Adaptive bandwidth framework for provisioning connection-level QoS for next-generation wireless cellular networks

Cadre adaptatif de la largeur de bande selon la qualité de service des connections dans les réseaux cellulaires sans fil de prochaine génération

Nidal Nasser and Hossam Hassanein*

The next generation of wireless cellular networks is expected to support real-time multimedia services with different classes of traffic and diverse bandwidth requirements. Bandwidth is a scarce resource in wireless networking that needs to be carefully allocated amid competing connections with different quality of service (QoS) requirements. In this paper, an adaptive framework for supporting multiple classes of multimedia services with different QoS requirements in wireless cellular networks is proposed. The framework combines the following components: (1) a threshold-based bandwidth allocation policy that gives priority to handoff calls over new calls and prioritizes among different classes of handoff calls by assigning a threshold to each class; (2) an efficient threshold-type call admission control (CAC) algorithm; and (3) a bandwidth adaptation algorithm (BAA) that dynamically adjusts the bandwidth of ongoing multimedia calls to minimize the number of calls receiving lower bandwidth than they had requested. Numerical results show that the proposed adaptive multimedia framework outperforms existing non-adaptive schemes in terms of the handoff call dropping probability and effective utilization.

On s'attend à ce que la prochaine génération de réseaux cellulaires sans fil supporte des services multimédia en temps réel avec différentes classes de trafic et différents besoins en largeur de bande. La largeur de bande, ressource rare dans les réseaux sans fil, a besoin d'être prudemment allouée parmi toutes les connections en compétition nécessitant différents niveaux de qualité de service. Cet article propose un cadre adaptatif supportant plusieurs classes de services multimédia nécessitant différentes qualités de service dans un réseau cellulaire sans fil. Le cadre combine les éléments suivants : (1) une politique d'allocation de la largeur de bande basée sur un seuil qui donne la priorité aux appels en transfert sur les nouveaux appels et qui priorise les différentes classes d'appels en transfert en leur assignant chacune un seuil, (2) un algorithme efficace de contrôle de l'admission des appels (CAC) basé sur un seuil, et (3) un algorithme d'adaptation de la largeur de bande (BAA) qui ajuste dynamiquement la largeur de bande d'appels multimédia en cours afin de minimiser le nombre d'appels recevant moins de largeur de bande que ce qu'ils ont demandés. Les résultats numériques montrent que notre cadre multimédia adaptatif est plus performant que les schèmes non adaptatifs existants sur le plan de la probabilité de perte d'appels en transferts et de l'utilisation effective.

Keywords: adaptive multimedia, bandwidth adaptation, call admission control, multiple classes, QoS provisioning, wireless cellular networks

I. Introduction

Wireless communications systems are becoming increasingly popular as they provide users the convenience of access to information and services anytime, anywhere and in any format. The upcoming wireless cellular infrastructures such as third generation (3G) and fourth generation (4G) are deemed to support new high-speed services. The expected services will include multimedia services that need real-time quality of service (QoS) guarantees. A wireless multimedia service enables the simultaneous transmission of voice, data, text and images through radio links by means of new wireless infrastructures. Examples of these services are mobile commerce, geographical and location information, Web services, cooperative group work, streaming media and entertainment, voice and gaming. Different wireless multimedia services have diverse bandwidth and QoS requirements, which need to be guaranteed by the wireless cellular networks. To achieve this goal, QoS provisioning in wireless multimedia networks is critical.

In wireless cellular networks, a mobile user's QoS requirements can be objectively expressed in terms of probabilistic connection-level QoS parameters related to connection establishment and management, such as new-call blocking probability (NCBP) and handoff call dropping probability (HCDP) [1]. The NCBP is the probability that a new call will be rejected—a measure of service connectivity. New-call blocking occurs when the entire bandwidth of the wireless system medium is busy upon a new call request. The HCDP is the probability that a handoff call will be rejected; it measures service continuity during handoffs. The procedure of moving from one cell to another, while the call is in progress, is called handoff. To fulfill handoff, the mobile requires that the base station in the cell that it moves into allocate the required bandwidth. If no bandwidth is available in the new cell, the handoff connection is dropped. This kind of dropping refers to blocking of ongoing connections due to the mobility of the users. Since call dropping of established connections is usually more annoying than rejection of a new connection request, it is widely believed that a wireless cellular network must give handoff connection requests a higher priority than is given to new connection requests.

Provisioning connection-level QoS in wireless cellular networks is complex due to the limited radio-link bandwidth, the highly fluctuating wireless environment and the user's mobility. The problem has

---

*Nidal Nasser and Hossam Hassanein are with the Telecommunications Research Laboratory, School of Computing, Queen's University, Kingston, Ontario K7L 3N6. E-mail: {nasser, hossam}@cs.queensu.ca become even more challenging as recent wireless cellular networks have been implemented based on small-sized cells (i.e., microcells or picocells [2]). These cells are intended to allow higher transmission capacity and thus to achieve better performance. However, they also increase the handoff rate and result in rapid changes in the network traffic conditions, making the assurance of QoS guarantees more difficult [3]. Therefore, the most important connection-level QoS issue that should be addressed is how to reduce/control handoff drops due to lack of available resources in the new cell, since mobile users should be able to continue their ongoing connections. Designing an efficient mechanism for management and sharing of bandwidth among different classes of traffic is another important issue that plays a major role in enhancing system performance.

Recently, in order to overcome the limitations of scarce, highly fluctuating link bandwidth in wireless multimedia networks, adaptive multimedia networking has been proposed (e.g., [4]–[5]). An adaptive multimedia paradigm can play an important role in mitigating the highly varying resource availability in wireless multimedia networks. Using the adaptive framework, the bandwidth of an ongoing call is variable and thus can be dynamically adjusted to adapt to various communication environments, especially in situations of overload. With the help of this adaptive framework and the aid of an efficient bandwidth allocation mechanism, the dropping probability of handoff calls is reduced to a negligible level in moderate traffic load. However, even though the HCDP drops to almost zero within the adaptive multimedia framework, QoS provisioning for multimedia in wireless cellular networks remains a challenge in overload situations.

