Performance Evaluation 68 (2011) 768–781

Contents lists available at ScienceDirect

Performance Evaluation

journal homepage: www.elsevier.com/locate/peva

Dynamic multiple-frame bandwidth provisioning with fairness and revenue considerations for Broadband Wireless Access Systems ☆

Bader Al-Manthari [a], Najah Abu Ali [b], Nidal Nasser [c], Hossam Hassanein [d,*]

[a] Center of Information Security, Information Technology Authority, Azaibah, P.O. Box: 1807, P.C: 130, Oman
[b] College of Information Technology, UAE University, Al-Ain, P.O. 17555, United Arab Emirates
[c] Department of Computing & Information Science, University of Guelph, Guelph, ON, Canada, N1G 2W1
[d] Telecommunications Research Laboratory, School of Computing, Queen's University, Kingston, ON, Canada, K7L 3N6

ARTICLE INFO

Article history:
Available online 21 December 2010

Keywords:
Multiple-frame bandwidth provisioning
Fairness
Opportunity cost
Dynamic weights

ABSTRACT

The increasing demand for wireless heterogeneous multimedia services presents a real challenge to mobile network operators. Even with the substantial increase in the supported bandwidth in emerging Broadband Wireless Access Systems (BWASs) such as 3.5G High Speed Downlink Packet Access (HSDPA) and the Worldwide Interoperability for Microwave Access (WiMAX), satisfying the bandwidth requirements of mobile users while increasing the revenues of network operators is still one of the major issues in these systems. Therefore, bandwidth provisioning will certainly play a decisive role in the success of such BWASs. In this paper, we propose a novel dynamic multiple-frame bandwidth provisioning scheme for BWASs. The proposed scheme spans multiple time frames and optimally allocates them to the different classes of traffic depending on their weights, the real-time bandwidth requirements of their users' connections, their channel quality conditions and the expected obtained revenues. To maximize fairness and still maintain service differentiation between classes, we provide a unique formulation for dynamically computing the class weights. Simulation results are provided to show the potential and effectiveness of our scheme.

1. Introduction

Emerging Broadband Wireless Access Systems (BWASs) such as High Speed Downlink Packet Access (HSDPA) [1] and the Worldwide Interoperability for Microwave Access (WiMAX) [2], pose a myriad of new opportunities for leveraging the support of a wide range of multimedia services with diverse bandwidth requirements. This is due to the remarkably high bandwidth that is supported by these systems, which was only available to wireline connections. To maximize their efficiency and reduce the cost of data delivery, these systems utilize high speed downlink channels that are shared among mobile users through packet-level bandwidth management, which is typically referred to as packet scheduling. Packet scheduling is one of the most prominent components of resource management that affects system capacity and potential bandwidth allocated to mobile users. A centralized downlink packet scheduler is implemented at the base stations of BWASs to control the allocation of the downlink shared channels to the mobile users by deciding which of them should transmit during a given time frame.

☆ An earlier much abbreviated appear in *Proceedings of the ACS/IEEE International Conference on Computer Systems and Applications (AICCSA)*, Doha, Qatar, pp. 999–1004, April 2008.

* Corresponding author.
E-mail addresses: bader.almanthari@ita.gov.om (B. Al-Manthari), najah@uaeu.ac.ae (N.A. Ali), nasser@cis.uoguelph.ca (N. Nasser), hossam@cs.queensu.ca (H. Hassanein).

0166-5316/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.peva.2010.11.001

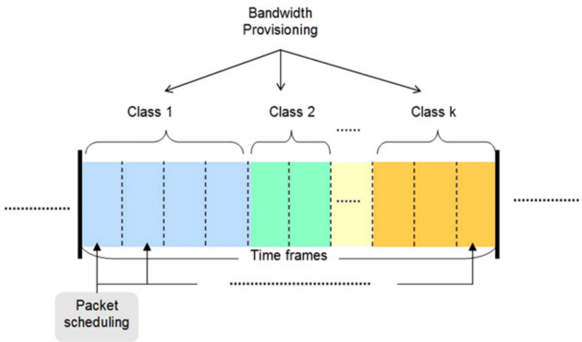


Fig. 1. Dynamic bandwidth provisioning.

Packet scheduling, however, cannot by itself achieve optimized bandwidth management. This is because it is a single-frame resource sharing scheme which only checks the current time frame to make its decision. Therefore, to maintain acceptable levels of Quality of Service (QoS) throughout the lifetime of users' connections, packet scheduling must be coupled with multiple-frame resource sharing or bandwidth provisioning scheme. The bandwidth provisioning scheme spans multiple time frames and decides how resources are shared among the different classes of traffic, and hence their corresponding users' connections. The bandwidth provisioning scheme must divide the bandwidth or the corresponding time frames between different classes of traffic in an efficient, prioritized and fair way. Thus, the scheme works as a multiple-frame inter-class selection algorithm as shown in Fig. 1.[1] Multiple-frame bandwidth provisioning can be interpreted as longer-term post admission bandwidth management that aims at satisfying the long-term bandwidth requirements of users during the lifetime of their connections, as opposed to packet scheduling, which only provisions over single time frames.

Most of the work on bandwidth provisioning has been done at the admission level [3–10] where these schemes implement Call Admission Control (CAC) and aim at maximizing the number of admitted users' connections while satisfying the bandwidth requirements of different classes of traffic. Bandwidth management at the admission level is very important in improving the performance of BWASs because of its role in maintaining the QoS of ongoing users' connections at acceptable levels especially during congested periods. There is a need, however, for bandwidth provisioning at the frame level (i.e., during the lifetime of users' connections). This is due to the varying bandwidth requirements of mobile users during the lifetime of their connections as a result of their traffic burstiness and also due to their varying channel quality conditions, which affect the capacity of the base station, and hence the amount of bandwidth that it can sustain to each one of them. Little research work has considered the problem of bandwidth provisioning at the frame level [11–15]. The scheme in [11] aims at minimizing the expected number of packets awaiting transmission for each user in order to reduce the overall system delay. It supports prioritization between users belonging to different classes of traffic. However, it does not support users with different bandwidth requirements. Therefore, users with a higher number of packets in their queues can get more bandwidth regardless of the bandwidth required by other users in the system. In addition, to increase the efficiency of the system, the scheme assigns higher priorities to users having higher "probability of connectivity" between them and the base station, where the probability of connectivity is used as a measure of the channel quality conditions of users. This measure, however, does not reflect the actual instantaneous data rates that the users can send or receive at, which depend on their instantaneous channel quality conditions. Using this measure, the scheme may consequently assign more/less bandwidth than what is actually needed by users.

The scheme in [12] divides the number of slots in each time frame between different classes of traffic so that the frame-level connection blocking probability of each class (i.e., the probability that connections within each class are blocked and not assigned time slots in the current frame) is minimized. Unlike the scheme in [11], the scheme in [12] considers the instantaneous channel quality conditions of users as well as their minimum bandwidth requirements in the slot allocation process.

The bandwidth provisioning schemes in [13–15] do not consider the varying channel quality conditions of mobile users. Hence, they cannot achieve optimized bandwidth provisioning. In addition, these schemes provide very limited QoS support, and hence they are incapable of supporting many multimedia services in BWASs.

We remark that the schemes in [11–15] are designed to allocate slots within one time frame. However, as aforementioned, to maintain the QoS of ongoing users at acceptable levels throughout the lifetime of their connections, there is a need for bandwidth provisioning over multiple time frames. In addition, these schemes lack support for fairness[2] between different classes. Hence, they may result in unfair allocation of bandwidth, where users with good channel quality conditions and/or high bandwidth requirements may monopolize the whole bandwidth. Furthermore, none of these

[1] Note that the frames allocated per class need not be consecutive and they are only depicted this way for illustration purposes.

[2] We define fairness in the context of this paper as the number of frames assigned to each class relative to their required ones. See Section 4.3 for the definition of the proportion of assigned frames, which is used as a fairness measure in the performance evaluation of our scheme.

