# Efficient Bandwidth Management in Broadband Wireless Access Systems Using CAC-based Dynamic Pricing

Bader Al-Manthari [1], Nidal Nasser [2], Najah Abu Ali [3], Hossam Hassanein [1]

[1] Telecommunications Research Laboratory
School of Computing, Queen's University
Kingston, ON, Canada K7L 3N6
{manthari, hossam}@cs.queensu.ca

[2] Department of Computing & Information Science
University of Guelph
Guelph, ON, Canada N1G 2W1
nasser@cis.uoguelph.ca

[3] College of Information Technology
UAE University, Al-Ain
P.O. 17555, UAE
najah@uaeu.ac.ae

***Abstract*— While the demand for mobile broadband wireless services continues to increase, radio resources remain scarce. Even with the substantial increase in the supported bandwidth in next generation Broadband Wireless Access Systems (BWASs), it is expected that these systems will severely suffer from congestion due to the rapid increase in demand of bandwidth intensive applications. Without efficient bandwidth management and congestion control schemes, network operators may not be able to meet the increasing demand of users for multimedia services, and hence they may suffer immense amount of revenue loss. In this paper, we propose an admission-level bandwidth management scheme consisting of Call Admission Control (CAC) and dynamic pricing. The main aim of our proposed scheme is to provide monetary incentives to users to use the wireless resources efficiently and rationally, hence, allowing efficient bandwidth management at the admission level. By dynamically determining the prices of units of bandwidth, the proposed scheme can guarantee that the arrival rates to the system are less than or equal to the optimal ones computed dynamically, hence, guaranteeing a congestion-free system. Simulation results show the effectiveness and strengths of our proposed approach.**

***Keywords*— *call admission control, congestion, dynamic pricing.***

## I. INTRODUCTION

The success of emerging next generation Broadband Wireless Access Systems (BWASs) such as 3.5G High Speed Downlink Packet Access (HSDPA) [1] and WiMAX [2] will depend, among other factors, on their ability to manage their shared wireless resources in the most efficient way. This is a complex task due to the expected increase in demand for multimedia services that have diverse and very high bandwidth requirements. Therefore, bandwidth management is crucial for the success of such communication systems.

To support as many users as possible while satisfying their bandwidth requirements, network operators typically employ Call Admission Control (CAC), which is an admission-level bandwidth management strategy. By limiting the number of admitted users' calls in the system, CAC can guarantee that the packet-level Quality of Service (QoS) (e.g., packet delay, average throughput, etc) of ongoing calls will not get degraded as a result of new incoming ones. CAC is very efficient in improving the packet-level QoS of ongoing calls especially during congestion periods. However, it may not be as efficient in improving the admission-level QoS (e.g., call blocking probabilities). This is because CAC, by itself cannot avoid congestion due to the fact that it does not provide incentives to the users to use the shared wireless system resources rationally and efficiently. Therefore, the call blocking probabilities can reach high levels during congestion periods. To overcome this problem, there has been some research work recently on integrating admission-level dynamic pricing with CAC in order to control call request arrivals to the system through monetary incentives [3], [4] and [5]. Hence, maintaining the call-level QoS at the desired thresholds. In admission-level dynamic pricing, the price for a unit of time or bandwidth is determined when the user initiates a call request before she is admitted to the system. The price in this case is fixed for the call duration. This price is dynamically determined according to the network load. Dynamic pricing can competently promote rational and efficient use of the shared wireless resources by influencing the users' behaviors. This is because it can discourage price-sensitive users from over using the wireless network when it is congested and encourage them on the other hand to increase their demand when the network is underutilized. Dynamic pricing is, therefore, a promising solution to traffic control problems, which can help alleviate the problem of congestion and provide efficient bandwidth management. In addition, it can ensure economic efficiency since it ensures that the wireless resources are given to those who value them the most. Furthermore, dynamic pricing is cost-effective and it can generate higher revenues.

Several CAC schemes with dynamic pricing have been proposed in the literature [3], [4] and [5]. The scheme in [3] dynamically computes the optimal price so that the price-affected call arrival rates maximize the summation of users'

This research is supported by the government of the Sultanate of Oman, Natural Sciences and Engineering Research Council of Canada and by a grant from Bell University Labs.

utilities (i.e., social welfare of the system), where the user utility is assumed to be a function of the call blocking probabilities, which are in turn a function of the arrival rates. However, the scheme is designed to prevent network congestion only, where a flat rate pricing is assumed when the network is underutilized. Therefore, users are not given any incentives to increase their usage of the network when it is underutilized, which results in resource wastage, and hence potential revenue loss. In addition, the scheme lacks the support of comprehensive QoS since it assumes that all calls require the same amount of resources, which makes it unsuitable for next generation BWASs.