A call admission control (CAC) algorithm is another key factor that enables efficient system resource utilization while ensuring that connection-level QoS requirements are satisfied. CAC is always performed when a mobile initiates communication in a new cell, either through a new call or a handoff. In this work, we introduce a simple, yet powerful, threshold-type CAC algorithm to provide acceptable QoS to users and to utilize the system bandwidth efficiently.

A. Related work and motivation

In recent years, QoS provisioning in wireless cellular networks has attracted significant attention. QoS provisioning is performed by two closely related algorithms: a call admission control algorithm and a bandwidth adaptation algorithm (BAA). Several CAC algorithms and BAAs have been proposed for wireless cellular networks using the adaptive framework [6]–[12]. In [6]–[9], it is assumed that all calls belong to a single class of adaptive multimedia traffic and, similarly to [5] and [13]–[14], receive varying bandwidth values from a discrete set of integer bandwidths. Chou et al. [6] derived an analytical model for wireless networks with adaptive bandwidth allocation and traffic-restriction CAC. In [7]–[8], new QoS parameters for adaptive multimedia in wireless networks are introduced. Kwon et al. [7] proposed the cell overload probability parameter, while Xiao et al. [8] proposed two QoS parameters: the degradation ratio (DR) and the degradation degree. Both schemes suggest a new CAC algorithm and BAA that satisfy the application QoS requirements in terms of the proposed parameters. Kwon et al. [9] identified the possible objectives of BAAs and, based on these, proposed two BAAs. The first algorithm considers the service-provider objectives of low cost and high revenue, while the second algorithm aims to provide good QoS and to maintain fairness from the standpoint of the service user.

Multiple classes of adaptive multimedia services in cellular wireless networks have rarely been introduced in the literature without consideration of the prioritization between a new call arrival and handoff arrival for each class of traffic [10]–[12]. This motivates us to provide a bandwidth allocation mechanism that takes into account the separation between incoming traffic for each class and prioritizes the handoff calls over the new calls. Xiao et al. in [10] and [11] proposed fair adaptive bandwidth algorithms for multiple classes of connections. Fairness among classes is achieved by partitioning the bandwidth according to arrival rates. The disadvantage of these algorithms is that they cause all ongoing connections to receive reduced bandwidth. In order to avoid this problem, we design a bandwidth adaptation algorithm that aims to minimize the number of ongoing degraded connections as well as to reduce the complexity of the bandwidth adaptation process where not all the connections are reduced. A prioritization in the process of bandwidth adaptation among multiple classes of multimedia services is presented in [12], where the bandwidth of calls with lower priority is preferably adapted. However, the authors of [12] assumed no handoff dropping, which makes their work impractical. This is not the case in our work; in order to make our contribution more realistic, a handoff call can be dropped if it does not satisfy the adaptation condition.

B. Contribution

Motivated by the above discussion, we present a novel adaptive multimedia framework for the next generation of wireless cellular networks at the connection level, where the bandwidth allocated to the ongoing calls can be dynamically adjusted. This framework supports multiple classes of adaptive multimedia services with diverse QoS requirements. The main objective of our work is to be able to prioritize between handoff calls and new calls for each class of traffic while minimizing the handoff call dropping probability and maximizing the effective bandwidth utilization. With this objective in mind, we design an adaptive multimedia framework that consists of the following related components: a threshold-based bandwidth allocation policy, a call admission control algorithm, and a bandwidth adaptation algorithm.

The rest of this paper is organized as follows: The adaptive multimedia model, the traffic model and the QoS parameters are described in Section II. The detailed adaptive multimedia framework with its main components is presented in Section III. Simulation results and performance comparisons are reported in Section IV. Finally, conclusions drawn from the paper and future work are discussed in Section V.

II. Multiclass adaptive multimedia system architecture

A. Adaptive multimedia model

We assume a fixed capacity in each cell. The fixed total capacity is B bandwidth units (in number of channels). Traffic arriving at the cell is partitioned into K separate classes based on bandwidth requirements. A multimedia call can dynamically change its bandwidth depending on the network load situation during its lifetime.

Similarly to [5], [14] and [15], we adopt a layered coding approach in which the bandwidth of a call can take a set of discrete values. Moreover, multiple classes of adaptive multimedia are taken into consideration. Each class i connection requires a discrete bandwidth value $b_{i,j}$, where $b_{i,j}$ belongs to the set $B_i = \{b_{i,1}, b_{i,2}, \ldots, b_{i,j}, \ldots, b_{i,K_i}\}$ for $i = 1, 2, \ldots, K$, and $b_{i,j} < b_{i,j+1}$ for $j = 1, 2, \ldots, K_i - 1$. Here, we express discrete bandwidth values in terms of basic bandwidth units. Thus, $b_{i,j}$ is an integer number of channels. The minimum and the maximum values that a class i connection can take are $b_{i,1}$ and b_{i,K_i}, respectively. Also, K_i denotes the number of possible values of bandwidth that can be allocated to a class i call. The requested bandwidth of a connection of class i is denoted as $b_{i,\text{request}}$, where $b_{i,\text{request}} \in B_i$. We assume that all connections in the same class use the same requested bandwidth. Hence, the requested bandwidth for each class is predetermined.

B. Traffic model

For traffic characterization, we assume a simple model from a cell's perspective. New call arrivals and handoff call arrivals of class i connections are assumed to follow a Poisson process with rates λ_{nc_i} and λ_{h_i}, respectively. The channel occupancy time is defined as the time during which a call occupies a channel. It is the minimum of the call holding time and the cell residence time. According to the classic traffic theory, the call holding time of class i connections is assumed to be exponentially distributed with mean $1/\mu_i$. The cell residence time is the amount of time that a mobile user stays in a cell before handoff. In a detailed study of mobility in wireless networks [16], it was found that this parameter follows a gamma distribution. However, approximating this distribution by the exponential distribution results in errors of less than 0.05% [17]. Considering the simplicity resulting from the assumption of exponentially distributed cell residence time, and noting the negligible nature of the resulting modelling errors, we assume that the cell residence time is exponentially distributed with the mean $1/h$. We also assume the cell residence time to be independent of the service class; hence, connections in any class follow the same cell residence time distribution. Note that the parameter h represents the call handoff rate. Recall that the channel occupancy time is the minimum of two exponentially distributed random variables: the call holding time and the cell residence time. Therefore, the channel occupancy time for new calls, as well as for handoff calls for class i, will be exponentially distributed with means $1/\mu_{\mathrm{nc}_i}$ and $1/\mu_{h_i}$, respectively, and is given by

$$f_i = (\mu_i + h)e^{-(\mu_i + h)t}, \tag{1}$$

where $\mu_{\mathrm{nc}_i} = \mu_{h_i} = \mu_i + h$.

