Region-based Call Admission Algorithms For Wireless Cellular Networks

K.K. Kasera & R.K. Sitaraman

CMPSCI TR 96-72

July 1997

Region-based Call Admission Algorithms For Wireless Cellular Networks*

Kamal K. Kasera   Ramesh K. Sitaraman

Department of Computer Science

University of Massachusetts at Amherst

Amherst, MA 01003, USA

E-mail: {kkasera,ramesh}@cs.umass.edu

Corresponding Author: Kamal K. Kasera

July 16, 1997

Abstract

Current advances in the area of wireless cellular networks, and the advent of real time services have mandated the need for an efficient network call admission controller. In this paper, we develop region-based call admission algorithms that can deal with multiple classes of prioritized real time traffic. We consider QoS metrics which arise due to mobility considerations, such as handoff and preemption call dropping probabilities. We show, using both analysis and simulations, that our call admission algorithm maintains all the QoS guarantees of the mobile users, while ensuring a high level of network utilization. Further, our call admission algorithm is computationally simple and can be efficiently implemented in modern wireless cellular architectures. The conceptual framework of characterizing call admission algorithms using admissibility sets is likely to be useful in a number of other contexts.

Keywords: wireless networks, personal communication services, admission control.

*The research was supported in part by NSF Grant CCR-94-10077 and in part by an NSF CAREER award CCR-9703017.

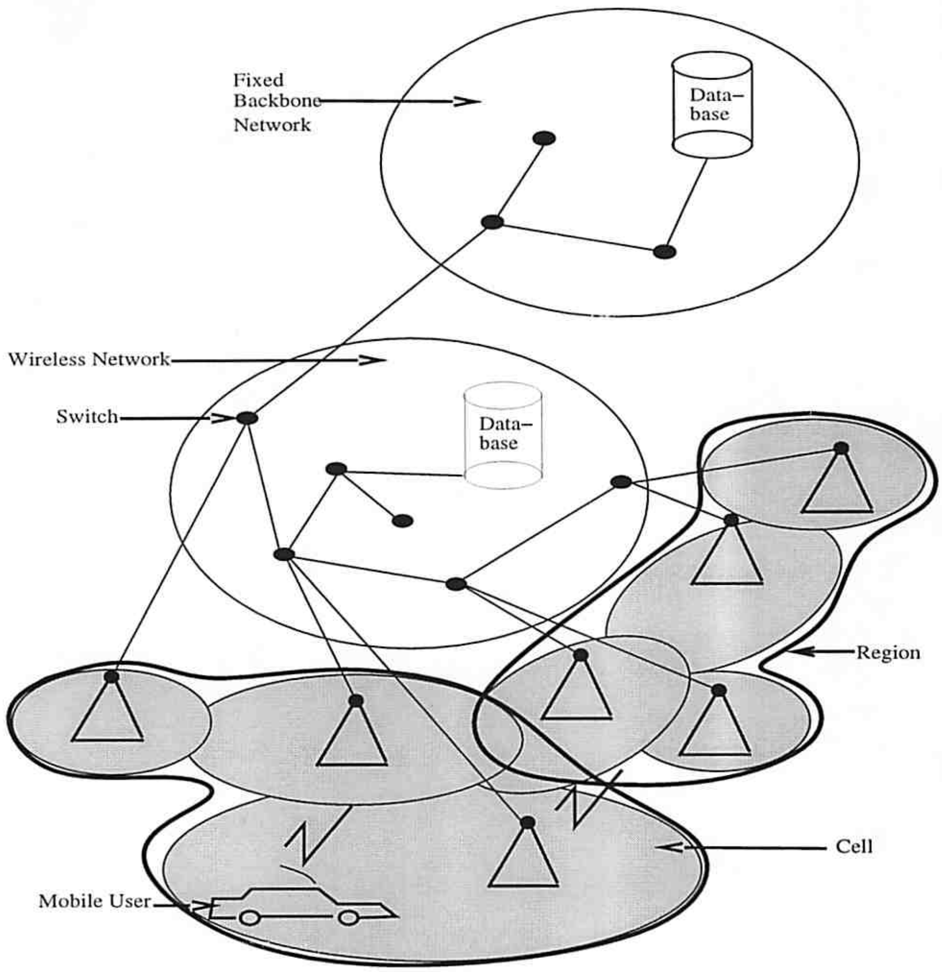

Figure 1: A cellular network architecture.

1 Introduction

The goal of the future wireless cellular network is to provide ubiquitous communication services to a large number of mobile users with varying Quality of Service (QoS) requirements [1] - [11]. This requires a wireless cellular network of high capacity to accommodate a large number of users and an efficient call admission controller to provide a guaranteed QoS.

The wireless cellular network architecture consists of a stationary and a mobile level (see Figure 1). The stationary level comprises of base stations or access points that are linked together via a fixed *backbone* network. The mobile level consists of radio links that connect base stations to the mobile users. Traditionally, the wireless cellular network service area is partitioned into *regions*, which are further hierarchically decomposed into *cells*. Each cell is serviced by a base station which acts as an interface between the mobile units and the fixed backbone network. Neighboring cells overlap with each other to provide communication continuity.

Since the users are mobile, the quality of the radio link may deteriorate as the user moves away from the base station. In such a situation, a new base station with an acceptable radio link quality is found and the call is handed over to it. This event is known as *handoff* [11].

Each base station can handle only a certain number of users. Thus, if the new base station has no spare capacity, the handoff call might be forced to terminate. The probability that a call is terminated during a handoff is known as the *handoff call dropping probability*. The current trend in wireless architectures is to have a large number of small capacity cells to provide maximum usage of the limited radio spectrum through frequency reuse [12]-[14]. Due to the small coverage area of such cells, a large number of handoffs may occur in the duration of a call. Hence, the handoff call dropping probability is an important QoS measure which arises due to mobility of users and must be kept below a desired level.

The limited call handling capacity of the base stations make it necessary to restrict the entry of new calls into the wireless system. The wireless network will be congested if the communication demands of the mobile users in an area exceeds the total call handling capacity of all the reachable base stations. The role of a call admission controller is to decide whether or not a new call can be admitted into the system, without violating the QoS guarantees of all the existing calls in the network. The new paradigm of high-speed wireless networks requires that the call admission controller be *simple* and *fast*. Further, the controller should be capable of dealing with *multiple classes* of traffic having different priority levels and varying QoS requirements.

1.1 QoS Metrics

In this paper, we focus on real time traffic. We consider only important QoS metrics arising specifically due to mobility considerations. We do not consider other important QoS metrics such as bandwidth availability, jitter, and delay. Real time connections have a low tolerance to abrupt terminations. Thus, the handoff call dropping probability is an important QoS measure for real time traffic. The real time traffic is often prioritized into different classes, with higher priority traffic requiring a smaller call dropping probability.

There are other situations besides handoff that lead to the termination of an active call. In the presence of multiple priority classes, a lower priority call can be dropped if its channel is preempted by a higher priority handoff call. This situation occurs only when no spare channel is available for the higher priority handoff call. We refer to the probability of dropping a call at a time step given that a higher priority call handoffs to its cell as the *preemption call dropping probability*. It is desirable to keep this probability below a required level. Clearly, preemption occurs only when multiple classes of traffic are present.

