Bounding the Performance of Dynamic Channel Allocation with QoS Provisioning for Distributed Admission Control in Wireless Networks

Xusheng Tian and Chuanyi Ji

***Abstract*—In this work, we investigate the performance of distributed admission control with quality of service (QoS) provisioning and dynamical channel allocation for mobile/wireless networks where co-channel reuse distance is considered as the only limiting factor to channel sharing. We first provide a QoS metric feasible for admission control with dynamically allocated channels. We then derive a criterion analytically using the QoS measure for distributed call admission control with dynamic channel allocation (DCA). When maximum packing is used as the DCA scheme, the results obtained are independent of any particular algorithm that implements dynamic channel assignments. Our results, thereby, provide the optimal performance achievable for the distributed admission control with the QoS provisioning by the best DCA scheme in the given setting.**

***Index Terms*—Distributed admission control, dynamic channel allocation, fixed channel allocation, land mobile radio cellular systems, quality of service.**

I. INTRODUCTION

ONE of the challenges in providing voice service over wireless networks is how to support the guarantees of quality of service (QoS) with the limited capacity. Call admission control is needed to meet this challenge. As various dynamic channel allocation (DCA) algorithms have been developed for admission control, little has been done on assessing the performance gain achievable by DCA with QoS provisioning. The goal of this work was to investigate this fundamental issue by providing answers to two questions: 1) what is the best achievable utilization under a given QoS constraint by distributed admission control with dynamically assigned channels and 2) how much gain can DCA provide compared to fixed channel allocations (FCAs)?

In cellular systems, a geographical region is split into cells, each containing one base station. When a new call request is made at a cell, a decision can be made on either accepting or rejecting the call at each base station and then a channel can be assigned to the call admitted. This results in a distributed admission control strategy that can be applied to every cell (base station). Such a strategy is suitable for large wireless systems with a changing topology. The admission control scheme we considered in this work falls into this category. Two factors determine an admission decision: 1) the availability of a channel at a cell and 2) the QoS constraints. Channels are made available at each cell by channel assignment schemes based on co-channel reuse constraints [1] which are considered as the only limitation to channel reuse in this work. Under such constraints, two classes of channel assignment algorithms have been widely investigated: 1) FCA and 2) DCA [2]. In an FCA scheme, a set of nominal channels are permanently assigned to each cell. An arriving call can only be accepted if there is a nominal channel available in that cell. Due to the temporal and spatial variations of the traffic in cellular systems, FCA schemes are not able to attain a high channel efficiency. To overcome this, DCA schemes have been studied. Unlike FCA, in DCA all channels are kept in a central pool to be shared by all calls in every cell. A channel is eligible for use in any cell provided the co-channel reuse constraint is satisfied. Many researchers [1]–[3] have given comprehensive overviews on existing algorithms for channel allocation.

In the previous work, DCA and QoS provisioning have been investigated in two separate contexts. On the one hand, a lot of channel allocation schemes have been investigated on how to assign channels specifically. In these cases, QoS has not been taken into consideration. On the other hand, channel allocation schemes that have included the QoS provisioning are for FCA alone [4], [5]. In particular, distributed admission control with FCA [5], namely FCA–QoS, has been analyzed with a QoS constraint on the handoff dropping probability for distributed admission control of cellular networks with homogeneous traffic. It remains an open question on how to analyze the performance of DCA with QoS provisioning in the context of distributed admission control. Moreover, FCA with QoS was considered under spatially uniform traffic [4], [5]. Nonuniform traffic patterns have not been taken into consideration.

The focus of this work was on providing answers to such an open question by analyzing the performance of distributed admission control with the QoS provisioning and DCA under both uniform and nonuniform traffic conditions. Two challenges accompany this investigation. The first challenge results from the fact that performance gain varies with respect to different DCA algorithms. Therefore, how to derive a general formulation to characterize the performance gain of DCA with QoS becomes difficult. Another challenge is the QoS measure used for FCA may not be feasible for DCA. How to define a QoS metric in the paradigm of DCA becomes an open problem.

To provide a general formulation on analyzing the performance gain of DCA, we focus on bounding the performance

Manuscript received June 5, 1999; revised April 27, 2000. This work was supported in part by NSF Grant CAREER IRI-9502518. An earlier version of this paper was presented in part at the IEEE ICC'98 Conference, Atlanta, GA. A short version of this paper was presented at IEEE INFOCOM'99, New York.

The authors are with the Department of Electrical, Computer, and System Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180 USA (e-mail: tianx@rpi.edu; jic@rpi.edu).

Publisher Item Identifier S 0018-9545(01)03223-6.

of DCA with the QoS provisioning for distributed admission control by considering the best DCA scheme rather than investigating a specific DCA algorithm. This is accomplished by using maximum packing (MP) [6] as the DCA scheme. As will be seen later, MP allows as many shared channels as possible under a reuse constraint. The results obtained provide the performance achievable by the best DCA scheme and are, therefore, independent of any particular algorithm which implements specific dynamic channel assignments. We will derive analytically an admission control policy with DCA under the QoS constraint. We will show that the derived admission control policy leads to 17%–30% increase in utilization compared to FCA–QoS under various traffic conditions. Thereby, the results provide the maximum possible increase in utilization by any DCA scheme compared to the FCA scheme in the same setting.

This paper is organized as follows. The cellular system considered in this paper is described in Section II. A distributed call admission control policy with the "best" DCA scheme, namely DCA–QoS, is then proposed in Section III. The implementing details of the admission policy are derived in Sections IV and V. In Section VI, we compare the performance of the DCA–QoS with the FCA–QoS analytically through examples. In Section VII, we compare the performance of our DCA–QoS with the FCA–QoS numerically and describe conclusions in Section VIII.

II. The Cellular System

A one-dimensional (1-D) cellular system was considered for analytical simplicity. Referring to Fig. 1, we denote C_j as cell j, and $n_j(t)$ as the number of users in cell j at time t, for $j \in \{i-2, i-1, i, i+1, i+2\}$. We assume that C_i is the cell where a call admission request is made. New call arrivals are assumed to be *Poisson* distributed with the arrival rate λ_i in C_i. Call lasting time is assumed to be exponentially distributed with the mean $1/\mu$. Interhandoff time is also assumed to be exponentially distributed with the mean $1/h$. Furthermore, new call arrivals, call terminations and call handoffs are assumed to be mutually independent both within a cell and among cells. There are M distinct channels in the system. The channel reuse distance is assumed to be two, i.e., a channel can be reused in every other cell. Two adjacent cells can support up to M calls simultaneously.

III. Optimal DCA Scheme: Maximum Packing

To bound the performance that is independent of specific channel assignment algorithms, we chose MP. This was motivated by the fact that MP is an idealized DCA algorithm. It was assumed that a new call would be blocked only if there was no possible re-allocation of a channel to the call. This would include re-allocating calls in progress, which would result in the call to be carried. MP does not depend on any algorithm that implements the assignments of channels, therefore, providing the best performance a DCA scheme can possibly achieve in a given setting.

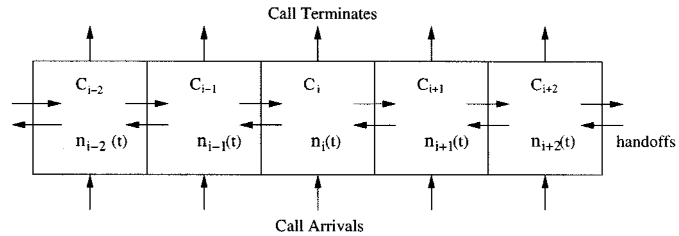

Fig. 1. One-dimensional cell array.