In [4], a CAC-based dynamic pricing scheme is proposed. In this scheme users are divided into two types, priority users and conventional users. When the network is underutilized, all users in this scheme are considered conventional users and are placed in the conventional queue awaiting admission, where they are charged a flat rate. During congestion periods, a dynamic price is computed and the users are given the option to choose between being priority users, where they are charged a higher dynamic price and are placed in the priority queue to be served faster or being conventional users, where they are charged a flat rate and they are served slower. The dynamic price is determined so that the maximum number of users that the network can accommodate, and yet conform to the delay the users can spend in the queue awaiting admission [4].

However, even though the scheme considers the delay the users experience in the admission queues, it does not take into account the new call blocking and handoff call dropping probabilities. This may not be practical since wireless network operators have a limit on the number of calls they can block, which is usually determined by regulations. In addition, similar to the scheme in [3], this scheme assumes that calls require the same amount of resources. Hence, it may not be unsuitable for Next generation BWASs. Moreover, the scheme is designed to prevent congestion only. Therefore, it does not provide incentives to users to increase their demand for the network services when the network is underutilized.

The CAC-based dynamic pricing scheme in [5] aims at reducing congestion and maximizing revenues in wireless cellular networks. The scheme considers the effects of prices on call arrivals, retrials (i.e., requesting the same service again after being blocked) and substitutions among services (i.e., substituting a service for another after being blocked). Using some assumptions about the new and handoff call arrival rates, the scheme dynamically determines the prices of network services as to encourage or discourage the arrival rates to the system in order to reserve some bandwidth for arriving handoff or higher revenue-generating users.

Even though the scheme considers different QoS classes, it assumes that users within each class request the same amount of bandwidth. This is still not practical in next generation BWASs since in these systems, each class can have a number of services each requesting a certain amount of bandwidth (e.g., audio streaming and video streaming in the streaming class). In addition, the scheme is complex and requires many calculations to determine the price.

Moreover, all the schemes in [3], [4] and [5] are based on certain assumptions about users' demand models and cannot, therefore, be generalized to work with different demand models without affecting the way prices are computed. This limits their scalability, since different network operators might have different demand models depending on their subscribers. The schemes in [6], [7], [8], [9] and [10] apply dynamic pricing at admission level without using CAC. These schemes, therefore, cannot achieve optimized call-level QoS.

Therefore, there is a need for a CAC-based dynamic pricing scheme that is able to support different QoS classes with different users having different bandwidth requirements, work with various demand models and compute the dynamic prices in a simple way

In this paper, we propose an admission-level bandwidth management scheme that consists of a CAC component and a dynamic pricing component. The proposed scheme aims at efficiently managing the bandwidth of BWASs in order to simultaneously satisfy the bandwidth requirements of users, maximize the utilization of BWASs and prevent congestion. By dynamically computing the prices of units of bandwidth, our scheme is able to force the arrival rates of call requests to the system towards the optimal ones as determined by the CAC component. The proposed scheme improves the scheme in [3] in the following aspects. First, unlike the scheme in [3], our proposed scheme employs dynamic pricing during all network conditions (i.e., whether the network is underutilized or congested). This way, our scheme is able to maximize the utilization of BWASs when these networks are underutilized as well as preventing congestion when they are over utilized. Second, our scheme supports multiple QoS classes with calls having multiple bandwidth requirements, which makes it more suitable for BWASs. Finally, congestion prices in the scheme in [3] are computed based on certain assumptions about the users' utilities not based on the amount of available network bandwidth. The scheme, therefore, is incapable of capturing the dynamics of the network (i.e., changes in available bandwidth) and the varying bandwidth requirements of users. This explains the inability of the scheme to achieve zero call blocking probabilities despite of assuming accurate users' demand model [3]. On the contrary, congestion prices in our scheme are computed dynamically based on the amount of available bandwidth in the network. As a result, our scheme is shown to achieve zero call blocking probabilities in case accurate users' demand model is assumed.

The rest of this paper is organized as follows. Section II provides an overview of the proposed scheme. Section III presents a description of the proposed scheme. Performance results are presented in Section IV. Finally, conclusions and future work are discussed in Section V.