C. QoS parameters

In this section, we describe the QoS parameters that will be used to evaluate the performance of our adaptive multimedia framework. In addition to the new-call blocking probability and the handoff call dropping probability, we introduce a degradation-ratio QoS parameter for adaptive multimedia to evaluate the performance of each individual traffic class. DR is defined in [8] for a single class of adaptive multimedia traffic. In this work, we redefine the DR for multiple classes of adaptive multimedia in order to effectively characterize the bandwidth degradation for different classes. Thus, DR_i represents the degradation ratio of class i and is equal to the average ratio of the number of degraded calls of class i to the number of ongoing calls of class i.

The state of the given cell at time t is defined by the vector $\mathbf{x}(t)$ as follows:

$$\mathbf{x}(t) = (x_1(t), x_2(t), \ldots, x_K(t)). \tag{2}$$

The non-negative integer $x_i(t)$ denotes the number of ongoing new and handoff class i connections at time t. Let $b_{i,\mathrm{request}}$ denote the predetermined requested bandwidth of an incoming call (new or handoff) of class i. We assume that any incoming call, either new or handoff, of class i uses the same $b_{i,\mathrm{request}}$ and that $b_{i,\mathrm{request}} > b_{i,1}$, where $b_{i,1}$ is the minimum bandwidth of class i. Let $b_{i,\mathrm{assigned}_j}$ denote the assigned bandwidth for call j, $1 \le j \le x_i(t)$, of class i at time t. Note that both $b_{i,\mathrm{request}}$ and $b_{i,\mathrm{assigned}_j}$ are in $B_i = \{b_{i,1}, b_{i,2}, \ldots, b_{i,j}, \ldots, b_{i,K_i}\}$. A call j of class i is called a “degraded call” if $b_{i,\mathrm{assigned}_j} < b_{i,\mathrm{request}}$.

Let $x_{d_i}(t)$ denote the number of degraded calls out of $x_i(t)$ calls of class i in a given cell at time t. Therefore, at time t, the instantaneous degradation ratio of class i ($\mathrm{IDR}_i(t)$) is defined as

$$\mathrm{IDR}_i(t) = \frac{x_{d_i}(t)}{x_i(t)}. \tag{3}$$

Then $\mathrm{DR}_i(t)$ is defined as the time average of $\mathrm{IDR}_i(t)$ and reflects the observed history of the cell bandwidth usage. The degradation ratio of class i (DR_i) is calculated every T time intervals as

$$\mathrm{DR}_i(\tau) = \frac{1}{T}\int_{\tau - T}^{\tau} \frac{x_{d_i}(t)}{x_i(t)}\, dt, \tag{4}$$

where τ is a time variable.

III. Adaptive multimedia framework

Our adaptive multimedia framework consists of three main components:

1. a threshold-based bandwidth allocation policy,
2. a call admission control algorithm, and
3. a bandwidth adaptation algorithm.

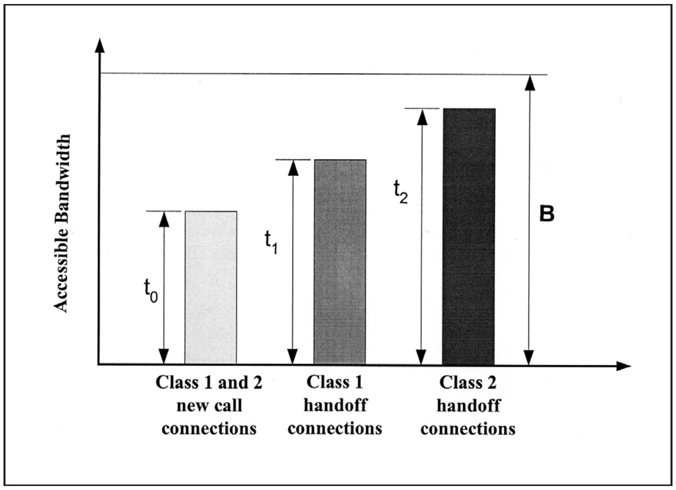

Figure 1: *Accessible bandwidth for a two-class system.*

The threshold-based bandwidth allocation policy prioritizes connections according to their QoS constraints by assigning maximum occupancy, i.e., a threshold, to each connection class. The CAC algorithm makes an admission/rejection decision based on the current load of the cell. In this work, we introduce a simple CAC algorithm, as will be explained below. The BAA performs two main procedures: reduction and expansion. The BAA for reduction applies to the case in which an incoming handoff call arrives in the given cell and the cell is overloaded, whereas the BAA for expansion may expand the bandwidth of calls having less than the predetermined request bandwidth (or more when there is an outgoing handoff call or a call completion in the given cell).

A. Threshold-based bandwidth allocation policy

Consider a cell that has a total capacity of B bandwidth units (in a number of channels). Traffic arriving at the cell is partitioned into K separate classes based on bandwidth requirements. The classes are indexed in an increasing order according to their bandwidth requirements, such that

$$b_{1,j} \le b_{2,j} \le \cdots \le b_{K,j}, \quad \text{for} \quad j = 1, 2, \ldots, K_i. \tag{5}$$

New calls and handoff calls are further segmented into separate subclasses, each representing connections with different QoS requirements.

Our bandwidth allocation policy is based mainly on the idea of reserving channels for aggregate handoff connections, thus giving them priority over new connections and providing them with lower handoff dropping probability. In addition, the policy prioritizes among different classes of handoff connections according to their QoS constraints by assigning a series of bandwidth thresholds $t_0, t_1, \ldots, t_K$, such that

$$t_0 \le \cdots \le t_i \le t_{i+1} \le \cdots \le t_K, \tag{6}$$

where t_0 denotes the maximum number of total bandwidth units that can be allocated to new connections, and t_i, $1 \le i \le K$, denotes the maximum number of total bandwidth units that can be allocated to class i handoff connections. It should be noted that if the different handoff connections were allowed to completely share the bandwidth ($t_1 = t_2 = \cdots = t_K$), then connections with lower bandwidth requirements would have a better chance of occupying the bandwidth than those with higher bandwidth requirements. Fig. 1 shows the accessible bandwidth regions for a two-class system.