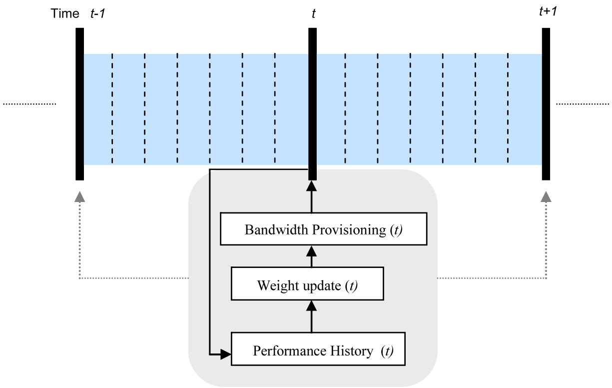


Fig. 2. Dynamic bandwidth provisioning with weight update.

schemes considers the revenues of network operators when allocating the time slots. As a result, these schemes may not be desired by network operators, who are certainly concerned with maximizing their revenues. Therefore, there is a need for a bandwidth provisioning scheme that is able to allocate multiple time frames and provide fairness between classes of traffic while considering the revenues of network operators.

In this paper, we propose a dynamic bandwidth provisioning scheme for BWASs. Our proposed scheme is designed to accommodate multi-class traffic with multiple users' connections having different bandwidth requirements and varying channel quality conditions. The main objective of our scheme is to optimally allocate bandwidth or the corresponding time frames for each class of traffic in order to satisfy the bandwidth requirements of their connections. In addition, the proposed scheme uniquely incorporates and bounds the cost of bandwidth provisioning (in terms of revenue loss) through an opportunity cost function. This provides greater flexibility to network operators to determine the levels of bandwidth provisioning to different classes that will guarantee a certain level of revenues. Moreover, the scheme allows prioritized bandwidth provisioning between classes through the use of class weights. To maximize inter-class fairness, we propose a weight update scheme to dynamically compute class weights based on performance history of classes. The weight update scheme distinctively ensures fairness while at the same time guarantees long-term service differentiation between classes.

The rest of this paper is organized as follows. Section 2 provides an overview of the proposed scheme. Section 3 presents a description of the proposed scheme. Performance results are presented in Section 4. Finally, conclusions drawn from the paper are discussed in Section 5.

2. Overview of dynamic bandwidth provisioning scheme

In this section, we provide an overview of our proposed dynamic bandwidth provisioning scheme. We consider that there are K classes of traffic, where class i has higher priority than class $i+1$, $1 \leq i$ and $i+1 \leq K$. We consider that connections within the same class may have different bandwidth requirements depending on the type of services they request. Suppose that the network operator wants to provision *NF* time frames between the K classes of traffic to satisfy the long-term bandwidth requirements of their users' connections. We assume that *NF* is given. In practice, simulation studies or real experiments can be used to determine empirically the appropriate value of *NF* that achieves the performance levels desired by the network operator. The proposed dynamic bandwidth provisioning scheme works as follows. At time t (where t is the beginning of *NF* frames), the base station will evaluate the performance history of the K classes and will use this information to update their weights to maximize inter-class fairness as described in Section 3. These weights are then used as input parameters to the proposed bandwidth provisioning scheme, which will partition the next *NF* frames among the K classes based on their weights, the bandwidth requirements of user connections, users' channel quality conditions and the expected revenues.

Once each class is assigned a number of frames, these frames will be distributed to connections within each class according to the packet scheduling scheme, which is executed every time frame. The network operator may utilize any existing packet scheduling scheme for distributing the partitioned frames among users' connections as our bandwidth provisioning scheme is independent from the scheduling algorithm being utilized. Fig. 2 shows an abstract timeline data flow chart of the proposed dynamic bandwidth provisioning scheme.

3. Dynamic bandwidth provisioning

We distinguish two cases of bandwidth provisioning. In the first no bandwidth guarantees are required for any class. We then extend our scheme to support minimum bandwidth guarantees. Next, we explain the weight update scheme to ensure fairness between classes.

3.1. Basic bandwidth provisioning

Before proceeding with describing the proposed bandwidth provisioning scheme, we make the following definitions. Let:

- N_i: number of class i user connections.
- $N = \sum_{i=1}^{K} N_i$: total number of user connections in the system.
- $S_{ij}^{\max}$: maximum data rate required by user j of class i, $j = 1, \ldots, N_i$.
- NF_i: number of frames allocated to class i.
- $\overline{R_i(t)}$: effective average estimated data rate (per second) that the base station can transmit to class i users during the next NF frames. This data rate will depend on the estimated instantaneous channel quality conditions of the users and their bandwidth requirements. $\overline{R_i(t)}$ can be roughly estimated using a moving average (i.e., $\overline{R_i(t)} = \alpha \cdot \overline{R_i(t-1)} + (1-\alpha) \cdot R_i(t)$, where $0 \le \alpha \le 1$) or using channel prediction schemes proposed in [16–18].
- p_{ij}: price per bit for user j of class i.
- $B_i^{\max}$: total required maximum data rate per frame of all users in class i at the beginning of the NF frames. Let $\sum_{j=1}^{N_i} S_{ij}^{\max}$ be the total required maximum data rate per second of all users in class i, and let D_{frame} be the frame duration in seconds, then $B_i^{\max} = \left(\left(\sum_{j=1}^{N_i} S_{ij}^{\max}\right)/1/D_{\text{frame}}\right)$. The quantity $B_i^{\max}$ determines the transmission rate the base station should be sending at per frame, in the next NF frames, in order to satisfy the maximum data rate requirements of class i users.
- $\overline{B_i} = \left((\overline{R_i(t)}/(1/D_{\text{frame}}))\right)$: actual (i.e., effective) total transmitted data rate per frame for class i users. That is, $\overline{B_i}$ determines the actual transmission rate per frame of the base station in the next NF frames for class i users.
- $\text{Rev}_i^{\max}$: total maximum revenue per frame of class i users at the beginning of the NF frames. Therefore, $\text{Rev}_i^{\max} = \left(\left(\sum_{j=1}^{N_i} p_{ij} \cdot S_{ij}^{\max}\right)/1/D_{\text{frame}}\right)$. The quantity $\text{Rev}_i^{\max}$ determines the revenue of the network operator per frame in the next NF frames if it grants all the users in class i their maximum required data rates. Therefore, $NF \cdot \text{Rev}_i^{\max}$ is the upper bound of the total revenue of the network operator during the next NF frames.
- $\overline{\text{Rev}_i} = \left(\left(\sum_{j=1}^{N_i} p_{ij} \cdot \overline{R_i(t)}\right)/1/D_{\text{frame}}\right)$: effective total revenue per frame actually generated from serving all users in class i. Therefore, $NF_i \cdot \overline{\text{Rev}_i}$ is the actual total revenue that the network operator earns from serving all users of class i provided that class i is allocated NF_i frames.
- $\{\overline{\mathbf{Rev}_i^z}\}_{z=1}^{K}$: descending ordered set of the actual effective total revenue per frame resulting from serving the K classes.

To satisfy all users, the base station should allocate a data rate of $NF \cdot B_i^{\max}$ per NF frame to class i, $1 \le i \le K$. However, this may not be possible in practice because of the high demand of services that have high bandwidth requirements and also because of the limitations of the base station's capacity, which is determined by the channel quality conditions of the mobile users. Therefore, the main objective of our bandwidth provisioning scheme is to allocate the NF frames among the K classes of traffic such that $\sum_{i=1}^{K} NF_i = NF$ and the satisfaction of different users is maximized. To this end, our bandwidth provisioning scheme will distribute the NF frames among the K classes of traffic so that it maximizes the ratio between the data rate allocated to class i users, given that it is assigned NF_i frames, to the data rate that the base station should transmit at during NF frames to satisfy the maximum data rate requirement of class i (i.e., $\frac{NF_i \cdot \overline{B_i}}{NF \cdot B_i^{\max}}$). The frames allocated to class i (i.e., NF_i) should guarantee that no class is allocated more than its maximum required data rate (i.e., $NF_i \cdot \overline{B_i} \le NF \cdot B_i^{\max}$).

