Adaptive bandwidth allocation and admission control for wireless integrated service networks with flexible QoS

Vittoria de Nitto Personè and Elisa Campagna
Department of Computer Science and Systems
Faculty of Engineering
University of Roma Tor Vergata
Via del Politecnico, 1
00133 Roma, Italy
+39 06 72597383
denitto@info.uniroma2.it

ABSTRACT

An analytical model of performance of a bandwidth allocation and admission control scheme in wireless mobile integrated services network is proposed. Due to the multimedia services and flexible QoS characteristics of the considered network, the proposed model includes two different classes of requests with priority, where each class allows different levels of QoS requirements. The proposed scheme considers an upgrade and degrade (u-d) mechanism, with the aim of keeping the requests at the highest possible level so as to give high quality whilst maintaining the performance of each class within the negotiated QoS. Instead of focusing only on the traditional metrics, we propose two important new metrics to evaluate the cost of the u-d mechanism and the QoS level provided. A comparison with other schemes in the literature concludes that our proposed scheme is mandatory if it is essential to respect the priority among classes and shows good performance from both user and service provider perspectives.

Categories and Subject Descriptors

C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design – *Wireless communication*; C.2.3 [**Computer-Communication Networks**]: Network Operations – *Network management, network monitoring*; C.4 [**Performance of Systems**]: *Modeling techniques*

General Terms

Management, Measurement, Performance

Keywords

Call admission control, handoff, bandwidth allocation, flexible QoS

1. INTRODUCTION

The rapid growth of multimedia service networks has forced us to consider flexibility as a key issue [1-22]. Managing multimedia streams with different QoS requirements on the one hand and minimizing resource wastage on the other, has renewed the importance of the resource allocation problem in the new context. This could also become more critical with the advent of service networks where service providers have SLAs to guarantee and the possibility of incurring penalties [16, 23].

In this scenario, analytical modeling can be viewed as a powerful tool to investigate different resource allocation algorithms to highlight future directions. It is well known that not all requests have the same "importance" to the system, so priority classes have to be included in the model.

Originally, the concept of adaptive multimedia networking was introduced in the wired network. It is well known that compared to wired networks, the fluctuation in resource availability in wireless/mobile networks is much more severe. Therefore the multimedia wireless mobile network is one of the most challenging application fields of the "adaptive paradigm". Bandwidth is the most important resource and an adaptive allocation is on demand in order to instantly adapt the network to changing conditions within the system while maintaining the negotiated QoS.

Over the past ten years and more the literature has been full of contributions with regards to flexible resource allocation, for both wireless mobile networks and fixed ones. See two recent surveys and references therein [24, 25].

Many papers deal with the single or, at most, two classes case [2, 3, 13, 14, 15, 16, 21], some papers deal with the flexible quality case [1, 2, 3, 4, 6, 11, 14, 15, 16], and most papers propose a simulation model [1, 4, 6, 11]. Very few papers deal with the multimedia and flexible quality case with an analytical modeling approach, and this is the direction of our contribution. As mentioned above, the analytical approach is important to manage the flexibility of the model itself and also to keep time complexity to an acceptable range.

In [19] the author considers a cellular network with three classes of calls. The author defines an analytical model of a threshold strategy. The "adaptive" attribute of the scheme is referred to the fractional guard channel and not to different QoS levels.

SIMUTools/QoSim 2009 March 2-6, Rome, Italy

In [20] the authors consider three classes of calls, real-time, non-real-time and best-effort traffic, with an adaptive channel allocation scheme where the different levels are independent of the class. The authors present an analytical approach that combines physical, radio link and network layer parameters.

In [22] the authors consider multiservice mobile wireless networks in which the bandwidth allocation on both uplink and downlink is adjustable in individual cells. They consider a multicell model to control also the performance of the neighboring cells. The authors propose a bandwidth reallocation scheme and use an analytical model to derive the admission thresholds. However, they do not consider adaptive bandwidth allocation. They conduct simulation experiments to validate the proposed scheme.

In [21] the authors analyze two CAC strategies for wireless mobile integrated services networks with two classes of requests. They propose an analytical model, but they do not include adaptive QoS levels.

In this paper we propose an adaptive bandwidth allocation and admission control scheme for wireless integrated service networks called MATS. As introduced above, the contribution of our proposed scheme in respect to the literature is of combining the multiclass characteristic with the adaptive bandwidth allocation. We define an analytical model to investigate the efficiency of the proposed scheme. To deal with the integrated services characteristic of the considered networks, the scheme allows the inclusion of different classes and each request of a given class can be served according to different service levels [6]. In this paper we consider two classes, the extension to any number of classes will be discussed in a forthcoming paper. For each class a reservation scheme is used to give preferential treatment to handoff calls in respect to new calls by means of a threshold defined for that class.

The proposed scheme assumes that each request will be treated at the higher level whenever possible. When an arrival occurs during a congestion period, a degradation mechanism is initiated to free bandwidth to admit the new arrival. When a departure occurs, an upgrade mechanism is initiated to increase the actual level of quality if possible. Beyond the traditional metrics such as the blocking and dropping probabilities for each class, we propose two important new metrics to evaluate the cost of the u-d mechanism and the QoS level provided.

The paper is organized as follows. The system model and the proposed MATS scheme are presented in the following section. The analytical model and the performance indices are defined in section 3. The main motivation of our work and the related papers are presented in section 4. Section 5 presents the comparisons and results and finally section 6 concludes the paper.

2. SYSTEM MODEL AND ASSUMPTIONS

A wireless network with a cellular infrastructure that carries two classes of traffic, for example voice and data traffic, or two classes of real-time traffic with different QoS requirements is considered. We assume the system uses Fixed Channel Allocation (FCA), which means each cell has a fixed amount of capacity that is equal to B basic bandwidth units. No matter which multiple access technology (FDMA, TDMA, or CDMA) is used, we could interpret system capacity in terms of bandwidth.

For simplicity, a homogenous cellular network is considered. In other words, all cells are assumed to be statistically identical so that we can focus on one particular cell. In the following we use call/request synonymously.

The infrastructure can support mobile users running multimedia services that demand a wide range of bandwidth allocations. The bandwidth of a multimedia call can be dynamically adjusted depending on the network load situation during the call's lifetime. Moreover, since the mobile user can freely roam within a network's coverage area, he/she may undergo a large number of handoff events during a typical session. In such a situation, the "adaptive" characteristic of bandwidth allocation allows the call to not be dropped but it will suffer bandwidth degradation.