In this paper, we are concerned with the two QoS metrics mentioned above - the handoff call dropping probability of the k^{th} traffic class, denoted by h_k, and the preemption call dropping probability of the k^{th} class, denoted by p_k. Each traffic class k has a pre-negotiated bound on the handoff call dropping probability, denoted by H_k, and a pre-negotiated bound on the preemption call dropping probability, denoted by P_k. The call admission algorithm must ensure that for every class k call admitted into the system, the QoS requirements are met, i.e., $h_k \leq H_k$, and $p_k \leq P_k$.

1.2 Previous Work

Call admission in wireless networks has received much attention in the recent past. Most of the known algorithms, like the guard-channel algorithm, are *cell-based*[17, 21]. These *algorithms* make admission decisions for a new call based only the local state of the cell to which the new call seeks entry. While cell-based algorithms have the advantage of extreme simplicity, it could lead to low network utilization. The reason is that calls may be rejected based on the fact that a certain cell has above-average load, although the system as a whole may be under-utilized. Recently, Acampora and Naghshineh [15, 20] have suggested the use of call admission algorithms that take into account the state of a cluster of contiguous cells in the same region[1]. In this paper, we develop a framework for the study of ***region-based*** call-admission algorithms, and show how to design optimal region-based call admission algorithms that provide high network utilization. Much of the previous work [15, 16, 17, 18, 19, 20] on call admission control pertains to a single traffic class. However, our framework and results apply to multiple traffic classes.

1.3 Our Results

In Section 2, we provide a framework for describing region-based call admission algorithms capable of dealing with multiple classes of traffic. A good call admission algorithm must satisfy two opposing requirements. First, the algorithm must ensure that the QoS requirements of all the admitted calls are met. Second, the utilization of network resources must be kept as high as possible, i.e., the algorithm should not reject more calls than necessary. Our framework provides formal means for measuring the "goodness" of a region-based call admission algorithm.

In Section 3, we devise an optimal call admission algorithm that is capable of dealing with a single class of real time traffic. The single-class call admission algorithm ensures that the handoff call dropping probability for each user is at most the pre-negotiated upper limit. Note that preemption call dropping probability is not a relevant QoS measure, since only one traffic class is present. We show via analyses and simulations that our call admission algorithm maintains a high level of network utilization.

In Section 4, we present an optimal call admission algorithm that is capable of dealing with multiple classes of real time traffic. This algorithm is a generalization of the single-class algorithm presented in Section 3. The multiple-class call admission algorithm ensures that both the handoff and preemption call dropping probabilities for each user are at most the respective pre-negotiated upper limits. As above, we show, both analytically and empirically, that our algorithm ensures a high level of network utilization. Our call admission algorithms are computationally simple and are easy to implement in modern wireless cellular networks.

Finally, in Section 5, we present our conclusions.

1.4 Network and Traffic Model

In this section, we describe the network and traffic model used in our analyses and simulations.

The *network model* assumes that the entire mobile service area is divided into regions, which are further subdivided into cells. We assume that a region consists of B contiguous

[1] Presently, the division of the cellular network service area into regions is done mostly according to commercial considerations.

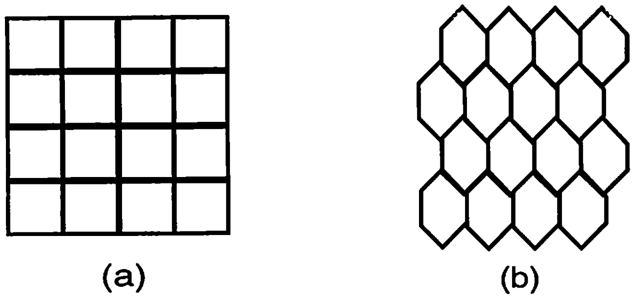

Figure 2: Symmetric layouts (a) square (b) hexagonal.

cells that are *identical* in every respect. The cells can have a number of different shapes, i.e., square, hexagonal etc. But the layout of the cells should be *symmetric*, i.e., each cell should have the same number of neighbors (except cells in the boundaries). Most standard cell layouts obey these properties (see Figure 2). Each cell is serviced by a base station. Channel assignment is assumed to be fixed [17], which implies that each base station can handle only a fixed[2] number of calls. We assume that each base station can handle C calls simultaneously. Consequently, each region can service a maximum of $B \cdot C$ calls simultaneously, denoted by N. A call can use any of the available call slots in a cell.

Our *traffic model* assumes that each cell has *identical statistical traffic characteristics.* More precisely, we assume the following.

- A large population of mobile users are *uniformly and randomly distributed* over the B cells of the service region.
- A new call is generated by a mobile user with *equal probability* in any of the B cells of the region.
- A mobile user in a cell is *equally likely* to go to any of its adjacent cells, in the event of a handoff.

Although our uniformity assumptions do not exactly represent a number of practical networks, it can be used to approximate many real systems [19]. Wireless networks with homogeneous distribution of users are more closely modeled by these assumptions, than networks with large disparities between cells within a region.

Handoff calls are assumed to have priority over all new calls, i.e., if handoff and new calls are queued for entry to the same cell, the handoff calls are considered first. In case of multiple classes, higher priority handoff calls can preempt lower priority active calls. That is, if a high priority call handoffs to a cell with no spare channel, it preempts one of the randomly chosen lowest priority active calls in the cell. This ensures that high priority calls are given preference over low priority ones.

[2]The number of calls that can be handled by a base station is strictly dependent on the radio link access protocol employed. Details of the radio link access protocols are beyond the scope of this paper.

Algorithm GENERIC
1. Input the current state-vector $\sigma = \langle \mu_1, \mu_2, \cdots, \mu_k, \cdots, \mu_m \rangle$ and the class k of the new call.
2. Compute $\sigma' = \langle \mu_1, \mu_2, \cdots, \mu_k + \frac{1}{N}, \cdots, \mu_m \rangle$, which is the new state-vector if the new call is admitted.
3. If $\sigma' \in S$ then **ADMIT** new call, else **REJECT** new call, where $S \subseteq [0,1]^m$ is the *admissibility set* of the algorithm.

Figure 3: A generic region-based call admission algorithm.

2 A Framework for Call Admission Algorithms

In this section, we develop a framework for describing region-based algorithms that use only the *state of a region* to make call-admission decisions. Recollect that a region consists of a cluster of contiguous cells. We assume that the state of the region is specified by the number of active calls in the region from each traffic class. For the sake of simplicity and easy implementability, we do not consider more complex algorithms that rely on more detailed state information about a region.

Formally, the state of a region is represented by an m-dimensional vector, $\sigma = \langle \mu_1, \mu_2, \cdots, \mu_m \rangle \in [0,1]^m$, where m is the total number of traffic classes and μ_i is the fractional utilization of the total network resources by calls of class i. That is, $\mu_i = N_i/N$, where N_i is the number of active calls of class i in the region, and N is the total number of possible calls that can be serviced simultaneously in the region. Since there are B cells in the region with each cell having a capacity of C, N equals $B \cdot C$.

A generic region-based call-admission algorithm is shown in Figure 3. A region-based call-admission can be specified completely by its *admissibility set* $S \subseteq [0,1]^m$. The admissibility set S is simply the set of the states that a region is allowed to get into, under the call admission algorithm. A new call is admitted by the algorithm if and only if the new state vector of the region including the new call, is an element of S.