## II. Overview of Bandwidth Management Scheme

In this section, we provide an overview of our proposed bandwidth management scheme. We assume that there are $K$ QoS classes, where class $i$ has higher priority than class $i$+1, $1 \le i$ and $i+1 \le K$. We consider that class $i$ calls request $b_i$ units of bandwidth. Our scheme consists of two components namely, the CAC component and the dynamic pricing component. Our scheme works as follows. The CAC component continuously monitors the amount of available bandwidth (i.e., unutilized bandwidth). When the amount of available bandwidth changes due to call completion or new admitted calls, the CAC component then computes the optimal arrival rate for each QoS class in order to maximize the utilization of the new available bandwidth in the system and achieve certain fairness levels between QoS classes. The actual arrival rates for the QoS classes are, however, different from the optimal rates determined by the CAC component. In this case, the dynamic pricing component dynamically determines the prices of units of bandwidth for each class based on the users' demands in order to force the actual arrival rates to be less than or equal to the optimal ones. The dynamic prices are computed independently from the optimal arrival rates, hence, simplifying the implementation of our scheme and providing the network operators the flexibility to use any CAC and users' demands functions without affecting the computation of prices.

It should be noted that in this paper, we focus on new calls only. This is because handoff calls are not affected by dynamic pricing when charging at admission level, since they have been already charged at the cell where the calls have been initiated. Therefore, the network operator cannot influence the behavior of handoff users by changing the price. In this case, the network operator can use a form of Guard Channel schemes in which a certain amount of bandwidth is exclusively reserved for handoff calls in order to maintain the handoff call dropping blocking probability below a certain threshold [11].

## III. Description of The Bandwidth Management Scheme Components

In this section, we describe each component of our proposed bandwidth management scheme. We make the following definitions. Let:

- $N_i$ : number of admitted calls in class $i$.
- $N = \sum_{i=1}^{K} N_i$ : total number of admitted calls.
- $B$: total bandwidth of the system
- $B_{free}(t)$ : total available bandwidth at time $t$. $B_{free}(t)$ can be computed as follows:

$$B_{free}(t) = B - \sum_{i=1}^{K}\sum_{j=1}^{N_i} b_{ij}(t) \qquad (1)$$

where $b_{ij}(t)$ is the bandwidth assigned to user $j$ in class $i$ at time t.

- $\lambda_i(t)$ : arrival rate of new calls to class $i$ at time $t$. Therefore, the total demand requested by class $i$ at time $t$ is equal to $\lambda_i(t) \cdot b_i$ , where $b_i$ is the bandwidth request of class $i$.
- $\lambda_{\max}$ : maximum total arrival rate to the system (i.e., of all QoS classes). $\lambda_{\max}$ can be easily computed from observed historical data of the network operator or it could be set to the total number of subscribers.
- $p_i(t)$: price in terms of units of money per unit of bandwidth for class $i$ services.
- $A_i$ : percentage of users who have sufficient *Willingness to Pay*[1] (WTP) to make call requests to class $i$. Clearly, $A_i$ is a function of the price (i.e., $A_i = f\left(p_i(t)\right) \to [0,1]$, where $p_i(t) = f^{-1}\left(A_i\right)$). $A_i$ can be constructed from the system's history by observing the users' responses to changes in the price. It is reasonable to assume that $A_i$ is monotonically decreasing function of the price. That is, when the price increases, $A_i$ either remains the same or it decreases. It should be noted that the computation of $A_i$ is a pure economic topic that is outside the scope of this paper. However, we utilize a well-known demand function in section IV to model $A_i$, although our scheme can work with any function for $A_i$ as explained next.

The main objective of our CAC component is to find the optimal arrival rate for each QoS class such that the utilization of available bandwidth is maximized. To achieve this objective, the CAC component will solve the following optimization problem every time it detects a change in the available bandwidth:

$$\text{Objective:} \quad \max_{\{\lambda_1,\lambda_2,\ldots,\lambda_K\}} \sum_{i=1}^{K} \lambda_i(t)\cdot b_i$$

$$\text{Subject to:} \quad \sum_{i=1}^{K} \lambda_i(t)\cdot b_i \le B_{free}(t),$$

$$\sum_{i=1}^{K} \lambda_i(t) \le \lambda_{\max}, \text{ and}$$

$$\left(\left(\sum_{j=1}^{N_i} b_{ij}(t)\right) + \lambda_i(t)\cdot b_i\right) / B \le \upsilon_i, \forall i, 1 \le i \le K \qquad (2)$$

where the first constraint ensures that the demand of all QoS classes does not exceed the total available bandwidth (i.e., supply). The second constraint ensures that the resulting total

[1] The monetary value users are willing to pay for a certain service.

arrival rate to the system (i.e., $\sum_{i=1}^{K}\lambda_i(t)$) is realistic and does not exceed the maximum arrival rate, which is limited by the number of subscribers. The last constraint is used to ensure fairness among QoS classes by restricting that the share of bandwidth for each class (i.e., the bandwidth of admitted calls + bandwidth of new calls) does not exceed a predefined value ($\upsilon_i$) determined by the network operator. For example, to achieve absolute fairness (i.e., an equal bandwidth share) between QoS classes, $\upsilon_i$ should be set to 1/*K*. Besides ensuring fairness, the second constraint can be used to promote certain services or increase revenues by assigning more bandwidth to QoS classes that are expected to yield higher revenues (e.g., higher priority classes). It should be noted that the objective function and the constraints in (2) do not include the call blocking probabilities. This is because our pricing component, as described below, can guarantee to force the actual arrival rates to be less than or equal to the optimal ones computed in (2). Hence, the system is guaranteed to be congestion-free. In addition, the objective function and the constraints in (2) are linear. Hence, the optimal arrival rates $\{\lambda_1^*(t),\lambda_2^*(t),...,\lambda_K^*(t)\}$ can be found using Linear Programming (LP) techniques.