User's QoS requirements can be quantitatively expressed in terms of probabilistic connection-level QoS parameters [26, 27] related to connection establishment and management: blocking probability P_b of a new call and dropping probability P_h of an handoff call. While minimizing these QoS parameters is very desirable from the user's point of view, this often comes at the expense of the resource utilization, which is extremely undesirable from the service provider's perspective. This proves the importance of providing a balance between the user's connection-level QoS satisfaction and system utilization.

Two major components in multimedia wireless networks (MWN) contribute in solving the above problems. The first is a Call Admission Control (CAC) algorithm. CAC is one of the most important components of the resource management affecting the bandwidth utilization and QoS guarantees provided to users. It is performed at the connection-level whenever a mobile initiates communication in a new cell, either through a new or handoff call. The CAC algorithm accepts or rejects an arriving request according to the amount of available resources versus call QoS requirements. The second component is the Adaptive Bandwidth Allocation (ABA) algorithm which is related to bandwidth management of ongoing calls in the system. In MWNs it is possible to overcome cell overload situations by dynamically adjusting the bandwidth of individual ongoing calls. In this process "fairness" amongst users is an important issue.

In the following section we define the proposed CAC-ABA algorithm for the considered mobile MWN.

2.1 The Proposed CAC Scheme

As introduced in the previous section, a single cell of a MWN with two classes of requests is considered. Class 2 is assumed to have priority over class 1.

The Base Station Controller (BSC) manages a fixed amount B of bandwidth in the cell. To give priority to class 2, the BSC adopts a bandwidth partition: $B_1<B$ basic bandwidth units (bbu) can be used for class 1 requests, while all available bandwidth $B_2=B$ can be used for class 2 requests.

It is well known that blocking a new call is tolerated more than dropping a handoff call. This is the motivation of the traditional Guard Channel (GC) scheme which uses thresholds to reserve bandwidth to the handoff calls.

Let us define $T_1<B_1$ the threshold for the class 1 new calls, and $T_2<B$ the threshold for the class 2 new calls. As a consequence, the class 1 handoff calls have B_1-T_1 bbu exclusively reserved in respect to the class 1 new calls. Similarly, the class 2 handoff calls have B-T_2 bbu exclusively reserved in respect to the all other calls, according to the high priority characteristic of class 2. Without loss of generality, in this paper we assume $T_1<T_2$ and $T_2=B_1$. The figure 1 illustrates the threshold reservation scheme.

Each class has flexible QoS requirements. Let us define $L_i=\{l_{1,i}, l_{2,i},\ldots l_{\max_i,i}\}$ the QoS levels for a call of class i, i=1, 2. We assume that, for class i, $l_{1,i}$ is the minimum bbu to maintain the connection, and $l_{\max_i,i}$ is the amount of bbu to give the connection

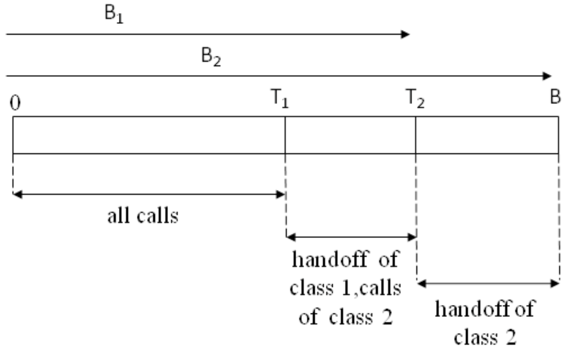


Figure 1. The threshold reservation scheme

the highest QoS level. The proposed Multimedia Adaptive Threshold Strategy (MATS) accepts new requests at the maximum possible level. So, if the bandwidth is available, $l_{\max_i,i}$ bbu are allocated to an incoming call (new or handoff) of class i, otherwise the available amount l_i^* is allocated, with $l_i^* = \max\{l \mid l \in L_i, l < l_{\max_i,i}\}$.

The threshold mechanism of the GC scheme yields a reduction in the dropping probabilities but this is obtained with an increase in the blocking probabilities. To overcome this problem, we propose an upgrade-degrade (u-d) mechanism which allows the acceptance of a class i incoming call at its minimum QoS level $l_{1,i}$ by degrading the quality level of the ongoing calls in the cell. The degradation mechanism is performed fairly among all the ongoing calls and by taking into account the priority characteristic. The description of the u-d mechanism is the object of the following section.

Figure 2 illustrates the proposed MATS scheme with reference to a call of class i. Note that the algorithm depends on the class and the type, new or handoff, of the incoming call.

Let us consider a new call arrival of class i. At first the BSC check if there are T_i or more bbu occupied. In this case, the second threshold B_i is checked. If it is entirely occupied, the system is in overload for that class and the new call is rejected. Otherwise, the BSC tries the degradation of the quality level of the ongoing calls to admit the new one within the limit threshold T_i. If the degradation is successful the new call is admitted at the minimum level $l_{1,i}$, otherwise it is refused.

If there are less than T_i bbu occupied, the BSC tries to allocate the call at the maximum possible level l_i^*, as defined above. If the free bbu are not enough to allocate the call even at its minimum level $l_{1,i}$, the degradation mechanism is activated and if it is successful the new call is admitted at the minimum level $l_{1,i}$ and within T_i, otherwise it is refused.

If the arrival call is a handoff, only the threshold B_i is checked. If the free bbu are not enough to admit the call even at its minimum level $l_{1,i}$, the degradation mechanism is activated. The scheme thus proceeds as for the new call case, but using the threshold B_i.

2.2 The upgrade-degrade mechanism

When a request for class i (new or handoff) arrives into the system and the available bandwidth B_i for that class is occupied (the system is in overload for that class), an u-d mechanism is activated to try to accept the request by degrading the QoS level of the ongoing calls, and thereby freeing bbu.