2.1 A Measure of Goodness for Call Admission Algorithms

We now evolve a formal measure of how well a region-based call admission algorithm performs. A "good" call admission algorithm must obey the following opposing requirements.

- The *correctness requirement* requires that the QoS guarantees of all admitted calls are met. In our situation, this means that the handoff call dropping probability, h_k and the preemption call dropping probability, p_k, are at most the respective pre-negotiated thresholds. This requirement necessitates the rejection of some new calls to ensure that the region does not get into an "overloaded" state.

- The *utilization requirement* requires that as many calls as possible be admitted, so as to keep the network utilization high.

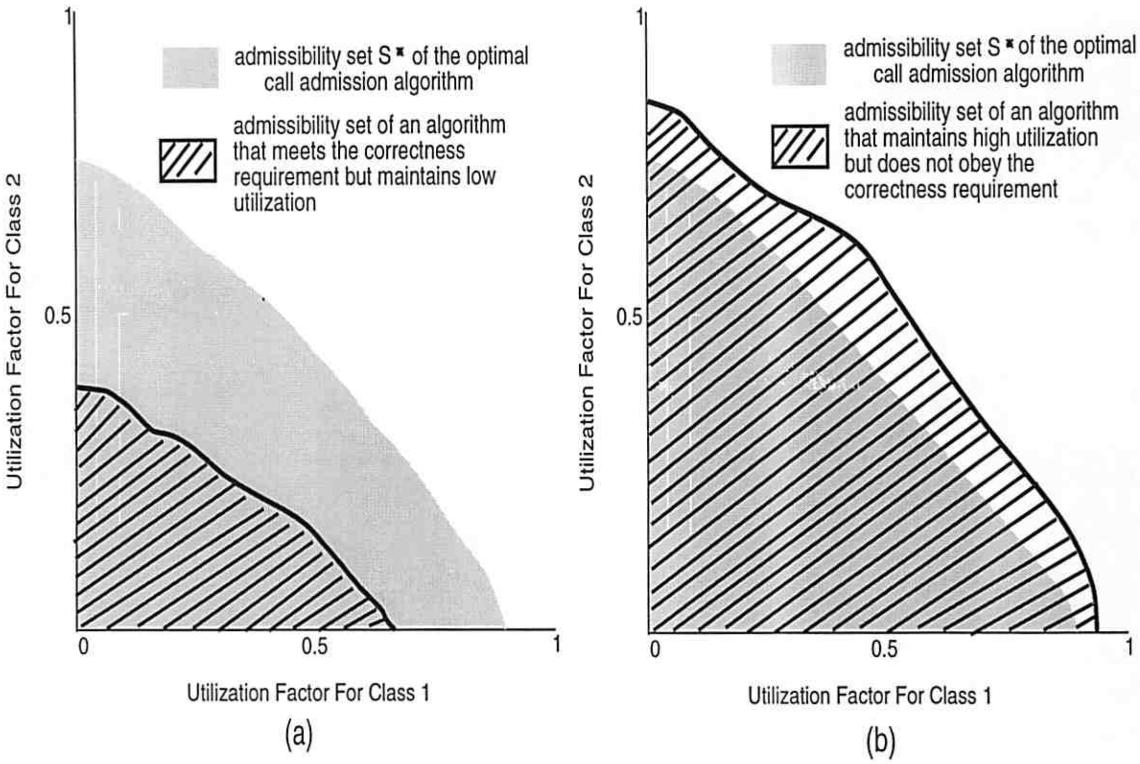

Figure 4: Comparison of the optimal call admission algorithm with (a) an algorithm that meets the correctness requirement, but maintains a low network utilization, and (b) an algorithm that does not meet the correctness requirement, but maintains a high network utilization.

Intuitively, an *optimal call admission algorithm* admits as many calls as possible, without jeopardizing the QoS guarantees of any of the admitted calls. Let the optimal admissibility set, $S^* \subseteq [0,1]^m$, be the set of *all* possible states σ such that the QoS requirements of each call in each traffic class is met, under our traffic and network model assumptions. The optimal call admission algorithm is one that has S^* as its admissibility set. Note that the optimal algorithm has the largest set of states of any algorithm that obeys the correctness requirement. It avoids exactly those states where the QoS will not be met, and includes exactly those states where the QoS will be met. The "goodness" of any call admission algorithm can be measured by the closeness of its admissibility set S to the optimal admissibility set S^*.

Note that there are call admission algorithms that violate one requirement but not the other. For instance, the call admission algorithm that accepts too few calls, obeys the correctness requirement, but fails to maintain a high network utilization. A call admission algorithm that accepts too many calls, maintains a high utilization of the network, but fails to meet the correctness requirement, since the QoS guarantees for admitted calls will not be met (see Figure 4). Neither of these algorithms are to be considered "good".

Our objective in the rest of the paper is to do the following.

- Given a set of QoS requirements for the traffic classes, we show how to analytically estimate the optimal admissibility set S^*. We use this analytic estimate of S^* to obtain an optimal call admission algorithm.

- Next, we show that our analytic estimate of the optimal admissibility set S^* agrees closely with our empirical estimate of S^*, using extensive simulations.

We first deal with a single traffic class in Section 3, and then extend our results to multiple traffic classes in Section 4.

3 Optimal Call Admission for a Single Traffic Class

In this section, we assume that all the calls belong to a single traffic class. The traffic class has a QoS requirement that the handoff call dropping probability h is at most some pre-negotiated value H. Note that we drop the subscripts, since there is only one traffic class. Further, note that for the same reason, the preemption call dropping probability is not a relevant metric.

Since there is only one traffic class, the optimal call admission algorithm has an admissibility set $S^* \subseteq [0,1]$. Clearly, S^* is an interval $[0,\mu^*]$, where μ^* is the maximum possible utilization such that the handoff call dropping probability h for each admitted call is at most the pre-negotiated upper bound H. The optimal call admission algorithm admits at most $\mu^* N$ calls into a region simultaneously (see Figure 5).

In Section 3.1, we show how to analytically estimate μ^*, as a function of the pre-negotiated upper bound H on the handoff call dropping probability. The optimal call admission algorithm uses this analysis to compute μ^*, given the QoS requirement of the traffic. In Section 3.2, we verify our analysis using simulations.

Algorithm OPTIMAL-SINGLE-CLASS

1. Input the current state-vector μ, which is the fractional utilization of the network, i.e., μ equals the number of active calls in the region divided by N.

2. Compute $\mu' = \mu + \frac{1}{N}$, which is the new state-vector, if the new call is admitted.

3. If $\mu' \leq \mu^*$ then **ADMIT** new call, else **REJECT** new call.

Figure 5: The optimal single class call admission algorithm.

3.1 Estimating μ^* Analytically

First, we derive a relationship between the utilization μ and the handoff call dropping probability h. That is, if there are μN active users in the region, we would like to estimate the probability h that an active user is dropped during a handoff, under the assumptions of our network and traffic model.

A cell is said to be *congested* if it already has C active calls, and therefore has no capacity for any other calls. An active user attempting to handoff to a congested cell is dropped. Since the users are distributed uniformly and randomly to each of the cells, the probability that a given user finds a cell congested during a handoff is evaluated via the following balls-and-bins problem.

