}(s \in \Gamma(\mathbf{x}', \mathbf{x}; d)). \end{aligned} \tag{A.11}$$

Since  $H_s(t) = F_e(t)$  (by Lemma A.4) and  $\hat{F}_e(\theta) = \mu_s(1 - \hat{F}(\theta))/\theta$ , we have

$$(K_1(\mathbf{x}, s) - \mathbf{1}(s \in \Gamma(\mathbf{x})) P[\mathbf{x}] \mu_s) \hat{F}_e(\theta) = \mu_s K_2(\mathbf{x}, s) \frac{1 - \hat{F}_e(\theta)}{\theta}.$$

If  $F(t)$  is not exponentially distributed, the above equation holds only if

$$K_1(\mathbf{x}, s) - \mathbf{1}(s \in \Gamma(\mathbf{x})) P[\mathbf{x}] \mu_s = K_2(\mathbf{x}, s) = 0^3.$$

Since  $H_d(t) = F_e(t)$  for  $d \in S_1$ ,

$$H'_d(0) = F'_e(0) = \lim_{t \rightarrow 0} \frac{1 - F(t)}{h} = \frac{1}{h} = \mu_d, \quad \text{for } d \in S_1,$$

where  $h$  is the average residual time in a cell. Similarly, we have  $H'_d(0) = \mu_d$  for  $d \in S_2$  or  $d \in S_3$ . Combining this fact with  $K_1(\mathbf{x}, s) - \mathbf{1}(s \in \Gamma(\mathbf{x})) P[\mathbf{x}] \mu_s$  yields (A.10). Note that  $K_2(\mathbf{x}, s) = 0$

<sup>3</sup>Since  $K_1(\mathbf{x}, s) - \mathbf{1}(s \in \Gamma(\mathbf{x})) P[\mathbf{x}] \mu_s = K_2(\mathbf{x}, s)$ , this condition is not required if  $F(t)$  is exponentially distributed.

follows from  $K_1(\mathbf{x}, s) - \mathbf{1}(s \in \Gamma(\mathbf{x}))P[\mathbf{x}, s]\mu_s = 0$  and (A.4). A similar argument holds for  $s \in S_2$  and  $s \in S_3$ .

It should be noted that when  $s \in S_2$ , in the summation of the right hand side of (A.10), the contribution of the type-0 site is cancelled out by that of the type-1 site. In addition, there is no contribution of the type-2 site. Thus, when  $s \in S_2$ , the summation on  $d$  on the right hand side of (A.10) is to be taken for type-3 sites. Similarly, when  $s \in S_3$ , the summation on  $d$  is to be taken for type-2 sites on the right hand side of (A.10).

By following the above discussion in reverse, we can confirm that (A.10) and (A.4) lead to product-from decomposability.  $\square$

Equation (A.10) allows the following interpretation; that is, the left-hand side of (A.10) corresponds to the expiration ratio of site  $s$ , and the right-hand side of (A.10) corresponds to the creation ratio of site  $s$ . Thus, (A.10) implies that the expiration and creation ratios of site  $s$  are balanced.

Now, we are ready to prove the theorem. Consider the case where  $\mathbf{x} = (n_1, m_1)$  and  $s \in S_2$ . Firstly, observe that

$$\text{left hand side of (A.10)} = \mu^{(h)} p(n_1, m_1)$$

Since  $p((n_1 - 1, m_1 + 1), (n_1, m_1); d) = 1/n_1$  for  $d \in S_3$  and the summation in the right-hand side of (A.10) is to be taken for sites in  $S_3$  when  $s \in S_2$ , then

$$\begin{aligned} \text{right hand side of (A.10)} &= \sum_{d \in S_3} p(n_1 - 1, m_1 + 1) \mu^{(nh)} p((n_1 - 1, m_1 + 1), (n_1, m_1); d) \\ &= p(n_1 - 1, m_1 + 1) \frac{m_1 + 1}{n_1} \mu^{(nh)} \end{aligned} \quad (\text{A.12})$$

