ELSEVIER

Computer Communications 25 (2002) 972–979

computer communications

www.elsevier.com/locate/comcom

Call admission control for voice and data traffic in wireless communications

Chi Wa Leong, Weihua Zhuang*

Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Ont., Canada N2L 3G1

Received 6 June 2001; revised 8 October 2001; accepted 16 October 2001

Abstract

Call admission control (CAC) in multimedia wireless communications is to guarantee quality of service (QoS) to mobile users and to achieve a high resource utilization efficiency. It is a challenging issue due to user mobility, limited radio spectrum, heterogeneous nature of multimedia traffic. This paper proposes a CAC policy for a wireless system supporting both constant-rate voice calls and available-rate data calls. The parameters of the proposed policy is determined in a systematic way by using an optimization approach. Soft QoS (relaxed target QoS) is introduced to make a compromise between the QoS levels of the voice and data calls, respectively, in a heavy traffic load situation. Numerical results show that the proposed CAC policy can guarantee the QoS of all the users under most traffic conditions considered.

Keywords: Call admission control; Quality-of-service; Wireless communications; Voice and data traffic

1. Introduction

The future personal communication system (PCS) will provide multimedia communication services that can be enjoyed by users *anytime* and *anywhere* [1]. It is expected that the PCS will guarantee a certain level of quality of service (QoS) to mobile users. In the system, call admission control (CAC) will be the first control function imposed on a user for QoS provisioning. When a user requests a new connection, what the CAC does is to calculate the amount of resources required by (i) the users already in the system, and (ii) the pending user. If the sum of the two is not larger than the total capacity, then the user's request will be acknowledged; otherwise, the request will be rejected. This is equivalent to first reserving resources for the admitted users and then checking if the remaining resources are sufficient to support the new connection. How to make the two calculations depends on the particular system under consideration. CAC is important because its result is irreversible. A CAC mechanism is usually defined as the detailed work involved in the CAC function. This includes the decision process, signalling, routing table establishment, etc. The decision process of CAC can often be formulated in a high level representation called the CAC policy.[1] Whenever a user requests a new connection, the CAC policy takes the call request as input, and based on the current traffic conditions of the system, decides whether or not to accept the user, as illustrated in Fig. 1.

Because of the limited radio spectrum, user mobility, and dynamic traffic of multimedia services, CAC in the PCS poses significant challenges. Using cellular structure for frequency reuse, there are two major types of calls that can arrive at any particular cell: new calls originated from within and handoff calls coming from adjacent cells. From the users' point of view, it is better to be blocked in the beginning rather than dropped in the middle of a connection. As a result, handoff calls should be given a higher priority than new calls. To do so, the usual approach is to reserve some capacity exclusively for them, e.g. by using the limited fractional guard channel policy (LFGCP) [2]. On the other hand, resource utilization is another important factor that needs to be considered, in addition to QoS provisioning. A highly utilized system that can provide a satisfactory level of QoS to the users is a desired solution. However, high resource utilization and QoS provisioning are always conflicting goals. As in the case of CAC, resources are set aside for active handoff users so that their QoS can be maintained, but unused resources result in low utilization. In order to have a balance in the two conflicting goals, the amount of reserved resources should

---

* Corresponding author. Tel.: +1-519-888-4567x5354; fax: +1-519-746-3077.

E-mail address: wzhuang@bbcr.uwaterloo.ca (W. Zhuang).

[1] A call admission control policy is synonymous to a call admission control algorithm.

0140-3664/02/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0140-3664(01)00434-0

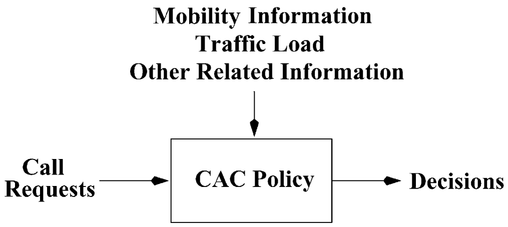

Fig. 1. The CAC decision process.

be calculated carefully. Due to the dynamic nature of multimedia traffic and the characteristics of a wireless mobile environment, the traffic load in a base station can change dramatically from time to time. As a result, simultaneously providing consistent QoS satisfaction to all the mobile users with heterogeneous requirements is not realistic. Certain compromise in QoS provisioning among different traffic classes has to be made. Adaptive CAC with soft QoS (i.e. relaxed target QoS) is one approach to make such a compromise. Normally, QoS requirement is specified by a value for each of the QoS parameters and is referred to as hard QoS requirement. On the other hand, when the QoS requirement is specified by a range (instead of a single value) for each of the QoS parameters, guaranteeing the QoS requirement is referred to as soft QoS provisioning. When the most stringent QoS requirement in soft QoS is the target service quality in the hard QoS, the soft QoS is actually relaxed hard QoS. It should be noted that provisioning soft QoS (or relaxed hard QoS) is different from setting hard QoS at a lower level (e.g. the least stringent requirement in the soft QoS). In provisioning hard QoS, there is no adaptive mechanism in resource allocation; e.g. when traffic load is low, available resources are not utilized for better QoS and are wasted. Even though it is quite obvious that relaxing QoS constraints can improve system resource utilization, in-depth investigation is necessary to evaluate the tradeoff between service quality and resource utilization.