Assume that there are B identical bins, each of capacity C. Consider the process of randomly tossing N' balls into the bins, where $N' \leq N = BC$. The balls are tossed independent of each other and each ball is equally likely to fall into any of the B bins, i.e., the probability that a ball falls in a particular bin is $\frac{1}{B}$. The probability that a given ball lands in a bin with C or more other balls is estimated by the following expression.

$$EXPR(N') \triangleq \sum_{j=C}^{N'-1} \binom{N'-1}{j} \frac{1}{B^j}\left(1-\frac{1}{B}\right)^{N'-1-j} \tag{1}$$

The process of randomly distributing μN identical active users into B cells in a region is equivalent to randomly tossing μN identical balls into B bins. Thus, by analogy, the probability h that a call is dropped because it handoffs into a congested cell can be estimated by $EXPR(\mu N)$. We can now use our analysis to estimate μ^*, given the pre-negotiated upper bound H. μ^* is simply the largest value for utilization μ such that $h \leq H$, i.e., μ^* is the largest value for μ such that the following inequality is satisfied.

$$EXPR(\mu N) \leq H \tag{2}$$

The *theoretical* curve in Figure 7 plots our theoretical estimate for h, $EXPR(\mu N)$, as a function of the utilization μ, when $B = 49$ and $C = 25$. The *theoretical* curve can be used to devise an optimal single class call admission algorithm. For instance, if $H = 0.0767$, the call admission algorithm uses the *theoretical* curve to find the corresponding value of μ^* to be 0.70 (see Figure 7).

No. of cells, B	Cell size, C	Figure no.
4	25	Figure 14
49	25	Figure 7
49	16	Figure 15
121	8	Figure 16
196	4	Figure 17

Figure 6: Types of networks simulated for single traffic class call admission algorithm

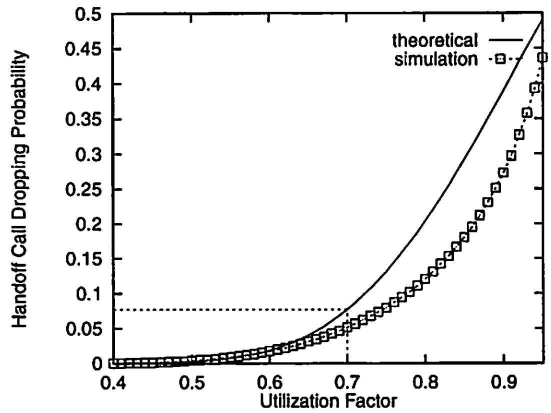

Figure 7: Handoff call dropping probability h vs utilization factor μ ($B = 49$, $C = 25$).

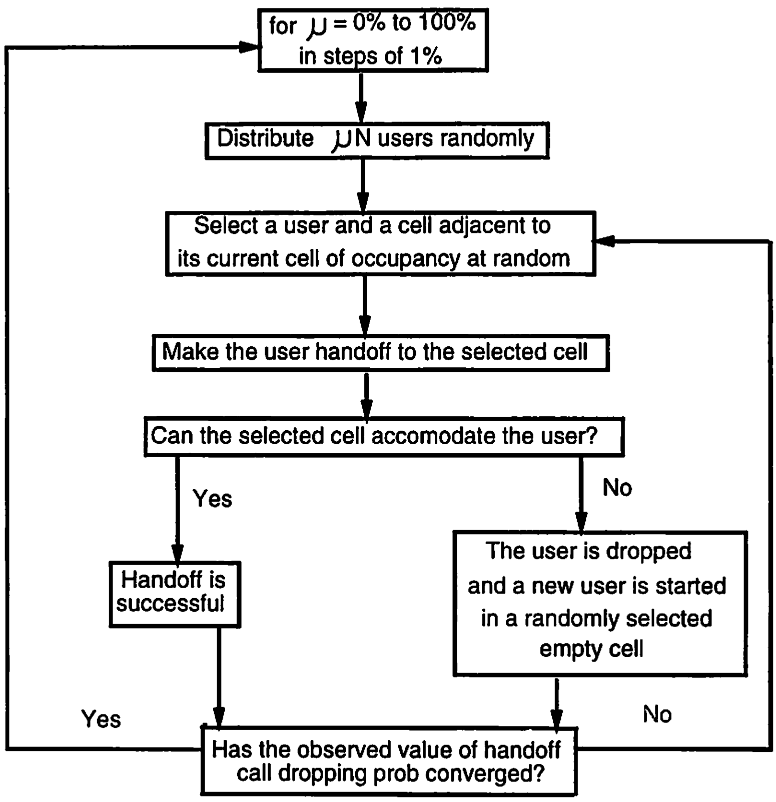

Figure 8: The simulation flowchart for single traffic class.

3.2 Simulation Results

We simulated our call admission algorithm on a variety of networks. Each network comprised of one region which was divided into a set of hexagonal-shaped cells. The regions ranged from those that contained a small number of large capacity cells to those that contained a large number of small capacity cells (see Figure 6). Mobile users wrapped around on hitting the region boundary. A region that contained B cells was laid out in a $\sqrt{B} \times \sqrt{B}$ matrix. Figure 2(b) shows the layout of a region containing 16 cells. The simulation results consisted of the observed values of the handoff call dropping probability h. We estimated h by the observed *ratio* of the number of active calls dropped during handoffs to the total number of handoffs in the region, for different values of the network utilization μ. A detailed description of the simulation process is given in Figure 8.

The *simulation* curves in Figures 7, 14, 15, 16 and 17 plot the observed values of handoff call dropping probability h against network utilization μ. The *theoretical* curves in these figures plot our theoretical estimate of h. The simulation results are in close agreement with the analytical estimates. Further, the *theoretical* curves form an upper bound of the *simulation* curves; hence, our call admission algorithm satisfies the QoS requirements of all the admitted calls while maintaining a high network utilization. Finally, note that the agreement between the analytical and empirical estimates of μ^* becomes better when the region contains a larger number of smaller capacity cells.

4 Optimal Call Admission for Multiple Traffic Classes

In this section, we assume that the calls belong to m different traffic classes. The traffic classes are prioritized with the k^{th} traffic class having higher priority than the $(k+1)^{th}$ traffic class. The k^{th} traffic class requires that the handoff call dropping probability, h_k, and preemption call dropping probability, p_k, for each admitted class k call, be at most the pre-negotiated values H_k and P_k respectively. Typically, higher priority classes have more stringent QoS requirements than the lower priority classes. Thus, $H_1 \leq H_2 \ldots \leq H_m$ and $P_1 \leq P_2 \ldots \leq P_m$.

Since there are m traffic classes, the optimal multiple-classes call admission algorithm has an m-dimensional admissibility set $S^* \subseteq [0,1]^m$. The optimal multiple-classes call admission algorithm admits a new call only if the new state vector σ' after admitting the new call is in S^* (see Figure 9).

In Section 4.1, we show how to analytically estimate S^*. The optimal call admission algorithm uses this analysis to compute S^*, given the QoS requirements of the m traffic classes. In Section 4.2, we verify our analysis using simulations.

4.1 Estimating S^* Analytically

First, we estimate the k^{th} class handoff call dropping probability, h_k, and the k^{th} class preemption call dropping probability, p_k, given that the state of the region is represented by the state vector $\sigma = \langle \mu_1, \mu_2, \cdots, \mu_m \rangle$. Then, we present the conditions under which the state vector σ belongs to the optimal admissibility set S^*. Finally, we show how to approximate S^* in a piece-wise linear fashion.