Thus, (A.10) holds if

$$n_1 \mu^{(h)} p(n_1, m_1) = (m_1 + 1) \mu^{(nh)} p(n_1 - 1, m_1 + 1) \quad (1 \leq n_1 \leq c_1),$$

which follows from (5.1). When  $\mathbf{x} = (n_1, m_1)$  and  $s \in S_3$ , (A.10) also holds if

$$m_1 \mu^{(nh)} p(n_1, m_1) = (n_1 + 1) \mu^{(h)} p(n_1 + 1, m_1 - 1) \quad (0 \leq n_1 < c_1, 1 \leq m_1),$$

which also follows from (5.1). Finally, when  $\mathbf{x} = (n_1, m_1)$  and  $s \in S_1$ , (A.10) holds if

$$\mu p(n_1, m_1) = \frac{\rho \lambda_{01}}{n_1 + m_1} p(n_1 - 1, m_1) + \frac{(1 - \rho) \lambda_{01}}{n_1 + m_1} p(n_1, m_1 - 1) \quad (1 \leq n_1 \leq c_1, 1 \leq m_1),$$

which also follows from (5.1). In addition, (A.4) is equivalent to the balance equation. Thus, it follows from Lemma A.5 that (5.1) is valid even when the residing time, active duration and inactive duration are non-exponentially distributed.

Note that the inter-arrival time of users should be exponentially distributed. This is because when  $\mathbf{x} = (0, 0)$  and  $s = 0$ , (A.10) becomes

$$\lambda p(0, 0) = 0,$$

which is not satisfied.

**Remark A.6.** When  $\mathbf{x} = (c_1, m_1)$  and  $s \in S_3$ , then (A.10) becomes

$$\mu^{(nh)} p(c_1, m_1) = \mu^{(nh)} p(c_1, m_1),$$

which is an identity and thus holds. This means that the occurrence of call blocking does not break the insensitivity.

## References

- [1] S. Asmussen: *Applied Probability and Queues (Second Edition)* (Springer, 2003).
- [2] F. Baskett, K. Chandy, R. Muntz and F. Palacios: Open, closed, and mixed networks of queues with different classes of customers. *J. ACM*, **22** (1975) 248–260.
- [3] H. Chen and D. Yao: *Fundamentals of Queueing Networks* (Springer-Verlag, 2001).
- [4] D. Hong and S. Rappaport: Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and nonprioritized hand-off procedures. *IEEE Trans. Vehic. Tech.*, **35** (1986) 77–92.
- [5] B. Jabbari: Teletraffic aspects of evolving and next-generation wireless communication networks. *IEEE Personal Communications*, **3** (1996) 4–9.
- [6] F. Kelly: *Reversibility and Stochastic Networks* (John Wiley & Sons, 1977).
- [7] F. Machihara: Mobile telecommunication systems and generalized erlang loss formula. *IEICE Trans. Commun.*, **E88-B** (2005) 183–189.
- [8] Ministry of Internal Affairs and Communications: WHITE PAPER Information and Communications in Japan. (2006).
- [9] M. Miyazawa: The derivation of invariance relations on complex queueing systems with stationary inputs. *Adv. Appl. Prob.*, **15** (1983) 874–885.
- [10] M. Miyazawa: Insensitivity and product-form decomposability of reallocatable GSMP. *Adv. Appl. Prob.*, **25** (1993) 415–437.
- [11] M. Ohmikawa and H. Takagi: Call loss probabilities in CDMA cellular mobile communication networks. *IEICE Trans. Commun.*, **J82-B** (1999) 2311–2319.
- [12] P. Orlik and S. Rappaport: A model for teletraffic performance and channel holding time characterization in wireless cellular communication with general session and dwell time distributions. *IEEE J. Select. Area. Commun.*, **16** (1998) 788–803.
- [13] P. Orlik and S. Rappaport: On the handoff arrival process in cellular communications. *Wireless Networks*, **7** (2001) 147–157.
- [14] S. Rappaport: The multiple-call hand-off problem in high-capacity cellular communication systems. *IEEE Trans. Vehic. Tech.*, **40** (1991) 546–557.
- [15] S. Su, J. Chen and J. Huang: Performance analysis of soft handoff in CDMA cellular networks. *IEEE J. Select. Areas Commun*, **14** (1996) 1762–1769.
- [16] T. Takahashi, T. Ozawa and Y. Takahashi: Bounds of performance measures in large-scale mobile communication networks. *Performance Evaluation*, **54** (2003) 263–283.

Shigeo Shioda  
Graduate School of Engineering  
Chiba University  
1-33 Yayoi, Inage, Chiba 263-8522, Japan  
E-mail: shioda@faculty.chiba-u.jp