In addition, it is imperative to realize that there is an opportunity cost[3] of frame allocation (i.e., bandwidth provisioning). The opportunity cost (in terms of revenue) of frame allocation is the maximum revenue that the network operator will earn if it serves the highest revenue generating classes minus the revenue that it will earn by allocating the frames otherwise. To compute the maximum revenue that the network operator could earn in the next NF frames, we first need to know the number of frames needed by each class (NF_i^{req}) in order to achieve the maximum required data rate by its users (i.e., $NF_i^{\text{req}} \cdot \overline{B_i} = NF \cdot B_i^{\max}$). Hence, $NF_i^{\text{req}} = \frac{NF \cdot B_i^{\max}}{\overline{B_i}}$. Therefore, the maximum revenue at the class level, Max Rev_c, is equal to

$$\text{Max Rev}_c = \sum_{z \in \{\text{Rev}_i^z\}_{z=1}^{K}} NF_i^{\text{req}}.\text{Rev}_i^z, \quad \text{given that} \sum_{z \in \{\text{Rev}_i^z\}_{z=1}^{K}} NF_i^{\text{req}} \le NF. \tag{1}$$

The maximum revenue is obtained by allocating the frames to the class with the highest actual revenue. If this class can be served by fewer than NF frames, the remaining frames are allocated to the class with the second-highest actual revenue, and so forth. Therefore, the opportunity cost ($\text{OC}(NF)$) of the frame allocation at the class level is equal to

$$\text{OC}(NF) = \text{Max Rev}_c - \left(\sum_{i=1}^{K} NF_i \cdot \overline{\text{Rev}_i}\right). \tag{2}$$

[3] The opportunity cost for a good is defined as the value of any other goods or services that a person must give up in order to produce or get that good [19].

This should be less than or equal to a predefined value H_c. For example, the network operator could restrict the revenue loss to be no more than 30% of the maximum obtainable revenue (i.e., $H_c = \zeta_c \cdot \text{Max Rev}_c$, where $\zeta_c = 0.3$).

To summarize, in our multiple-frame bandwidth provisioning scheme, the following optimization problem will be solved

$$\begin{aligned} &\text{Objective}: \max_{NF_i, 1\le i\le K} \sum_{i=1}^{K} w_i \cdot \left(\frac{NF_i \cdot \overline{B_i}}{NF \cdot B_i^{\max}}\right) \\ &\text{Subject to:} \quad \sum_{i=1}^{K} NF_i = NF, \\ &\qquad NF_i.\overline{B_i} \le NF.B_i^{\max}, \quad 1 \le i \le K, \quad \text{and} \\ &\qquad \text{OC}(NF) \le H_c \end{aligned} \tag{3}$$

where w_i is a weight assigned to class i to give it priority over class $i+1$ in the frame allocation process. Since the objective function and the constraints are linear, our bandwidth provisioning scheme can be solved using Linear Programming (LP) techniques.

The proposed bandwidth provisioning scheme is adaptive to the varying requirements of different classes of traffic, since the objective function is evaluated every NF frame. Therefore, if the required bandwidth (or frames) of class i changes during the current frames (due to new admitted connections and completed ones or bandwidth adaptive requests as it is the case in WiMAX), its new total required bandwidth will be reflected in the next NF frames.

3.2. Bandwidth provisioning with minimum guaranteed bandwidth

Even though the dynamic bandwidth provisioning scheme in Section 3.1 aims at maximizing the satisfaction of the different users, it does not provide bandwidth guarantees to traffic classes. The network operator may want to provide such guarantees. Therefore, the bandwidth provisioning scheme should consider such a case. Here, we extend our scheme to support minimum bandwidth guarantees. Let:

- $S_{ij}^{\min}$: minimum required data rate of user j of class i, $j = 1, \ldots, N_i$.
- $B_i^{\min}$: total required minimum data rate per frame of all users in class i at the beginning of the NF frames. That is, $B_i^{\min} = \left(\left(\sum_{j=1}^{N_i} S_{ij}^{\min}\right) / 1/D_{\text{frame}}\right)$.

Therefore, in this case, the dynamic bandwidth provisioning scheme should guarantee that no class of traffic is allocated less than its minimum required data rate (i.e., $NF \cdot B_i^{\min} \le NF_i \cdot \overline{B_i}$) or allocated more than its maximum required data rate (i.e., $NF_i \cdot \overline{B_i} \le NF \cdot B_i^{\max}$). Therefore, the same problem in Eq. (3) will be solved except that the bandwidth constraint changes to

$$NF.B_i^{\min} \le NF_i.\overline{B_i} \le NF.B_i^{\max}, \quad 1 \le i \le K. \tag{4}$$

It is imperative to point out that if the network operator wants to provide minimum bandwidth guarantees to some classes, the optimization problem in Eq. (3) may not have a feasible solution. This is because the bandwidth provisioning scheme may have to allocate a certain number of time frames to certain classes of traffic in order to satisfy their minimum bandwidth requirements even though they do not satisfy the opportunity cost constraint. Therefore, to satisfy both constraints, the bound on the opportunity cost (i.e., H_c) has to be dynamically computed in order to ensure the existence of a feasible solution of Eq. (3) as follows. Let:

- $\text{Rev}_{\mathbf{K}^*} = \sum_{i\in\mathbf{K}^*} \overline{\text{Rev}_i}$, where $\mathbf{K}^* \in \{1, 2, \ldots, K\}$ is the set of classes that require minimum bandwidth guarantees. That is, $\text{Rev}_{\mathbf{K}^*}$ is the revenue the network operator will earn from serving these classes.

In this case, the opportunity cost of serving the classes in $\mathbf{K}^*$ is given by $\text{OC}_{\mathbf{K}^*}(NF) = \text{Max Rev}_c - \text{Rev}_{\mathbf{K}^*}$, where Max Rev_c is defined in Section 3.1. Therefore, to avoid infeasibility in Eq. (3), we must have $H_c \ge \text{OC}_{\mathbf{K}^*}(NF)$. The network operator could, for example, set a predefined value for H_c, call it ϑ_c and use it only when $H_c \ge \text{OC}_{\mathbf{K}^*}(NF)$ is satisfied as follows

$$H_c = \begin{cases} \text{OC}_{\mathbf{K}^*}(NF), & \text{if } \vartheta_c \le \text{OC}_{\mathbf{K}^*}(NF) \\ \vartheta_c, & \text{otherwise.} \end{cases} \tag{5}$$

3.3. Weight update scheme

The weights in our bandwidth provisioning scheme determine the priority of each class, and hence they have a great impact on the frame allocation process and user satisfaction. In this section, we show how to update the weights to increase inter-class fairness of the bandwidth provisioning scheme while maintaining a long-term service differentiation between them. Let:

- $U_i(t) = \frac{NF_i.\overline{B_i^{\text{effec}}}}{NF.B_i^{\max}}$: utility of class i at the beginning of time t, where $\overline{B_i^{\text{effec}}}$ is the actual average data rate of class i (i.e., the effective rate class i was sending at during the previous frames) and t is the time at the end of the previous NF frames and the beginning of new ones (i.e., the beginning of a new bandwidth provisioning period). The higher the data rate assigned to class i, the higher its utility.
- $\overline{U_i(t)} = \alpha \cdot \overline{U_i(t-1)} + (1-\alpha) \cdot U_i(t)$: average utility of class i at time t, computed as a moving average, where $0 \le \alpha \le 1$.
- $w_i(t)$: weight of class i at time t.
- $\overline{w_i(t)} = \alpha \cdot \overline{w_i(t-1)} + (1-\alpha) \cdot w_i(t)$: average weight of class i, where $0 \le \alpha \le 1$.
- $\{\mathbf{LA}^z(t)\}_{z=1}^{h_i} = \{\text{LA}^1(t), \text{LA}^2(t), \dots, \text{LA}^{h_i}(t)\}$: set of average utilities that are larger than the average utility of class i at time t, where h_i is the number of classes whose average utilities are larger than the average utility of class i.
- $\{\mathbf{LO}^q(t)\}_{q=1}^{l_i} = \{\text{LO}^1(t), \text{LO}^2(t), \dots, \text{LO}^{l_i}(t)\}$: set of average utilities that are lower than the average utility of class i at time t, where l_i is the number of classes whose average utilities are lower than the average utility of class i.

