This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the ICC 2008 proceedings.

Dynamic Bandwidth Provisioning with Fairness and Revenue Considerations for Broadband Wireless Communication

Bader Al-Manthari [1] , Najah Abu Ali [2] , Nidal Nasser [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] College of Information Technology
UAE University, Al-Ain
P.O. 17555, UAE
najah@uaeu.ac.ae

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

***Abstract*— The success of emerging wireless broadband communication systems such as 3.5G wireless cellular systems and 802.16 broadband wireless systems (WiMAX) 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 heterogeneous nature and, hence, diverse bandwidth requirements of applications that these communication systems support and the reliance on high speed shared channels for data delivery instead of dedicated ones. Therefore, bandwidth provisioning is crucial for the success of such communication systems. In this paper, we propose a novel dynamic bandwidth provisioning scheme for broadband wireless communication. The proposed scheme spans multiple time slots/frames and optimally allocates them to the different classes of traffic depending on their weights, the real-time bandwidth requirements of their connections, their channel quality conditions and the expected obtained revenues. Simulation results are provided to show the potential and effectiveness of our scheme.**

***Keywords* - Bandwidth provisioning; fairness; opportunity cost.**

I. INTRODUCTION

Emerging wireless broadband communication systems such as High Speed Downlink Packet Access (HSDPA) [1] and 802.16 broadband wireless access systems (WiMAX) [2], pose a myriad of new opportunities for leveraging the support of a wide range of multimedia applications that have different bandwidth requirements. This is due to the high data rates that are supported by these systems which were only available to wireline users. To maximize their efficiency and reduce the cost of data delivery, these systems utilize high speed downlink channels that are shared among mobile connections through packet scheduling. Packet scheduling is one of the most important components of resource management that affects system capacity and potential bandwidth allocated to mobile connections. A centralized downlink packet scheduler is implemented at the base stations of broadband wireless communication systems to control the allocation of the downlink shared channels to the mobile connections by deciding which of them should transmit during a given time slot/frame. Therefore, packet scheduling can be thought of as a short-term resource sharing scheme since it only checks the current time frame to make its decision.

However, to enhance the scheduling performance and satisfy the bandwidth requirements of different applications in broadband wireless systems, packet scheduling must be coupled with a long-term resource sharing or bandwidth provisioning scheme to span multiple time frames and decide how resources are shared among the different traffic classes and, hence, their corresponding 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 long-term inter-class selection algorithm.

Most of the work on bandwidth provisioning has been done at the admission level [3, 4, 5, 6, 7], where these schemes implement the call admission control function (CAC) and aim at reducing the call blocking and dropping probabilities while satisfying the bandwidth requirements of different QoS classes. There is a need, however, for bandwidth provisioning at the frame level. This is due to the varying bandwidth requirements of mobile connections during their calls 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. Few works, have considered the problem of bandwidth provisioning at the frame level [8, 9, 10]. Nevertheless, none of which has considered the varying channel quality conditions of mobile connections. Hence, they cannot achieve optimized bandwidth provisioning.

In this paper, we propose a dynamic bandwidth provisioning scheme for future broadband wireless systems. Our proposed scheme is designed to accommodate multi-class traffic with multiple 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

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.

connections. In addition, the proposed scheme uniquely incorporates and bounds the cost (in terms of revenue loss) of bandwidth provisioning through an opportunity cost function. This provides greater flexibility to service providers to determine the levels of bandwidth provisioning to different traffic classes that will guarantee a certain level of revenues.

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 DYNAMIC BANDWIDTH PROVISIONING SCHEME

In this section, we provide an overview of our proposed dynamic bandwidth provisioning scheme. We assume that there are *K* classes of traffic, where class *i* has higher priority than class $i+1$, $1 \le i$ and $i+1 \le K$. We consider that connections within the same class may have different bandwidth requirements depending on the type of applications they are running. Suppose that the service provider wants to provision *SL* time frames between the *K* classes of traffic to satisfy their long-term bandwidth requirements. We assume that *SL* is given. The proposed dynamic bandwidth provisioning scheme works as follows. At time *t* (where *t* is the beginning of *SL* frames), the dynamic bandwidth provisioning scheme will partition the next *SL* frames among the *K* classes of traffic based on the bandwidth requirements of their connections, their weights, their channel quality conditions and the expected revenues. It should be noted that the partitioned frames can be assigned to connections of each class in any order. For example, connections may be served based on the delay or packet loss requirements of the traffic 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, and so on.