The actual arrival rates to the system are, however, different from the optimal arrival rates (i.e., $\{\lambda_1(t),\lambda_2(t),...,\lambda_K(t)\}\neq\{\lambda_1^*(t),\lambda_2^*(t),...,\lambda_K^*(t)\}$. Therefore, the dynamic pricing component will adjust the prices of units of bandwidth for each QoS class such that the actual arrival rates are less than or equal the optimal ones computed in (2) (i.e., $\{\lambda_1(t),\lambda_2(t),...,\lambda_K(t)\}\leq\{\lambda_1^*(t),\lambda_2^*(t),...,\lambda_K^*(t)\}$ as follows. We know from the arrival rate to class *i* at time *t* (i.e., $\lambda_i(t)$) that it constitutes the following ratio of the total users that could request the service

$$\frac{\lambda_i(t)}{\lambda_{\max}} \tag{3}$$

From (3) we know that the ratio of users that have sufficient WTP to make call requests to class *i* is at least $\frac{\lambda_i(t)}{\lambda_{\max}}$ (there could be other users who have sufficient WTP, but choose not to make such requests at this time). However, the optimal ratio should equal to

$$\frac{\lambda_i^*(t)}{\lambda_{\max}} \tag{4}$$

Therefore, when $\frac{\lambda_i(t)}{\lambda_{\max}}\neq\frac{\lambda_i^*(t)}{\lambda_{\max}}$, the price of class *i* services will be adjusted so that

$$A_i = f\left(p_i(t)\right)=\frac{\lambda_i^*(t)}{\lambda_{\max}}, \forall i, 1\leq i\leq K \tag{5}$$

There are two cases, and hence two implications of price setting. The first case is when $\frac{\lambda_i(t)}{\lambda_{\max}}>\frac{\lambda_i^*(t)}{\lambda_{\max}}$, according to (5), the price should be increased so that $A_i=\frac{\lambda_i^*(t)}{\lambda_{\max}}\Rightarrow\frac{\lambda_i(t)}{\lambda_{\max}}=\frac{\lambda_i^*(t)}{\lambda_{\max}}$. In this case, if $A_i$ is accurate in modeling the users' WTP, then the ratio of incoming users who have sufficient WTP to make call requests is guaranteed to equal the optimal ratio. The second case is when $\frac{\lambda_i(t)}{\lambda_{\max}}<\frac{\lambda_i^*(t)}{\lambda_{\max}}$, the price should be lowered such that $A_i=\frac{\lambda_i^*(t)}{\lambda_{\max}}\Rightarrow\frac{\lambda_i(t)}{\lambda_{\max}}\leq\frac{\lambda_i^*(t)}{\lambda_{\max}}$. In this case, the price is lowered so that enough users have sufficient WTP to make call requests. It should be noted that users with sufficient WTP may not make call requests at the time depending on their preferences. Using our scheme they are, however, encouraged to make such calls due to lower prices. In this case, the incoming arrival rate is guaranteed to be less than or equal to the optimal ratio.

Based on the above discussion and from (5), the dynamic pricing component will set the new prices as follows

$$\mathbf{p(t)}=\left\{f\left(\frac{\lambda_i^*(t)}{\lambda_{\max}}\right)^{-1}, \forall i, 1\leq i\leq K\right\} \tag{6}$$

As it can be seen, the price equation in (6) is very simple to compute and is independent from the objective function in (2). Such independence allows the network operator to use any objective function in (2) without affecting the computations of prices and vice versa. In addition, based on the aforementioned discussion, the actual arrival rates are guaranteed to be less than or equal to the optimal ones computed in (2). Hence, using our pricing scheme, the system is guaranteed to be congestion-free.

## IV. PERFORMANCE EVALUATION

In this section, we evaluate the performance of our proposed scheme by means of call-level dynamic discrete event simulation. We test our scheme on High-Speed Downlink Packet Access (HSDPA) system. HSDPA is a 3.5G wireless system that has been introduced by the 3rd Generation Partnership Project (3GPP) as an extension to the 3G cellular system Universal Mobile Telecommunication System (UMTS). More information on HSDPA can be found in [1].