B. Call admission control algorithm

The objectives of any CAC algorithm are to satisfy the QoS requirements and to utilize the system resources in an efficient manner [18]–

[20]. In this section, we provide a simple threshold-type CAC algorithm. The algorithm uses the threshold values that are assigned in the bandwidth allocation policy, described in the previous section, to make a decision as to whether to admit or reject an incoming connection request as follows.

Let m_i denote the number of class i new connections, and let n_i denote the number of class i handoff connections that are present in the system at the time of the call request. Also, let B_{nc_i} denote the total capacity allocated for class i new connections, and let B_{h_i} denote the total capacity allocated for class i handoff connections. Obviously,

$$B_{\text{nc}_i} = \sum_{j=1}^{m_i} b_{i,\text{assigned}_j} \tag{7}$$

and

$$B_{h_i} = \sum_{j=1}^{n_i} b_{i,\text{assigned}_j}. \tag{8}$$

The free available capacity, B_{Tot_A}, is given by

$$B_{\text{Tot}_A} = B - \sum_{i=1}^{K} (B_{\text{nc}_i} + B_{h_i}). \tag{9}$$

A class i new connection is accepted if

$$b_{i,\text{request}} + \sum_{i=1}^{K} B_{\text{nc}_i} \le t_0 \quad \text{and} \quad b_{i,\text{request}} \le B_{\text{Tot}_A}; \tag{10}$$

that is, a newly arriving call of class i is blocked if its requested bandwidth plus the current total bandwidth of ongoing new connections for all classes is greater than t_0, or if no more bandwidth is available in the given cell to accommodate the new connection. On the other hand, we always allow an incoming handoff connection of class i to be temporarily accepted regardless of its bandwidth requirements. If necessary, bandwidth adaptation is performed. Final acceptance of a handoff connection is constrained by the following condition: *The bandwidth adaptation algorithm, described in Section III.C, must be able to allocate enough bandwidth for the "accepted" handoff request.* If this condition is not satisfied, then the handoff connection will be dropped.

C. Bandwidth adaptation algorithm

A BAA increases or decreases assigned bandwidth of ongoing connections in the cell according to the network load situation. The BAA performs two main procedures: reduction and expansion. The reduction procedure is activated when an incoming handoff call arrives at an overloaded cell. On the other hand, the expansion procedure is activated when there is an outgoing handoff call or a call completion in the given cell.

With respect to different QoS objectives, several bandwidth adaptation algorithms [18]–[21] have been proposed and studied. In this work, our BAA seeks to satisfy the following objectives at any time instance:

1. minimize the number of degraded connections;
2. minimize the number of calls of class i with assigned bandwidth that is lower than requested.

We assume that the requested bandwidth of a class i incoming call, $b_{i,\text{request}}$, is predetermined.

We do not apply bandwidth adaptation to incoming new calls, as our main objective is to minimize the handoff dropping probability. Therefore, the bandwidth adaptation reduction will be performed only for the incoming class i handoff calls, $b_{i,\text{request}}$, according to the following cases:

Table 1
Bandwidth reduction procedure decisions

Case 1	Case 2	Action
False	False	The call is accepted
False	True	Execute single-class reduction algorithm
True	False	Execute multiple-classes reduction algorithm
True	True	Execute MCRA if SCRA is successfully executed

Table 2
Simulation parameters

Class index	Bandwidth set	$b_{i,\text{request}}$	$1/h$	$1/\mu_{b_i}$			
1	$\{1,3,5,7\}$	5	100	500			
2	$\{2,4,6,8\}$	4	100	700			
3	$\{8,16,24,32\}$	16	100	900			
Other parameters							
B	T_{sim}	T	t_0	t_1	t_2	t_3	New-call arrival rate (λ)
100	1000 s	4 s	50	100	100	100	$\lambda_{\text{nc}_1} = \lambda_{\text{nc}_2} = \lambda_{\text{nc}_3} = \lambda$

Case 1: No more bandwidth is available in the cell to accommodate this connection:

$$b_{i,\text{request}} + \left(\sum_{i=1}^{K} \sum_{j=1}^{m_i} b_{i,\text{assigned}_j} + \sum_{i=1}^{K} \sum_{j=1}^{n_i} b_{i,\text{assigned}_j} \right) > B. \tag{11}$$

Case 2: No more bandwidth is available under class i to accommodate this connection, i.e., the current total bandwidth of ongoing class i handoff connections is greater than t_i, $1 \le i \le K$:

$$b_{i,\text{request}} + \sum_{j=1}^{n_i} b_{i,\text{assigned}_j} > t_i \ \forall i. \tag{12}$$

Based on these two cases the system will decide which reduction algorithm should be invoked, as shown in Table 1. The main objective of both algorithms is to accommodate incoming handoff calls and to assign the maximum bandwidth (b_{i,K_i}) for such connections.

The selection of the algorithm will reduce the complexity of the reduction procedure since in the single-class reduction algorithm (SCRA) only the handoff connections of class i are affected in the reduction process, while in the multiple-classes reduction algorithm (MCRA) all handoff connections in the system are affected. The detailed descriptions of the SCRA and the MCRA are shown in Fig. 2. Note that B_{A_i} represents the free available bandwidth of class i, while B_{Tot_A} represents the free available bandwidth in the given cell.

As a call of class i leaves the cell, whether it is an outgoing handoff call or a call completion, the total available bandwidth of class i, B_{A_i}, increases. The system will invoke the expansion procedure to increase the bandwidth for one or more of the degraded class i calls, starting from class one and ending with class K. The detailed description of the expansion algorithm is also shown in Fig. 2, where b_{upgrade} denotes the bandwidth required to upgrade the bandwidth of the corresponding call and $b_{i,\text{current}}$ denotes the currently allocated bandwidth of the corresponding call.