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 service provider may utilize any existing packet scheduling scheme for distributing the partitioned frames among connections as our bandwidth provisioning scheme is independent of the scheduling algorithm being utilized. Figure 1 shows an abstract timeline data flow chart of the proposed dynamic bandwidth provisioning scheme.

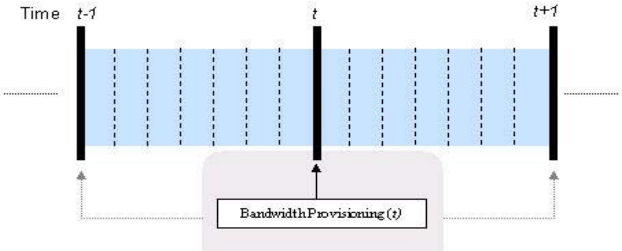

Figure 1. Dynamic Bandwidth Provisioning

III. DYNAMIC BANDWIDTH PROVISIONING

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

- N_i : number of connections in class *i*.
- $N = \sum_{i=1}^{K} N_i$, total number of connections in the system.
- $S_{ij}^{\max}$: maximum required bandwidth of connection *j* of class *i*, j = 1, ..., N_i.
- $S_{ij}^{\min}$: minimum required bandwidth of connection *j* of class *i*, j = 1, ..., N_i.
- SL_i : number of frames assigned to class *i*.
- $\overline{R_i}(t)$: effective average estimated data rate (per second) that the base station can transmit to connections of class *i* during the next *SL* frames. This data rate will depend on the estimated instantaneous channel quality conditions of the connections. $\overline{R_i}(t)$ can be computed using a moving average (i.e., $\overline{R_i}(t) = \alpha \cdot \overline{R_i}(t-1) + (1-\alpha) \cdot R_i(t)$) or using channel prediction schemes proposed in [11], [12] and [13].
- c_{ij} : charge per bit for connection *j* of class *i*.
- $B_i^{\max}$: total required maximum bandwidth per frame of all connections in class *i* at the beginning of the *SL* frames. Let $\sum_{j=1}^{N_i} S_{ij}^{\max}$ = the total required bandwidth per second of all connections in class *i*, and let D_{frame} be the frame duration in second, then $B_i^{\max} = \left(\sum_{j=1}^{N_i} S_{ij}^{\max} / 1 / D_{frame} \right)$. $B_i^{\max}$ determines how much the required transmission rate base station should be sending at per frame in the next *SL* frames in order to satisfy the connections of class *i*.
- $B_i^{\min}$: total required minimum bandwidth per frame of all connections in class *i* at the beginning of the *SL* frames. Therefore, $B_i^{\min} = \left(\sum_{j=1}^{N_i} S_{ij}^{\min} / 1 / D_{frame} \right)$.
- $\overline{B_i} = \left((\overline{R_i(t)} / (1/D_{frame})) \right)$: actual (i.e. effective) total transmitted bit rate per frame for connections in class *i*. That is, $\overline{B_i}$ determines the actual transmission rate per frame of the base station in the next *SL* frames for the connections of class *i*.
- $Rev_i^{\max}$: total maximum revenue per frame of connections in class *i* at the beginning of the *SL* frames. Therefore, $Rev_i^{\max} = \left(\sum_{i=1}^{N_i} (c_{ij} \cdot S_{ij}^{\max} / 1 / D_{frame}) \right)$. That is, $Rev_i^{\max}$

determines the revenue of the service provider per frame in the next SL frames if it grants all the connections in class i their maximum required data rates. Therefore, $SL^* \mathrm{Re}\, v_i^{\max}$ is the upper bound of the total revenue of the service provider during the SL frames.

- $\overline{\mathrm{Re}\, v_i} = \left(\sum_{i=1}^{N_i} (c_{ij} \cdot \overline{R_i(t)} / 1 / D_{frame}) \right)$: actual (i.e. effective) total revenue per frame actually generated from serving all connections of class i. Therefore, $SL_i * \overline{\mathrm{Re}\, v_i}$ is the actual total revenue that the service provider attains from serving all connections of class i provided that class i is assigned SL_i frames.
- $Rev = \{\overline{\mathrm{Re}\, v_i}\}_{i=1}^{K}$: ordered set of the actual effective total revenue per frame for the K classes in descending order.