Due to MP policy, a call will be accepted in cell C_i at time t provided

$$n_{j-1}(t) + n_j(t) \leq M, \quad \text{for } j = i, i+1 \tag{1}$$

is satisfied after the call is accepted under the assumption of reuse distance equal to two. If the above condition is not satisfied, i.e., if by accepting the call (1) is violated, channel resource is going to be overloaded. The probability for such an event to occur is denoted as the *overload probability* $P_i^{(O)}(t)$ at cell C_i at time t, i.e.[1]

$$P_i^{(O)}(t) = \Pr\left\{O_i^{(L)}(t) \cup O_i^{(R)}(t)\right\}$$

where $O_i^{(L)}(t) = \{N_{i-1}(t) + N_i(t) > M\}$, $O_i^{(R)}(t) = \{N_{i+1}(t) + N_i(t) > M\}$. $N_j(t)$ is the number of users in cell C_j at time t, $j = i-1, i, i+1$. Since (1) only involves the number of calls at the current and neighboring cells, it is distributed in nature.

A. Distributed Call Admission Control with the "Best" DCA

We defined our admission control policy to be consistent with FCA–QoS but extended for dynamically assigned channels. Referring to Fig. 1, a new call is admitted to cell C_i at time t_0 if and only if the following admission conditions are satisfied.

1) At the current time t_0, the number of calls in cell C_i and its adjacent cells will not exceed the total number of channels in the system. In other words, (1) for MP strategy need to be satisfied at time t_0. This ensures that there is a channel available for the new call. If this condition is considered alone without QoS constraints, no QoS is enforced and we end up with the original MP.
2) At the future time t_1 $(=t_0 + T)$, the predicted overload probability of cell $C_j, \forall j \in \{i-1, i, i+1\}$ is bounded by a given QoS threshold, i.e.,

$$P_j^{(O)}(t_1 \,|\, t_0) \leq P_{\text{QoS}}, \quad \forall j \in \{i-1, i, i+1\} \tag{2}$$

where

$$P_j^{(O)}(t_1 \,|\, t_0) \triangleq \Pr\left\{ O_j^{(L)}(t_1) \cup O_j^{(R)}(t_1) \middle| A_j(t_0)\right\} \tag{3}$$

$$A_j(t_0) \triangleq \{N_{j-1}(t_0) = n_{j-1}, N_j(t_0) = n_j, N_{j+1}(t_0) = n_{j+1}\}. \tag{4}$$

$N_j(t_1)$ is the number of users in cell C_j at time t_1. Then some channels will be reserved for handoff calls in the

[1]Random variables and constants are denoted by capital letters. Observed values are denoted by small letters.

future. The number of channels reserved by the policy is not a constant, but determined by the number of users in cells C_j, where $j \in \{i-2, i-1, i, i+1, i+2\}$. This ensures that QoS will be maintained in cell C_i.

A handoff call is accepted by cell C_i if and only if (1) is satisfied. This means a handoff call is accepted as long as the channel reuse constraint is not violated.

The admission policy was implemented in two steps. The first step was to derive a QoS threshold $P_{\rm QoS}$. The next step was to evaluate the overload probabilities $P_j^{(O)}(t_1 \,|\, t_0)$, for $j \in \{i-1, i, i+1\}$.

It should be noticed that the idea of our distributed admission control algorithm is not limited to 1-D system with re-use distance two. Equations (1), (2), and (3) should be understood as an implementation of our distributed admission control algorithm for the special case we considered. The general idea behind the algorithm is based on the concept of neighborhood, which includes all the cells that are within the channel reuse distance. Then the admission condition can be stated as follows:

1) at the current time t_0, the number of calls in cell C_i and its re-use neighbors will not exceed the total number of channels in the system;
2) at the future time t_1 $(=t_0 + T)$, the predicted overload probability of cell $C_j, \forall j \in \{i, \text{neighbor}(i)\}$ is bounded by a given QoS threshold, where $\text{neighbor}(i)$ is the set of cells within the channel re-use distance.

IV. A QoS Measure on DCA for Admission Control

A commonly-used QoS measure is the handoff dropping probability. Such a measure was used in the FCA–QoS as a threshold on the highest tolerable handoff dropping probability. Since the handoff dropping probability was difficult to evaluate directly, a threshold was used on the overload probability instead. As will soon become clear, this measure is not able to provide the QoS for the case of DCA, where the contribution of new calls to the overload probability is too significant to be neglected. Although a weighted sum of the blocking probability and the handoff dropping probability has been used to circumvent this problem [7], a weighting factor has to be determined, which can be somewhat arbitrary.

We took a different approach to derive a QoS measure for the case of DCA through analyzing the contributions of the blocking and the handoff dropping probabilities to the overload probability. Using our QoS measure, we will show that a QoS threshold on the handoff dropping probability alone will not be optimal as long as the blocking probability to new calls is not negligible. We will also show empirically that the QoS measure we obtained leads to an almost guaranteed QoS on the handoff dropping probability under both uniform and nonuniform traffic conditions.

A. Deriving a QoS Measure for DCA

What should be an appropriate QoS measure $(P_{\rm QoS})$ for the distributed admission control with dynamically assigned channels? We claim that a logical choice for $P_{\rm QoS}$ is

$$P_{\rm QoS} = (1-\beta)P_{\rm BQoS} + \beta P_{\rm DQoS} \tag{5}$$

where

$$\beta = \frac{1}{1+\alpha'}, \quad \alpha' = \frac{1/h}{(1/\mu - 1/h)(1-P_{\rm BQoS})} \tag{6}$$

- $P_{\rm BQoS}$ and $P_{\rm DQoS}$: given QoS thresholds on the blocking and handoff dropping probabilities;
- μ and h: call service rate and call handoff rate, respectively;
- β: weighting factor defined in (6).

To explain how such a $P_{\rm QoS}$ is derived, we note that since $P_{\rm QoS}$ is a bound on the overload probability, a meaningful expression for $P_{\rm QoS}$ should be obtained through analyzing how the overload probability relates to the blocking and handoff dropping probabilities.

Let $n_{\rm new}$ and $n_{\rm hd}$ be the number of new call and handoff call requests. Let n_B and n_D be the number of blocked new call requests, and the number of dropped handoff call requests, respectively. Then the frequency $\hat{P}_o$ of the overload probability P_o can be expressed as

$$\hat{P}_o = \frac{n_B + n_D}{n_{\rm new} + n_{\rm hd}} = \frac{(n_B/n_{\rm new})(n_{\rm new}/n_{\rm hd}) + (n_D/n_{\rm hd})}{(n_{\rm new}/n_{\rm hd}) + 1}.$$

$n_B/n_{\rm new}$ is the frequency and $\hat{P}_B$ of the blocking probability P_B for new call arrivals. Similarly, $n_D/n_{\rm hd}$ is the frequency $\hat{P}_D$ of the handoff dropping probability P_D. If we let $\hat{\alpha} = n_{\rm new}/n_{\rm hd}$, we can obtain

$$\hat{P}_o = \frac{\hat{\alpha}\hat{P}_B + \hat{P}_D}{\hat{\alpha}+1}. \tag{7}$$

When the number of calls defined is large, the frequencies will approach the corresponding probabilities. If we assume also that $\hat{\alpha}$ approaches a quantity α, we can obtain

$$P_o = \frac{\alpha P_B + P_D}{\alpha + 1}. \tag{8}$$

We can then choose a QoS threshold to have a similar expression, i.e.,

$$P_{\rm QoS} = (1-\beta)P_{\rm BQoS} + \beta P_{\rm DQoS} \tag{9}$$

where $\beta = 1/(1+\alpha')$. The above derivation shows that intuitively

- β: ratio of handoff call requests over total call requests;
- $P_{\rm BQoS}$: desired blocking probability threshold;
- $P_{\rm DQoS}$: commonly used QoS threshold for the handoff dropping probability.

These two quantities are usually given as the QoS requirements in practice. Here we use a different quantity α' rather than α, since the parameter α in (8) was difficult to be measured directly and needs to be approximated.

To find a reasonable approximation for α, we replaced $n_{\rm new}$ and $n_{\rm hd}$ by their expected values $\mathbf{E}[n_{\rm new}]$ and $\mathbf{E}[n_{\rm hd}]$, respectively. Then $\mathbf{E}[n_{\rm new}] = \lambda t$, where $t > 0$ is the observing time.