This paper proposes a CAC policy for a wireless system supporting both constant-rate voice and available-rate data traffic based on the LFGCP. This work is based on the previous research reported in Ref. [2], where only constant-rate traffic was considered. Although the same mixed traffic types have been considered in Ref. [3], the mobility model considered is more general. Studies related to admission control for integrated voice and data services in a code division multiple access (CDMA) system have been carried out in Refs. [4–8]. The results of Refs. [4–7] are specific to the system model and can be applied to CDMA systems only. In Ref. [8], a general queuing model is proposed to describe the flow of voice and data traffic in the system. However, both the voice and data services considered are constant-rate services. On the other hand, a concept similar to soft QoS was introduced in Ref. [9], where an 'adaptive reserved service' framework was proposed to support mobile connections. However, no in-depth investigation on CAC is presented. The contribution of this paper lies in the development of the CAC policy for the integrated voice and data services, which share the total resources without partitioning, where handoff calls are given with a high priority than new calls. The objective of the service provider is formulated in a mathematical manner and an optimization approach is then used to obtain the CAC policy parameters that optimize such objective. In this way, the search for the CAC policy can be performed systematically rather than arbitrarily. The proposed policy is optimal in the sense that, given a traffic load, the QoS requirements from both constant- and available-rate users are guaranteed with the minimum system resources. In addition, soft QoS levels are specified in the objective so that none of the QoS of the competing traffic types is discriminated against indefinitely. The CAC policy incorporates the concept of soft QoS in order to make a good compromise in resource allocation between the two traffic types. If mobile users can tolerate a certain degree of fluctuations in the service quality, then the system can exploit the QoS flexibility to improve resource utilization by adaptive resource allocation based on the instantaneous traffic load, wireless link quality, and user application characteristics. Consequently, the results of this paper are expected to provide some insights into how the optimal CAC policy can be found systematically in the multimedia wireless environment of the PCS.

The rest of this paper is organized as follows. Section 2 describes the system model under consideration. Section 3 presents the proposed CAC policy for both constant- and available-rate traffic. Section 4 discusses how to determine the parameters of the proposed CAC policy. Numerical results are given in Section 5 to demonstrate the performance of the proposed CAC policy with target and soft QoS, respectively. Finally, Section 6 concludes this research.

2. The system model

A wireless system with hexagonal radio cells and fixed channel allocations is considered. Although there are many types of multimedia traffic in the PCS, we consider only two of the most important ones, voice and data. Voice traffic carries real-time information and can be classified as constant-rate traffic; data traffic carries non-real-time information and can be classified as available-rate traffic. The traffic load is uniform throughout the coverage area. Both the system and the traffic are assumed to have attained stationary states. All new and handoff call arrivals to a cell are Poisson, and all call connection periods and cell residence times are exponentially distributed and are independent of each other. The handoff rates of the same traffic type at a cell from the neighboring cells can be lumped into one single parameter. After admitted to a cell, a handoff call is treated in the same way as a new call of the same traffic type, i.e. the service rates for both handoff and new calls of

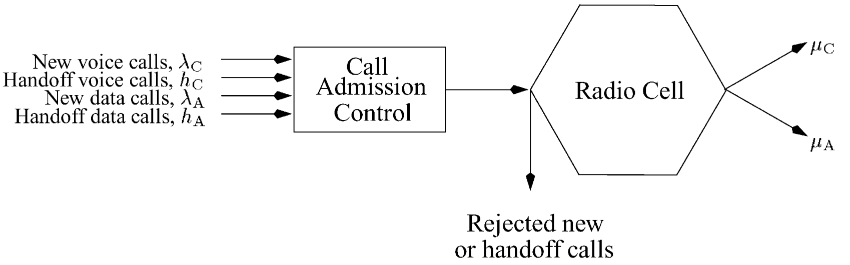

Fig. 2. The system traffic model under CAC.

the same traffic type are the same. In such a system, the focus can be placed on a single test cell described by the new and handoff call arrival rates, service rates and the total cell capacity. The service time in the test cell is the minimum of call duration and the cell residence time. Therefore, the service rate is the sum of the call service rate and the handoff rate of the admitted users. The following assumptions are made: (i) arrivals and services of constant-rate traffic are independent to those of available-rate traffic; (ii) the effects of channel fading are completely mitigated by the physical and link layers, and the problem of capacity fluctuation in a wireless environment is eliminated; and (iii) the amount of resources occupied by control signals is negligible. Fig. 2 shows the traffic modeling of the system under CAC, where λ_C, h_C, and μ_C denote the new call arrival rate, handoff call arrival rate and service rate for constant-rate traffic, respectively, the same notations with subscript A are used to represent the corresponding quantities for available-rate traffic. The QoS parameters of interest are new call blocking probability and handoff call dropping probability for both constant- and available-rate calls.

Since there are two types of traffic competing for the resources in the network, a resource allocation scheme is needed to specify the priority between the different traffic types. Here, the restricted access (RA) scheme [3] is used. This scheme is preferred over others because low priority traffic will never be deprived of resources indefinitely. The RA scheme is depicted in Fig. 3, where C is the total capacity (radio resources) of the test cell. Here, the transmission capacity is assumed to be discretized into small but equal portions. Each portion is called a capacity unit (CU). The real-time constant-rate traffic has preemptive priority over the non-real-time available-rate traffic, and can occupy up to C_C CUs out of the total capacity. The bandwidth requirement of each admitted constant-rate call in CU is denoted by γ_C. The leftover resources in the test cell are equally shared among all the admitted available-rate users. The amount of resources received by each available-rate user in CU, denoted by γ_A, is therefore time-varying, depending on the numbers of the admitted constant- and available-rate users, respectively, at the instant. When the system has a high traffic load, this amount may drop to such a low level that any available-rate connection will suffer dramatic performance degradation. This phenomenon is called overload event. Let γ_{crA} denote the critical amount of capacity below which severe performance degradation occurs to the admitted available-rate users. The overload probability for available-rate users, P_{oA}, which is defined as the probability that γ_A is less than γ_{crA}, is a QoS measure that should be considered for the available-rate calls, in addition to the call blocking and dropping probabilities.