To satisfy all connections, the base station should assign $SL^* B_i^{\max}$ bandwidth per SL frames. However, this may not be possible in practice due to the high demand of applications that have high bandwidth requirements and also due to the limitations of the base station's capacity, which is determined by the channel quality conditions of the mobile connections. Therefore, the main objective of our bandwidth provisioning scheme is to divide the SL frames between the K classes of traffic such that $\sum_{i=1}^{K} SL_i = SL$ and the satisfaction of different classes of traffic is maximized. To do so, our bandwidth provisioning scheme will divide the SL frames between the K classes of traffic such that it maximizes $\frac{SL_i * \overline{B_i}}{SL * B_i^{\max}}$ (i.e., the ratio between the requested bandwidth of class i connections given that SL_i frames are assigned to class i to the bandwidth that the base station should transmit at during SL frames to satisfy the maximum bandwidth requirements of class i). The frames assigned to class i (i.e., SL_i) should guarantee that no class of traffic is assigned less than its minimum required bandwidth (i.e., $SL * B_i^{\min} \le SL_i * \overline{B_i}$) or assigned more than its maximum required bandwidth (i.e., $SL_i * \overline{B_i} \le SL * B_i^{\max}$).

In addition, it is important to realize there is an opportunity cost of frame assignment[1]. The opportunity cost (in terms of revenue) of frame assignment is the maximum revenue that the service provider will get if it serves the highest revenue generating classes minus the revenue that it would get by assigning the frames otherwise. To compute the maximum revenue that the service provider could get in the next SL frames, we first need to know the number of frames needed for each class in order to achieve its maximum required bandwidth (i.e., $SL_i^{req} * \overline{B_i} = SL * B_i^{\max}$). Hence, $SL_i^{req} = \frac{SL * B_i^{\max}}{\overline{B_i}}$.

Therefore, the maximum revenue, *Max Rev*, is equal to

$$Max\ \mathrm{Re}\, v = \sum_{i=1}^{K} SL_i^{req} \cdot \mathrm{Re}\, v_i, \sum_{i=1}^{K} SL_i^{req} \le SL \qquad (1)$$

That is, the maximum revenue is obtained, by assigning the frames to the class with the highest actual revenue. If this class can be served by less number of frames than SL frames, the available frames are assigned to the class with the second highest actual revenue and so on.

Therefore, the opportunity cost ($OC(SL)$) of the frames assignment is equal to

$$OC(SL) = Max\ \mathrm{Re}\, v - \left(\sum_{i=1}^{K} SL_i \cdot \overline{\mathrm{Re}\, v_i} \right) \qquad (2)$$

This should be less than or equal to a predefined value H. For example, the service provider could restrict the revenue loss to be no more than 50% of the maximum obtainable revenue (i.e., $H = \zeta \cdot Max\ \mathrm{Re}\, v$ where $\zeta = 0.5$).

To summarize, in our bandwidth provisioning scheme, the following optimization problem will be solved:

Objective: $\max_{SL_i, 1 \le i \le K} \sum_{i=1}^{K} w_i \left(\frac{SL_i * \overline{B_i}}{SL * B_i^{\max}} \right)$

Subject to:

$$\sum_{i=1}^{K} SL_i = SL,$$

$$SL * B_i^{\min} \le SL_i * \overline{B_i} \le SL * B_i^{\max}, \text{ and}$$

$$OC(SL) \le H \qquad (3)$$

where w_i is a weight assigned to class i to give it priority over class i+1 in the frame assignment process. Therefore, the weights in our bandwidth provisioning scheme can be used to determine the appropriate level of inter-class fairness (i.e., fairness between different classes) according to the requirements of the service providers. Since the objective function and the constraints are linear, our bandwidth provisioning scheme could be solved using Linear Programming (LP) techniques.

The proposed bandwidth provisioning scheme is adaptive to the varying requirements of different classes since the objective function is evaluated every SL frames and, therefore, if the required bandwidth (or frames) for 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 SL frames.

[1] 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 [14].