If we assume that both $\mathbf{E}[n_{\rm new}]$ and $\mathbf{E}[n_{\rm hd}]$ are the expected number of new and handoff requests at the entire system[2]

$$\mathbf{E}[n_{\rm hd}] \approx \mathbf{E}[n_{\rm new}](1-P_B)\frac{1/\mu - 1/h}{1/h} \tag{10}$$

where $\mathbf{E}[n_{\rm new}](1-P_B)$ is the expected number of accepted new calls, $(1/\mu - 1/h)/(1/h)$[3] of handoffs for a call. α' can be chosen as

$$\alpha' \approx \frac{\mathbf{E}[n_{\rm new}]}{\mathbf{E}[n_{\rm hd}]} \approx \frac{1/h}{(1/\mu - 1/h)(1-P_{\rm BQoS})}. \tag{11}$$

Inserting (11) into (9), we have the QoS measure (threshold) $P_{\rm QoS}$ completely specified.

B. Examples

To examine whether such a QoS measure is meaningful, we consider three examples.

- *Example 1:* $P_{\rm BQoS} \ll 1$.

  Our QoS measure shown in (5) is a weighted sum of desired thresholds on blocking probability ($P_{\rm BQoS}$) and handoff dropping probability ($P_{\rm DQoS}$). The weighting factors $(1-\beta)$ and β, however, are nonlinear in terms of $P_{\rm BQoS}$ in general. When $P_{\rm BQoS} \ll 1$, the system is expected to run in a lightly loaded environment. In this case, the weighting factor reduces to a constant $1-\beta \approx (1/h)/(1/\mu)$, where β can be interpreted as the probability for a call to be handed off. A large β (and, thus, a small $1-\beta$) indicates that calls tend to be handed off frequently. That is, handoff calls are a major factor to cause the load at a cell. Therefore, $P_{\rm DQoS}$ is weighted more than $P_{\rm BQoS}$ in $P_{\rm QoS}$. Otherwise, new calls would contribute more significantly than handoff calls in overloading a cell, and $P_{\rm BQoS}$ will be weighted more heavily.

  This example shows that our derived QoS measure is more general than a commonly used weighted sum of $P_{\rm BQoS}$ and $P_{\rm DQoS}$, where the weighting factors were usually chosen in an ad hoc fashion.

- *Example 2:* $P_{\rm BQoS} = 0$.

  When $P_{\rm BQoS} = 0$, no call should be blocked. Then $P_{\rm QoS} = P_{\rm DQoS}/((1/h)/(1/\mu - 1/h) + 1)$. For this case, since average call lasting time $(1/\mu)$ is greater than average call sojourn time $(1/h)$ in a cell, $P_{\rm QoS} < P_{\rm DQoS}$. If we choose a threshold $P_{\rm DQOS}$ on the handoff dropping probability alone as the QoS threshold, we would have a looser control over the desired QoS. $P_{\rm QoS} \approx P_{\rm DQoS}$, only when the handoff happens more frequently than a new call arrival, i.e., $1/h \ll 1/\mu$ so that $(1/h)/(1/\mu - 1/h) \approx 0$. In other words, the threshold $P_{\rm DQoS}$ on the handoff dropping probability alone is a good QoS measure when $P_{\rm BQoS}$ is very small and the handoff occurs much more frequently than a new call arrival.

- *Example 3:* $P_{\rm BQoS} \approx 1$.

  In this example, no constraint is enforced on the blocking probability. This results in $1-\beta \approx 1$, thus, $P_{\rm QoS} \approx 1$. This simply means that no QoS constraint is enforced.

All these examples show that our QoS measure is consistent with existing ones. However, our approach is better as will be discussed later.

V. Evaluating the Overload Probability

The second step toward implementing the admission policy was to evaluate the overload probabilities, $P_j^{(O)}(t_1 \,|\, t_0)$, for $j \in \{i-1, i, i+1\}$. Since these probabilities have a similar form, we only need to evaluate $P_i^{(O)}(t_1 \,|\, t_0)$. Using the model given in Section II, we obtained an expression of the overload probability for DCA–QoS in Theorem 1.

Theorem 1: Let M be the number of channels in the system. Let $n_i(t_0)$ the number of users at cell C_i at time t_0. At time t_1 ($\triangleq t_0 + T$), let $N_i(t_1)$ be the number of users in cell C_i. Define $F_i(k) \triangleq \Pr\{N_i(t_1) \le k \,|\, N_i(t_0) = n_i(t_0)\}$ to be the cumulative distribution function of $N_i(t_1)$ with initial value $N_i(t_0) = n_i(t_0)$. Then with DCA, the overload probability at cell C_i is

$$P_i^{(O)}(t_1 \,|\, t_0) = 1 - \sum_{k=0}^{M} F_{i-1}(M-k) \cdot F_{i+1}(M-k) \cdot (F_i(k) - F_i(k-1)). \tag{12}$$

The proof of the theorem can be found in Appendix A. Intuitively, the second term of (12) sums up all possible nonoverloading combinations, thus, the overload probability can be expressed in (12).

Since $P_i^{(O)}(t_1 \,|\, t_0)$ depends only on $F_{i-1}(k)$, $F_i(k)$, and $F_{i+1}(k)$, we focused on evaluating $F_i(k)$ alone. $F_{i-1}(k)$ and $F_{i+1}(k)$ can be evaluated similarly. $F_i(k)$ is determined by its probability mass function $\Pr\{N_i(t_1) = k \,|\, N_i(t_0) = n_i(t_0)\}$. Such a probability can be evaluated based on the given traffic model as given in Theorem 2.

Theorem 2:

$$\begin{aligned}\Pr\{N_i(t_1) = k \,|\, N_i(t_0) = n_i(t_0)\} &= e^{-\left(\lambda_i^{(I)}(t_0) + \lambda_i^{(O)}(t_0)\right)T} \left(\frac{\lambda_i^{(I)}(t_0)}{\lambda_i^{(O)}(t_0)}\right)^{\frac{k - n_i(t_0)}{2}} \\ &\quad \cdot I_{|k - n_i(t_0)|}\left(2T\sqrt{\lambda_i^{(I)}(t_0)\lambda_i^{(O)}(t_0)}\right)\end{aligned} \tag{13}$$

where $I_k(x)$ is the modified *Bessel* function of the first kind of order k. $\lambda_i^{(I)}(t_0)$ and $\lambda_i^{(O)}(t_0)$ can be regarded as the equivalent arrival and departure rate to and from cell C_i

$$\lambda_i^{(I)}(t_0) = \lambda_i + (n_{i-1}(t_0) + n_{i+1}(t_0))h/2 \tag{14}$$

$$\lambda_i^{(O)}(t_0) = n_i(t_0)(\mu + h). \tag{15}$$

[2]This assumption is made by the motivation that a QoS threshold is defined for the system rather than for each cell. Then each base station can try to perform the best admission control possible to achieve the QoS.

[3]Assuming $1/\mu > 1/h$, i.e., handoff rate is higher than call completion rate.

The proof of the theorem can be found in Appendix B. When T, the prediction interval, is small, the *Bessel* function $I_k(x)$ can be approximated as $I_k(x) \approx (x^k/2^k\Gamma(k+1))$, for $x \to 0, k > 0$. We have a simpler expression for (13)

Corollary 1: For $T \to 0$

$$\Pr\{N_i(t_1)=k \mid N_i(t_0)=n_i(t_0)\} \approx \begin{cases} \alpha_i \left(\lambda_i^{(I)}T\right)^{k-n_i(t_0)} \Big/ [k-n_i(t_0)]! & k \ge n_i(t_0) \\ \alpha_i \left(\lambda_i^{(O)}T\right)^{n_i(t_0)-k} \Big/ [n_i(t_0)-k]! & k < n_i(t_0) \end{cases} \tag{16}$$

where

$$\alpha_i = \left(e^{\lambda_i^{(I)}T} + e^{\lambda_i^{(O)}T} - 1\right)^{-1}$$

is a normalizing factor. $\lambda_i^{(I)}$, and $\lambda_i^{(O)}$ are shown in (14) and (15), respectively.