The mechanism starts to degrade the calls of the minimum priority class, and only if this is not enough does the system degrade calls of class i itself, with respect to the hierarchy among classes. Let us consider the two classes case and, for the sake of simplicity, where class 2 shows a very stringent QoS level, it is only one QoS level $l_{1,2} = l_{\max_2,2}$. Assume that a class 1 request arrives. The *degrade mechanism* is performed in two steps:

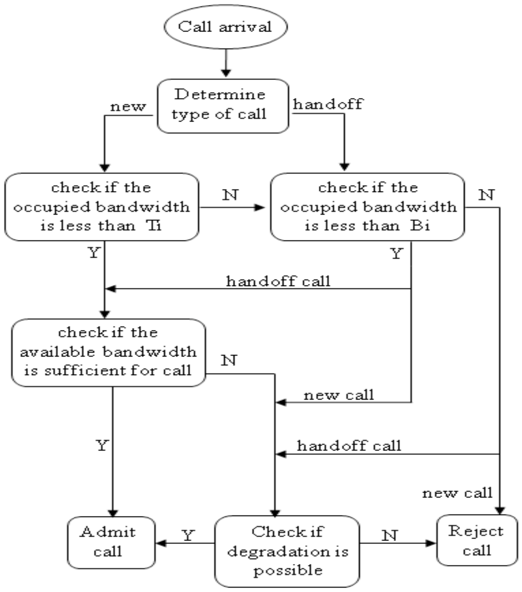


Figure 2. The MATS scheme for a class i arrival

1d. A class 1 ongoing call at the maximum QoS level $l_{j,1} > l_{1,1}$ is selected. For fairness purposes the call is selected in a random way. The QoS level is degraded to $l_{(j-1),1}$.

2d. If the freed bbu are enough to allocate the incoming call and the thresholds T_1 and B_1 are preserved (for the new and handoff incoming calls respectively), the algorithm is stopped. Otherwise, step 1 is repeated on the other class 1 ongoing calls until enough bbu are freed.

Note that if all class 1 ongoing calls are at the minimum level, the degradation is not possible and the incoming call is lost (blocked or dropped for new or handoff respectively).

When a class 1 departure occurs, an upgrade mechanism is activated. The *upgrade mechanism* is performed in two steps:

1u. A class 1 ongoing call at the minimum QoS level $l_{j,1} < l_{\max_1,1}$ is selected. For fairness purposes the call is selected in a random way. The QoS level is upgraded to $l_{(j+1),1}$ whenever the freed bbu are enough and the threshold B_1 is preserved. Otherwise the upgrade is not possible and the mechanism is stopped.

2u. If the bbu freed from the departure are not completely utilized, step 1 is repeated until all the freed bbu are allocated or the $l_{\max_1,1}$ level is reached for all the ongoing class 1 calls.

Note that from when a call is accepted into the system and so by becoming an ongoing call, the distinction between new or handoff is meaningless. As a consequence, in step 1 of the upgrade mechanism, we use only the threshold B_1.

It is worth noting that when class 2 has more than one QoS level, a class 2 ongoing call could also be degraded and upgraded according to similar fairness and priority properties considered above.

3. ANALYTICAL MODEL

In this section the analytical model we use to evaluate the proposed MATS scheme is defined.

The arrivals for each class are assumed to be Poisson distributed with mean rates $\lambda_{n,i}$ and $\lambda_{h,i}$ for new and handoff calls respectively. The service time for each class i is assumed to be exponentially distributed with a mean rate μ_i. Note that the service time represents both the call holding time and the cell residence time. For the memoryless property, if both the parameters are exponentially distributed, it is easy to extend the model to include two different mean rates.

Under these assumptions, the model is a continuous Markov process with a finite state space.

The system state space E is defined as follows:

$$E=\{\ \mathbf{s}=\{(n_1, \ldots, n_{\max_1}), (m_1, \ldots, m_{\max_2})\} \mid 1\le n_i \le B_1, 1\le m_i \le B\}$$

In a generic system state **s**, the subset $(n_1, \ldots, n_{\max_1})$ represents the number of class 1 ongoing calls for each QoS level, that is n_j being the number of class 1 ongoing calls at QoS level $l_{j,1}$ with $l_{j,1} \in L_1=\{l_{1,1}, l_{2,1},\ldots l_{\max,1}\}$. The subset $(m_1, \ldots, m_{\max_2})$ represents the number of class 2 ongoing calls for each QoS level, and is defined in a similar manner.

For the sake of clarity, in the following example we illustrate a very small case.

Example

Let us assume a system with two kinds of calls and B=4 bbu. B_1=3 bbu can be used by both classes, while $B-B_1$=1 bbu is strictly reserved for class 2, so giving priority to class 2. The reservation scheme allows the allocation of new class 1 calls within T_1=2 bbu, while the handoff class 1 calls can be allocated within B_1=3. The new class 2 calls can use T_2=3 bbu, while the handoff class 2 calls can use all the available bandwidth B=4 bbu, with respect to priority. As assumed above, $T_1<T_2$ and $T_2=B_1$.

Class 1 has two acceptable QoS levels $L_1=\{1, 2\}$, while class 2 has a more stringent QoS requirement with a unique acceptable level, that is $L_2=\{2\}$.

The state $\mathbf{s}=\{(2, 0), (0)\}$ means that no class 2 call is in the system, while two class 1 calls are in the system. Since threshold T_1=2, these calls are receiving service at the minimum quality level $l_{1,1}$=1. If a new class 1 arrival occurs, no degradation is possible, so the arrival is lost. If a handoff class 1 arrival occurs, this can be accepted at the minimum QoS level, so a system transition occurs into the state $\mathbf{s}'=\{(3, 0), (0)\}$. In this state any class 1 arrival will be lost due to saturation of threshold B_1. If from state **s** a departure occurs, the remaining ongoing class 1 call can be upgraded to the second QoS level $l_{\max_1,1}$=2. So the state process transition is in $\mathbf{s}''= \{(0, 1), (0)\}$.

Let $\pi(\mathbf{s})$ denote the equilibrium probability of state **s** and $\boldsymbol{\pi}=[\pi(\mathbf{s}_1), \ldots, \pi(\mathbf{s}_{|E|})]$ the steady state probability vector. The system solution can then be expressed as:

$$\boldsymbol{\pi}\,\mathbf{S} = \mathbf{0}, \qquad \sum_{\forall \mathbf{s}:\ \mathbf{s}\in E} \pi(\mathbf{s}) = 1$$

where **S** is the generator matrix of the markovian process.

3.1 The performance indices

In this section, the performance indices considered in the paper to evaluate the proposed scheme are defined.