Three design features are taken into consideration in developing our weight update scheme. First, the weight of each class is gradually increased or decreased depending on its performance history and all other classes' performance histories. In particular, the weight of class i is increased or decreased depending on the difference between its average utility, the average utilities that are larger than it (i.e., set $\{\mathbf{LA}^z(t)\}_{z=1}^{h_i}$) and the average utilities that are smaller than it (i.e., set $\{\mathbf{LO}^q(t)\}_{q=1}^{l_i}$). To achieve this, the new weight of class i at time t is updated as follows

$$w_i(t) = w_i(t-1) + \Delta w_i \tag{6}$$

where

$$\Delta w_i = \frac{\sum\limits_{z\in\{\mathbf{LA}^z(t)\}_{z=1}^{h_i}} \left(\text{LA}^z - \overline{U_i(t)}\right) - \sum\limits_{q\in\{\mathbf{LO}^q(t)\}_{q=1}^{l_i}} \left(\overline{U_i(t)} - \text{LO}^q\right)}{\sum\limits_{i=1}^{K} \overline{U_i(t)}}. \tag{7}$$

Note that Δw_i is the normalized difference between the average utilities that are larger than the average utility of class i and the average utilities that are less than it. The quantity Δw_i can be thought as a performance measure. It increases as the difference between the average utilities in $\{\mathbf{LA}^z(t)\}_{z=1}^{h_i}$ and average utility of class i increases and it decreases as the difference between the average utility of class i and the average utilities in $\{\mathbf{LO}^q(t)\}_{q=1}^{l_i}$ increases. Note that Δw_i is negative when $\sum_{z\in\{\mathbf{LA}^z(t)\}_{z=1}^{h_i}} \left(\text{LA}^z - \overline{U_i(t)}\right) < \sum_{q\in\{\mathbf{LO}^q(t)\}_{q=1}^{l_i}} \left(\overline{U_i(t)} - \text{LO}^q\right)$. Δw_i is negative when the difference between class i and the classes of lower average utilities is higher than the difference between the classes of higher average utilities and class i. In this case, it is best to decrease the weight of class i to allow classes of lower average utilities to be allocated more bandwidth, and hence increase inter-class fairness.

The second design feature is that the weights of lower priority classes are allowed to be temporarily higher than those of higher priority ones to further increase inter-class fairness. However, to ensure service differentiation between classes, we require that the ratio between the average weight of each class and the next class that has a higher priority does not exceed a certain threshold $0 < \tau_i < 1$ (i.e., $\frac{\overline{w_i(t)}}{\overline{w_{i-1}(t)}} \le \tau_i$, where $\overline{w_1(t)} > \overline{w_2(t)} > \cdots > \overline{w_K(t)}$). This guarantees a long-term service differentiation between classes by ensuring that the long-term average weight of class i is less than or equal to $\tau_i \cdot \overline{w_{i-1}(t)}$.

An additional design feature is to restrict the weights to fall within a certain range as determined by the network operator (i.e., $W_{\min} \le w_i(t) \le W_{\max}$) in order to ensure that the weight update does not result in extremely high or low weight values. Following our design features, the weight of each class is updated as follows

$$w_i(t) = \max\left(\min((w_i(t-1) + \Delta w_i), W_{\max}), W_{\min}\right). \tag{8}$$

That is, $w_i(t) = w_i(t-1) + \Delta w_i$ as long as $W_{\min} \le w_i(t-1) + \Delta w_i \le W_{\max}$. If $w_i(t-1) + \Delta w_i < W_{\min}$, then $w_i(t) = W_{\min}$. On the other hand, if $w_i(t-1) + \Delta w_i > W_{\max}$, then $w_i(t) = W_{\max}$. Note that Eq. (8) satisfies only the condition $W_{\min} \le w_i(t) \le W_{\max}$. Therefore, once $w_i(t)$ is computed, the condition $\frac{\overline{w_i(t)}}{\overline{w_{i-1}(t)}} \le \tau_i$ is checked. If it is not satisfied, then $w_i(t)$ is recomputed as follows

$$w_i(t) = \frac{\tau_i \cdot \left(\alpha \cdot \overline{w_{i-1}(t-1)} + (1-\alpha) \cdot w_{i-1}(t)\right) - \alpha \cdot \overline{w_i(t-1)}}{(1-\alpha)}. \tag{9}$$

That is, $w_i(t)$ is computed such that $\frac{\overline{w_i(t)}}{\overline{w_{i-1}(t)}} = \tau_i$ as follows

$$\frac{\overline{w_i(t)}}{\overline{w_{i-1}(t)}} = \tau_i, \quad \therefore \frac{\alpha \cdot \overline{w_i(t-1)} + (1-\alpha) \cdot w_i(t)}{\alpha \cdot \overline{w_{i-1}(t-1)} + (1-\alpha) \cdot w_{i-1}(t)} = \tau_i. \tag{10}$$

Table 1
Simulation parameters.

Simulation time	400 s
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

Rearranging the terms

$$\alpha \cdot \overline{w_i(t-1)} + (1-\alpha) \cdot w_i(t) - \tau_i \cdot \left(\alpha \cdot \overline{w_{i-1}(t-1)} + (1-\alpha) \cdot w_{i-1}(t)\right) = 0. \tag{11}$$

Therefore,

$$w_i(t) = \frac{\tau_i \cdot \left(\alpha \cdot \overline{w_{i-1}(t-1)} + (1-\alpha) \cdot w_{i-1}(t)\right) - \alpha \cdot \overline{w_i(t-1)}}{(1-\alpha)}. \tag{12}$$

3.4. Packet scheduling

Once each class is assigned a number of frames, these frames will be shared among users' connections within each class according to the packet scheduling scheme, which is executed every time frame. These frames can be served in any order. For example, they could be served based on the delay or packet loss requirements of the service classes. In this paper, however, the frames of the class with the highest priority are served first, then those of the class of the second highest priority, etc. In the performance evaluation, which is the subject of next section, we utilize a packet scheduling scheme that we proposed in [20], though our proposed dynamic bandwidth provisioning scheme can utilize any other packet schedulers as aforementioned. This scheduler is used because it has been shown to increase the system's capacity and provide intra-class fairness simultaneously.

4. Performance evaluation

In this section, we evaluate the performance of our proposed scheme by means of dynamic discrete event simulation. We tested our scheme on an 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) [1].

4.1. Simulation model

Since our proposed scheme is implemented at every base station in the network, its performance within one base station is independent of the performance of other base stations. Therefore, we considered a single-cell scenario with one base station. The base station is located at the center of the cell. The cell's radius is 1 km and the base station's transmission power is 38 dBm. At initialization, N connections are uniformly distributed in the cell. The Pedestrian A environment is used in our experiments where every mobile connection moves inside the cell with a constant speed of 3 km/h [21]. This speed is the recommended value for The Pedestrian A environment by the 3GPP [21]. Unless otherwise specified, call arrivals are modeled as a Poisson process with a mean value of 0.5 calls per second. The LP problem of Eq. (3) is solved using *lp-solve*, which is a free Linear/Integer Programming solver [22].

Connections are uniformly distributed in the cell. We choose $NF = 20$ time frames (i.e. 20×2 ms) and we use a moving average to compute $\overline{R_i(t)}$ (i.e., $\overline{R_i(t)} = \alpha \cdot \overline{R_i(t-1)} + (1-\alpha) \cdot R_i(t)$), where $\alpha = 0.99$. We adopt the same channel model as in [23]. The simulation time step is one time frame, which is 2 ms in HSDPA [1], and the simulation time is 400 s. Other simulation parameters are listed in Table 1.