The validity of *Poisson*-like approximation is further discussed in Appendix C.

Combining Theorem 1 and Theorem 2, we have the expression for the overload probability $P_i^{(O)}(t_1 \mid t_0)$ at cell C_i. The other two overload probabilities can be obtained similarly.

In general, these overload probabilities can not be simplified to closed forms. But they can be evaluated numerically, and compared with the QoS threshold P_{QoS}. If $P_j^{(O)}(t_1 \mid t_0) \le P_{\text{QoS}}$ are satisfied for all $j \in \{i-1, i, i+1\}$, a new call is admitted. Otherwise, the call is rejected.

VI. Comparison with FCA: Analytical Results

In order to understand intuitively the advantage of using DCA as opposed to FCA, we first simplified the overload probabilities for FCA–QoS and DCA–QoS for some special cases so that the results can be more intuitive to understand.

A. Overload Probability for the FCA–QoS

Using the definition of $F_i(k)$, we can easily obtain the overload probability for the FCA–QoS.

Lemma 1: Let M be the number of channels in the system, each cell is assigned $M/2$ channels. Let $n_i(t_0)$ be the number of users at cell C_i at time t_0. At time t_1 ($\triangleq t_0 + T$), let $N_i(t_1)$ be the number of users in cell C_i. Define $F_i(k) \triangleq \Pr\{N_i(t_1) \le k \mid N_i(t_i) = n_i(t_0)\}$ to be the cumulative distribution function of $N_i(t_1)$ with initial value $N_i(t_0) = n_i(t_0)$. Then with FCA policy, the overload probability at cell C_i is

$$\begin{aligned} P_i^{(O)}(t_1 \mid t_0) &= \Pr\left\{ N_i(t_1) > \frac{M}{2} \;\middle|\; N_i(t_0)=n_i(t_0)\right\} \\ &= 1 - F_i\left(\frac{M}{2}\right). \end{aligned} \tag{17}$$

The result given by (17) means that under our system model, the overload probability for FCA is determined by the availability of channels in the cell and traffic parameters, such as call arrival rate, handoff rate, and service rate. It is noted that such an overload probability is exact whereas the one derived in [5] is a special case of Lemma 1 which approximates $F_i(M/2)$ by a Gaussian distribution.

B. A Special Case: Light Traffic

When the system is lightly loaded, i.e., the number $n_j(t_0)$ of existing calls in a cell is small, for $j \in \{i-1, i, i+1\}$. Equation (16) is reduced to

$$\Pr\{N_i(t_1)=k \mid N_i(t_0)=n_i(t_0)\} \approx \alpha_i \frac{\left(\lambda_i^{(I)}T\right)^k}{k!}, \quad \text{for } k = 0,1,2,\ldots. \tag{18}$$

For the sake of simplicity, we assume the arrival rate of new calls at all three cells is the same, i.e., $\lambda_i = \lambda$. Since $n_j(t_0)$'s $(j = i - m1, i, i+1)$, are small, the influence of handoffs can be neglected. Therefore, it is reasonable to assume that $\alpha_j \triangleq \alpha$, for $j = i-1, i, i+1$. The overload probability can be simplified as the following form:

$$\text{FCA-QoS: } P_i^{(O)}(t_1 \mid t_0) = \alpha \frac{(\lambda T)^{\frac{M}{2}+1}}{(\frac{M}{2}+1)!} + o\left(T^{\frac{M}{2}+1}\right) \tag{19}$$

$$\begin{aligned} \text{DCA-QoS: } P_i^{(O)}(t_1 \mid t_0) = {} & 3\alpha \frac{(\lambda T)^{M+1}}{(M+1)!} \\ & + 2\alpha^2 \frac{(2\lambda T)^{M+1} - (\lambda T)^{M+1}}{(M+1)!} \\ & + o(T^{M+1}). \end{aligned} \tag{20}$$

The proof of (19) and (20) can be found in Appendix D. From (19) and (20), it is easy to conclude that the DCA–QoS has a smaller overload probability than the FCA–QoS, since it is in a much higher order of T than that of the FCA–QoS. Hence, the DCA–QoS performs better than the FCA–QoS in light traffic. Intuitively, when a cell is lightly loaded, there is a lot of room for the DCA–QoS to adapt itself to traffic variations and hence, outperforms the FCA–QoS. It should be noted that if no QoS is considered, the results shown in (19) and (20) will reduce to those given by Kelly [8].

C. A Special Case: Heavy Traffic

Does the DCA–QoS always perform better than the FCA–QoS? To answer this question, we compared the overload probabilities between the FCA–QoS and the DCA–QoS under such an extreme heavy traffic condition that if accepting a call, the cell C_i would use up almost all of its available channels.

If T is small, (16) can be approximated as

$$\Pr\{N_i(t_1)=k \mid N_i(t_0)=n_i(t_0)\} \approx \begin{cases} \lambda_i^{(I)}T + o(T) & k = n_i(t_0)+1 \\ 1 - \lambda_i^{(I)}T - \lambda_i^{(O)}T + o(T) & k = n_i(t_0) \\ \lambda_i^{(O)}T + o(T) & k = n_i(t_0) - 1. \end{cases} \tag{21}$$

In the case of the FCA–QoS, when $n_i(t_0) = M/2$, substituting (21), (14) and (15) into (17), we have

$$\begin{aligned} P_i^{(O)}(t_1 \mid t_0) &= \lambda_i^{(I)}T + o(T) \\ &= \left(\lambda + \frac{h}{2}(n_{i-1}(t_0) + n_{i+1}(t_0))\right) T + o(T) \end{aligned} \tag{22}$$

where we assume $\lambda_i = \lambda$, for all i, for simplicity.

In the case of the DCA–QoS, when $n_{i-1}(t_0) + n_i(t_0) = M$ and $n_{i+1}(t_0) + n_i(t_0) = M$, by inserting (21), (14), and (15) into (12), we have

$$\begin{aligned} P_i^{(O)}(t_1 \mid t_0) &= \lambda_{i-1}^{(I)} T + \lambda_i^{(I)} T + \lambda_{i+1}^{(I)} T + o(T) \\ &= 3\left(\lambda + \frac{h}{6}(n_{i-2}(t_0) + n_{i-1}(t_0) + 2n_i(t_0) \right. \\ &\quad \left. + n_{i+1}(t_0) + n_{i+2}(t_0))\right) T + o(T). \end{aligned} \tag{23}$$

If we regard $(1/2)(n_{i-1}(t_0) + n_{i+1}(t_0))$ in (22) and $(1/6)(n_{i-2}(t_0) + n_{i-1}(t_0) + 2n_i(t_0) + n_{i+1}(t_0) + n_{i+2}(t_0))$ in (23) as two estimates to the sample mean of number of existing calls in each cell, and denote the mean as $\bar{n}$, then (22) and (23) can be rewritten as follows:

$$\text{FCA-QOS: } P_i^{(O)}(t_1 \mid t_0) = (\lambda + h\bar{n})T + o(T) \tag{24}$$

$$\text{DCA-QoS: } P_i^{(O)}(t_1 \mid t_0) = 3(\lambda + h\bar{n})T + o(T). \tag{25}$$

By observing (24) and (25), it is obvious that the DCA–QoS has a higher overload probability than that of the FCA–QoS. This is in agreement with the result for pure FCA and MP in [8]. An intuitive explanation is that under heavy traffic conditions, as there is little room for dynamically assigning channels, DCA–QoS disrupts the more optimal packing of FCA–QoS.

VII. Numerical Results and Discussions

As special cases shed light on the advantage of DCA over FCA, we evaluated the exact performance of DCA–QoS by solving (2) numerically and compared it with the performance of FCA.

A. Experimental Setup

We used the same 1-D simulation system and the experimental setup as in [5] to obtain our results for the purpose of comparison. The system consists of ten cells arranged on a circle so that the boundary effect can be ignored. The number (M) of distinguishable channels in the system is 40. We assume that the call duration ($1/\mu = 500$ s) is exponentially distributed. The time a call stays in a cell before handing off to another ($1/h = 100$ s) is also exponentially distributed. We used the same performance measure as the FCA–QoS, i.e., the new call blocking probability P_B and handoff dropping probability P_D. The desired maximum tolerable handoff dropping probability P_{DQoS} is assumed to be 0.01.