IV. Simulation experiments

In order to evaluate the performance of our proposed adaptive multimedia framework, we developed a simulation model for the wire-

```
SCRA( int $b_{i,\mathrm{request}}$, int i) // i: class index
{
  if ($B_{A_i} \geq b_{i,\mathrm{request}}$)
  {
    allocateBandwidth($b_{i,j}, b_{i,\mathrm{request}}, b_{i,K_i}$);
    allocationIsDone = 1; // the call is accepted and allocated a bandwidth
  }

  else
  {
    reduceAll($b_{i,\mathrm{request}}$);
    if ($B_{A_i} \geq b_{i,\mathrm{request}}$)
    {
      allocateBandwidth($b_{i,j}, b_{i,\mathrm{request}}, b_{i,K_i}$);
      allocationIsDone = 1;
    }
    else
         {
           reduceAll($b_{i,1}$)
           if($B_{A_i} \geq b_{i,1}$)
           {
             allocateBandwidth($b_{i,j}, b_{i,1}, b_{i,\mathrm{request}}$);
             allocationIsDone = 1;
           }
           else allocationIsDone = 0; // the call is dropped
         }
  }
return allocationIsDone;
}// end of SCRA

allocateBandwidth(int assignedValue, int minBW, int maxBW)
{
  allocate assignedValue to the incoming call such that
  (assignedValue $\leq B_{A_i}$) and (minBW $\leq$ assignedValue $\leq$ maxBW)
}// end of allocateBandwidth function

reduceAll(int minLevel)
{
  order the calls decreasingly by bandwidth;
  while($B_{A_i} < b_{i,\mathrm{request}}$ and exist a call's bandwidth > minLevel)
  {
    reduce the calls with more than minLevel to minLevel
    starting with the largest bandwidth;
    add the extra bandwidth to $B_{A_i}$;
  }
}// end of reduceAll function

MCRA( )
{
  classIndex = 1;

  while (classIndex $\leq K$ and $B_{\mathrm{Tot}_A} < b_{i,\mathrm{request}}$)
  {
    tryNextClass = SCRA($b_{i,\mathrm{request}}$,classIndex);
    if (tryNextClass == 0) // need to degrade calls of other classes
    {
      $B_{\mathrm{Tot}_A} = B_{\mathrm{Tot}_A} + B_{A_i}$;
      classIndex ++;
    }
    else break; // the call is accepted
  }

if(classIndex $> K$) drop the call;
} // end of MCRA

Expansion Algorithm()
{
  for (classIndex = 1; classIndex $\leq K$; classIndex ++)
  {
    while(exist degraded calls of class classIndex to upgrade)
    {
      $i$ = classIndex;
      order the most degraded calls increasingly by bandwidth;
      starting with the most degraded call from above step;
        $b_{\mathrm{upgrade}} = b_{i,\mathrm{request}} - b_{i,\mathrm{current}}$;
        if ($B_{A_i} \geq b_{\mathrm{upgrade}}$)
          allocateBandwidth($b_{i,j}, b_{i,\mathrm{request}}, b_{i,K_i}$);
    }// end of while
  }// end of for
}// end of expansion algorithm
```

Figure 2: *Bandwidth reduction algorithms and expansion algorithm.*

less cellular network. The experimental results here are based on the simulation of a network consisting of seven cells arranged in a circle with identical mobility and traffic conditions, a common topology used in many previous studies [22]–[23]. We simulated an environment in which there are three classes of adaptive multimedia services. The simulation parameters, consisting of the discrete set of bandwidth values for each class and the traffic parameters described in Section II, are shown in Table 2. B is the total bandwidth capacity in each cell, T_{sim} is the simulation time, and T is the measurement time period. We assume that the handoff call arrival rate of class i is proportional to the new-call arrival rate of class i according to $\lambda_{h_i} = (h/\mu_i)\lambda_{\mathrm{nc}_i}$ [24].

The performance measures are connection-level QoS parameters, the new-call blocking probability and the handoff call dropping probability, the degradation ratio, and effective utilization. We define the effective utilization as the ratio of the bandwidth used by completely serviced calls to the total bandwidth capacity [6]. These performance measures are plotted as a function of the offered load (new-call arrival rate). We compare our adaptive multimedia framework with the non-adaptive case. The bandwidth of a call of class i $(i = 1, 2, 3)$ in the non-adaptive multimedia networking paradigm is fixed throughout the simulation lifetime and is equal to the requested bandwidth value of class i $(i = 1, 2, 3)$. Before we proceed with evaluating the performance of our adaptive multimedia framework, we first validate the threshold-bandwidth allocation policy in the case of non-adaptive multimedia paradigm networking.

Fig. 3 shows the effect of varying the new-call arrival rate on the NCBP and the HCDP. It can be seen that the NCBP and the HCDP of the three evaluated classes increase as the new-call arrival rate increases. However, the HCDP is always lower than the NCBP as a result of the 50 bandwidth units $(B - t_0)$ reserved exclusively for the handoff connections. Moreover, the HCDP increases as the class index increases. This behaviour is due to the complete sharing between all three classes' handoff connections $(t_1 = t_2 = t_3 = 100)$, which results in a higher dropping probability for the higher-bandwidth class. This demonstrates the importance of giving priority to higher-bandwidth connections by adjusting the values of t_1, t_2, t_3 in order to give all classes their required QoS.