The blocking and dropping probabilities for each class are traditionally used as a measure of the QoS received by the user. They are defined simply as the sum of the steady state probabilities in which the request is blocked or dropped:

$$P_{x,i} = \sum_{\forall \mathbf{s}:\ \mathbf{s}\in E \wedge I^{x,i}(\mathbf{s})=0} \pi(\mathbf{s}) \qquad i= 1, 2$$

where x is b in case of blocking probability and is d for dropping probability. The function $I^{x,i}(\mathbf{s})$ verifies that in state **s** the call is lost ($I^{x,i}(\mathbf{s})$=0). According to MATS, this happens since there are not enough bbu to admit the call even through the application of the degradation algorithm. In particular, for the new and handoff calls of class i respectively, the function is defined as follows:

$$I^{b,i}(\mathbf{s}) = f^{n,i}(\mathbf{s}, T_i, l_i^{*}) + d(\mathbf{s}) f_d^{n,i}(\mathbf{s}, T_i)$$

$$I^{d,i}(\mathbf{s}) = f^{h,i}(\mathbf{s}, B_i, l_i^{*}) + d(\mathbf{s}) f_d^{h,i}(\mathbf{s}, B_i) \qquad i= 1, 2 \tag{1}$$

where $l_i^{*}\in L_i$ is the maximum available amount of bbu in state **s** to allocate the incoming call and the functions d(**s**), $f^{x,i}(\mathbf{s}, X, l_{j,i})$ and $f_d^{x,i}(\mathbf{s}, X)$ are defined in the following. Informally, the first term considers the case that the call can be admitted without invoking the degrade mechanism. The second term takes into account the case that the call is admitted after the degradation algorithm is applied.

First we simply define d(**s**)=1 if in state **s** there are calls at a QoS level superior to the minimum, the degradation is possible. If the degradation is not possible, d(**s**)=0.

Let us denote $f^{x,1}(\mathbf{s}, X, l_{j,i})$ the function that verifies if the free bbu in state **s** are enough to satisfy the incoming call, of type x and class i where x=n,h and i=1,2, at level $l_{j,i}$. X is the threshold for class i and calls of type x. The function is defined as follows:

$$f^{x,i}(\mathbf{s}, X, l_{j,i}) = \begin{cases} 1 & \text{if } (X - g(\mathbf{s})) \ge l_{j,i} \\ 0 & \text{otherwise} \end{cases} \qquad i=1, 2$$

where the function g(s) is the number of used bbu in state s and given $s=\{(n_1, \ldots, n_{\max_1}), (m_1, \ldots, m_{\max_2})\}$, is defined as follows:

$$g(\mathbf{s}) = \sum_{j=1}^{\max_1} n_j\, l_{j,1} + \sum_{j=1}^{\max_2} m_j\, l_{j,2}$$

Analogously, $f_d^{x,i}(\mathbf{s}, X)$ is the function that verifies if from state **s** the bbu freed through the degrade mechanism, are enough to satisfy the incoming call at the minimum quality level $l_{1,i}$. The function is defined as follows:

$$f_d^{x,i}(\mathbf{s}, X) = \begin{cases} 1 & \text{if } (X - g(\mathbf{s'})) \ge l_{1,i} \\ 0 & \text{otherwise} \end{cases} \qquad i=1, 2$$

with **s**' the successor state of **s** generated by the degrade mechanism application.

For the sake of example, we consider $I^{b,1}(\mathbf{s})$, which is the case of a class 1 new incoming call. According to (1), $I^{b,1}(\mathbf{s})$=0 if $f^{n,i}(\mathbf{s}, T_1, l_1^{*})$=0 and $d(\mathbf{s}) f_d^{n,i}(\mathbf{s}, T_1)$=0

According to MATS, as explained in the previous section, in state **s** if the level $l_1^{*}=\max\{l \mid l \in L_1, l < l_{\max_1,1}\}$ is determined such that $f^{n,1}(\mathbf{s}, T_1, l_1^{*})$=1, the new incoming call is admitted at l_1^{*} quality level. This case is considered by the first term in (1).

On the contrary, if the free bbu are not enough to admit the call, according to MATS the system activates the degrade mechanism if there are calls to degrade (d(**s**)=1). If the bbu freed by the degradation algorithm are enough, the call is admitted at the minimum quality level $l_{1,1}$ and $f_d^{n,1}(\mathbf{s}, T_1)$=1. However, if the free bbu (in state **s**') are not enough the new incoming call is lost. This case is considered by the second term in (1).

The system utilization is a QoS measure from the service provider's point of view. It is defined as follows:

$$U = \frac{1}{B} \sum_{\forall \mathbf{s}:\ \mathbf{s} \in E} \pi(\mathbf{s})\, g(\mathbf{s})$$

The following two less traditional metrics are defined in order to capture the cost of the upgrade-degrade proposed mechanism and the QoS delivered to the users.

For each class i, QL_i is a measure of the probabilities in which calls are served at the maximum QoS level. This is defined as follows:

$$QL_i = \sum_{\forall \mathbf{s} \in E:\ y_{max_i} > 0} \pi(\mathbf{s}) \frac{y_{max_i}}{y_{tot}} \qquad i = 1, 2$$

where given $s=\{(n_1,\ n_{max_1}), (m_1,\ m_{max_2})\}$, y is the state component related to class i, that is, for example, for class 1: $y_{max_1} = n_{max_1}$ and $y_{tot} = \sum_{j=1}^{max_1} n_j$.

The cost of the upgrade-degrade mechanism is a measure of the frequency of application of the u-d mechanism. It is defined as follows:

$$C_{u-d} = \sum_{\forall \mathbf{s}:\ \mathbf{s} \in E} \pi(\mathbf{s}) \frac{t_{u-d}}{t_{tot}}$$

where t_{u-d} is the transition rate from state **s** for effect of the application of u-d mechanism, t_{tot} is the total out-transition rate from state **s**.

4. RELATED WORK

As stated in the introduction, whilst there is much research about the CAC scheme in different frameworks and perspectives [1-22], very few papers [3, 20] combine the multiclass characteristic with the adaptive issue. This is the main motivation of our work.

The traditional guard channels scheme is unfair for multiclass application, since calls with lower bandwidth requirements fill up any free channels (bbu) whilst calls with higher bandwidth requirements cannot be accepted until all its required bbu are available. Therefore different thresholds T_i have to be considered in the scheme so as to guarantee fairness to different types of requests [19, 21].