B. Testing the QoS Threshold P_{QoS}

The first purpose of the experiments is to test the validity of the derived QoS measure. Fig. 2 compares the handoff dropping probability P_D among the FCA–QoS, MP and the DCA–QoS under different QoS thresholds. For the case of the FCA–QoS, P_{QoS} is chosen to satisfy $P_{\mathrm{QoS}} = P_{\mathrm{DQoS}}$. Then the resulting handoff dropping probability P_D is much smaller than P_{DQoS}. This is because our QoS threshold (P_{QoS}) should include both P_{BQoS} and P_{DQoS} as explained in Section IV. If not, the actual

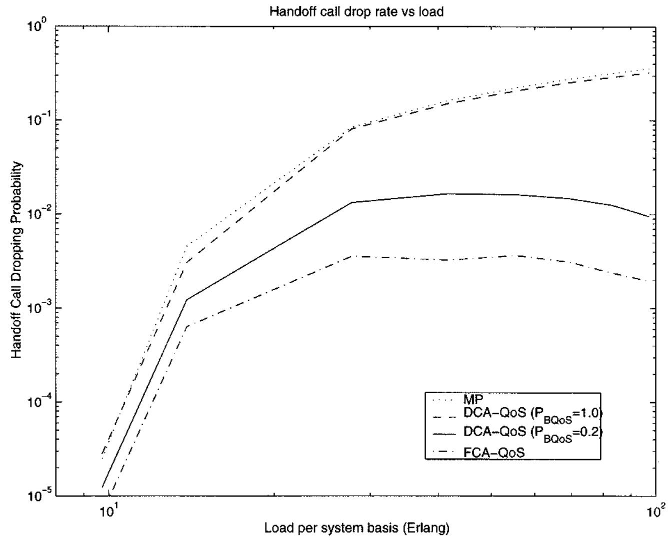

Fig. 2. Comparison of QoS Threshold ($P_{\mathrm{DQoS}} = 0.01$). "$-\cdot-$": FCA–QoS; "—": DCA–QoS ($P_{\mathrm{BQoS}} = 0.2$); "$- - -$": DCA–QoS ($P_{\mathrm{BQoS}} = 1.0$); "$\cdots$": MP.

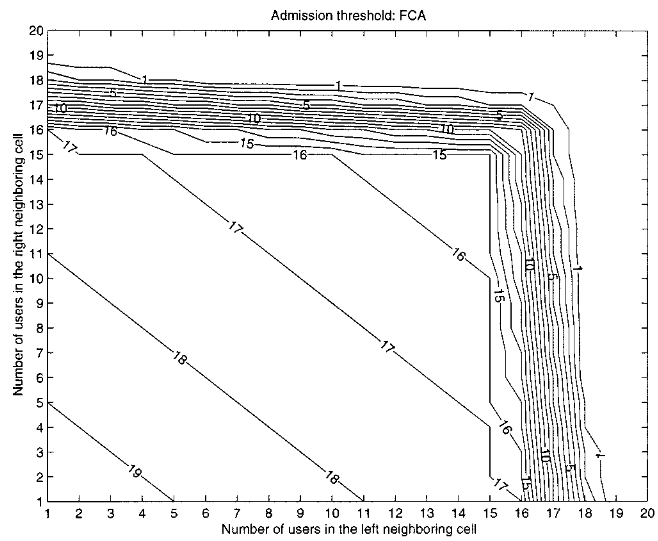

Fig. 3. Contour of admission hreshold for the FCA–QoS ($M = 40$).

QoS threshold set on the handoff dropping probability is much smaller than it should be. This results in a very low handoff dropping probability P_D and is undesirable since the capacity is not being fully utilized.

For the DCA–QoS with P_{BQoS} chosen to be one, the handoff dropping probability P_D is almost the same as that for MP. This is consistent with example three we considered in Section IV-B, which means that the QoS is not enforced. When $P_{\mathrm{BQoS}} = 0.2$, DCA–QoS maintains the handoff dropping probability P_D to be close to P_{DQoS} more consistently than FCA–QoS where P_{DQoS} is being used alone.

C. Admission Thresholds for DCA–QoS and FCA–QoS

One way to compare the performance of the DCA–QoS and the FCA–QoS is to show the admission thresholds obtained from (2) for the DCA–QoS and the FCA–QoS, respectively.

The admission threshold for the FCA–QoS is shown in Fig. 3, which reproduces the results in [5]. The X-axis and Y-axis of

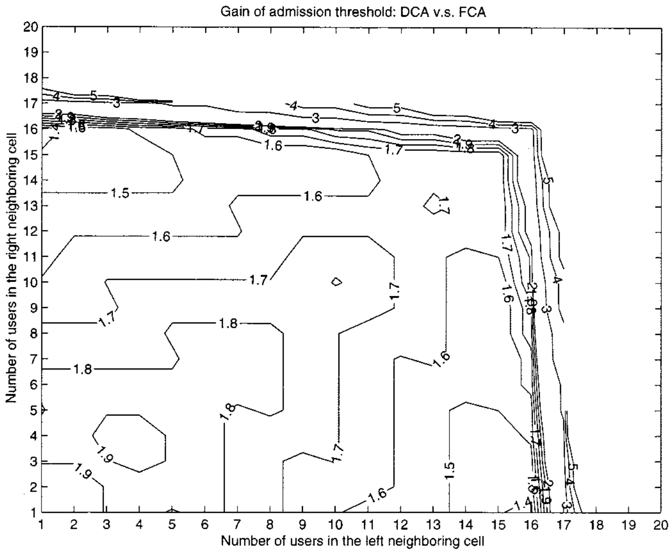

Fig. 4. Contour of Admission Threshold Gain for the DCA–QoS over the FCA–QoS $(M = 40)$.

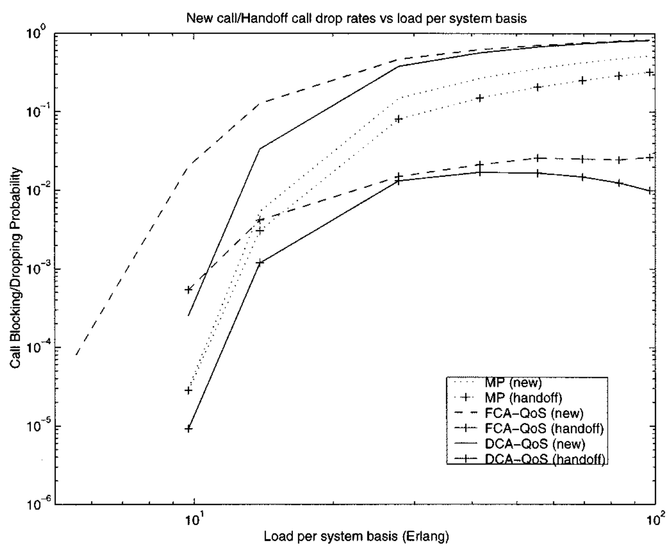

Fig. 5. Comparison of DCA–QoS to FCA–QoS and MP: uniform traffic. "—": DCA–QoS; "– – –": FCA–QoS; "···": MP. Three curves from the top without "+": new call blocking probability P_B. Three curves with "+": handoff dropping probability P_D.

Fig. 3 are the number of users in the left and the right neighboring cells, respectively. The admission threshold is the value of the nearest upper-right line to the reference point. For example, if the number of users in the left and the right neighboring cells are two and three, since the nearest line to the upper-right of the point (2, 3) has value 19, the admission threshold is 19. Under heavily loaded traffic, i.e., when the number of users in the left and the right cells approaching the number of channels assigned to those cells, the admission threshold will approach zero.