Algorithm OPTIMAL-MULTIPLE-CLASSES

1. Input the current state-vector $\sigma = \langle \mu_1, \mu_2, \cdots, \mu_k, \cdots, \mu_m \rangle$ and the class k of the new call.
2. Compute $\sigma' = \langle \mu_1, \mu_2, \cdots, \mu_k + \frac{1}{N}, \cdots, \mu_m \rangle$, which is the new state-vector if the new call is admitted.
3. If $\sigma' \in S^*$ then **ADMIT** new call, else **REJECT** new call, where $S^* \subseteq [0,1]^m$ is the *admissibility set* of the algorithm.

Figure 9: The optimal multiple classes call admission algorithm.

Estimating h_k: Let C be an active class k call. Call C will be dropped during a handoff if it handoffs to a congested cell containing only active calls from classes 1 to k. Note that if active calls belonging to classes more than k are present in the congested cell, they are of lower priority than call C and thus, can be preempted by call C. Proceeding as in Section 3.1, the probability that call C is dropped during a handoff can be estimated by $EXPR(\sum_{i=1}^{k} \mu_i \cdot N)$. Thus, the k^{th} class handoff call dropping probability h_k, which is the probability of dropping call C during a handoff, can be estimated by $EXPR(\sum_{i=1}^{k} \mu_i \cdot N)$.

Estimating p_k: Since class 1 is the highest priority class, class 1 calls are never preempted and therefore $p_1 = 0$. Let C be an active class k call such that k is greater than 1. Call C can be preempted only if one of the following two *disjoint* events occur.

- *Event 1:* An active call belonging to class i, $1 \leq i \leq k-1$, handoffs to a congested cell that contains call C and the other calls in this cell belong to classes 1 to $k-1$.
- *Event 2:* An active call belonging to class i, $1 \leq i \leq k-1$, handoffs to a congested cell that contains call C and at least one more active class k call, and the other calls in this cell belong to classes 1 to $k-1$.

Proceeding as in Section 3.1, the probability that *Event 1* occurs can be estimated by $EXPR((\sum_{i=1}^{k-1} \mu_i \cdot N) + 1)$. Given that *Event 1* occurs, the probability that call C is preempted is 1, since it is the only class k call in the congested cell. Proceeding as in Section 3.1, the probability that either *Event 1* or *Event 2* occurs can be estimated by $EXPR(\sum_{i=1}^{k} \mu_i \cdot N)$. Thus, the probability that *Event 2* occurs can be estimated by $EXPR(\sum_{i=1}^{k} \mu_i \cdot N) - EXPR((\sum_{i=1}^{k-1} \mu_i \cdot N) + 1)$. Given that *Event 2* occurs, the probability that call C is preempted is at most $\frac{1}{2}$, since the congested cell contains at least one more class k call and the preempted class k call is chosen randomly among all the class k calls in the congested cell. Thus, the k^{th} traffic class preemption call dropping probability p_k, which is the probability of preempting call C, can be estimated by

$$1 \cdot EXPR((\sum_{i=1}^{k-1} \mu_i \cdot N) + 1) + \frac{1}{2} \cdot (EXPR(\sum_{i=1}^{k} \mu_i \cdot N) - EXPR((\sum_{i=1}^{k-1} \mu_i \cdot N) + 1))$$
$$= \frac{EXPR((\sum_{i=1}^{k-1} \mu_i \cdot N) + 1) + EXPR(\sum_{i=1}^{k} \mu_i \cdot N)}{2}$$

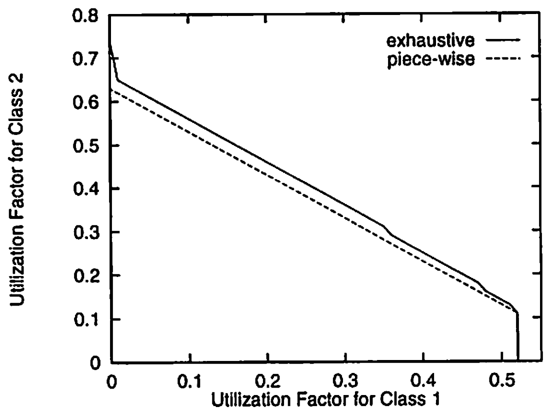

Figure 10: Comparison of the optimal admissibility set S^* obtained using the PIECE-WISE method (area below the *piece-wise* curve) and by exhaustive sampling (area below the *exhaustive* curve).

Conditions for σ to be in S^*: Recollect that a state vector σ belongs to the optimal admissibility set S^* if the handoff and preemption call dropping probabilities for each admitted user is at most the respective pre-negotiated values, i.e., $h_k \leq H_k$ and $p_k \leq P_k$. Thus, $\sigma = \langle \mu_1, \mu_2, \cdots, \mu_m \rangle$ belongs to S^* if the following inequalities are satisfied.

$$EXPR(\sum_{i=1}^{k} \mu_i \cdot N) \leq H_k, \quad for\ k = 1, 2, \ldots, m \tag{3}$$

$$\frac{EXPR((\sum_{i=1}^{k-1} \mu_i \cdot N) + 1) + EXPR(\sum_{i=1}^{k} \mu_i \cdot N)}{2} \leq P_k, \quad for\ k = 2, 3, \ldots, m \tag{4}$$

Piece-wise linear approximation of S^*: The optimal admissibility set S^* consists of all the state vectors σ that satisfy the inequalities given by Equations 3 and 4. Given the pre-negotiated upper bounds H_k and P_k for each of the m classes, our multiple classes call admission algorithm (see Figure 9) can use Equations 3 and 4 to test if the new state vector after admitting the new call is in S^*. Since evaluating Equations 3 and 4 is complex, we approximate S^* in a piece-wise linear fashion using the PIECE-WISE method detailed in Figure 11. The PIECE-WISE method evaluates the constants $\mu_1^*, \mu_2^*, \ldots, \mu_m^*$, using the pre-negotiated upper bounds H_k and P_k, $1 \leq k \leq m$. The values $\mu_1^*, \mu_2^*, \ldots, \mu_m^*$, need to be computed and stored only once. Our call admission algorithm accepts a new call only if, after admitting the new call, the total network utilization due to active calls from classes 1 to k is at most μ_k^*, for $k = 1, 2, \ldots, m$. In other words, our call admission algorithm always ensures that the state vector $\sigma = \langle \mu_1, \mu_2, \cdots, \mu_m \rangle$ satisfies the following set of linear inequalities.

$$\sum_{i=1}^{k} \mu_i \leq \sum_{i=1}^{k} \mu_i^*, \quad for\ k = 1, 2, \ldots, m$$

Approximation PIECE-WISE

1. Choose μ_1^* to be the largest value of μ_1 that satisfies $EXPR(\mu_1 \cdot N) \leq H_1$.
2. For $k = 2, 3, \ldots, m$ choose μ_k^* to be the largest value of μ_k that satisfies $EXPR(((\sum_{i=1}^{k-1} \mu_i^*) + \mu_k) \cdot N) \leq H_k$ and $\frac{EXPR((\sum_{i=1}^{k-1} \mu_i^* \cdot N)+1)+EXPR(((\sum_{i=1}^{k-1} \mu_i^*)+\mu_k)\cdot N)}{2} \leq P_k$.
3. S^* is the set of all state vectors $\sigma = \langle \mu_1, \mu_2, \cdots, \mu_m \rangle \subseteq [0,1]^m$ that satisfies the set of m linear inequalities $\sum_{i=1}^{k} \mu_i \leq \sum_{i=1}^{k} \mu_i^*$, $for\ k = 1, 2, \ldots, m$.