For the aim of comparison, we consider among the proposed schemes those that are more consistent with our approach. In particular, we consider the Multi-Guard Channel Scheme (MGC) proposed in [19] and the Guard Channels with Fixed Reservation (GCFR) recently proposed in [21]. It is worth noting that while the MGC scheme uses the same criterion of priority as we consider, the GCFR uses a less stringent concept of priority since it allows handoff calls of all types to completely share the reserved guard channels group.

Note that both MGC and GCFR schemes consider no adaptive bandwidth allocation algorithm, in other words they assume only one QoS level for each class of requests. On the other hand, this is the main contribution of our work. We can say that the proposed MATS scheme is the extension of the MGC scheme including the u-d mechanism.

For the aim of consistency and completeness, we extend the GCFR scheme to include the ABA algorithm and so completing the comparisons. We denote the extension GCFR-D scheme.

In the following section we present the numerical results.

5. RESULTS AND COMPARISONS

In this section we present two sets of experiments.

5.1 A comparison with literature schemes

In the first set of experiments, the proposed MATS scheme is compared with the MGC and GCFR schemes. A single-cell with B=10 bbu is considered. Two classes of calls are assumed, with class 2 having strict priority over class 1. Class 1 calls have flexible QoS level with $L_1 = \{1, 2\}$, while class 2 calls have more stringent QoS requirements with $L_2 = \{2\}$.

The bandwidth partition allows class 1 calls to use B_1=8 bbu, while the highest priority class 2 calls can use the whole bandwidth available in the cell, that is B_2=B=10 bbu. To guarantee that handoff calls dropping probabilities are as low as possible, the following thresholds are assumed: T_1= 6 and T_2= 8.

For each class, we assume $\lambda_{h,1} = \lambda_{n,1}/2$ [20, 21], with $\lambda_i = \lambda_{n,1} + \lambda_{h,1}$. Moreover, by assuming $\lambda_1 = \lambda_2$, in all our experiments we consider $\lambda = \lambda_1 + \lambda_2 = 0.05$, 0.1, 0.15, 0.2, 0.25, 0.3 calls/min. Finally, we assume μ_i.=0.05 calls/min.

Figures 3 and 4 show the blocking and dropping probabilities of the two classes respectively. These are the QoS measures from the user perspective.

Let us consider figure 3. (a): the MATS scheme performs better than the other two schemes for a class 1 request. We believe the reason is the u-d mechanism of the MATS: class 1 requests, that are probably blocked according to the other schemes, may be accepted according to the MATS by means of the degradation mechanism application. This seems to be confirmed from the results for class 2 in figure 4. (a), where the improvement of the blocking probabilities is very limited with the MATS, since the unique QoS level for class 2 does not allow for degradation to be applied.

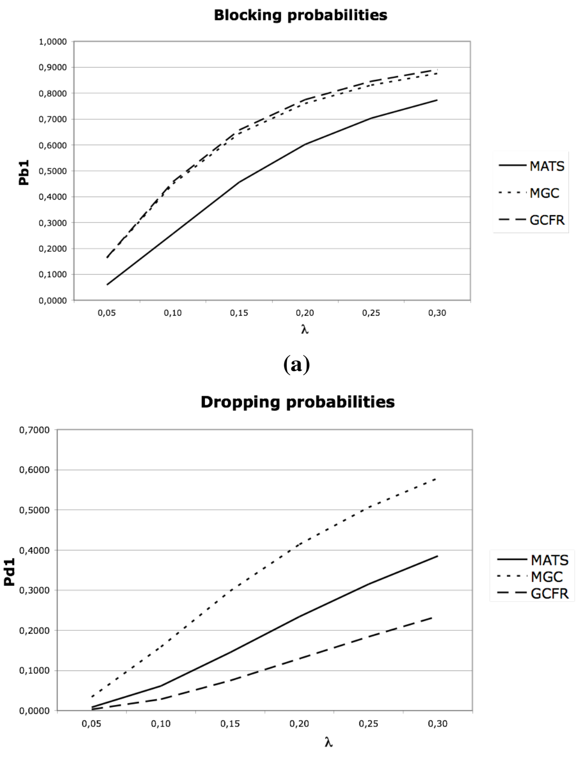
For this reason and to obtain a more consistent comparison, in the following section we present a second set of results regarding the extension of GCFR with the u-d mechanism.

Regarding the dropping probabilities, figures 3-4 (b) demonstrate, as already observed for class 1, that the MATS scheme shows better performance than the MGC scheme for class 1 and the same performance as the MGC scheme for class 2. The GCFR scheme shows the best performance for class 1 but the worst performance for class 2. This is probably the consequence of allowing class 1 handoff calls to fully share with class 2 the bandwidth region reserved for handoff. It is easy to see that this yields a gain for class 1 requests but a penalty for class 2 requests. Moreover, as we highlighted above, this is in contrast with a stringent priority criterion that should give a preferential treatment to the highest priority class. This is the assumption for our MATS scheme.

Figure 5 shows the bandwidth utilization under the three schemes.

The MGC scheme shows the lowest utilization, the GCFR scheme performs better than the MGC scheme and the proposed MATS scheme shows the highest utilization. We believe this is due to the u-d mechanism that, by means of flexibility, allows optimal utilization of bandwidth.

As stated in section 3.1, we define a less traditional metric to measure the amount of high QoS delivered to the users. Figure 6 shows the QL_i delivered to class i calls. It is not surprising that, while all schemes obviously deliver the same level to class 2 calls, the proposed MATS scheme yields the worst treatment for class 1 calls in respect to the other two schemes. We believe this behavior is also due to the application of the u-d mechanism.



(a)

(b)

Figure 3. Class1 blocking and dropping probabilities, according to the MATS, MGC and GCFR schemes.