In Fig. 4, we demonstrate the effect of varying the new-call threshold, t_0, on the NCBP and the HCDP. At low new-call threshold values, the NCBP for the three evaluated classes is high, while the HCDP is low. As the threshold value, t_0, increases, the NCBP for the three classes decreases while the HCDP increases until they converge to almost the same value. This is because the new calls are given more access to the available bandwidth. While the HCDP increases as a result of the higher degree of sharing between the new and the hand-

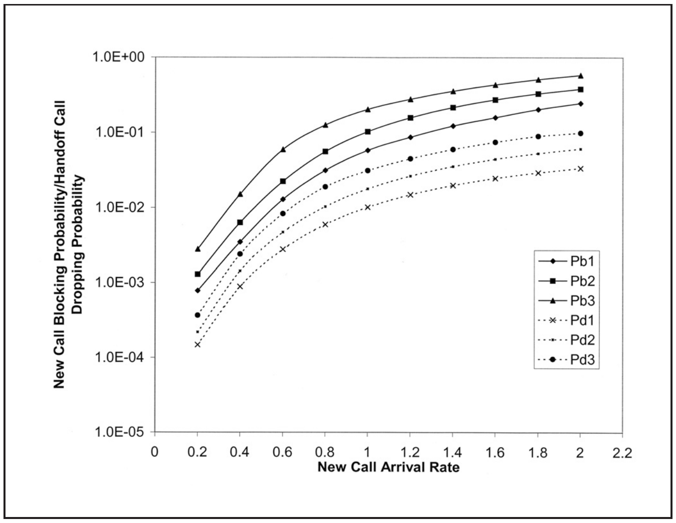

Figure 3: *Effect of varying the new-call arrival rate on NCBP and HCDP for non-adaptive multimedia.*

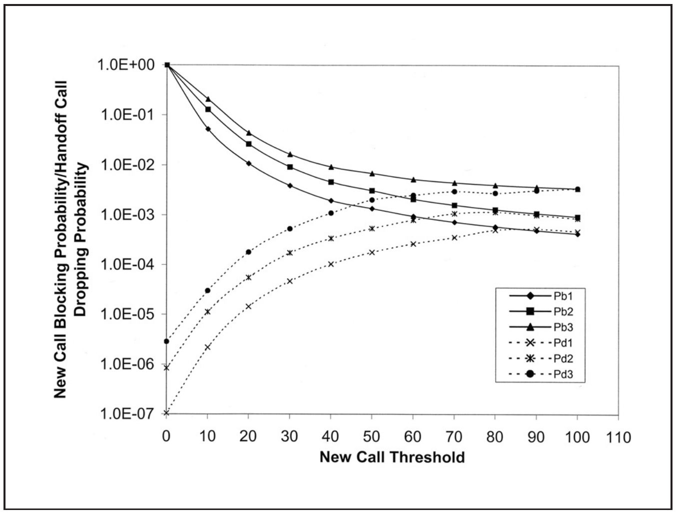

Figure 4: *Effect of varying the new-call threshold on NCBP and HCDP for non-adaptive multimedia.*

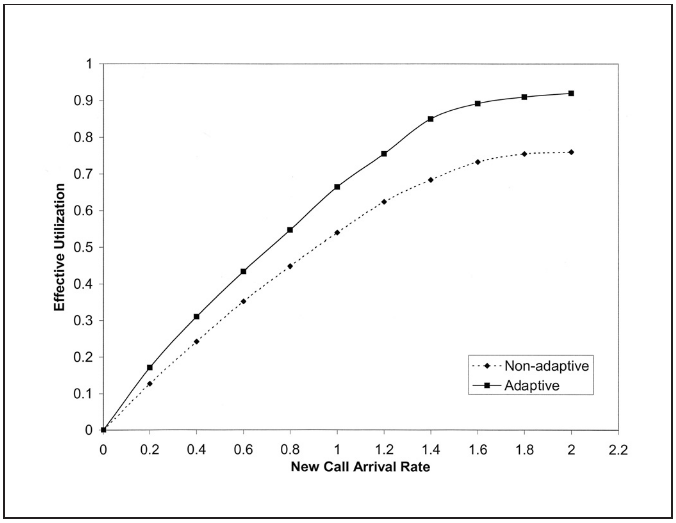

Figure 5: *Effect of varying the new-call arrival rate on effective utilization for non-adaptive/adaptive multimedia.*

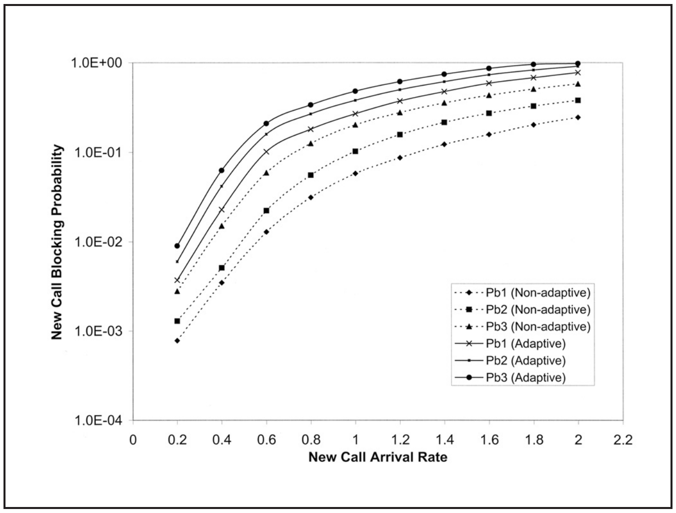

Figure 6: *Effect of varying the new-call arrival rate on NCBP for non-adaptive/adaptive multimedia.*

off calls, it is always lower than the corresponding new-call blocking probability.

In the above-mentioned figures, the results illustrate the ability of the threshold-bandwidth allocation policy to accurately capture the behaviour of the system and to achieve differentiation among different traffic classes with respect to their QoS constraints.

Fig. 5 shows the effective utilization versus the offered load (new-call arrival rate) for an adaptive multimedia framework as opposed to a non-adaptive multimedia framework. Clearly, an effectively utilized adaptive multimedia framework outperforms the non-adaptive multimedia framework. As the offered load increases, the advantage is more evident. This follows, since a highly restrictive threshold-bandwidth allocation policy (non-adaptive case) can result in high call dropping before completion of the calls, whereas for the adaptive multimedia framework, use of the bandwidth adaptation algorithm allows the system to offer services whenever there is sufficient amount of bandwidth by intelligently adjusting bandwidth allocation to achieve a near-zero HCDP. Therefore, more calls are able to complete their services, and as a result better effective utilization is obtained.