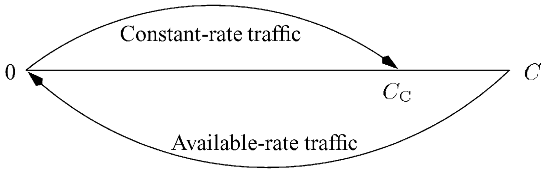

Fig. 3. The resource allocation scheme.

3. The proposed CAC policy

The proposed CAC policy consists of two LFGCPs, one for the constant-rate calls and the other for the available-rate calls. The LFGCP is chosen because (i) the policy is simple to implement, and (ii) given certain QoS requirements, the CAC policy parameters can be determined in a systematic way. The proposed policy can be denoted by $\{g^{\beta_C}_{T_C,M_C}, g^{\beta_A}_{T_A,M_A}\}$, where $T_C < M_C$, $T_A < M_A$, M_C (M_A) is the maximum number of total constant-rate (available-rate) users allowed in the cell, T_C (T_A) is the number of admitted constant-rate (available-rate) users over which no new constant-rate (available-rate) calls are accepted, β_C (β_A) is a constant denoting the probability of accepting a new constant-rate (available-rate) call when the number of constant-rate (available-rate) users admitted in the cell is T_C (T_A). Let i (j) denote the current number of admitted constant-rate (available-rate) users. The policy can be summarized as follows: (i) a new constant-rate (available-rate) call is always accepted if $i < T_C$ ($j < T_A$), is accepted with probability β_C (β_A) if $i = T_C$ ($j = T_A$), and is always rejected if $i > T_C$ ($j > T_A$), and (ii) a handoff constant-rate (available-rate) call is always accepted if $i < M_C$ ($j < M_A$). The two LFGCPs are related to each other as the constant- and available-rate users share the total capacity C. The correlation of the policies is captured in the calculation of the overload probability for the admitted available-rate users and in the process of determining the policy parameter set $\{M_C, T_C, \beta_C, M_A, T_A, \beta_A\}$.

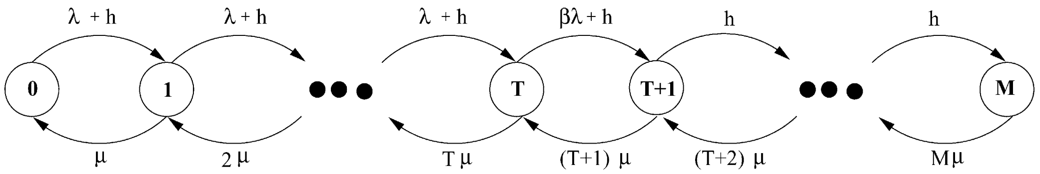

Fig. 4. The state diagram for a system with LFGCP.

Given the parameter set of the CAC policy, the traffic model for the numbers of constant- and available-rate users, respectively, in the cell can be decoupled, each being represented by a Markov chain, as shown in Fig. 4, where the subscript C or A is omitted. Let B_{nC} (B_{nA}) and B_{hC} (B_{hA}) denote the new call blocking probability and handoff call dropping probability, respectively, for the constant-rate (available-rate) users. From the Markov chain shown in Fig. 4, it can be derived that the steady-state probabilities of the constant-rate calls are

$$p_{\mathrm{C}}(i) = \begin{cases} \dfrac{\rho_{\mathrm{C}}^i}{i!} p_{\mathrm{C}}(0), & 0 \le i \le T_{\mathrm{C}} \\ \dfrac{\rho_{\mathrm{C}}^i[\alpha_{\mathrm{C}} + (1-\alpha_{\mathrm{C}})\beta_{\mathrm{C}}]}{i!} p_{\mathrm{C}}(0), & i = T_{\mathrm{C}} + 1 \\ \dfrac{\rho_{\mathrm{C}}^i[\alpha_{\mathrm{C}} + (1-\alpha_{\mathrm{C}})\beta_{\mathrm{C}}]\alpha_{\mathrm{C}}^{i-T_{\mathrm{C}}-1}}{i!} p_{\mathrm{C}}(0), & T_{\mathrm{C}} + 2 \le i \le M_{\mathrm{C}} \end{cases}, \tag{1}$$

where $\rho_{\mathrm{C}} = (\lambda_{\mathrm{C}} + h_{\mathrm{C}})/\mu_{\mathrm{C}}$, $\alpha_{\mathrm{C}} = h_{\mathrm{C}}/(\lambda_{\mathrm{C}} + h_{\mathrm{C}})$, and $p_{\mathrm{C}}(0)$ is a normalization constant given by

$$p_{\mathrm{C}}(0) = \left[\sum_{k=0}^{T_{\mathrm{C}}} \frac{\rho_{\mathrm{C}}^k}{k!} + \frac{\rho_{\mathrm{C}}^{T_{\mathrm{C}}+1}[\alpha_{\mathrm{C}} + (1-\alpha_{\mathrm{C}})\beta_{\mathrm{C}}]}{(T_{\mathrm{C}}+1)!} + \sum_{k=T_{\mathrm{C}}+2}^{M_{\mathrm{C}}} \frac{\rho_{\mathrm{C}}^k[\alpha_{\mathrm{C}} + (1-\alpha_{\mathrm{C}})\beta_{\mathrm{C}}]\alpha_{\mathrm{C}}^{k-T_{\mathrm{C}}-1}}{k!} \right]^{-1}.$$

The new call blocking probability and handoff call dropping probability for constant-rate calls in the cell can then be derived as

$$B_{\mathrm{nC}} = (1 - \beta_{\mathrm{C}}) \cdot p_{\mathrm{C}}(T_{\mathrm{C}}) + \sum_{i=T_{\mathrm{C}}+1}^{M_{\mathrm{C}}} p_{\mathrm{C}}(i), \tag{2}$$

$$B_{\mathrm{hC}} = p_{\mathrm{C}}(M_{\mathrm{C}}). \tag{3}$$

Eqs. (1)–(3) can be used to calculate the steady-state probabilities and blocking performance of available-rate calls by simply replacing i with j and the subscript C with A.