Fig. 4 shows the ratio of admission thresholds for the DCA–QoS to the FCA–QoS. From Fig. 4, we can find that the gain of admission threshold for the DCA–QoS to the FCA–QoS is about two under light traffic. Under heavily loaded traffic, the gain of admission threshold for the DCA–QoS to the FCA–QoS illustrated in the figure is getting bigger and bigger. This does not mean that the DCA–QoS outperforms the FCA–QoS in heavy traffic, since the de-numerator, the admission threshold of the FCA–QoS, is approaching zero then.

D. Performance of DCA–QoS and FCA–QoS

To further assess the optimal performance of the distributed admission control policy with the DCA–QoS, and the advantage of using the DCA–QoS compared with the FCA–QoS, we compare the performance of the DCA–QoS with that of MP and the FCA–QoS under various traffic conditions. $P_{\mathrm{BQoS}} = 0.2$ and $P_{\mathrm{DQoS}} = 0.01$ are used to set our QoS threshold given in (5), and $P_{\mathrm{DQoS}} = 0.01$ is used as the QoS threshold for the FCA–oS.

The comparison of the DCA–QoS to MP and the FCA—QoS under uniform traffic is shown in Fig. 5. In terms of the blocking probability (P_B), it increases with the load for all three policies. MP always has the lowest blocking probability among the three policies, since no QoS is enforced to control the handoff dropping probability. The FCA–QoS has a higher P_B than the DCA–QoS. This is because DCA makes more effective use of channels. In terms of the handoff dropping probability (P_D), the DCA–QoS always has the lowest value among the three policies. The P_D for MP goes up with the load, since no QoS is enforced for MP. For both the FCA–QoS and the DCA–QoS, P_D saturates around the pre-defined value P_{DQoS}. The DCA–QoS always has a lower handoff dropping probability P_D than the FCA–QoS. In other words, QoS is better maintained by the DCA–QoS. We think the reason for this is that a Gaussian approximation is used to approximate the overload probability for the FCA–QoS [5], whereas a more accurate model is used in this work to derive the overload probabilities in the DCA–QoS as shown in Appendix C.

To compare the performance of the DCA–QoS, MP, and the FCA–QoS under nonuniform traffic, the same simulation was used. The new call arrival rate is set to be the same every other cell so that the neighboring cells have a different new arrival rate, and the ratio is two between the heavy and the light load.[4] Figs. 6 and 7 show the new call blocking probability (P_B) and the handoff call dropping probability (P_D) of the DCA–QoS, the FCA–QoS, and MP under nonuniform traffic, respectively. Similar to the results under the uniform traffic, MP has the lowest P_B and the FCA–QoS has the highest P_B in both lightly loaded and heavily loaded cells. Referring to Fig. 7, in lightly loaded cells, P_D has almost the same trend for both the DCA–QoS and the FCA–QoS. In heavily-loaded cells, however, the DCA–QoS maintains a much better control on P_D than the FCA–QoS does. Intuitively, this is because DCA has the ability to adapt to traffic variations automatically.

To assess how much has been gained by the DCA–QoS as compared to the FCA–QoS in terms of capacity, Table I compares *Erlang* load between the DCA–QoS and the FCA–QoS when $P_B = 0.2$, $P_{\mathrm{DQoS}} = 0.01$ and $1/\mu = 500$ s. We observed that the DCA–QoS consistently has a higher capacity than the FCA–QoS under both uniform and nonuniform traffic. In particular, compared with the FCA–QoS, the DCA—QoS has the

[4]The "heavy" load here refers to the cells with a heavier load in the nonuniform traffic, which is different from the heavy load discussed in Section VI-C.

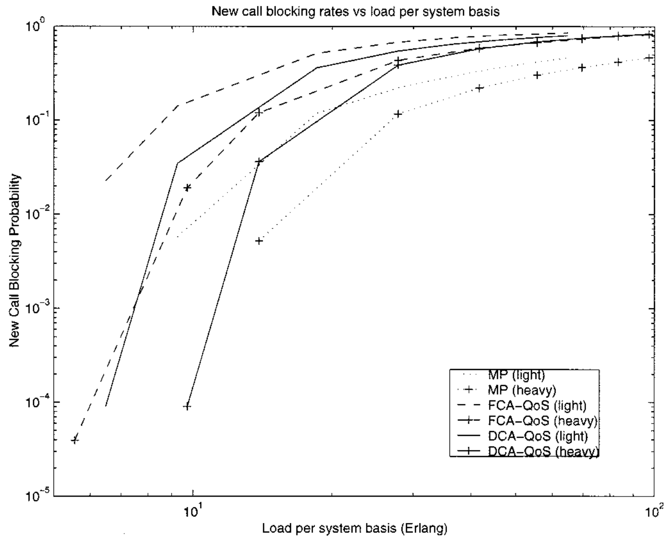

Fig. 6. Comparison of DCA–QoS to FCA–QoS and MP: Nonuniform Traffic. "—": DCA–QoS; "— — —": FCA–QoS; "· · ·": MP. Curves with "+": Heavily loaded; curves without "+": Lightly loaded.

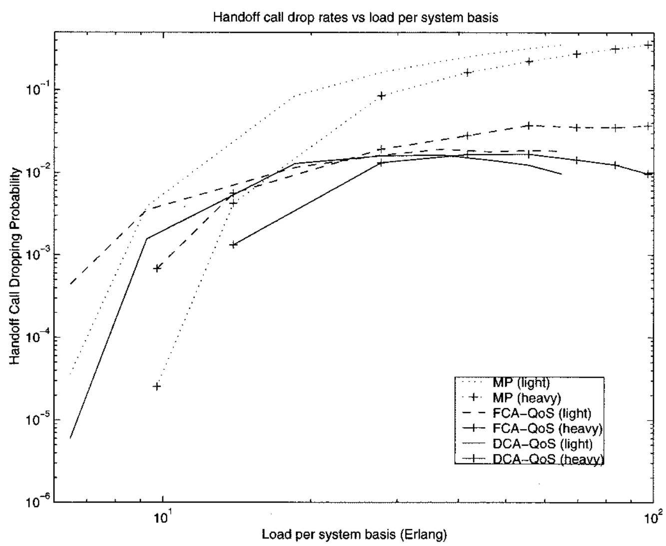

Fig. 7. Comparison of DCA–QoS to FCA–QoS and MP: Nonuniform traffic. "—": DCA–QoS; "— — —": FCA–QoS; "· · ·": MP. Curves with "+": Heavily loaded; curves without "+": Lightly loaded.

highest gain of 30% in capacity for nonuniform traffic at lightly loaded cells. This is because the DCA is effective to adapt to channel allocation to nonuniform traffic and works the best under light load. When the load is heavy, the gain reduces to 17% since the heavy load leaves a little for dynamically allocating the channels.

VIII. Conclusion

In this work, we have analyzed the performance of distributed admission control with DCA and QoS provisioning in cellular environment, where co-channel reuse distance is considered as the only limiting factor to channel reuse. We have first derived a novel QoS threshold that maintains the QoS on handoff dropping probabilities consistently under both uniform and nonuniform conditions. We have then investigated the DCA–QoS in

TABLE I
Comparison of *Erlang* Load When the Overload Probability Is 0.01

Traffic Type	DCA-QoS	FCA-QoS	Gain (%)
Uniform	20.57	16.84	22.15
Non-uniform (Light)	13.94	10.71	30.16
Non-uniform (Heavy)	20.37	17.41	17.00

such a way that a performance bound is provided on how well DCA can possibly do under the given QoS constraint in the given setting. Under the special cases, we found analytically that the DCA is better than the FCA–QoS in light traffic conditions. We have found empirically that the capacity (in *Erlang*) gain due to using the DCA is 17% to 30% under various traffic conditions. Although the results are derived from a 1-D system, they serve the goal of providing a systematic way of bounding the performance of DCA with QoS. The approach we have used in this work can be readily extended to 2-D systems with increased computational complexity.