The effect of varying the new-call arrival rate on the connection-level QoS parameter NCBP for both adaptive multimedia and non-adaptive multimedia frameworks is illustrated in Fig. 6. The NCBP for both frameworks increases as the new-call arrival rate increases. However, the NCBP in the case of the non-adaptive multimedia framework is lower than that of the adaptive multimedia framework. This is because there are already many calls in each cell in the adaptive multimedia paradigm, as there is almost no handoff dropping (HCDP is near zero). Note that among the evaluated classes, class one has the lowest NCBP values, while class three has the highest NCBP values. This is due to the fact that the bandwidth requirement decreases as the class index decreases. Thus, there is a better chance to accept calls of the smallest indexed class.

Fig. 7 illustrates the performance of the proposed adaptive multimedia framework and non-adaptive multimedia framework in terms of the HCDP. It is apparent that by applying the CAC algorithm and BAA in the manner described above, the HCDP of the adaptive multimedia framework is minimized (lower than $0.000\,48$) and, therefore, it surpasses the non-adaptive multimedia framework. However, as the traffic load increases, the HCDP increases. The reason is that even the CAC algorithm will accept the handoff calls all the time, but when the traffic load becomes larger, the threshold-bandwidth allocation policy can still reject some handoff arrivals due to the restrictive threshold occupancy. In the non-adaptive multimedia framework, an incoming handoff connection of class i is immediately dropped if the current total bandwidth of ongoing class i connections is greater than t_i, $1 \leq i \leq K$, or no more bandwidth is available in the cell to ac-

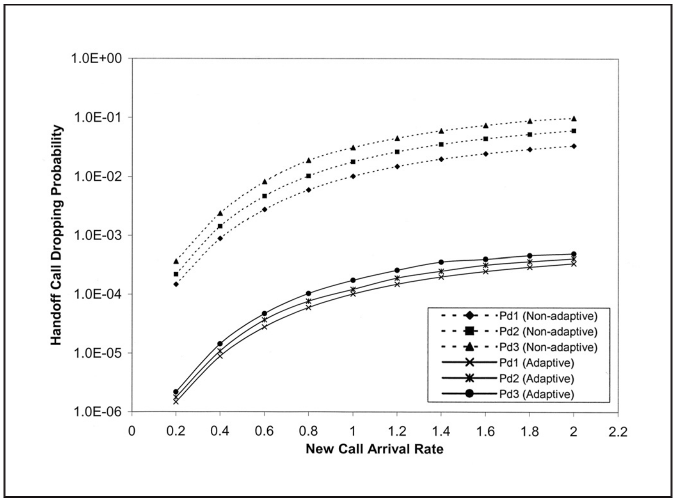

Figure 7: *Effect of varying the new-call arrival rate on HCDP for non-adaptive/adaptive multimedia.*

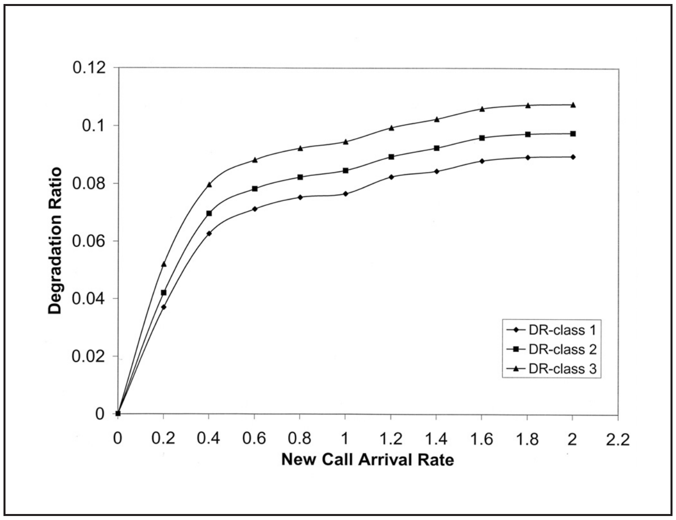

Figure 8: *Effect of varying the new-call arrival rate on degradation ratio for adaptive multimedia.*

commodate this connection. This results in an increasing HCDP for class i as the new-call arrival rate increases. Furthermore, the results in Fig. 7 show that the HCDP increases as the class index increases for both cases. This behaviour is due to the complete sharing between the handoff connections ($t_1 = t_2 = t_3 = 100$). Hence, prioritization among different classes is achieved.

Fig. 8 shows the degradation ratio versus the offered load (new-call arrival rate) for all classes. As expected, DR increases for all three evaluated classes as the call arrival rate increases. However, the degradation ratio of the different classes can be clearly distinguished, where class one has the lowest DR values while class three has the highest DR values, i.e., $\mathrm{DR}_1 < \mathrm{DR}_2 < \mathrm{DR}_3$. This behaviour can be explained as follows: the requested bandwidth of class i increases as the class index increases, and since all three classes share the same bandwidth threshold, $t_1 = t_2 = t_3 = 100$, there are more calls to share the degradation under class i as the new-call arrival rate increases.

V. Conclusions and future work

An adaptive multimedia paradigm can play an important role in mitigating the highly fluctuating availability of bandwidth resources in wireless cellular networks. In this work, a novel adaptive multimedia framework for wireless cellular networks has been presented. The proposed framework considers multiple classes of adaptive multimedia services with different QoS requirements. Three related components comprise the main building blocks of the framework: (1) a threshold-based bandwidth allocation policy, (2) a threshold-based call admission control algorithm, and (3) a bandwidth adaptation algorithm.

Simulation results show that the handoff call dropping probability is always lower than the new-call blocking probability. Moreover, the handoff connections with lower bandwidth requirements always have lower dropping probability than those with higher bandwidth requirements. The performance of the adaptive multimedia framework has been compared to that of the existing non-adaptive multimedia framework. The results show that although the new-call blocking probability of the adaptive multimedia framework is greater than that of the non-adaptive multimedia framework, the overall performance of the adaptive multimedia networking is very attractive in that the handoff call dropping probability is near zero (negligible) and the effective utilization increases as the offered traffic load increases. Observing the reported results, we conclude that this work is a very powerful and novel contribution to wireless cellular network design, as it can be used to support high-bandwidth multimedia services in a manner that enhances the overall system performance. In addition, our framework is cell-oriented, meaning that all its components are implemented on a cell-by-cell basis. It thus has an extremely low complexity, making it practical for real wireless cellular networks.