The other QoS of interest is the overload probability P_{oA} for available-rate calls. By following the approach similar to that used in Ref. [3], the overload probability can be derived as

$$P_{\mathrm{oA}} = \sum_{r \in A_{\mathrm{R}} \cap r < \gamma_{\mathrm{crA}}} \Pr[\gamma_{\mathrm{A}} = r], \tag{4}$$

where

$$\Pr[\gamma_{\mathrm{A}} = r] = \frac{\sum_{(i,j) \in A_{\mathrm{S}}} p_{\mathrm{C}}(i) \cdot j \cdot p_{\mathrm{A}}(j)}{\sum_{k=0}^{M_{\mathrm{A}}} k \cdot p_{\mathrm{A}}(k)},$$

A_{R} is the set of distinct values in

$$\left\{ \frac{C - i \cdot \gamma_{\mathrm{C}}}{j} : \quad 0 \le i \le M_{\mathrm{C}},\ 0 \le j \le M_{\mathrm{A}} \right\},$$

and A_{S} is the set of all combinations of i ($0 \le i \le M_{\mathrm{C}}$) and j ($0 \le j \le M_{\mathrm{A}}$) such that $(C - i\gamma_{\mathrm{C}})/j = r$.

4. Search for the CAC policy parameters

The parameter set $\{M_{\mathrm{C}}, T_{\mathrm{C}}, \beta_{\mathrm{C}}, M_{\mathrm{A}}, T_{\mathrm{A}}, \beta_{\mathrm{A}}\}$ will be determined based on not only the optimization approach but also the specification of soft QoS levels. Let Q_{nC} (Q_{nA}) and Q_{hC} (Q_{hA}) denote the required upper bounds of the new call blocking probability for the constant-rate (available-rate) users, handoff call dropping probability for the constant-rate (available-rate) users, respectively; let Q_{oA} denote the required maximum overload probability of available-rate calls. The optimization problem can be formulated as follows:

Given: the new call arrival rates $\{\lambda_{\mathrm{C}}, \lambda_{\mathrm{A}}\}$, handoff arrival rates $\{h_{\mathrm{C}}, h_{\mathrm{A}}\}$, service rates $\{\mu_{\mathrm{C}}, \mu_{\mathrm{A}}\}$, QoS requirements $\{Q_{\mathrm{nC}}, Q_{\mathrm{nA}}, Q_{\mathrm{hC}}, Q_{\mathrm{hA}}\}$;
Objective: find the minimum M_{C} and M_{A} such that

$$B_{\mathrm{nC}} \le Q_{\mathrm{nC}}, \qquad B_{\mathrm{hC}} \le Q_{\mathrm{hC}}, \qquad B_{\mathrm{nA}} \le Q_{\mathrm{nA}},$$

$$B_{\mathrm{hA}} \le Q_{\mathrm{hA}}.$$

Minimizing the values of M_{C} and M_{A} under the QoS constraints helps to bring down the service cost since the cost of building a base station is proportional to the numbers of the constant- and available-rate users that it can support. In addition, a certain level of overload probability for admitted available-rate users should be guaranteed. For a given total capacity C, because the available-rate calls are of low priority in the system, the suggested upper bound for the overload probability cannot be guaranteed if the constant-rate traffic load is high. In this case, it may be necessary to make a compromise in service quality between

Table 1
Experimental parameters used in Example 1

μ_C	μ_A	λ_C	h_C	λ_A	h_A	C_C	C	η_n	η_h	γ_A	Q_{nC}^T	Q_{hC}^T	Q_{nA}	Q_{hA}	Q_{oA}
10	10	[20,400]	[20,400]	130	70	5	250	1.25	1.25	1	10^{-3}	10^{-4}	0.2	0.02	0.01

the constant- and available-rate traffic. One possible approach is to introduce soft QoS to constant-rate traffic. By specifying a soft QoS level as well as a target one for the new and handoff calls of the constant-rate traffic, a certain overload probability can also be guaranteed to any available-rate call for most of the time. Especially, let

$$Q_{\mathrm{nC}}^{\mathrm{S}} = \eta_{\mathrm{n}} \cdot Q_{\mathrm{nC}}^{\mathrm{T}}, \tag{5}$$

$$Q_{\mathrm{hC}}^{\mathrm{S}} = \eta_{\mathrm{h}} \cdot Q_{\mathrm{hC}}^{\mathrm{T}}, \tag{6}$$

where $\eta_n > 1$ and $\eta_h > 1$, and the superscript S or T is used to indicate whether the QoS requirement is the soft or the target one for the constant-rate calls. With both the optimization approach and soft QoS for the constant-rate calls, the parameters of the CAC policy can be determined using `Algorithm CAR` given in the following, where MinM(·) represents a call to `Algorithm Min M`:

`Algorithm CAR`

```
1. $(M_C^T, T_C^T, \beta_C^T) = \mathrm{MinM}(Q_{nC}^T, Q_{hC}^T)$;
2. $(M_C^S, T_C^S, \beta_C^S) = \mathrm{MinM}(Q_{nC}^S, Q_{hC}^S)$;
3. $(M_A, T_A, \beta_A) = \mathrm{MinM}(Q_{nA}, Q_{hA})$;
4. if $M_C^S \cdot \gamma_C > C$
       return ('Error: The QoS requirements can
          never be fulfilled with current resources.')
       else
       if $Q_{oA} < P_{oA}(g_{T^T}^{\beta_C^T}, M_C^T, g_{T_A}^{\beta_A}, M_A)$
          {return$(M_C^S, T_C^S, \beta_C^S, M_A, T_A, \beta_A)$}
       else {return$(M_C^T, T_C^T, \beta_C^T, M_A, T_A, \beta_A)$};
```