4.2. Traffic model

To demonstrate the ability of our scheme to support connections having different QoS requirements, we assume three different classes with four different types of traffic namely VoIP (class 1), audio streaming (class 2), video streaming (class 2) and FTP (class 3). In addition, to demonstrate the ability of our scheme to prioritize different classes of traffic (i.e., inter-class prioritization), we assume that class 1 has the highest priority and class 3 has the lowest priority. Moreover, we assume that audio streaming has a higher priority than video streaming. Furthermore, for demonstration purposes, we assume that $p_{ij} = 6, 4, 2$ and 1 units of money for VoIP, audio streaming, video streaming and FTP connections, respectively.

For VoIP traffic, we adopt the model in [24], which assumes the use of an Adaptive Multi-Rate (AMR) codec. In this model, packets are generated using a negative exponentially distributed ON-OFF traffic source to simulate the talk and silence spurts where the mean duration of both ON and OFF periods is 3 s. During the ON periods, a voice packet of 244 bits is generated every 20 ms, corresponding to a source bit rate of 12.2 kbps, which is comparable to one of the AMR bit rates [25]. The compressed IP/UDP/RTP header increases the bit rate to 13.6 kbps [26]. The ITU E-model [27] states that when the one-way mouth-to-ear delay exceeds 250 ms, the voice quality rating rapidly deteriorates. About 80–150 ms remains for the base station processing and connection reception when the delay induced by the voice encoder/decoder and other components in the system is subtracted [28]. Therefore, we set the maximum delay threshold for VoIP traffic to a value between 80 and 150 ms, specifically, 100 ms.

Audio streaming is modeled with a minimum rate of 12 kbps, mean rate of 38 kbps, maximum rate of 64 kbps, maximum packet delay of 150 ms and a packet size uniformly distributed between 244 and 488 bits. These values are chosen from within the range of specific QoS requirements defined by 3GPP in order to provide an adequate service to mobile users [29–31]. Video streaming is modeled with a minimum data rate of 64 kbps, mean rate of 224 kbps, maximum data rate of 384 kbps and a packet size uniformly distributed between 1200 and 2400 bits [29–31]. FTP traffic is simulated by a constant rate of 128 kbps and a fixed packet size of 1200 bits. Durations of VoIP, audio and video streaming users' connections are modeled by an exponential distribution with a mean value of 30 s. Whereas, in the case of FTP users it is assumed that each user requests one file of size 50 MB and terminates its connection after the file download is complete.

4.3. Performance metrics

Three test cases are considered in our experiments. In the first case, we evaluate the performance of packet scheduling with and without dynamic bandwidth provisioning. This case is designed to show the advantage of using dynamic bandwidth provisioning along with packet scheduling. In the second case, we evaluate our dynamic bandwidth provisioning scheme (with packet scheduling) under different fixed class weights and opportunity cost values for H_c in order to show their role in the bandwidth allocation process. In the third case, we evaluate our dynamic bandwidth provisioning scheme using our proposed weight update scheme. For this case, we set the minimum weight (i.e., $W_{\min}$) and maximum weight (i.e., $W_{\max}$) values to 1 and 10, respectively.

The following performance metrics are used to evaluate the performance of the proposed scheme:

- Proportion of assigned frames ($\overline{P_i}$): the average ratio of assigned frames to class *i* to the total number of frames needed to satisfy its maximum bandwidth requirements. If different classes of traffic achieve similar values for $\overline{P_i}$ then the bandwidth provisioning scheme is fair. Otherwise, it is unfair.
- Service coverage: percentage of connections, which achieve their required QoS with a certain outage level. For VoIP and audio streaming, a connection's call is considered an outage and, therefore, is dropped if its packet loss (due to packet discarding, transmission errors and/or buffer overflow) exceeds 5% [32–34]. For video streaming, a connection's call is considered an outage if its achieved average throughput is less than its minimum required rate. Finally, for FTP traffic, a connection's call is considered an outage if its achieved average throughput is less than 9.6 kbps [35].
- Per-class weights: we report the temporal and average values of the dynamic weights per class, as well as the 10th and 90th percentile values, defined as the values where 10% and 90% of measured weights are lower, respectively. These values are computed from the weights resulting from our weight update scheme.

4.4. Simulation results

4.4.1. Case 1: scheduling with and without dynamic bandwidth provisioning

Figs. 3 and 4 show the percentage of service coverage for VoIP and audio streaming before and after implementing our dynamic bandwidth provisioning scheme, with $w_1 = 6, w_2 = 4$ and $w_3 = 2$. As expected, the performance of VoIP and audio streaming is clearly improved when dynamic provisioning is implemented along with packet scheduling. This is because dynamic bandwidth provisioning aims at satisfying the users' bandwidth requirements over longer time intervals than packet scheduling, which only works in very small time intervals (i.e., single time frames). In other words, dynamic bandwidth provisioning improves the management of network resources, which results in more users meeting their minimum bandwidth requirements; hence, improving the overall service coverage. In addition, the figures show that, as the arrival rate to the system increases, the performance difference between the case of packet scheduling only and the case of packet scheduling with dynamic bandwidth provisioning increases. This is because at high arrival rates, more users compete for resources, and hence the performance difference between different bandwidth management techniques is clearly revealed.

The service coverage of video streaming is improved with dynamic bandwidth provisioning as shown in Fig. 5. Such an improvement is observed at arrival rates between 0.1 and 0.7 connections per second. At higher arrival rates, however, the service coverage of video streaming degrades to values below those of packet scheduling only. This is because, at high arrival rates, the demand for bandwidth is high and it is difficult to satisfy the bandwidth requests of all classes of traffic. In such a case, more time frames are allocated to VoIP and audio streaming than video streaming, since the former have higher priority. This leaves fewer frames for lower priority traffic, which negatively affects their performance in terms of service coverage. We remark, however, that the network operator would typically employ a CAC scheme in order to improve the satisfaction of users and prevent performance degradation at high arrival rates.

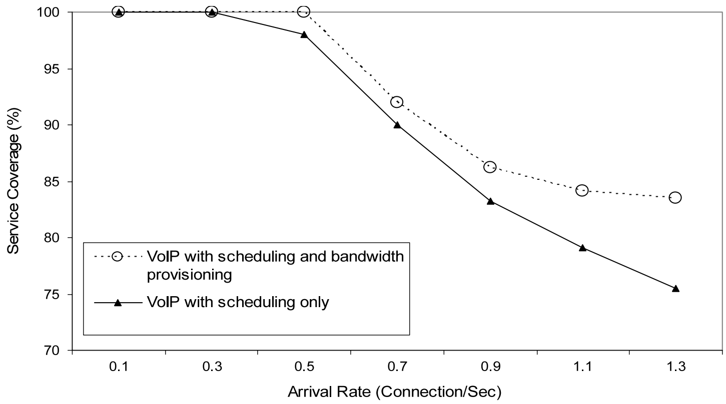


Fig. 3. Service coverage for VoIP with/without bandwidth provisioning.

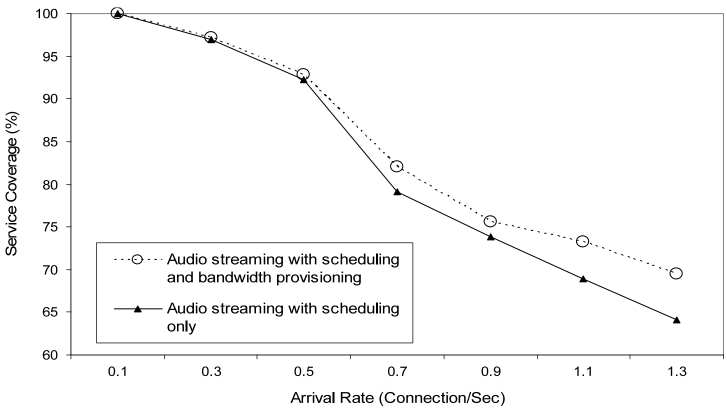


Fig. 4. Service coverage for audio streaming with/without bandwidth provisioning.