### A. Simulation Model

For simplicity, we simulate a single-cell scenario. We ignore handoff call, since they are not affected by dynamic pricing as aforementioned. The base station is located at the center of the cell. Therefore, only one base station is involved in allocating the radio resources. The cell radius is 1 Km and the base station's transmission power is 38 dBm. Two QoS classes are considered (i.e. $K$=2). For demonstration purposes, we let class 1 and class 2 calls request a bandwidth of 128 Kbps and 64 Kbps, respectively. In addition, we set $\upsilon_i$ in (2) to 1/K in order to achieve an equal share of bandwidth between the two classes.

Actual arrival rates to the system are normally time varied, and therefore, we adopt a 24 hour model for the arrival rates. In this model, the day is divided into 24 hours starting at midnight, where different arrival rates are assigned to different hours of the day based on observation of the call arrivals in a typical business day [12], [13]. It is observed in [13] that the peak hours (maximum call arrivals) occur around 11:00 AM and 16:00 PM. In our simulation, each hour of the day is simulated by 400 s and the performance results are collected at end of the each simulated hour. Call arrivals are modeled by a Poisson process where the mean total arrival rates to the system for each hour of the day are shown in Figure 1. The total arrival rate to the system is equally divided between the two classes. The arrival rates in Figure 1 constitute the actual arrival rates before dynamic pricing is implemented. When dynamic pricing is implemented, the actual arrival rates will depend on the prices. In this case, during congestion periods, our pricing component guarantees that the actual arrival rate will match the optimal one as discussed in Section III. On the other hand, when the network is underutilized, which occurs in early morning hours (00:00-05:00 AM) and at night (21:00-24:00 PM), our pricing component guarantees to provide incentives to users to use the network services while simultaneously preventing congestion. However, as discussed in Section III, not every user who has a sufficient WTP to make a call at a certain time is willing to make such a call at that time. In this case, the arrival rate to the system may stay at its low level or it may increase up to the optimal one depending on the preferences of users. To evaluate such a case when the network is underutilized, we test our proposed scheme with 0% increase in the arrival rate (i.e., the actual arrival rate stays at its low value and does not increase as a result of lower prices) and with a 10%, 30% and 50% increase of the optimal arrival rate, respectively (i.e., 10%, 30% and 50% of the users who have sufficient WTP to make call requests as a result of reducing the prices will make such calls, respectively).

The call duration of each call is modeled by an exponential distribution with a mean value of 30 s. Users are uniformly distributed in the cell. Pedestrian A (Ped A) environment is used in our simulation, which is recommended by 3GPP. Mobile users in Ped A environment move at a fixed speed of 3 Km/hr. We adopt the same channel model as in [14]. The simulation time step is one time frame, which is 2 ms in HSDPA [1]. Other simulation parameters are listed in Table I.

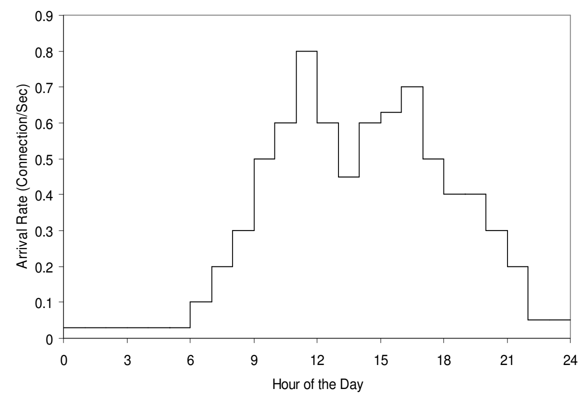


Figure 1. Arrival Rates in a Typical Business Day

TABLE I
SIMULATION PARAMETERS

| | |
|---|---|
| Simulation time per hour | 400s |
| Base Station Transmission power | 38 dBm |
| Antenna gain | 17 dBi |
| Base Station buffer size | 30 MB |
| Shadowing | Lognormal distribution |
| Intra-cell interference | 30 dBm |
| Inter-cell interference | -70 dBm |

### B. Demand Model

As aforementioned, our pricing scheme is general and can work with any demand model. To test our scheme, however, we utilize the following well-known demand model [6], [15]:

$$A_i = f\left(p_i(t)\right) = a_i(t) \cdot e^{-c_i(t) \cdot p_i(t)} \qquad (7)$$

where $a_i(t)$ is the demand shift constant for class $i$ users at time $t$ and $c_i(t)$ is the price elasticity of demand (i.e., the change in demand for a certain product or service due to a change in its price). The reason for using this particular demand model is that it can support different QoS classes and different user behaviors by considering their price elasticity of demand and their demand shift constants, which can assume different values for different times of the day. To ensure that $A_i = f\left(p_i(t)\right) \to [0,1]$, we set $a_i(t)$ to 1. In addition, for demonstration purposes, we set $c_i(t)$ to 1 and 2 for classes 1 and 2, respectively. These two values are chosen so that class 1 users are less responsive to price changes than those of class 2. This way, class 1 users are charged higher than class 2 users because class 1 users have higher priority 2. It should be noted that the actual values of $c_i(t)$ should be determined by market studies on real demand behaviors for the different users.