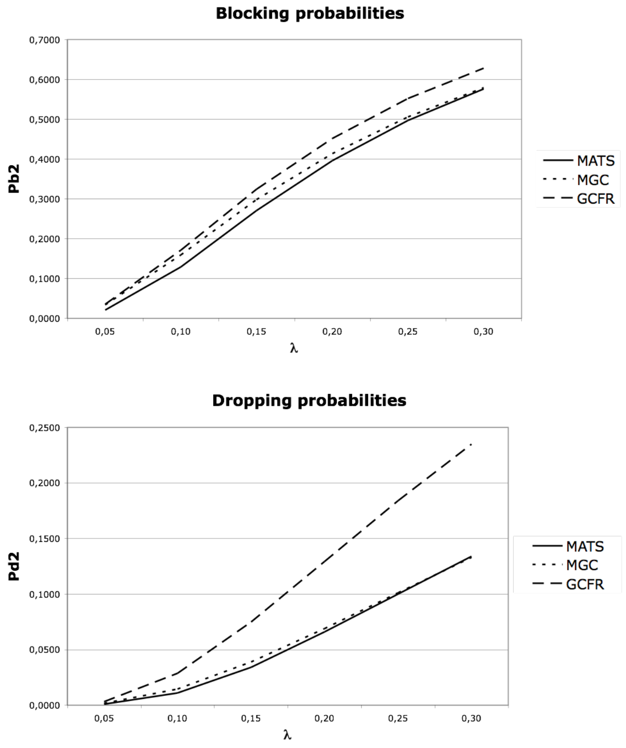


Figure 4. Class 2 blocking and dropping probabilities, according to the MATS, MGC and GCFR schemes.

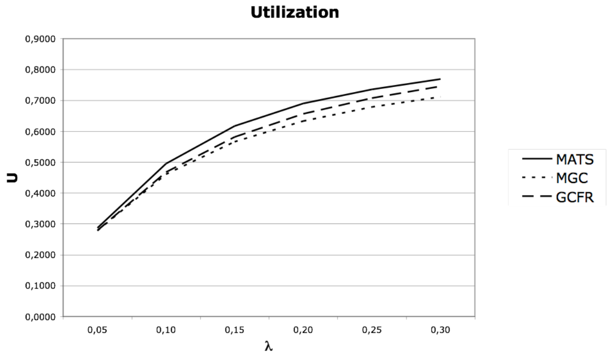


Figure 5. Bandwidth utilization

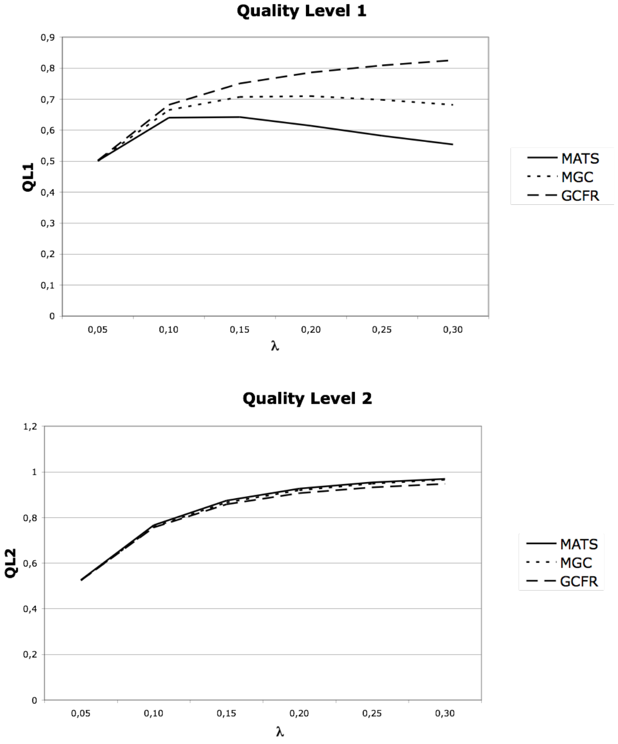


Figure 6. Quality level for each class

5.2 An extension of the GCFR scheme

The results of the previous section, yield to the consideration that, in the case of multimedia services with flexible QoS, an adaptive bandwidth algorithm could not only improve the system utilization, but also the performance received by the users at the expense of a degraded service.

As a consequence, we have extended the GCFR scheme to include the u-d mechanism (GCFR-D) as defined for our MATS scheme.

In this second set of experiments we consider B=100 bbu, L_1= {5, 10}, L_2= {10}.

The bandwidth partition is: B_1=90 bbu, B_2=B=100 bbu. The thresholds are: T_1= 80 and T_2= 90, and the arrival rates are: λ = 0.2, 0.4, 0.6, 0.8, 1, 1.2 calls/min. As in the first set of experiments, we assume μ_i.=0.05 calls/min, i=1, 2. All the other assumptions still hold.

The following figures 7 and 8 show blocking and dropping probabilities for classes 1 and 2 respectively.

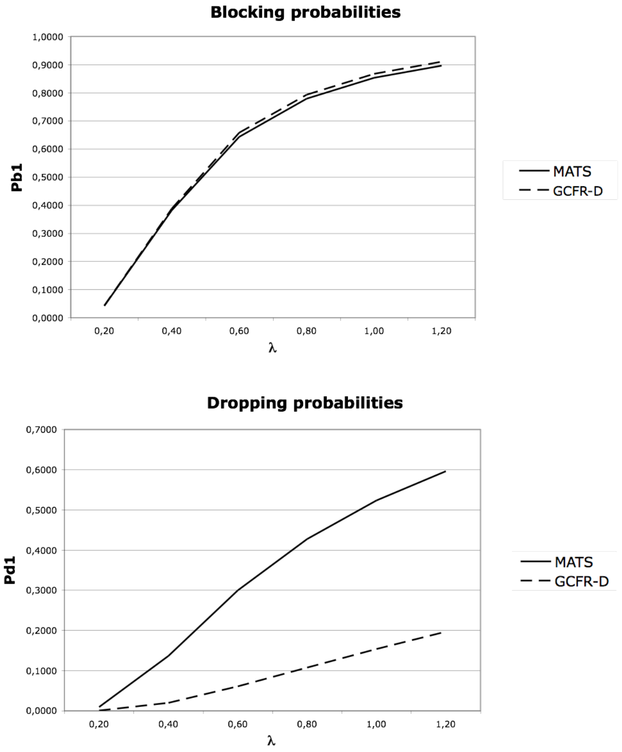


Figure 7. Class1 blocking and dropping probabilities, according to the MATS and GCFR-D schemes.