This work can be extended in several ways. One straightforward extension is to develop an adaptive threshold-based bandwidth allocation policy, which dynamically computes and changes the threshold values based on the traffic and mobility parameters.

References

[1] S.S. Rappaport, "The multiple-call handoff problem in high-capacity cellular communications systems," *IEEE Trans. Veh. Technol.*, vol. 40, no. 3, Aug. 1991, pp. 546–557.

[2] J. Padgett, C. Gunther, and T. Hattori, "Overview of wireless personal communications," *IEEE Commun. Mag.*, vol. 33, no. 1, Jan. 1995, pp. 28–41.

[3] B. Jabbari, G. Colombo, A. Nakajima, and J. Kulkarni, "Network issues for wireless communications," *IEEE Commun. Mag.*, vol. 33, no. 1, Jan. 1995, pp. 88–99.

[4] A. Alwan et al., "Adaptive mobile multimedia networks," *IEEE Pers. Commun.*, vol. 3, no. 2, Apr. 1996, pp. 34–51.

[5] M. Naghshineh and M. Willbeek-LeMair, "End-to-end QoS provisioning in multimedia wireless/mobile networks using an adaptive framework," *IEEE Commun. Mag.*, vol. 35, no. 11, Nov. 1997, pp. 72–81.

[6] C. Chou and K. Shin, "Analysis of combined adaptive bandwidth allocation and admission control in wireless networks," in *Proc. IEEE INFOCOM*, New York, June 2002, pp. 676–684.

[7] T. Kwon, Y. Choi, C. Bisdikian, and M. Naghshineh, "QoS provisioning in wireless/mobile multimedia networks using an adaptive framework," *Kluwer Wireless Networks*, vol. 9, no. 1, 2003, pp. 51–59.

[8] Y. Xiao, C.L.P. Chen, and B. Wang, "Bandwidth degradation QoS provisioning for adaptive multimedia in wireless/mobile networks," *Comput. Commun.*, vol. 25, no. 13, Aug. 2002, pp. 1153–1161.

[9] T. Kwon, Y. Choi, and S. Das, "Bandwidth adaptation algorithms for adaptive multimedia services in mobile cellular networks," *Kluwer Wireless Pers. Commun.*, vol. 22, no. 3, Sept. 2002, pp. 337–357.

[10] Y. Xiao and C. Chen, "Improving degradation and fairness for mobile adaptive multimedia wireless networks," in *Proc. 10th IEEE Int. Conf. Comput. Commun. and Networks (ICCCN)*, Phoenix, Ariz., Oct. 2001, pp. 598–601.

[11] Y. Xiao, C. Chen, and Y. Wang, "Fair bandwidth allocation for multi-class of adaptive multimedia services in wireless/mobile networks," in *Proc. IEEE Veh. Technol. Conf. (VTC '01)*, Rhodes, Greece, May 2001, pp. 2081–2085.

[12] S. Kim, T. Kwon, and Y. Choi, "Call admission control for prioritized adaptive multimedia services in wireless/mobile networks," in *Proc. IEEE Veh. Technol. Conf. (VTC '00)*, Tokyo, Japan, May 2000, pp. 1536–1540.

[13] V. Bharghavan, K. Lee, S. Lu, S. Ha, D. Dwyer, and J. Li, "The TIMELY adaptive resource management architecture," *IEEE Pers. Commun.*, vol. 5, no. 8, Aug. 1998, pp. 20–31.

[14] A. Talukdar, B. Badrinath, and A. Acharya, "Rate adaptation schemes in networks with mobile hosts," in *Proc. 4th ACM/IEEE MobiCom Conf.*, Dallas, Tex., Oct. 1998, pp. 169–180.

[15] T. Kwon, Y. Choi, C. Bisdikian, and M. Naghshineh, "Call admission control for adaptive multimedia in wireless/mobile networks," in *Proc. ACM Workshop Wireless Mobile Multimedia (WOWMOM)*, Dallas, Tex., Oct. 1998, pp. 111–116.

[16] M. Zonoozi and P. Dassanayake, “User mobility modeling and characterization of mobility pattern,” *IEEE J. Select. Areas Commun.*, vol. 15, no. 7, Sept. 1997, pp. 1239–1252.

[17] K. Yeung and S. Nanda, “Channel management in microcell/macrocell cellular radio systems,” *IEEE Trans. Veh. Technol.*, vol. 45, no. 4, Nov. 1996, pp. 601–612.

[18] Y. Xiao, C. Chen, and Y. Wang, “An optimal distributed call admission control for adaptive multimedia in wireless/mobile networks,” in *Proc. 8th Int. Symp. Modeling, Analysis and Simulation of Comput. and Telecommun. Syst. (MASCOTS)*, San Francisco, Calif., Aug. 2000, pp. 477–482.

[19] Y. Xiao, C. Chen, and Y. Wang, “Quality of service provisioning framework for multimedia traffic in wireless/mobile networks,” in *Proc. 9th IEEE Int. Conf. Comput. Commun. and Networks (ICCCN)*, Las Vegas, Nev., Oct. 2000, pp. 644–648.

[20] ———, “Optimal admission control for multi-class of wireless adaptive multimedia services,” *IEICE Trans. Commun.*, vol. E84-B, no. 4, Apr. 2001, pp. 795–804.

[21] T. Kwon, S.l Das, I. Park, and Y. Choi, “Bandwidth adaptation algorithms with multi-objectives for adaptive multimedia services in wireless/mobile networks,” in *Proc. ACM Workshop Wireless Mobile Multimedia (WOWMOM)*, Seattle, Wash., Aug. 1999, pp. 51–58.

[22] C. Ho and C. Lea, “Improving call admission policies in wireless networks,” *ACM J. Wireless Networks*, vol. 5, no. 4, July 1999, pp. 257–265.

[23] M. Sidi and D. Starobinski, “New call blocking versus handoff blocking in cellular networks,” *ACM J. Wireless Networks*, vol. 3, no. 1, Mar. 1997, pp. 15–27.

[24] N. Nasser and H. Hassanein, “Multi-class bandwidth allocation policy for 3G wireless networks,” in *Proc. IEEE Conf. Local Comput. Networks (LCN)*, Bonn, Germany, Oct. 2003, pp. 203–209.