Figure 11: Piece-wise approximation method for S^*.

No. of cells, B	Cell size, C	H_1	H_2	P_2	Figure no.
49	16	0.015	0.200	0.050	Figure 18
49	16	0.015	0.050	0.050	Figure 19
49	16	0.005	0.150	0.050	Figure 20
49	16	0.000001	0.015	0.050	Figure 21
121	8	0.060	0.250	0.180	Figure 22
121	8	0.080	0.150	0.100	Figure 23
196	4	0.021	0.200	0.150	Figure 24
196	4	0.050	0.200	0.150	Figure 25

Figure 12: Types of networks simulated for multiple traffic classes call admission algorithm.

We compared our estimate of S^* obtained using the PIECE-WISE method with the estimate of S^* obtained by exhaustively finding values of σ that satisfy the Equations 3 and 4. Figure 10 shows one such comparison when $B = 49$, $C = 16$, $H_1 = 0.02$, $H_2 = 0.20$ and $P_2 = 0.05$. Since the estimate of S^* obtained using the PIECE-WISE method is a good approximation of the estimate of S^* obtained by exhaustively finding values of σ that satisfy the Equations 3 and 4, we will use the PIECE-WISE method throughout the remainder of the paper.

4.2 Simulation Results

We simulated our multiple classes call admission algorithm using two traffic classes and using different QoS requirements for each class (see Figure 12). The network model used for the simulations was the same as that used for simulating the single traffic class call admission algorithm, described in Section 3.2. Simulation results for each network was obtained by varying the state vector σ. The simulation results consisted of the observed values of h_1, h_2 and p_2. Recall that $p_1 = 0$. We estimated the class 1 handoff call dropping probability h_1 by the observed *ratio* of the number of active class 1 calls dropped during handoffs to the total number of class 1 handoffs in the network. Similarly, we estimated the class 2 handoff call dropping probability h_2 by the observed *ratio* of the number of active class 2 calls dropped during handoffs to the total number of class 2 handoffs in the network. The

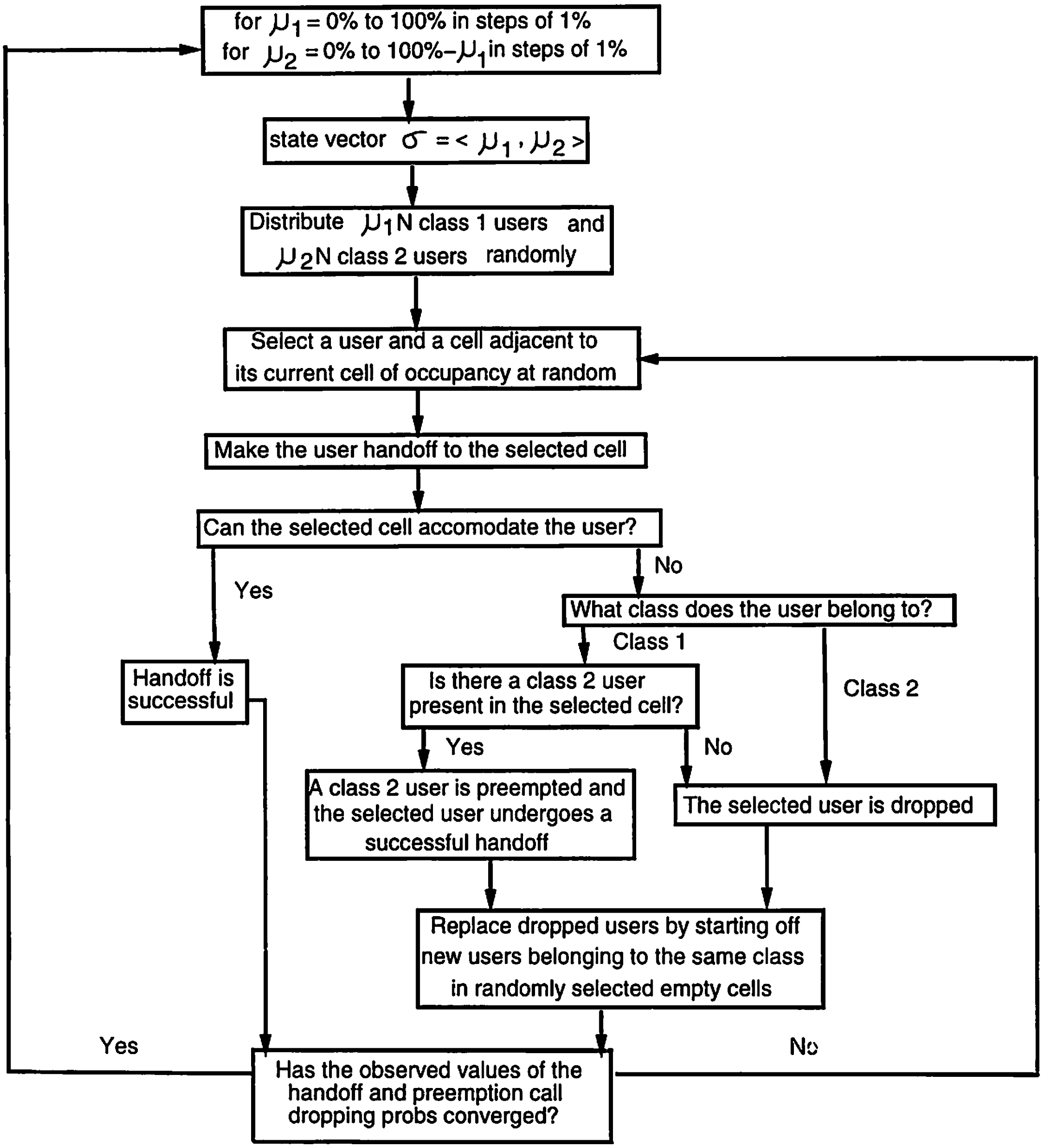

Figure 13: The simulation flowchart for two traffic classes.

class 2 preemption call dropping probability p_2 was estimated by the observed *ratio* of the number of active class 2 calls preempted to the total number of successful class 1 handoffs. A detailed description of the simulation process is given in Figure 13.

The region labelled *simulation* in Figures 18, 19, 20, 21, 22, 23, 24 and 25 represents our empirical estimate of the optimal admissibility set S^*. The empirical estimate of S^* consists of all the state vectors σ such that the observed values of h_1, h_2 and p_2 for all the calls were found to be at most the pre-negotiated values H_1, H_2 and P_2 respectively. The area below the *theoretical* curves in these figures represents our theoretical estimate of the optimal admissiblity set S^*, found using the PIECE-WISE approximation method.

The analytical estimate of the optimal admissibility set S^* used by our algorithm agrees closely with the empirical estimate of S^*, ensuring that our algorithm maintains a high network utilization while meeting the QoS requirements for each admitted call.