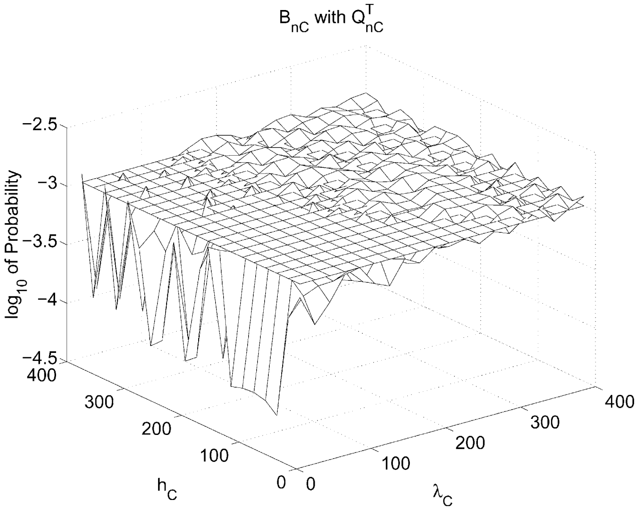

Fig. 5. New call blocking probability for constant-rate calls with target QoS.

The algorithm is to find a set of suitable parameters of each LFGCP which can satisfy the specified new call blocking probability and handoff call dropping probability requirements with the minimum M_C or M_A value. The function $P_{oA}(g_{T_C,M_C}^{\beta_C}, g_{T_A,M_A}^{\beta_A})$ returns the overload probability when $g_{T_C,M_C}^{\beta_C}$ and $g_{T_A,M_A}^{\beta_A}$ are chosen as the CAC policy for constant- and available-rate calls, respectively. Finally, `Algorithm CAR` returns a hextuple $(M_C, T_C, \beta_C, M_A, T_A, \beta_A)$, which completely specifies the CAC policy. The use of the soft QoS is signified in step 4, which indicates that if the overload probability bound is violated by the achievement of the target QoS requirements for the constant-rate calls, soft QoS requirements will be adhered to.

`Algorithm Min M` [2], e.g. for constant-rate calls, can be represented as

`Algorithm Min M`

```
1. $M_C = 1$;
2. while $B_{nC}(M_C, M_C, 0) > Q_{nC}$ {$M_C = M_C + 1$};
3. if $M_C > M_{C,MAX}$ {return('Error')};
4. if $B_{hC}(M_C, M_C, 0) \le Q_{hC}$ {return$(M_C, M_C, 0)$};
5. $U = M_C$; $L = 0$; $x = (U + L)/2$;
6. while $B_{nC}(M_C, \mathrm{Int}(x), \mathrm{Frac}(x)) > Q_{nC}$ XOR
            $B_{hC}(M_C, \mathrm{Int}(x), \mathrm{Frac}(x)) > Q_{hC}$
      {if $B_{nC}(M_C, \mathrm{Int}(x), \mathrm{Frac}(x)) > Q_{nC}$
         {$L = x$; $x = (U + L)/2$}
      else if $B_{hC}(M_C, \mathrm{Int}(x), \mathrm{Frac}(x)) > Q_{hC}$
         {$U = x$; $x = (U + L)/2$}};
7. if $B_{nC}(M_C, \mathrm{Int}(x), \mathrm{Frac}(x)) \le Q_{nC}$ AND
            $B_{hC}(M_C, \mathrm{Int}(x), \mathrm{Frac}(x)) \le Q_{hC}$
      {return$(M_C, \mathrm{Int}(x), \mathrm{Frac}(x))$}
8. else {$M_C = M_C + 1$; goto step 3};
```

where $B_{nC}(M_C, T_C, \beta_C)$ ($B_{hC}(\cdot)$) is the function that calculates the new call blocking probability (handoff call dropping probability) of LFGCP $g_{T_C,M_C}^{\beta_C}$, and the functions Int(x) and Frac(x) return, respectively, the integer part and the fractional part of the number x. Note that step 3 of the algorithm returns an error, when no practical value of M_C can be found to fulfill the QoS requirements specified by Q_{nC} and Q_{hC}. The value of $M_{C,MAX}$ in step 3 is the physical limit of the number of calls that any radio cell can support.

5. Numerical examples

The performance of the proposed CAC policy with `Algorithm CAR` will be demonstrated by the following two examples under various traffic conditions.

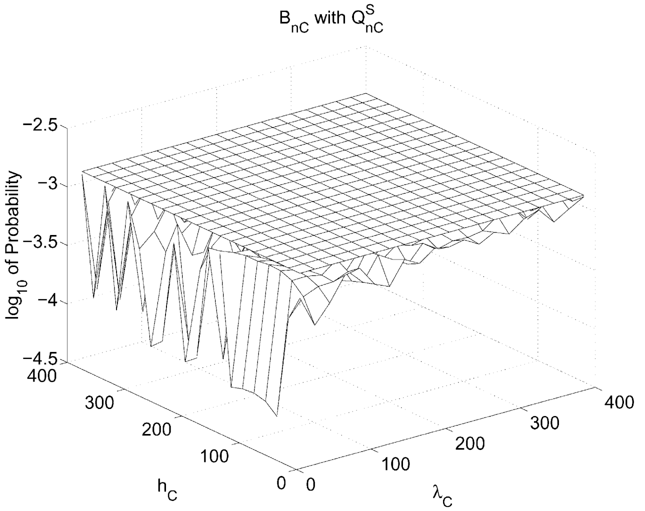

Fig. 6. New call blocking probability for constant-rate calls with soft QoS.