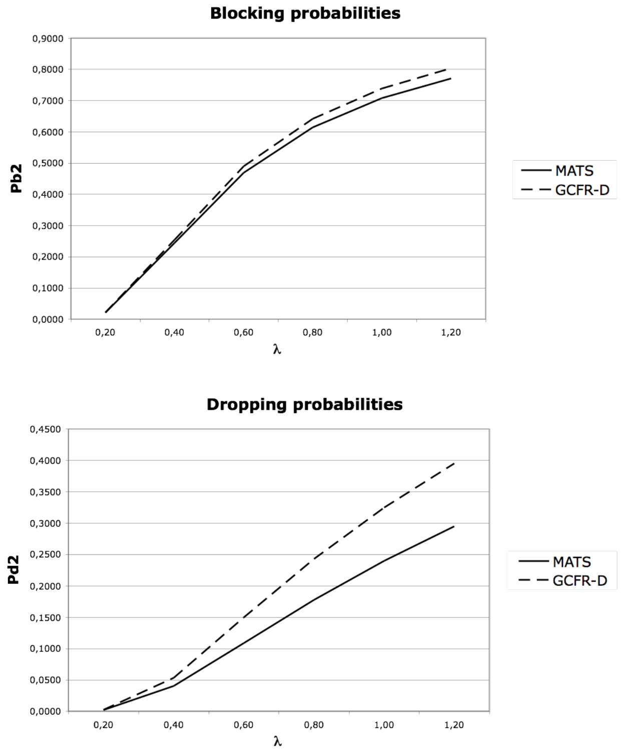


Figure 8. Class 2 blocking and dropping probabilities, according to the MATS and GCFR-D schemes.

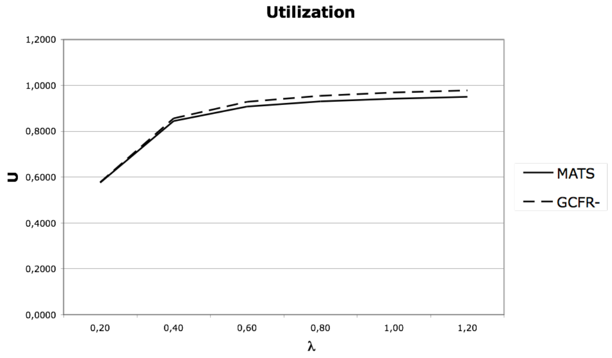


Figure 9. Bandwidth utilization

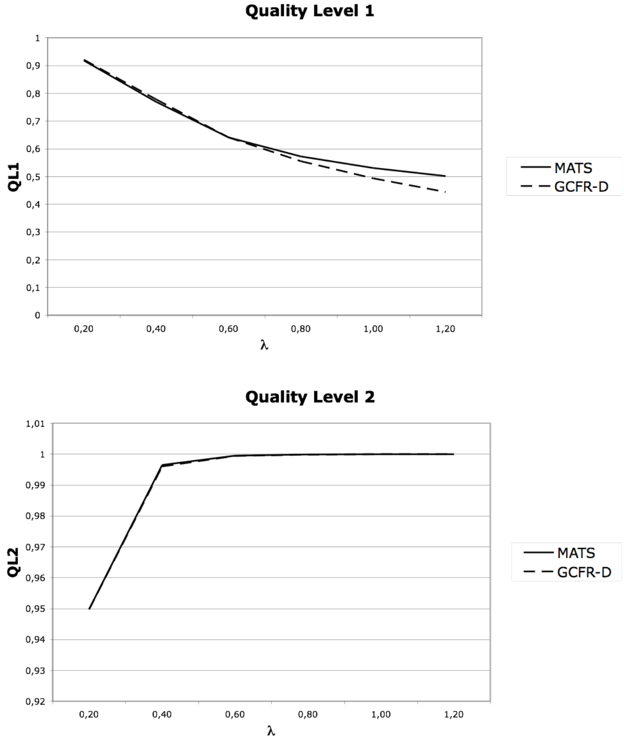


Figure 10. Quality level for each class

Figure 9 shows system utilization and figure 10 shows the quality level for class 1 and class 2.

The results seem to confirm our hypothesis: the application of an adaptive mechanism is useful for various perspectives. When the priority among classes has to be strictly preserved, the proposed MATS scheme performs better: at a modest decrease of bandwidth utilization, the highest priority class experiments an increase of the performance received from the user's point of view, mainly in regard to dropping probability that, as well known, is the more critical aspect. It is worth noting that surprisingly the quality level for class 1 also increases in respect to GCFR-D. We believe this is due to the use of a threshold for class 1 handoff calls and to the particular level values for each class. Indeed, the reserved bandwidth portion for the class 2 handoff calls most likely yields a higher probability that a departure will occur for a call at a high level, thereby generating the activation of the upgrade mechanism. On the contrary, for the GCFR-D scheme the reserved bandwidth portion for handoff calls is fully shared by the two classes, so the

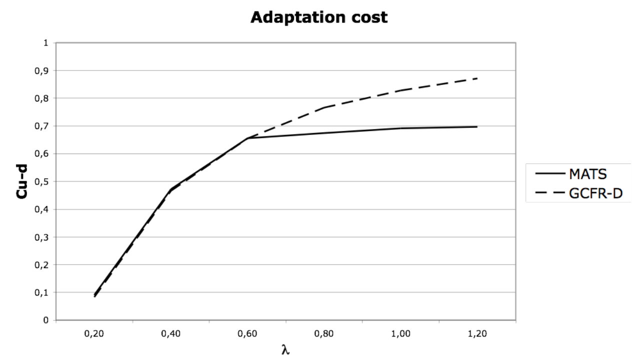


Figure 11. Activation of the u-d mechanism cost

probability for a departure to be of a high level will probably be lower, therefore decreasing the activation of the upgrade and maintaining a lower quality level for class 1 requests.

Since we are comparing two schemes with adaptive bandwidth allocation, we now evaluate the cost of this mechanism activation. For this purpose, a cost measure has been defined in section 3.1. Figure 11. shows the cost of the activation of the u-d mechanism.

In conclusion, the proposed MATS scheme seems to perform better than GCFR-D: it is more consistent with the priority criterion and delivers better quality to both classes, without resource wastage.

6. CONCLUSIONS

In this paper, we considered the resource allocation problem. This is a very critical issue in the new framework of differentiated services in heterogeneous networks. In particular, we considered wireless mobile integrated service networks and proposed a CAC and ABA algorithm , that we called MATS, for a multiclass environment with flexible QoS requirements and priority. Whilst there is much research about the admission control and resource allocation in wireless networks, very few papers combine the multiclass characteristic with the adaptive issue. This is the main contribution of this paper. We developed a general analytical model to include the characteristics of the considered environment. We considered traditional performance metrics and also defined two metrics to measure the cost of the ABA algorithm activation and to measure the amount of high level service received by users.