5 Conclusion

We have presented call admission algorithms that can deal with multiple classes of prioritized real time traffic. Our call admission algorithm makes admission decisions for a new call based on the state of a region, instead of just the cell, to which the call seeks entry. The conceptual framework of characterizing call admission algorithms using admissibility sets is likely to be useful in a number of other contexts. The QoS metrics guaranteed for each admitted user includes handoff and preemption call dropping probabilities. Simulations performed using a wide variety of networks and using different QoS requirements show that our call admission algorithm meets the QoS requirements for all the calls, while ensuring a high level of network utilization.

References

[1] D. Goodman, "Cellular Packet Communications", *IEEE Trans. on Commun.*, vol. 28, no. 8, Aug. 1990.

[2] D. Goodman, "Trends in Cellular and Cordless Communications", *IEEE Commun. Mag.*, Feb. 1991.

[3] D. Goodman, G. Pollini and K. -Hellstern, "Network Control for Wireless Communications", *IEEE Commun. Mag.*, Dec. 1992.

[4] D. Cox, "Wireless Network Access for Personal Communications", *IEEE Commun. Mag.*, Dec. 1992.

[5] R. J. Lynch, "PCN: Son of Cellular? The Challenges of Providing PCN Service", *IEEE Commun. Mag.*, Feb. 1991.

[6] I. M. Ross, "Wireless Network Directions", *IEEE Commun. Mag.*, Feb. 1991.

[7] R. Steele, "Deploying Personal Communication Network", *IEEE Commun. Mag.*, Sep. 1990.

[8] A. D. Kucar, "Mobile Radio: An Overview", *IEEE Commun. Mag.*, Nov. 1991.

[9] R. Steele, "The Cellular Environment of Lightweight Handheld Portables", *IEEE Commun. Mag.*, July 1989.

[10] D. Cox, "Portable Digital Radio Communications - An Approach to Tetherless Access", *IEEE Commun. Mag.*, July 1989.

[11] A. Bar-Noy and I. Kessler, "Tracking Mobile Users in Wireless Communications Network", *IEEE Trans. on Information Theory*, pp. 1877-1886, Nov. 1993.

[12] J. Sarnecki *et. al.*, "Microcell Design Principles", *IEEE Commun. Mag.*, April 1993.

[13] W. C. Y. Lee, "Smaller Cells for Greater Performance", *IEEE Commun. Mag.*, Nov. 1991.

[14] L. I. Greenstein *et. al.*, "Microcells in PCS", *IEEE Commun. Mag.*, Dec. 1992.

[15] A. S. Acampora and M. Naghshineh, "Control and Quality-of-Service Provisioning in High Speed Networks", *IEEE Personal Commun.*, Second Quarter 1994.

[16] B. Eklundh, "Channel Utilization and Blocking Probability in a Cellular Mobile Telephone System with Directed Entry", *IEEE Trans. on Vehicular Tech.*, vol. 35, no. 3, Aug. 1987.

[17] S. Tekinay and B. Jabbari, "Handover and Channel Assignment in Mobile Cellular Networks", *IEEE Commun. Mag.*, Nov. 1991.

[18] 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. on Vehicular Tech.*, vol. 35, no. 3, Aug. 1986.

[19] J. Zander, "Performance of Optimum Transmitter Power Control in Cellular Radio Systems", *IEEE Trans. on Vehicular Tech.*, vol. 41, no. 1, Aug. 1992.

[20] A. S. Acampora and M. Naghshineh, "An Architecture and Methodology for Mobile-Executed Handoff in Cellular ATM Networks", *IEEE Journal on Sel. Areas of Commun.*, Vol. 12, No. 8, Oct. 1994.

[21] S. -Oh and D. Tcha, "Prioritized Channel Assignment in a Cellular Radio Network", *IEEE Trans. on Vehicular Tech.*, vol. 40, no. 7, Aug. 1992.

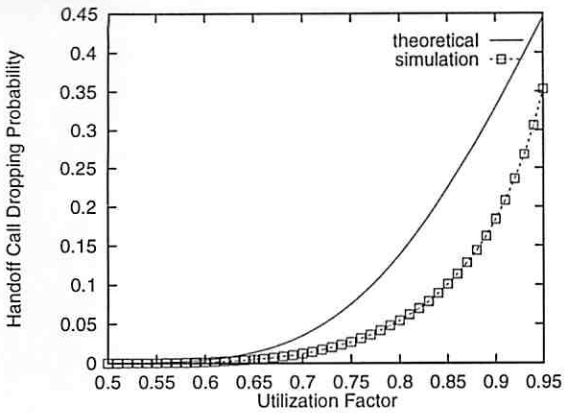

Figure 14: Handoff call dropping probability h vs utilization factor μ ($B = 4$, $C = 25$).

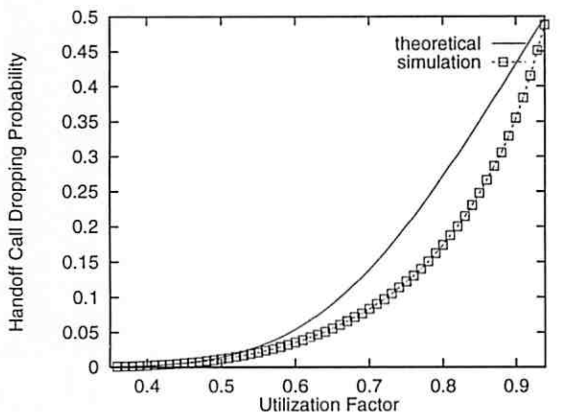

Figure 15: Handoff call dropping probability h vs utilization μ ($B = 49$, $C = 16$).

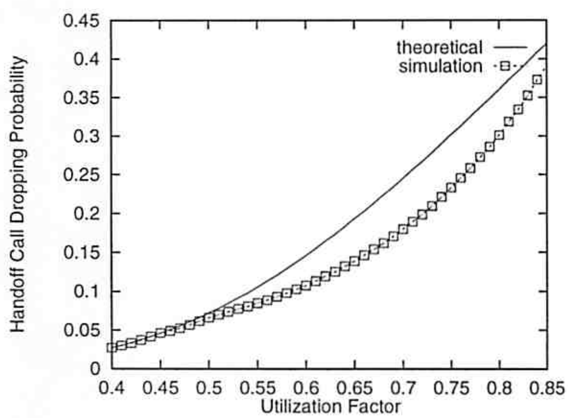

Figure 16: Handoff call dropping probability h vs utilization μ ($B = 121$, $C = 8$).

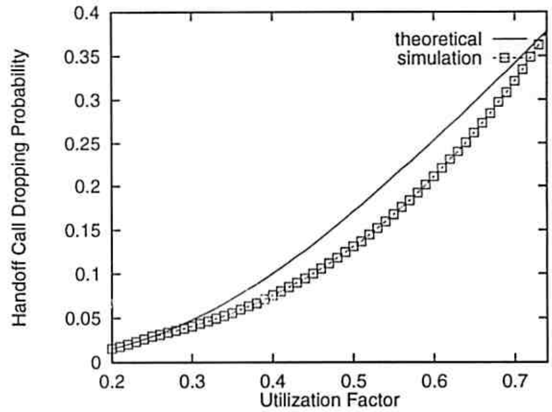

Figure 17: Handoff call dropping probability h vs utilization μ ($B = 196$, $C = 4$).