Example 1. In this example, the parameters of the available-rate traffic load (λ_A and h_A) are fixed, while the parameters of the constant-rate traffic load (λ_C and h_C) are allowed to vary. Table 1 gives the system parameters used. Numerical analysis indicates that B_{nA} is fixed at 0.18216 and B_{hA} at 0.01104. The probabilities are constant because λ_A and h_A are constant. Note that both blocking probabilities meet the QoS requirements. This is for $g^{\beta_A}_{T_A,M_A}$ with $M_A = 23$, $T_A = 20$ and $\beta_A = 0.8438$. As the values of λ_C and h_C change, the LFGCP parameters for the constant-rate calls also change. Other system QoS parameters under the control of the proposed CAC policy are plotted in Figs. 5–9. It can be observed that:

(i) For the constant-rate calls, the new call blocking probability is below the specified target upper bound Q^{T}_{nC} when the traffic load is low. However, as the traffic load (h_C and λ_C) increases, the target upper bound for the new call blocking probability no longer holds. On the other hand, with the introduction of the soft QoS, the CAC policy can guarantee the target upper bound, Q^{T}_{nC}, in a low traffic load condition and the soft upper bound, Q^{S}_{nC}, in a heavy traffic load condition. Note that without introducing the soft QoS to the CAC policy, the soft upper bound, Q^{S}_{nC}, cannot be guaranteed automatically.

(ii) Similar to the new call blocking probability, for the constant-rate calls, the target upper bound, Q^{T}_{hC}, for handoff call dropping probability is guaranteed in a low traffic load condition and is violated when traffic gets heavier, although the soft QoS (Q^{S}_{hC}) is always precisely maintained.

(iii) For the available-rate calls, the overload probability changes with h_C and λ_A, even though λ_A and h_A are fixed.

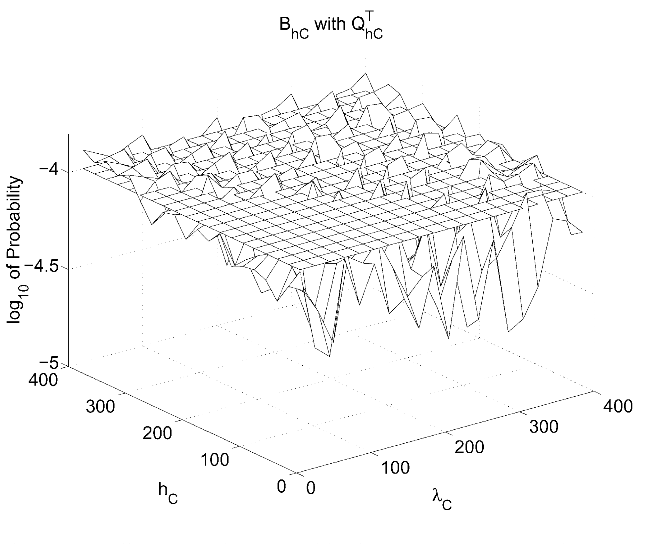

Fig. 7. Handoff call dropping probability for constant-rate calls with target QoS.

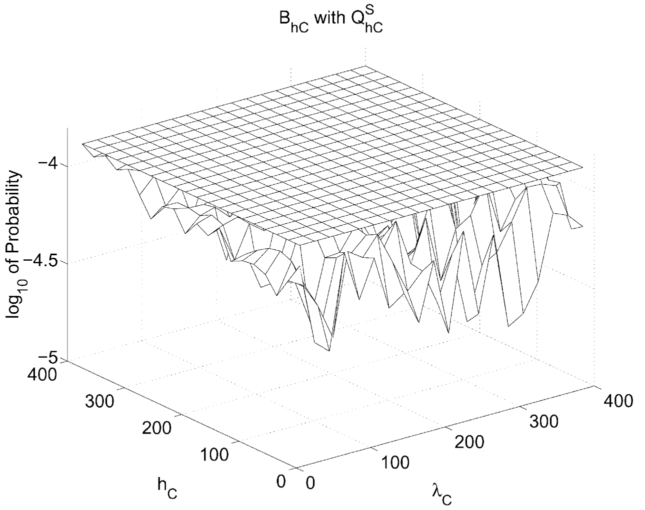

Fig. 8. Handoff call dropping probability for constant-rate calls with soft QoS.

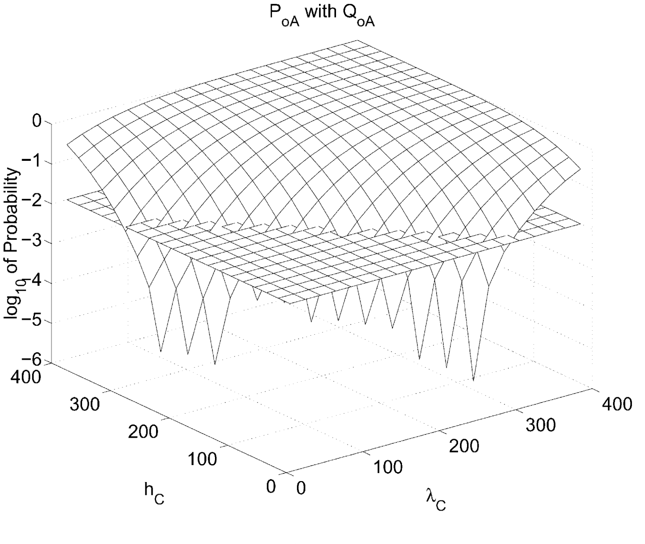

Fig. 9. The overload probability for available-rate calls and the required upper bound.