We compared our proposed scheme with the most recent schemes proposed in the literature that are consistent with our assumptions. It is worth noting that these schemes do not include an adaptive mechanism for flexible QoS. From the first set of experiments one can conclude that an adaptive mechanism is mandatory for efficient resource utilization and to respect the priority criterion.

As a consequence, we extended the literature schemes to include the ABA algorithm and presented a second set of experiments. These results further confirm that our proposed MATS scheme shows good performance both from user and service provider perspectives.

Acknowledgement

The authors are thankful to the anonymous referees for the improvements their comments have earned to the quality of the presentation and to the completeness of the paper.

7. REFERENCES

[1] A.R. Aljadhai, T.F. Znati, A Bandwidth Adaptation Scheme to Support QoS Requirements of Mobile Users in Wireless Environments, Proc. of the IEEE 9th ICCCN, October 16-18, 2000, Las Vegas, Nevada, pp.34-39.

[2] C.T. Chou, K.G. Shin, “Analysis of Adaptive Bandwidth Allocation in Wireless Networks with Multilevel Degradable Quality of Service”, IEEE T Mobile Comput, vol.3, No. 1, pp. 5-17, January-March 2004

[3] C. Lindemann, M. Lohnmann and A. Thummler, ”Adaptive call admission control for QoS/revenue optimization in CDMA cellular networks”, Wirel Netw, vol.10, Issue 4, pp. 457-472, 2004

[4] G.V. Zaruba, I. Chlamtac, S.K. Das, A Prioritized Real-Time Wireless Call Degradation Framework for Optimal Call Mix Selection, Mobile Netw Appl 7, 143-151, 2002.

[5] S.K. Das, R. Jayaram, N.K. Kakani, S.K. Sen, A call admission control scheme for quality-of-service (QoS) provisioning in next generation wireless networks, Wirel Netw 6 (2000) 17-30.

[6] M.R. Sherif, I.W. Habib, M. Nagshineh, P. Kermani, Adaptive Allocation of Resources and Call Admission Control for Wireless ATM Using Genetic Algorithms, IEEE J Sel Area Comm, Vol. 18, No. 2, February 2000.

[7] V. B. Bharghavan, K.-W. Lee, S. Lu, S. Ha, J.-R. Li, D. Dwyer, “The TIMELY Adaptive Resource Management Architecture”, IEEE Pers Commun, pp. 20 – 31, August 1998

[8] S. Sen, J. Javanda, K. Basu, S. Das, “Quality-of-Service Degradation Strategies in Multimedia Wireless Networks”, Proc. of the IEEE VTC’98, pp. 1884 – 1888, 1998

[9] A. K. Talukdar, B. R. Badrinath, A. Acharya, “Rate Adaptation Schemes in Networks with Mobile Hosts”, Proc. of the ACM MobiCom 98, Dallas Texas USA, 1998

[10] S. K. Das, R. Jayaram, N. K. Kakani, S. K. Sen, “A call admission and control scheme for quality-of-service (QoS) provisioning in next generation wireless networks”, Wirel Netw, vol. 6 , Issue 1, pp. 17 – 30, 2000

[11] Y. Xiao, Chen C. L. P., Yang Wang, “Fair bandwidth allocation for multi-class for adaptive multimedia services in wireless/mobile networks” Proc. of the IEEE VTC’01, vol. 3, Issue , 2001 pp. 2081 - 2085, 2001

[12] C.T. Chou, K.G. Shin, “Analysis of combined adaptive bandwidth allocation and admission control in wireless networks”, Proc. of the IEEE INFOCOM 2002, vol.2, pp.676-684, 2002

[13] J. Wang, Q. Zeng, Agrawal D.P ,“Performance analysis of a preemptive and priority reservation handoff scheme for integrated service-based wireless mobile networks”, IEEE T Mobile Comput, vol. 2, Issue 1, pp. 65 – 75, 2003

[14] T. Kwon, Y. Choi, C. Bisdkian, M. Naghshineh, “QoS provisioning in wireless / mobile multimedia networks using an adaptive framework” , Wirel Netw, vol. 9, Issue 1, 2003

[15] N. Nasser, “Real-time service adaptability in multimedia wireless networks”, Proc. of the ACM Q2SWinet’05, Montreal, Quebec, Canada, 2005

[16] S.T. Cheng, C.M. Chen, I.R.Chen, “Performance evaluation of an admission control algorithm with dynamic threshold with negotiation”, Perform Evaluation 52, pp.1-13, 2003

[17] M. Ghaderi, R. Boutaba, “Call admission control for Voice/Data Integration in Broadband Wireless Networks”, IEEE T Mobile Comput, vol.5, No. 3, pp. 193-207, March 2006

[18] M. Salamah, I. Candan, "Performance analysis of a time-threshold based multi-guard bandwidth al location scheme in cellular networks ", Proc. of the IEEE ICWMC '06, 2006

[19] M. Salamah, "An adaptive multi-guard channel scheme for multi-class traffic in cellular networks", Proc. of the ACS/IEEE AICCSA-06, pp.716-723, 2006

[20] D. Niyato, E. Hossain, "A novel analytical framework for integrated cross-layer study of call-level and packet-level QoS in wireless mobile multimedia networks", IEEE T Mobile Comput, vol. 6, no. 3, pp. 322-335, 2007

[21] N. A. Mohamed, D. Z. Deniz, "Fixed and dynamic bandwidth allocation strategies for wireless mobile integrated services networks", Wirel Netw, vol.14 Issue 1, pp.121-131, 2008

[22] X. Yang, G. Feng, "Bandwidth reallocation for bandwidth asymmetry wireless networks based on distributed multiservice admission control", IEEE T Mobile Comput, vol. 7, no. 11, 2008

[23] ITU-T E.860 Framework of a service level agreement, 2002

[24] M. Ghaderi, R. Boutaba, "Call admission control in mobile cellular networks: a comprehensive survey", Wirel Commun Mob Com, vol.6, Issue 1, pp. 69-93, 2006

[25] O. E. Falowo, H. A. Chan, "Joint call admission control algorithms: requirements, approaches and design considerations", Comput Comun, pp.1200-1217, 2008

[26] International Standard "Information Technology-Quality of Service: Framework", Reference number ISO/IEC 13236:1998 (E)

[27] ITU-T Recommendation G.1000 - Communications quality of service: a framework and definitions, 2001