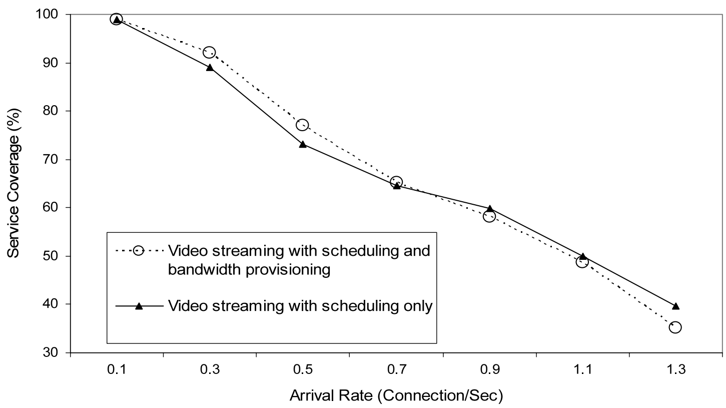


Fig. 5. Service coverage for video streaming with/without bandwidth provisioning.

Similarly, the service coverage of FTP traffic is slightly lower with dynamic bandwidth provisioning as shown in Fig. 6, since class 3 is assigned the lowest priority. Therefore, more time frames are allocated to classes 1 and 2 at the expense of class 3. The performance of class 3 can certainly be improved by increasing its priority as discussed in the next sections.

4.4.2. *Case 2: fixed weights*

Table 2 shows the proportion of assigned frames for each traffic class with the corresponding fixed weights. We note that the proportion of assigned frames, and hence fairness, for class 3 can be increased by increasing its priority (through increasing its weight and decreasing the weights of higher priority classes). However, this occurs at the expense of

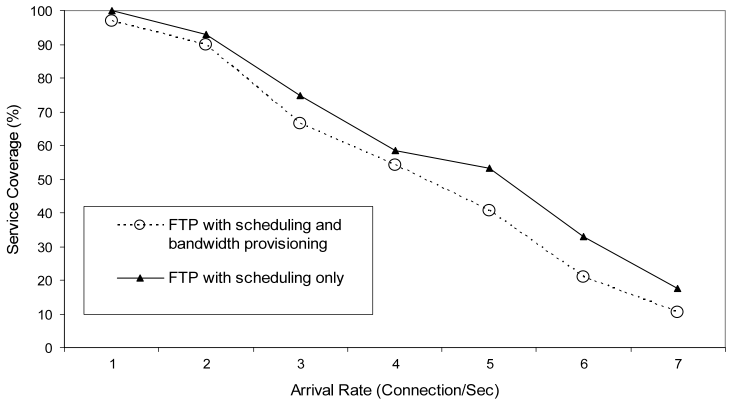


Fig. 6. Service coverage for FTP with/without bandwidth provisioning.

Table 2
Proportion of assigned frames with different fixed weights (case 2).

w_1	w_2	w_3	$\overline{P_1}$ (%)	$\overline{P_2}$ (%)	$\overline{P_3}$ (%)	ζ_c
7	5	1	100	62.1	9.6	1
6	4	2	100	59.4	16.5	1
5	4	3	95.2	51.7	28.3	1
1	1	1	91.7	37.2	45.8	1

Table 3
Proportion of assigned frames with different opportunity cost values (case 2).

w_1	w_2	w_3	$\overline{P_1}$ (%)	$\overline{P_2}$ (%)	$\overline{P_3}$ (%)	ζ_c
1	1	1	91.7	37.2	45.8	1
1	1	1	97.1	50.3	31.2	0.66
1	1	1	99.4	59.1	10.9	0.33
1	1	1	100	66.6	0	0

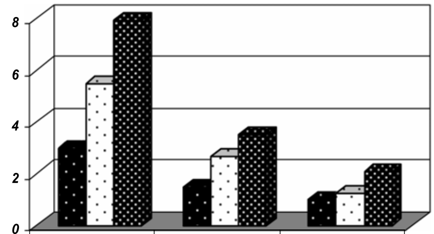
Table 4
Proportion of assigned frames with dynamic weights (case 3).

$W_{\min}$	$W_{\max}$	$\overline{P_1}$ (%)	$\overline{P_2}$ (%)	$\overline{P_3}$ (%)	ζ	τ_i
1	10	97.5	63.4	8.3	1	0.5
1	10	86.8	53.9	28.7	1	0.75
1	10	59.8	42.3	45.6	1	1

decreasing the proportion of assigned frames for classes 1 and 2. In addition, the effect of opportunity cost of bandwidth provisioning can be controlled by controlling H_c, where we let $H_c = \zeta_c \cdot \text{Max}\,\text{Rev}_c$. When $\zeta_c = 1$, this implies that the network operator can tolerate a revenue loss as high as the maximum revenue that could be obtained. That is, in this case, the opportunity cost of bandwidth provisioning is ignored. However, as ζ_c is decreased, then the network operator can tolerate less revenue loss, and thus more frames are given to the highest-revenue-generating classes (i.e., higher priority classes) as shown in Table 3. When $\zeta_c = 0$, then the network operator cannot tolerate any revenue loss, and hence only the classes that have the maximum revenue (i.e., $\text{Max}\,\text{Rev}_c$ in Eq. (2)), which are classes 1 and 2 are assigned frames. Therefore, the network operator can choose the level at which it can tolerate revenue loss as a result of bandwidth provisioning by controlling ζ_c.

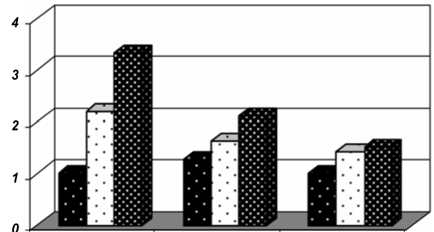
4.4.3. Case 3: dynamic weights

The proportion of assigned frames for each class in case of dynamic weights is shown in Table 4. The weight ranges are chosen to be between 1 and 10 (i.e., $W_{\min} = 1$ and $W_{\max} = 10$). The importance of the weight update scheme is that it allows service differentiation between classes while at the same time it ensures inter-class fairness (in terms of proportion of assigned frames). The resulting fairness is more adaptive to the performance of classes since it is based on their performance history. Therefore, inter-class fairness can be better achieved using this scheme instead of setting fixed weights. The network operator can achieve different fairness levels by controlling τ_i, where small τ_i values result in less fairness. This is not possible with fixed weights since the performance of each class is not fixed because of the varying bandwidth requirements and channel quality conditions.



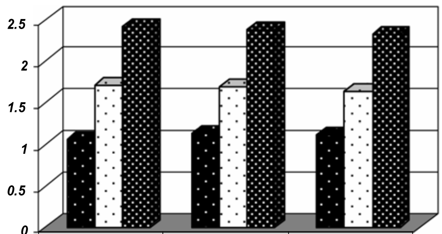
	class 1	class 2	class 3
10th percentile	2.97	1.49	1
avg	5.43	2.68	1.23
90th percentile	7.9	3.5	2.05

Fig. 7. 10th, average and 90th percentile of dynamic weights with $\tau_i = 0.5$.



	class 1	class 2	class 3
10th percentile	1	1.28	1
avg	2.19	1.62	1.42
90th percentile	3.32	2.1	1.5

Fig. 8. 10th, average and 90th percentile of dynamic weights with $\tau_i = 0.75$.



	class 1	class 2	class 3
10th percentile	1.05	1.13	1.11
avg	1.7	1.68	1.63
90th percentile	2.41	2.38	2.32

Fig. 9. 10th, average and 90th percentile of dynamic weights with $\tau_i = 1$.