### *C. Simulation Results*

In this section, we compare the performance of our Bandwidth Management Scheme denoted by (BMS+*x*%, where *x* %= 0%, 10%, 30%, 50% increase in user calls when the network is underutilized as a result of lower prices as discussed in the previous section) with a Conventional CAC scheme denoted by (CCAC). In CCAC, no dynamic pricing is implemented. Instead, users are charged fixed prices and call requests are always accepted as long as there is enough bandwidth to support them. In this case, we fix the prices to 0.35 and 0.17 units of money per units of bandwidth for classes 1 and 2, respectively. These two values are chosen so that at least %70 of users have sufficient WTP to make call requests according to the demand model in (7). In practice, fixed prices are determined so that the majority of people have sufficient WTP to make calls, which is one of the main causes of congestion.

To demonstrate the ability of our scheme to maximize the utilization of the system while preventing congestion and increasing revenues, it suffices to show:

- Percentage of bandwidth utilization: the percentage of the utilized bandwidth to the total bandwidth.
- Call blocking probability: the probability that a call is blocked due to insufficient bandwidth to meet its requirements.
- Percentage of bandwidth share: the percentage of used bandwidth for each class to the total utilized bandwidth. This metric is used to test our fairness measure in (2).
- Revenue: the amount of money earned during the day.

Figure 2 shows the percentage of bandwidth utilization for our scheme and the CCAC scheme. The figure shows that our scheme can significantly increase the bandwidth utilization of the system as more users (i.e., 10%, 30% and 50%) decide to make call requests as a result of lower prices during off-peak hours. In case users are not effected by lowering the prices (i.e., case with 0% increase), the bandwidth utilization of our scheme is the same as CCAC, which is expected since our scheme is distinguished by its ability to increase the usage of the network when it is underutilized. We remark, however, that since most users are price-sensitive, they will try to make their calls when the prices are lower. Hence, the case of 0% can rarely occur in practice. Therefore, using our scheme, the network operator can increase the usage of the network when it is underutilized, hence, increasing its revenues. In addition, our scheme can efficiently prevent network congestion, and hence achieving 0% blocking probabilities as shown in Figure 3. The reason for this is that our scheme optimally determines the prices of units of bandwidth as to encourage enough users to make call requests, hence, ensuring that the system is never congested. This figure confirms the superiority of our scheme compared to CCAC scheme where users are not provided any incentives to regulate their usage of the network. Such scheme can result in very high blocking probabilities during peak hours leading to user unsatisfaction and potential revenue loss. For instance, Figure 3 shows that at peak hours (e.g., 11 AM), the blocking probability of CCAC can reach up to 11.4%.

Table II shows the percentage of bandwidth share for each class. As aforementioned, we set $\upsilon_i$ to 1/K in our objective function so that each class gets an equal share of bandwidth. The table shows that our scheme achieves better bandwidth share than CCAC. The reason for the unfair bandwidth share in CCAC is that according to our traffic model, the actual arrival rate (before dynamic pricing is implemented) is equally divided between the two classes and since class 1 users request double the amount of bandwidth compared to class 2 users, this results in higher bandwidth share for class 1. Our scheme, on the other hand, determines the dynamic prices of units of bandwidth so that the arrival rates for the QoS classes achieve the maximum possible bandwidth utilization while maintaining a certain fairness level (i.e., absolute fairness in this case). Hence, our scheme achieves better fairness as shown in Table II. An interesting result revealed from Table II is that even though we set $\upsilon_i$ to 1/K in our scheme, the bandwidth share of class 1 is still higher than that of class 2. This is due mainly to the high bandwidth requests of class 1 users in off-peak hours where users are not affected by our dynamic prices. This is clearly shown in Figure 4, which depicts the bandwidth share of class 1 in each hour of the day. In peak hour periods, users are more affected by dynamic prices, and hence their call requests arrival rates equal the optimal computed ones, hence, achieving 1/K bandwidth share. In fact, as more users decide to make call requests as a result of lowering the prices, the bandwidth share of both classes approaches 1/K because as aforementioned, prices are designed to achieve an equal share of bandwidth in our experiments. This explains the increased fairness of our scheme in Figure 4 as more users tend to make call requests during off-peak hours.