Table 2
Experimental parameters used in Example 2

μ_C	μ_A	λ_C	h_C	λ_A	h_A	C_C	C	η_n	η_h	γ_A	Q_{nC}^T	Q_{hC}^T	Q_{nA}	Q_{hA}	Q_{oA}
10	10	200	100	[20,300]	[20,300]	5	250	1.6	1.6	1	10^{-3}	10^{-4}	0.2	0.02	0.01

When the traffic load is low (small λ_C and h_C values), P_{oA} can stay below the QoS bound (Q_{oA}) most of the time. However, as λ_C and/or h_C increases, the QoS requirement for the data calls cannot be met. Other numerical results show that the overload probability is indeed reduced by introducing the soft QoS to the constant-rate calls in the heavy traffic load condition. Given C and C_C, as the traffic load increases, the available-rate calls definitely will experience degraded QoS (e.g. the increased overload probability) because the available-rate calls have a lower priority than the constant-rate calls.

Example 2. In this example, the traffic load parameters λ_C and h_C are fixed while λ_A and h_A are allowed to vary. The system parameters are given in Table 2. The system QoS parameters under the control of the proposed CAC policy are plotted in Figs. 10–14. The following observations can be made:

(i) Because λ_C and h_C are constant, there are only two possible values for the new call blocking probability of the constant-rate calls, B_{nC}, one to satisfy the target QoS (Q_{nC}^T) and the other, the soft QoS (Q_{nC}^S). The target upper bound for the blocking probability ($\log_{10} Q_{nC}^T = -3$) is satisfied when the traffic load is low. As the traffic load increases, the upper bound cannot be guaranteed any more and the CAC policy tries to satisfy the soft upper bound. The introduction of soft QoS works well as the soft upper bound ($\log_{10} Q_{nC}^S \approx -2.8$) is guaranteed in the heavy traffic load situation. Similarly, the handoff call dropping probability for the constant-rate calls has two possible values. The target upper bound for the handoff call dropping probability ($\log_{10} Q_{hC}^T = -4$) is guaranteed in the light traffic condition and the soft upper bound ($\log_{10} Q_{hC}^S \approx -3.8$) is ensured in the heavy traffic load condition.

(ii) For the available-rate calls, the QoS requirements for both new and handoff calls are always maintained under the traffic load considered.

(iii) As to the overload probability of the available-rate calls, in the low to medium traffic conditions, the QoS requirement ($\log_{10} Q_{oA} = -2$) can be met most of time. It is only when both the available-rate arrival rates (λ_A and h_A) have a large value that the QoS requirement is violated. Under the traffic load condition, the overload probability would have a higher value if soft QoS were not introduced to the constant-rate calls. As the constant-rate traffic has a higher priority than the available-rate traffic, given C and C_C, the overload probability increases with the traffic load. The soft QoS to the constant-rate calls can mitigate the QoS degradation of the available-rate calls to a certain extent. If the traffic load further increases, more resources (higher C and C_C values) will be required to meet the QoS requirements of both the constant- and available-rate calls.

6. Conclusions

A CAC policy for wireless networks with constant-rate voice and available-rate data traffic has been proposed. The

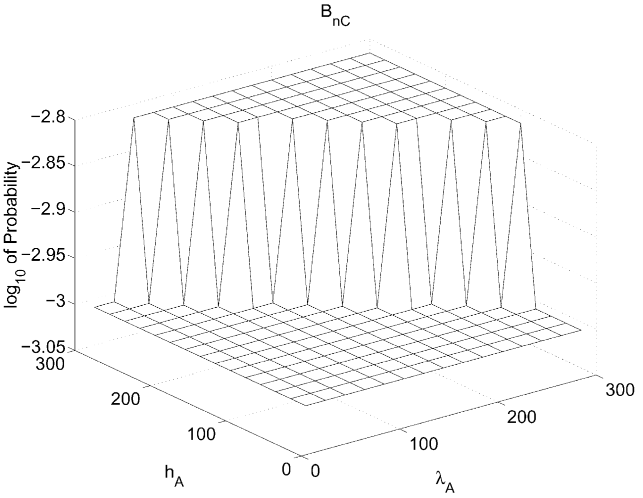

Fig. 10. New call blocking probability for constant-rate calls.

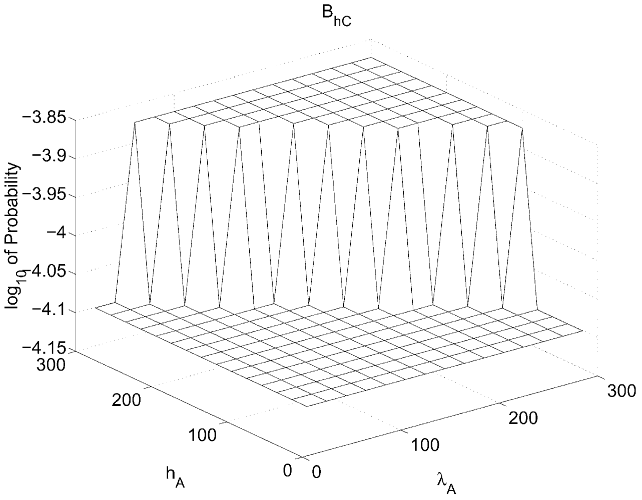

Fig. 11. Handoff call dropping probability for constant-rate calls.