The role of τ_i in controlling inter-class fairness is shown in Figs. 7–9, which depict the 10th, average and 90th percentile of the dynamic weights of each class for $\tau_i = 0.5$, 0.75 and 1, respectively. The figures show that by increasing τ_i, the dynamic weight values for different classes are allowed to get closer to each other; hence, improving inter-class fairness. This behavior is also confirmed in Figs. 10–12, which show the instantaneous weights during the simulation time with $\tau_i = 0.5$, 0.75 and 1 for all classes. The figures show that the instantaneous weights of low priority classes could be temporarily higher than those of higher priority classes. However, there is a clear separation, on average, between the weight of each class and that of the class of higher priority. This separation is due to the long-term service differentiation between classes that is achieved through the condition $\frac{\overline{w_i(t)}}{\overline{w_{i-1}(t)}} \leq \tau_i$.

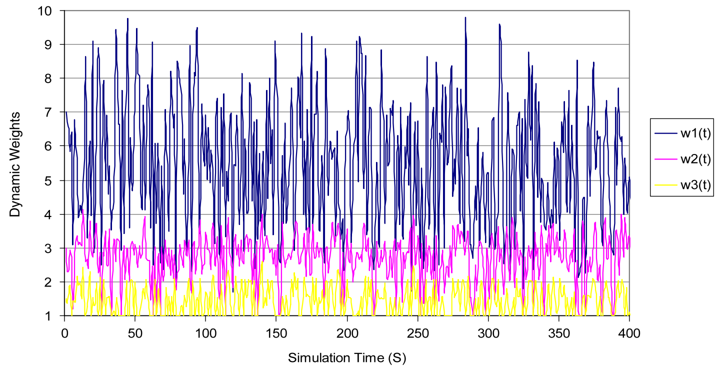


Fig. 10. The instantaneous weights with $\tau_i = 0.5$.

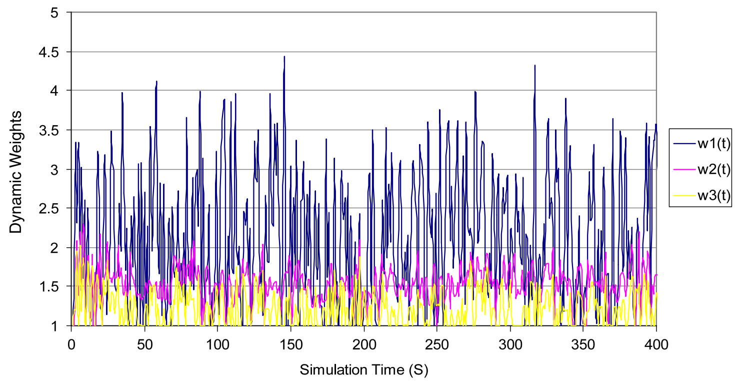


Fig. 11. The instantaneous weights with $\tau_i = 0.75$.

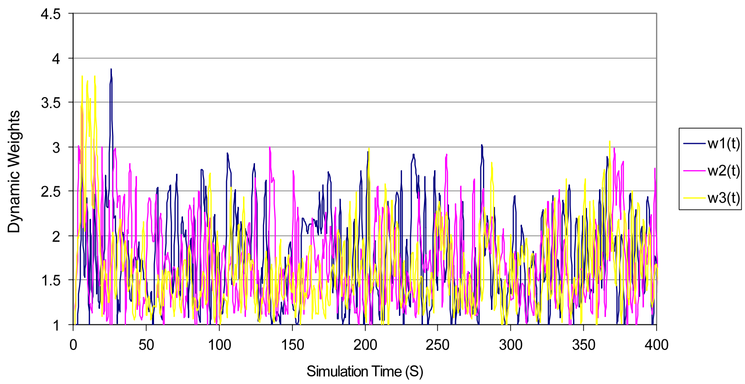


Fig. 12. The instantaneous weights with $\tau_i = 1$.

5. Conclusion

Emerging Broadband Wireless Access Systems (BWASs) will enhance the mobile users' wireless experience by supporting a wide range of multimedia services. However, to satisfy the bandwidth requirements of such services, bandwidth provisioning is critical. In this paper, a novel bandwidth provisioning scheme for BWASs is proposed. The proposed scheme allows for prioritized bandwidth provisioning to different classes of traffic supporting multiple users' connections with different bandwidth requirements. It also incorporates a unique opportunity cost function to bound the cost of allocating bandwidth to different classes so as to maintain a certain revenue level to the network operator. To maximize inter-class fairness, a weight update scheme is integrated with the bandwidth provisioning scheme to dynamically configure the weights of different classes as to achieve a certain level of fairness as desired by the service provider. Simulation results show that our proposed dynamic provisioning scheme along with the weight update scheme can significantly improve the performance of BWASs.

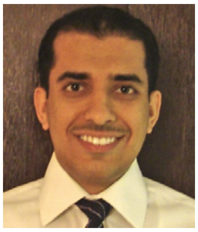
We are currently working on integrating our proposed schemes with Call Admission Control (CAC) in order to provide minimum bandwidth guarantees to users' connections. The CAC scheme will also utilize our proposed frame-level opportunity cost function, however, at the admission level, since different connections requests can generate different revenues to network operators.

Acknowledgements

This work is supported by a grant from the Natural Science and Engineering Research Council of Canada (NSERC) and by a grant from Qatar Foundation (QF) under the National Priorities Research Program (NPRP).