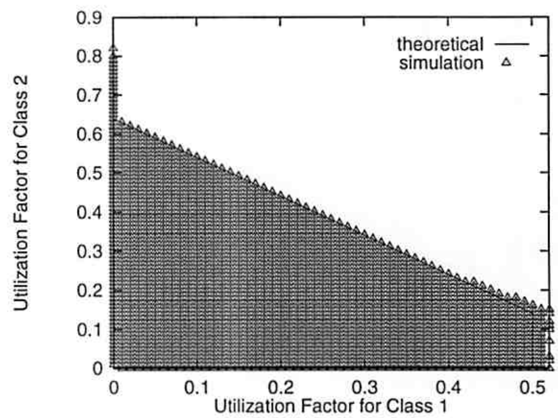

Figure 18: $B = 49$, $C = 16$, $H_1 = 0.015$, $H_2 = 0.200$, $P_2 = 0.050$.

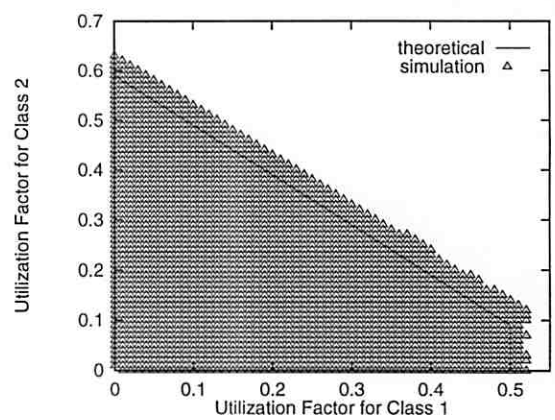

Figure 19: $B = 49$, $C = 16$, $H_1 = 0.015$, $H_2 = 0.050$, $P_2 = 0.050$.

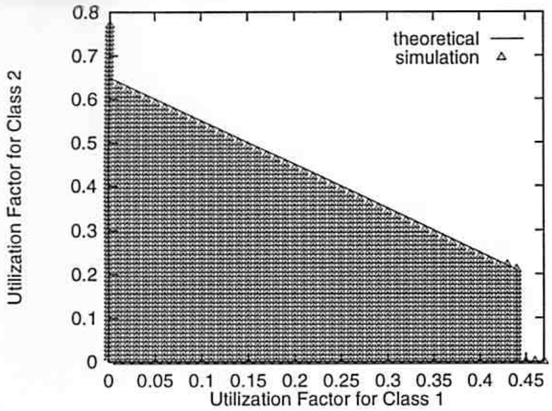

Figure 20: $B = 49$, $C = 16$, $H_1 = 0.005$, $H_2 = 0.150$, $P_2 = 0.050$.

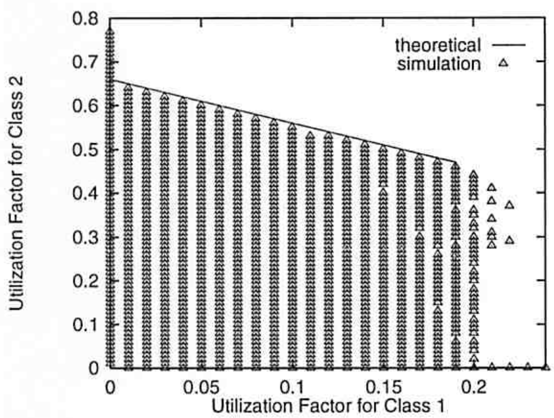

Figure 21: $B = 49$, $C = 16$, $H_1 = 0.000001$, $H_2 = 0.15$, $P_2 = 0.050$.

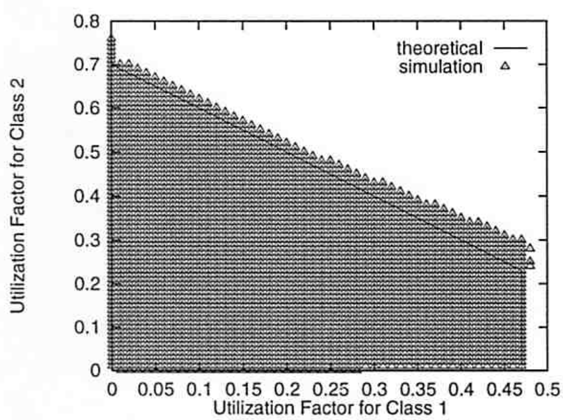

Figure 22: $B = 121$, $C = 8$, $H_1 = 0.06$, $H_2 = 0.25$, $P_2 = 0.18$.

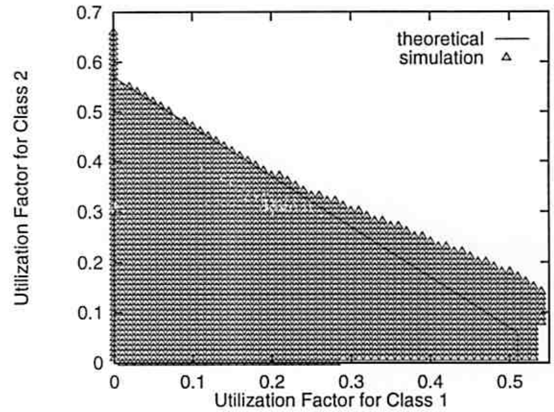

Figure 23: $B = 121$, $C = 8$, $H_1 = 0.08$, $H_2 = 0.15$, $P_2 = 0.10$.

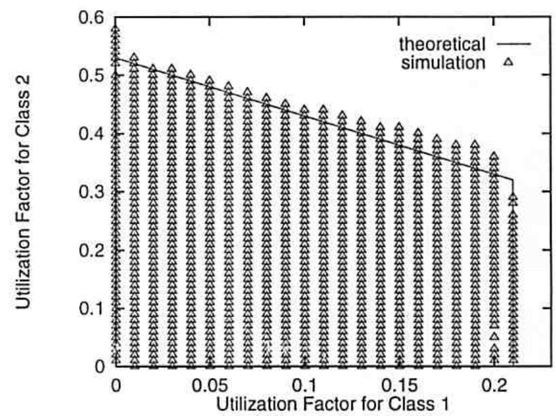

Figure 24: $B = 196$, $C = 4$, $H_1 = 0.021$, $H_2 = 0.200$, $P_2 = 0.150$.

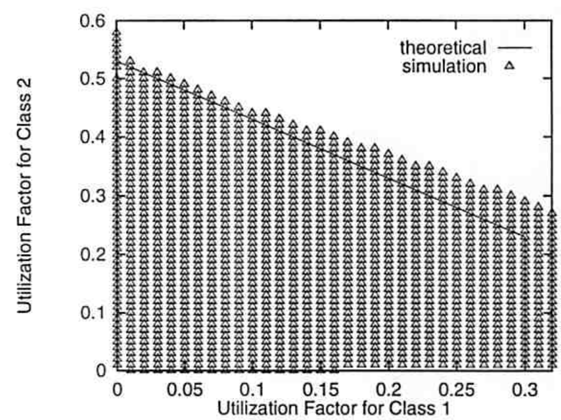

Figure 25: $B = 196$, $C = 4$, $H_1 = 0.05$, $H_2 = 0.20$, $P_2 = 0.15$.