Table III shows the total revenue collected throughout the day for our scheme and CCAC. Our scheme clearly outperforms CCAC in terms of revenues due to charging users higher prices during peak hours. In addition, as more users decide to make call requests, more revenues can be collected. The revenue collected from class 1 users is higher than that from class 2 users because the formers pay higher prices for class 1 services in addition to requesting double the amount of bandwidth. It should be noted that more revenue can be earned if more bandwidth is assigned to class 1 (i.e., if class 1 is assigned more than 1/K bandwidth share). Therefore, the fairness constraint in (2) can be used also to increase revenues by assigning more bandwidth to classes that are expected to yield higher revenues.

## V. Conclusions and Future Work

Wireless network services have encountered an enormous demand in the past few years. This trend is expected to continue in the future due to the emergence of new broadband wireless technologies that are able to support high data rates, hence, offering a wide range of multimedia services. Due to scarcity of radio resources, wireless bandwidth must be managed in a way that maximizes the efficiency of the wireless network and meet the requirements of both network operators and users.

In this paper, a novel admission-level bandwidth management scheme is proposed for next generation broadband wireless access systems. The proposed scheme consists of two components namely, the Call Admission Control (CAC) component and the dynamic pricing component. Our scheme aims at providing monetary incentives to users to use the wireless bandwidth efficiently and rationally. By dynamically computing the prices of wireless services according to the network load, our scheme is able to prevent congestion while increasing the utilization of the network. Our scheme is simple to compute and can work with any CAC and dynamic pricing functions due to the separation of the CAC function and dynamic price computation.

Dynamic pricing, however, can guarantee to prevent congestion only if the users' demand models are accurate. Therefore, we are currently investigating the effect of inaccurate demand models on the system performance and how such inaccuracy can be efficiently dealt with.

## References


[1] 3GPP TS 25.308, "High Speed Downlink Packet Access (HSDPA); Overall Description", Release 5, March 2003.

[2] IEEE 802.16 Working Group, "IEEE 802.16-2005e Standard for Local and Metropolitan Area Networks: Air interface for fixed broadband wireless access systems – amendment for physical and medium access control layers for combined fixed and mobile operation in licensed bands", December 2005.

[3] J. Hou, J. Yang and S. Papavassilliou, "Integration of Pricing with Call Admission Control to Meet QoS Requirements in Cellular Networks", *IEEE Transactions on Parallel and Distributed Systems*, vol. 13, no. 9, pp. 898-910, September 2002.

[4] S. Yaipairoj and F.C. Harmantzis, "Congestion Pricing with Alternatives for Mobile Networks", *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, Atlanta, U.S.A, vol. 4, pp. 671-676, March 2004.

[5] S.L. Hew and L. B. White, "Optimal Integrated Call Admission Control and Congestion Pricing with Handoffs and Price-Affected Arrivals", *Proceedings of the Asian-Pacific Conference on Communications (APCC)*, Perth, Australia, pp. 396-400. October 2005.

[6] S. Mandal, D. Saha and A. Mahanti, "A Technique to Support Congestion Pricing Strategy for Differentiated Cellular Mobile Services", *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM)*, St. Louis, U.S.A, vol. 6, pp. 3388-3392, December 2005.

[7] S. Yaipairoj and F.C. Harmantzis, "Auction-based Congestion Pricing for Wireless Data Services", *Proceedings of the IEEE International Conference on Communications (ICC)*, Istanbul, Turkey, pp. 1059-1064, June 2006.

[8] S. Mandal, D. Saha and M. Chatterjee, "Pricing Wireless Network Services Using Smart Market Models", *Proceedings of the IEEE Consumer Communications and Networking Conference (CCNC)*, Las Vegas, U.S.A., vol. 1, pp. 574-578, January 2006.

[9] S. Mandal, D. Saha and M. Chatterjee, "Dynamic Price Discovering Models for Differentiated Wireless Services", *Journal of Communications*, vol. 1, no. 5, pp. 50-56, August 2006.

[10] E. Viterbo and C.F. Chiasserini, "*D*ynamic Pricing for Connection-Oriented Services in Wireless Networks", *Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, San Diego, U.S.A, vol. 1, pp. 68-72, September 2001.

[11] D. Hong and S.S. Rappaport, "Traffic Model and Performance Analysis for Cellular Mobile Radio Telephone Systems with Prioritized and None-prioritized Handoff Procedures," *IEEE Transactions of Vehicular Technology*, vol. 35, no. 3, pp. 77-92, August 1986.

[12] D. Liu and Y. Zhang, "A Self-Learning Adaptive Critic Approach for Call Admission Control in Wireless Cellular Networks", *Proceedings of the IEEE International Conference on Communications (ICC)*, Anchorage, U.S.A., vol. 3, pp.1853-1857, May 2003.