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] B. Rong, Y. Qian, H.H. Chen, Adaptive power allocation and call admission control in multiservice WiMAX access networks, IEEE Wireless Communications Magazine 14 (1) (2007) 14–19.
[4] N. Nasser, H. Hassanein, Optimized bandwidth allocation with fairness and service differentiation in multimedia wireless networks, Journal of Wireless Communications and Mobile Computing 8 (4) (2008) 501–511.
[5] B. Rong, Y. Qian, K. Lu, Integrated downlink resource management for multiservice WiMAX networks, IEEE Transactions on Mobile Computing 6 (6) (2007) 621–632.
[6] N. Zorba, A.I. Perez-Neira, CAC for multibeam opportunistic schemes in heterogeneous WiMax systems under QoS constraints, in: Proceedings of the IEEE Global Telecommunications Conference, GLOBECOM, November 2007, pp. 4296–4300.
[7] E.B. Rodrigues, J. Olsson, Admission control for streaming services over HSDPA, in: Proceedings of the Advanced Industrial Conference on Telecommunications/Service Assurance with Partial and Intermittent Resources Conference, AICT/SAPIR, Lisbon, Portugal, July 2005, pp. 255–260.
[8] M. Kazmi, P. Godlewski, C. Cordier, Admission control strategy and scheduling algorithms for downlink packet transmission in WCDMA, in: Proceedings of the IEEE Vehicular Technology Conference, VTC, Boston, USA, vol. 2, September 2000, pp. 674–680.
[9] D. Kim, Efficient interactive call admission control in power-controlled mobile systems, IEEE Transactions on Vehicular Technology 49 (2000) 1017–1028.
[10] B. Li, L. Li, B. Li, K.M. Sivalingam, X.-R. Cao, Call admission control for voice/data integrated cellular networks: performance analysis and comparative study, IEEE Journal on Selected Areas in Communications 22 (4) (2004) 706–718.
[11] J. Gadze, N. Pissinou, K. Makki, G. Crosby, On optimal slot allocation for reservation TDMA MAC protocol in shadow fading environment, in: Proceedings of the EEE International Symposium on Wireless Communication Systems, ISWCS, Trondheim, Norway, October 2007, pp. 809–813.
[12] N.A. Ali, H. Hayajneh, H. Hassanein, Cross layer scheduling algorithm for IEEE 802.16 broadband wireless networks, in: Proceedings the IEEE International Conference on Communications, ICC, Beijing, China, May 2008, pp. 3858–3862.
[13] B. Kraimeche, M. Schwartz, Bandwidth allocation strategies in wide-band integrated networks, IEEE Journal on Selected Areas in Communications 4 (6) (1986) 869–878.
[14] Y. Afek, M. Cohen, E. Hallman, Y. Mansour, Dynamic bandwidth allocation policies, in: Proceedings of the Annual Joint Conference of the IEEE Computer Societies, INFOCOM, San Francisco, USA, vol. 2, March 1996, pp. 880–887.
[15] X. Xu, D. Liang, H. Jiang, X. Lin, Dynamic bandwidth allocation in fixed BWA systems, in: Proceedings of the International Conference on Communication Technology, ICCT, Beijing, China, vol. 2, April 2003, pp. 1000–1003.
[16] A. Duel-Hallen, S. Hu, H. Hallen, Long-range prediction of fading signals, IEEE Signal Processing Magazine 17 (2000) 62–75.
[17] R. Vaughan, P. Teal, R. Raich, Short-term mobile channel prediction using discrete scatterer propagation model and subspace signal processing algorithms, in: Proceedings of the IEEE Vehicular Technology Conference, VTC, Boston, MA, USA, September 2000.
[18] T. Ekman, Prediction of mobile radio channels, modeling and design, Ph.D. Thesis, Uppsala University, Sweden, 2002.
[19] H. Varian, Intermediate Microeconomics: A Modern Approach, 6th ed., W.W. Norton & Company, 2003.
[20] Bader Al-Manthari, Najah Abu Ali, Nidal Nasser, Hossam Hassanein, QoS-based resource management scheme for multimedia traffic in high speed wireless networks, in: Proceedings of the Annual IEEE Global Telecommunications Conference, Globecom, Washington, DC, USA, November 2007, pp. 5236–5241.
[21] Deliverable D3. 2v2, End-to-end network model for enhanced UMTS. Available: http://www.ti-wmc.nl/eurane/.
[22] http://lpsolve.sourceforge.net/5.5/.
[23] B. Al-Manthari, N. Nasser, H. Hassanein, Packet scheduling in 3.5G high-speed downlink packet access networks: breadth and depth, IEEE Network Magazine 21 (1) (2007) 41–46.
[24] Y.S. Kim, Capacity of VoIP over HSDPA with frame bundling, IEICE Transaction on Communications E89-B (12) (2006) 3450–3453.
[25] 3GPP TS 26.071 V6, AMR speech codec; general description, December 2004.
[26] C. Bormann, C. Burmeister, M. Degermark, H. Fukushima, H. Hannu, L. Jonsson, R. Hakenberg, T. Koren, K. Le, Z. Liu, A. Martensson, A. Miyazaki, K. Svanbro, T. Wiebke, T. Yoshimura, H. Zheng, RObust header compression (ROHC): framework and four profiles: RTP, UDP, ESP, and uncompressed, in: IETF RFC Standards Track 3095, July 2001.
[27] ITU-T, one-way transmission time G.114, May 2003.
[28] B. Wang, K.I. Pedersen, T.E. Kolding, P.E. Mogensen, Performance of VoIP over HSDPA, in: Proceedings of the IEEE Vehicular Technology Conference, VTC, vol. 4, May 2005, pp. 2335–2339.
[29] R. Lloyd-Evan, QoS in Integrated 3G Networks, 1st ed., Artech House, 2002.
[30] 3GPP TS 23.107 V5.12.0, Quality of service (QoS) concept and architectures, Release 5, March 2004.
[31] 3GPP TS 22.105 V 6.4.0, Services and service capabilities, Release 6, September 2005.
[32] G. Rittenhouse, H. Zheng, Providing VoIP service in UMTS-HSDPA with frame aggregation, in: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP, Philadelphia, PA, USA, vol. 2, March 2005, pp. 1157–1160.
[33] P. Lunden, M. Kuusela, Enhancing performance of VoIP over HSDPA, in: Proceedings of the IEEE Vehicular Technology Conference, VTC, Dublin, Ireland, April 2007, pp. 825–829.
[34] P. Lunden, J. Aijanen, K. Aho, T. Ristaniemi, Performance of VoIP over HSDPA in mobility scenarios, in: Proceedings of the IEEE Vehicular Technology Conference, VTC, Marina Bay, Singapore, May 2008, pp. 2046–2050.
[35] P. Jose, Packet scheduling and quality of service in HSDPA, Ph.D. Dissertation, Aalborg University, Denmark, October 2003.



Bader Al-Manthari received his B.Sc. with Honors, M.Sc. and Ph.D. from Queen's University, Kingston, Canada in 2004, 2005, and 2009 respectively. He is currently working as an information security specialist at the Center of Information Security, Information Technology authority in the Sultanate of Oman.

His research interests include information security, economic-based radio resource management in next generation wireless cellular networks, wireless ad hoc and sensor networks, performance evaluation of communication protocols and schemes.

Najah Abu Ali received her B.S. and M.S. degrees in Electrical Engineering in 1989 and 1995 respectively from University of Jordan, Amman, Jordan and her Ph.D. degree in 2006 in Computer Networks in Electrical Engineering department at Queen's University, Kingston, Canada. She joined the College of Information Technology, United Arab Emirates University (Al Ain, UAE), as an Associate Professor with the Computer Networks Engineering track. She had a postdoctoral fellowship at the School of Computing, Queen's University from January 2006 to August 2006. She worked as an instructor and the head of the Engineering Department at Queen Noor College in Jordan from 1995 to 2003. Her research interests comprise wired and wireless communication networks. Specifically, analytical and measurement based network performance management and Quality of Service and resource management of single and multihop wireless networks. Dr. Abu Ali is an expert on Broadband Wireless Networks architecture, design, QoS provisioning and performance and has published extensively in the area. She delivered several tutorials before including one on Resource Management in WiMax Networks at ICC 2008, an overview of IEEE 802.16/WiMAX at CCNC 2009, and two on IMT-Advanced standardization and technologies, presented in Globecom 2009 and VTC-Fall 2010.

Nidal Nasser received his B.Sc. and M.Sc. degrees with Honors in Computer Engineering from Kuwait University, State of Kuwait, in 1996 and 1999, respectively. He completed his Ph.D. in the School of Computing at Queen's University, Kingston, Ontario, Canada, in 2004. He is currently an Associate Professor in the School of Computer Science at University of Guelph, Guelph, Ontario, Canada. He has authored several journal publications, refereed conference publications and nine book chapters. He has also given tutorials in major international conferences. He is an associate editor of the Journal of Computer Systems, Networks, and Communications, Wiley's International Journal of Wireless Communications and Mobile Computing and Wiley's Security and Communication Networks Journal. He has been a member of the technical program and organizing committees of several international IEEE conferences and workshops. Dr. Nasser is a member of several IEEE technical committees. He received Fund for Scholarly and Professional Development Award in 2004 from Queen's University. He received the Best Research Paper Award at the ACS/IEEE International Conference on Computer Systems and Applications (AICCSA'08) and at the International Wireless Communications and Mobile Computing Conference (IWCMC'09).

Hossam Hassanein is with the School of Computing at Queen's University working in the areas of broadband, wireless and variable topology networks architecture, protocols, control and performance evaluation. Dr. Hassanein obtained his Ph.D. in Computing Science from the University of Alberta in 1990. He is the founder and director of the Telecommunication Research (TR) Lab http://www.cs.queensu.ca/~trl in the School of Computing at Queen's. Dr. Hassanein has more than 350 publications in reputable journals, conferences and workshops in the areas of computer networks and performance evaluation. He has delivered several plenary talks and tutorials at key international venues, including Unconventional Computing 2007, IEEE ICC 2008, IEEE CCNC 2009, IEEE GCC 2009, IEEE GIIS 2009, ACM MSWIM 2009 and IEEE Globecom 2009. Dr. Hassanein has organized and served on the program committee of numerous international conferences and workshops. He also serves on the editorial board of a number of International Journals. He is a senior member of the IEEE, and is currently chair of the IEEE Communication Society Technical Committee on Ad hoc and Sensor Networks (TC AHSN). Dr. Hassanein is the recipient of Communications and Information Technology Ontario (CITO) Champions of Innovation Research award in 2003. He received several best paper awards, including at IEEE Wireless Communications and Network (2007), IEEE Global Communication Conference (2007), IEEE International Symposium on Computers and Communications (2009), IEEE Local Computer Networks Conference (2009) and ACM Wireless Communication and Mobile Computing (2010). Dr. Hassanein is an IEEE Communications Society Distinguished Lecturer.