Appendix A
Proof of Theorem 1

Since both events $O_i^{(L)}(t_1)$ and $O_i^{(R)}(t_1)$ depend on $N_i(t_1)$, we can use the conditional probabilities to

$$\begin{aligned}
&P_i^{(O)}(t_1 \mid t_0)\\
&\quad= \Pr\left\{ O_i^{(L)}(t_1) \cup O_i^{(R)}(t_1) \,\Big|\, A_i(t_0) \right\}\\
&\quad= \sum_{k=0}^{M} \Pr\{\{N_{i-1}(t_1) > M-k\}\\
&\qquad \cup \{N_{i+1}(t_1) > M-k\} \mid A_i(t_0), N_i(t_1)=k\}\\
&\qquad \cdot \Pr\{N_i(t_1)=k \mid A_i(t_0)\}\\
&\quad= \sum_{k=0}^{M} \Pr\{\{N_{i-1}(t_1) > M-k\}\\
&\qquad \cup \{N_{i+1}(t_1) > M-k\} \mid A_i(t_0)\}\\
&\qquad \cdot \Pr\{N_i(t_1)=k \mid N_i(t_0)=n_i(t_0)\}\\
&\quad= 1 - \sum_{k=0}^{M} \Pr\{\{N_{i-1}(t_1) \le M-k\}\\
&\qquad \cap \{N_{i+1}(t_1) \le M-k\} \mid A_i(t_0)\}\\
&\qquad \cdot \Pr\{N_i(t_1)=k \mid N_i(t_0)=n_i(t_0)\}.
\end{aligned} \tag{26}$$

Since the number of calls in different cells are assumed to be independent[5]

$$\begin{aligned}
&\Pr\{\{N_{i-1}(t_1) \le M-k\} \cap \{N_{i+1}(t_1) \le M-k\} \mid A_i(t_0)\}\\
&\quad= \Pr\{N_{i-1}(t_1) \le M-k \mid A_i(t_0)\}\\
&\qquad \cdot \Pr\{N_{i+1}(t_1) \le M-k \mid A_i(t_0)\}\\
&\quad= \Pr\{N_{i-1}(t_1) \le M-k \mid N_{i-1}(t_0) = n_{i-1}(t_0)\}\\
&\qquad \cdot \Pr\{N_{i+1}(t_1) \le M-k \mid N_{i+1}(t_0) = n_{i+1}(t_0)\}\\
&\quad= F_{i-1}(M-k)\cdot F_{i+1}(M-k)
\end{aligned} \tag{27}$$

[5]Since we use infinite number of channels to estimate overload probability, where product form solution is available in spite of the existence of handoff, it is reasonable to assume that number of calls in different cells are independent.

and

$$\Pr\{N_i(t_1) = k \mid N_i(t_0) = n_i(t_0)\} = F_i(k) - F_i(k-1). \tag{28}$$

By putting (27) and (28) into (26), we have

$$P_i^{(O)}(t_1 \mid t_0) = 1 - \sum_{k=0}^{M} F_{i-1}(M-k) \cdot F_{i+1}(M-k) \cdot (F_i(k) - F_i(k-1)).$$

Q.E.D.

Appendix B
Proof of Theorem 2

The theorem can be proved based on our assumed traffic model.

Since $\Pr\{N_i(t_1) = k \mid N_i(t_0) = n_i(t_0)\}$ can be written as

$$\begin{aligned}\Pr\{N_i(t_1) = k \mid N_i(t_0) = n_i(t_0)\} &= \Pr\{N_i(t_1) - N_i(t_0) = k - n_i(t_0) \mid N_i(t_0) = n_i(t_0)\} \\ &= \Pr\{\Delta N_i(t_0) = k - n_i(t_0) \mid N_i(t_0) = n_i(t_0)\}\end{aligned}$$

where $\Delta N_i(t_0) \triangleq N_i(t_1) - N_i(t_0)$ is the number of user changed from t_0 to t_1. Meanwhile, $\Delta N_i(t_0)$ is the difference between the number of users entering and leaving cell C_i during the period T, which, based on the *Poisson* and independent assumptions, are independently distributed *Poisson*, with rate expressed in (14) and (15), respectively.

A result is given in [9] for a subtraction of two *Poisson* random variables. Let X and Y be two independent *Poisson* random variables with the mean pt and qt. Then

$$\Pr\{X - Y = k\} = e^{-(p+q)t}(p/q)^{\frac{k}{2}} I_{|k|}(2t\sqrt{pq})$$

for k being an integer, where

$$I_k(x) = \sum_{m=0}^{\infty} \frac{1}{m!\Gamma(k+m+1)} \left(\frac{x}{2}\right)^{2m+k}$$

is the modified *Bessel* function of the first kind. For the case we consider that X and Y correspond to the number of users entering and leaving the cell C_i during the period T. This completes the proof of Theorem 2. Q.E.D.

Appendix C
The Accuracy of *Poisson*-Like Approximation

To illustrate the validity of *Poisson*-like approximation to $\Pr\{N_i(t_1) = k \mid N_i(t_0) = n_i(t_0)\}$ expressed in Corollary 1, we compared the cumulative distribution functions obtained through Theorem 2 $(F_i(k))$, Corollary 1 $(F_i^{(P)}(k))$, and Gaussian approximation $(F_i^{(G)}(k))$ [5]. Since $F_i(k)$, $F_i^{(P)}(k)$ and $F_i^{(G)}(k)$ are similar for different values of λ_i, μ, h, and $N_j(t_0)$'s $(j = i-1, i, i+1)$, we will take $\lambda_i/\mu = 30$, $N_{i-1}(t_0) = N_{i+1}(t_0) = 16$ in the rest of this section.

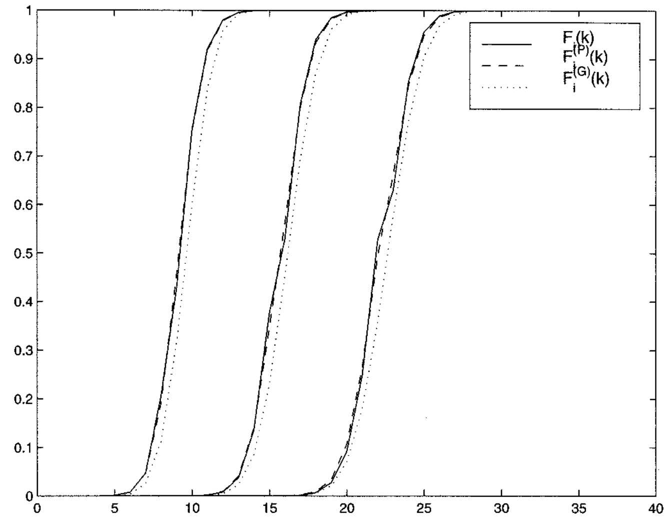

Fig. 8. Comparison of $F_i(k)$, $F_i^{(P)}(k)$ and $F_i^{(G)}(k)$ $(M = 40)$. "—": $F_i(k)$; "– – –": $F_i^{(P)}(k)$; "· · ·": $F_i^{(G)}(k)$. From left to right: $N_i(t_0) = 9, 16, 23$.

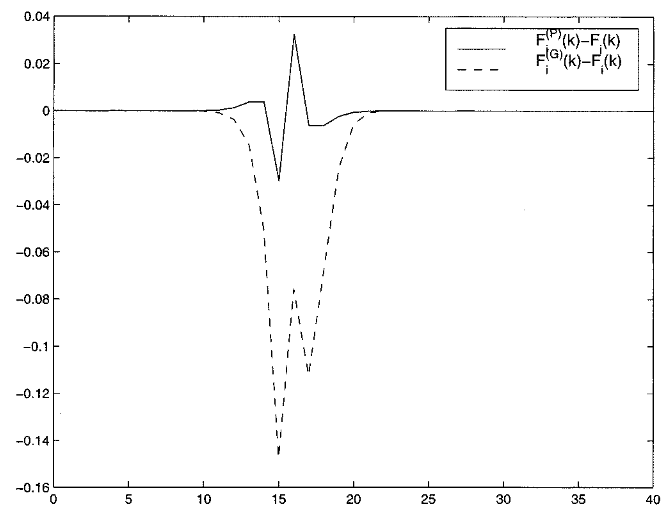

Fig. 9. Approximation error to $F_i(k)$ $(M = 40, N_i(t_0) = 16)$. "—": $F_i^{(P)}(k) - F_i(k)$; "– – –": $F_i^{(G)}(k) - F_i(k)$.