[13] R. L. Freeman, "Telecommunication System Engineering", 3rd edition, Wiley, 1996.

[14] B. Al-Manthari, N. Nasser and H. Hassanein, "Packet Scheduling in 3.5G High-Speed Downlink Packet Access Networks: Breadth and Depth", *IEEE Network Magazine*, vol. 21, no. 1, pp. 41-46, January 2007.

[15] E. D. Fitkov-Norris, A Khanifar, "Congestion pricing in Cellular Networks, A Mobility Model with a Provider-Oriented Approach", *Proceedings of the IEEE International Conference on 3G Mobile Communication Technologies*, London, UK, pp. 63-67, March 2001.


TABLE II
PERCENTAGE OF BANDWIDTH SHARE

| Scheme | Class 1 | Class 2 |
|---|---|---|
| BMS+0% | 59.956% | 40.044% |
| BMS+10% | 57.104% | 42.896% |
| BMS+30% | 53.420% | 46.580% |
| BMS+50% | 51.536% | 48.464% |
| CCAC | 69.344% | 30.656% |

TABLE III
TOTAL EARNED REVENUE DURING THE DAY (UNITS OF MONEY)

| Scheme | Total Revenue | Class 1 | Class 2 |
|---|---|---|---|
| BMS+0% | 542 x10 $^{4}$ | 353 x10 $^{4}$ | 189 x10 $^{4}$ |
| BMS+10% | 562 x10 $^{4}$ | 411 x10 $^{4}$ | 151 x10 $^{4}$ |
| BMS+30% | 591 x10 $^{4}$ | 387 x10 $^{4}$ | 204 x10 $^{4}$ |
| BMS+50% | 627 x10 $^{4}$ | 402 x10 $^{4}$ | 225 x10 $^{4}$ |
| CCAC | 523x10 $^{4}$ | 355 x10 $^{4}$ | 168 x10 $^{4}$ |

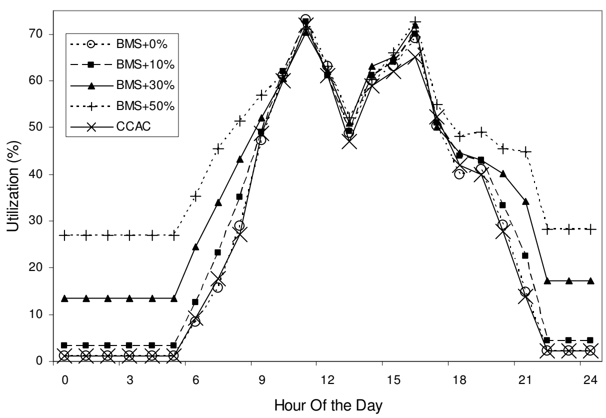


Figure 2. Percentage of Bandwidth Utilization at Different Hours of the Day

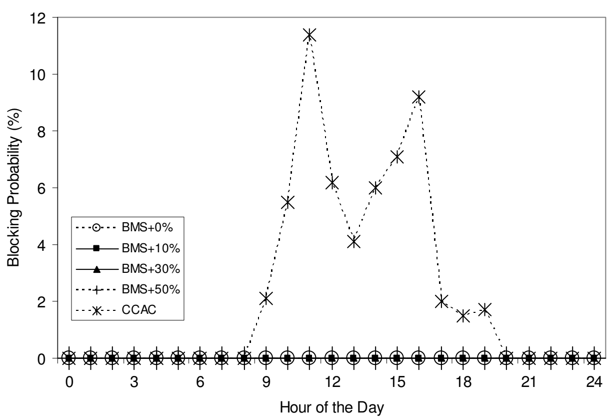


Figure 3. Call Blocking Probability at Different Hours of the Day

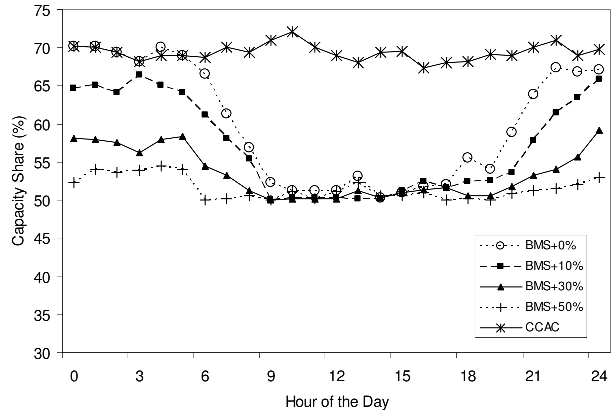


Figure 4. Percentage of Bandwidth Share for Class 1 at Different Hours of the Day