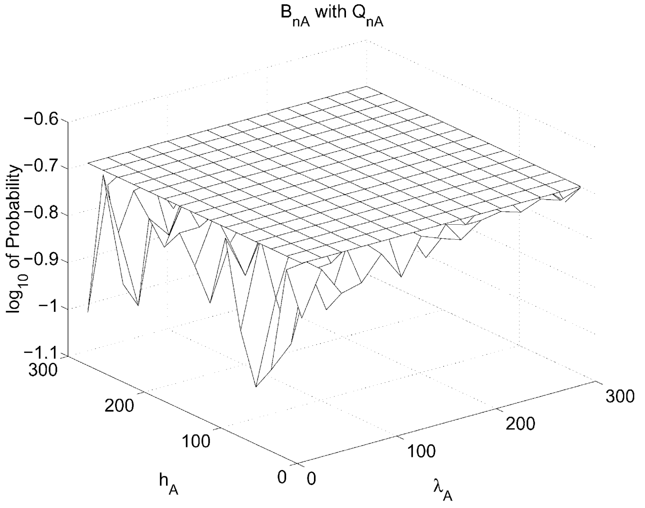

Fig. 12. New call blocking probability for available-rate calls with required QoS.

optimal parameters of the policy can be found using the proposed `Algorithm CAR`, which incorporates soft QoS for the constant-rate calls in order to improve the QoS of the available-rate calls in a heavy traffic load environment. Numerical results demonstrate that the proposed CAC policy can meet all the QoS requirements under most traffic conditions. This indicates the robustness of the proposed CAC policy and `Algorithm CAR`, and demonstrates the benefit of their deployment in a network supporting both voice and data traffic.

The objectives of CAC in wireless systems are to guarantee QoS to mobile users and to achieve a high resource utilization efficiency. To achieve the two objectives simultaneously is technically very challenging due to limited radio spectrum, user mobility, and heterogeneous nature of multimedia traffic. Soft QoS is one approach for a compromise between the two objectives. In-depth numerical analysis is necessary to demonstrate the benefits of soft QoS and to ensure the balance between QoS provisioning and resource utilization. This research presents an example in such investigation.

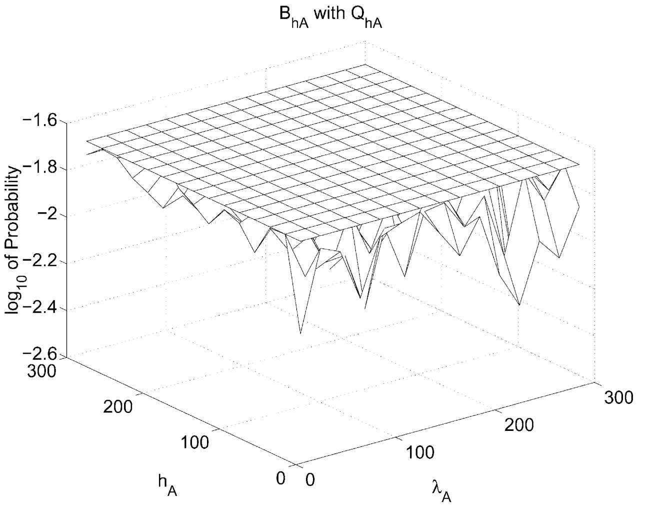

Fig. 13. Handoff call dropping probability for available-rate calls with required QoS.

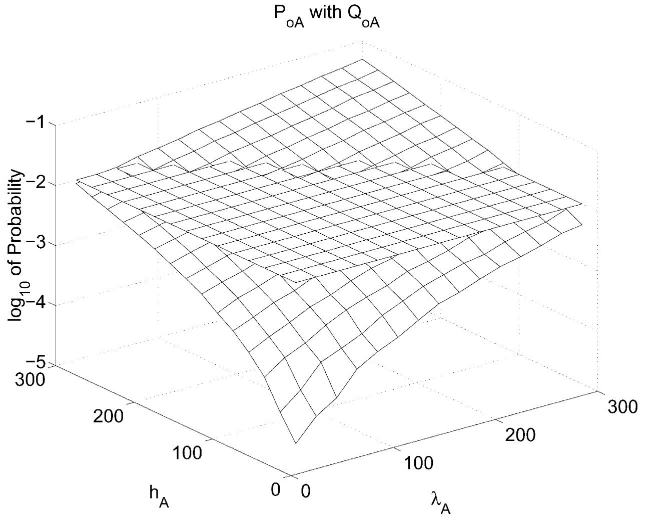

Fig. 14. The overload probability for available-rate calls and the required upper bound.

Acknowledgements

This work was supported by the Communications and Information Technology Ontario (CITO) and by the Canadian Institute for Telecommunications Research (CITR).

References

[1] V.O.K. Li, X. Qiu, Personal communication system (PCS), Proc. IEEE 83 (1995) 1210–1243.

[2] R. Ramjee, D. Towsley, R. Nagarajan, On optimal call admission control in cellular networks, Wireless Networks 3 (1997) 29–41.

[3] M. Naghshineh, A. Acampora, QoS provisioning in micro-cellular networks supporting multiple classes of traffic, Wireless Networks 2 (1996) 195–203.

[4] D. Ayyagari, A. Ephremides, Cellular multicode CDMA capacity for integrated (voice and data) services, IEEE J. Select. Areas Commun. 17 (5) (1999) 928–938.

[5] T. Liu, J. Silvester, Joint admission/congestion control for wireless CDMA systems supporting integrated services, IEEE J. Select. Areas Commun. 16 (6) (1998) 845–857.

[6] A. Sampath, J. Holtzman, Access control of data in integrated voice/data CDMA systems: benefits and tradeoffs, IEEE J. Select. Areas Commun. 15 (8) (1997) 1511–1526.

[7] W. Yang, E. Geraniotis, Admission policies for integrated voice and data traffic in CDMA packet radio networks, IEEE J. Select. Areas Commun. 12 (4) (1994) 654–664.

[8] C. Wu, Y. Tsai, J. Chang, A quality-based birth-and-death queuing model for evaluating the performance of an integrated voice/data CDMA cellular system, IEEE Trans. Veh. Technol. 48 (1) (1999) 83–89.

[9] K. Lee, Supporting mobile multimedia in integrated services networks, Wireless Networks 2 (3) (1996) 205–217.