Fig. 8 illustrates the cumulative distribution functions of $F_i(k)$, $F_i^{(P)}(k)$ and $F_i^{(G)}(k)$ when $N_i(t_0) = 9, 16$, and 23 (from left to right). The solid lines, which are obtained through Theorem 2, are exact cumulative distribution functions $(F_i(k))$. The dashed lines are through *Poisson*-like approximation $(F_i^{(P)}(k))$. The dotted lines are from the *Gaussian* approximation $(F_i^{(G)}(k))$. Fig. 9 gives the errors of *Poisson*-like approximation and *Gaussian* approximations for $N_i(t_0) = 16$. It is observed from the two figures that *Poisson*-like approximation is more accurate than *Gaussian*, which always underestimates the actual cumulative distribution functions. *Poisson*-like approximation is very accurate except for a couple of points close to the initial point $(N_i(t_0) = 16)$. The nice property of such an approximation is that the cumulative errors made close to the initial point is complementary.

APPENDIX D
PROOF OF EXAMPLE OF LIGHT TRAFFIC

For the FCA–QoS, by inserting (18) into (17), we have

$$P_i^{(O)}(t_1 \mid t_0) \approx \alpha_i \sum_{k=\frac{M}{2}}^{\infty} \frac{\left(\lambda_i^{(I)} T\right)^k}{k!} = \alpha_i \left(\frac{\left(\lambda_i^{(I)} T\right)^{\frac{M}{2}+1}}{(\frac{M}{2}+1)!} + o\left(T^{\frac{M}{2}+1}\right) \right). \quad (29)$$

For the DCA–QoS, we can rewrite (12) as

$$\begin{aligned} P_i^{(O)}(t_1 \mid t_0) &= \Pr\{N_i(t_1) > M \mid N_i(t_0) = n_i(t_0)\} \\ &\quad + \Pr\{N_{i-1}(t_1) > M \mid N_{i-1}(t_0) = n_{i-1}(t_0)\} \\ &\quad + \Pr\{N_{i+1}(t_1) > M \mid N_{i+1}(t_0) = n_{i+1}(t_0)\} \\ &\quad + \sum_{j=0}^{M} \Pr\{N_i(t_1) = j \mid N_i(t_0) = n_i(t_0)\} \\ &\quad \cdot \Pr\{N_{i+1}(t_1) > M + j \mid N_{i+1}(t_0) = n_{i+1}(t_0)\} \\ &\quad + \sum_{j=0}^{M} \Pr\{N_i(t_1) = j \mid N_i(t_0) = n_i(t_0)\} \\ &\quad \cdot \Pr\{N_{i-1}(t_1) > M + j \mid N_{i-1}(t_0) = n_{i-1}(t_0)\} \\ &\quad + o(T^{M+1}). \end{aligned} \quad (30)$$

Similar to (29)

$$\Pr\{N_j(t_1) > M \mid N_j(t_0) = n_j(t_0)\} \approx \alpha_j \left(\frac{\left(\lambda_j^{(I)} T\right)^{M+1}}{(M+1)!} + o\left(T^{M+1}\right) \right) \quad (31)$$

for $j = i-1, i, i+1$, and

$$\begin{aligned} &\sum_{j=0}^{M} \Pr\{N_i(t_1) = j \mid N_i(t_0) = n_i(t_0)\} \\ &\quad \cdot \Pr\{N_{i+1}(t_1) > M + j \mid N_{i+1}(t_0) = n_{i+1}(t_0)\} \\ &\quad \approx \sum_{j=0}^{M} \alpha_i \frac{\left(\lambda_i^{(I)} T\right)^j}{j!} \alpha_{i+1} \frac{\left(\lambda_{i+1}^{(I)} T\right)^{M-j+1}}{(M-j+1)!} + o(T^{M+1}) \\ &\quad = \frac{\alpha_i \alpha_{i+1}}{(M+1)!} \left(\left(\lambda_i^{(I)} + \lambda_{i+1}^{(I)}\right)^{M+1} - \left(\lambda_i^{(I)}\right)^{M+1} \right) \\ &\quad \times T^{M+1} + o(T^{M+1}). \end{aligned} \quad (32)$$

For the sake of simplicity, we assume the arrival rate of new calls at all three cells is the same, i.e., $\lambda_i = \lambda$. Since $n_j(t_0)$'s ($j = i-1, i, i+1$), are small, the influence of handoffs can be neglected. Therefore, it is reasonable to assume $\alpha_j \triangleq \alpha$, for $j = i-1, i, i+1$. Equations (29) and (30) will become (19) and (20).

REFERENCES

[1] I. Katzela and M. Naghshineh, "Channel assignment schemes for cellular mobile telecommunication systems," *IEEE Personal Commun.*, vol. 3, pp. 10–31, June 1996.

[2] V. O. K. Li and X. Qiu, "Personal communication systems (PCS)," *Proc. IEEE*, vol. 83, pp. 1210–1243, Sept. 1995.

[3] S. Tekinay and B. Jabbari, "Handover and channel assignment in mobile cellular networks," *IEEE Commun. Mag.*, pp. 42–46, Nov. 1991.

[4] R. Ramjee, R. Nagarajan, and D. Towsley, "On optimal call admission control in cellular networks," in *Proc. IEEE INFOCOM'96*, 1996, pp. 43–50.

[5] M. Naghshineh and M. Schwartz, "Distributed call admission control in mobile/wireless networks," *IEEE J. Select. Areas Commun.*, vol. 14, pp. 711–716, May 1996.

[6] D. E. Everitt and N. W. Macfadyen, "Analysis of multi-cellular mobile radiotelephone systems with loss," *Brit. Telecom Technol. J.*, vol. 1, no. 2, pp. 37–45, 1983.

[7] D. Hong and S. S. Rappaport, "Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and nonprioritized handoff procedures," *IEEE Trans. Veh. Technol.*, vol. VT-35, pp. 77–92, Aug. 1986.

[8] F. P. Kelly, "Stochastic models of computer communication systems," *J. R. Statist. Soc.*, vol. 47, no. 3, pp. 379–395, 1985.

[9] W. Feller, *An Introduction to Probability Theory and Its Applications*, 3rd ed. New York: Wiley, 1968, vol. 2.

Xusheng Tian received the B.S. degree (with honors) from Southeast University, Nanjing, China, in 1991 and the M.S. degree from Tsinghua University, Beijing, China, in 1994, both in electrical engineering. He is currently working toward the Ph.D. degree at Rensselaer Polytechnic Institute, Troy NY.

His research interests are in the areas of wireless communications and computer communication networks. Specifically, he has been involved in resource management of wireless networks, measurement-based network management, and network traffic modeling using wavelet.

Chuanyi Ji received the B.S. degree (with honors) from Tsinghua University, Beijing, China, in 1983, the M.S. degree from the University of Pennsylvania, Philadelphia, and the Ph.D. degree from the California Institute of Technology, Pasadena, both in electrical engineering, in 1986 and 1992, respectively.

In 1991, she joined the Rensselaer Polytechnic Institute, Troy, NY, where she is currently an Associate Professor of Electrical, Computer, and Systems Engineering. She has held visiting positions with Mathematics of Networks and Systems Bell Laboratories, Murray Hill, NJ, and the Laboratory of Information Decision Systems, Massachusetts Institute of Technology, Cambridge. Her research interests include computer communication networks with a focus on network management and adaptive systems. She has worked in developing a framework for measurement-based network monitoring and QoS inference, the wavelet models for heterogeneous traffic, intelligent agents for proactive network anomaly detection, resource management of wireless networks, performance, and efficiency of adaptive learning systems and applications of adaptive algorithms and information theory in networking and communication.

Dr. Ji received the National Science Foundation's CAREER Award in 1995 and an Early Career Award from Rensselaer Polytechnic Institute in 2000.