IV. 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 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) [1].

A. Simulation Model

For simplicity, we simulated one-cell and ignored handoffs. 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. Four classes of traffic are considered (i.e. K=4). For demonstration purposes, we assume that connections within each class request a bandwidth of 128 Kbps and there are 10 connections in each class. We also assume that c_{ij} = 4, 3, 2 and 1 units of money for connections of classes 1, 2, 3, and 4, respectively. These values are chosen such that connections of higher priority classes are charged more than those of lower priority classes. Packets for connections are generated as a Poisson process with a mean value of 0.5 s. The packet size is fixed at 1000 bytes. The call duration of each connection is modeled by an exponential distribution with a mean value of 30 s. Connections are uniformly distributed in the cell. We choose SL=20 time frames (i.e. 20 x 2ms) and we use a moving average to compute $\overline{R_i(t)}$ (i.e., $\overline{R_i(t)} = \alpha \cdot \overline{R_i(t-1)} + (1-\alpha) \cdot R_i(t)$). We adopt the same channel model as in [15]. The simulation time step is one time frame, which is 2 ms (i.e., the time frame of HSDPA [1]), and the simulation time is 200 s. Other simulation parameters are listed in Table I.

TABLE I
SIMULATION PARAMETERS

Simulation time	200s
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
α	0..99

B. Performance Metrics

To demonstrate the ability of our scheme to satisfy the dynamic bandwidth requirements of different classes and achieve inter-class fairness, it suffices to show the:

- Average utility ($\overline{U_i}$): the average of the received bandwidth to the maximum requested one by class i.
- Proportion of assigned frames ($\overline{P_i}$): the average percentage of assigned frames to class i of the total number of frames assigned to all classes.

Another important metric is the amount of earned revenue. However, and for lack of space, we do not report such results in this paper. The service charge values are used to validate our scheme and show the effect of opportunity cost on bandwidth provisioning.

C. Simulation Results

Table II shows the average utility of each traffic class with the corresponding weights. In general, all classes achieve acceptable performance levels. As it can be seen however, the average utilities of low priority classes (i.e., classes 3 and 4) can be increased by increasing their weights. However, this occurs at the expense of decreasing the average utilities of higher priority classes (i.e., classes 1 and 2). This shows the important role of class weights and how they can be used to ensure inter-class fairness. In addition, the effect of opportunity cost on bandwidth provisioning can be controlled by controlling H where we let $H = \zeta \cdot Max\ \mathrm{Re}v$ as shown in Table III. When $\zeta = 1$, this means that the service provider 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 is decreased, then the service provider can tolerate less revenue loss and, hence, more frames are given to the highest revenue generating classes (i.e., higher priority classes). When $\zeta = 0$, then the service provider cannot tolerate any revenue loss and, hence, only the classes that make the maximum revenue (i.e., *Max Rev* in Eq. 2.) which are classes 1 and 2 are assigned frames. Therefore, the service provider can choose the level at which it can tolerate revenue loss as a result of bandwidth provisioning by controlling ζ and hence, controlling H. These results are also confirmed in Tables IV and V which show the proportion of assigned frames for each class for different class weights and ζ values.

These results show that the proposed dynamic provisioning scheme can satisfy the bandwidth requirements of different classes of traffic while allowing the service provider to bound the cost of allocating bandwidth to different classes. Results also show that by controlling the class weights, the service provider can achieve its desired levels in inter-class fairness.

V. Conclusions and Future Work

Future broadband wireless systems will enhance the mobile connections' wireless experience by supporting a wide range of multimedia applications. However, to satisfy the bandwidth requirements of such applications, bandwidth provisioning is critical. In this paper, a novel bandwidth provisioning scheme for broadband wireless network is proposed. The proposed scheme allows for prioritized bandwidth provisioning to different classes of traffic supporting multiple connections with different bandwidth requirements. It also incorporates a unique opportunity cost function to bound the cost of allocating bandwidth to different classes as to maintain a certain revenue level to the service provider. Simulation results show that the proposed dynamic provisioning scheme can satisfy the bandwidth requirements of different classes of traffic while allowing the service provider to bound the cost of allocating bandwidth to different classes.

Fixed class weights allow the service provider to control inter-class fairness as demonstrated in our experiments. They however, cannot achieve optimized fairness since the performance of each class is not fixed due to the varying bandwidth requirements and channel quality conditions. Therefore, in our future work, we will work on a dynamic weight update scheme to dynamically compute the weights of different classes of traffic based on their performance history in order to maximize inter-class fairness. This way, the resulting fairness is more adaptive to the performance of classes since it is based on their performance history.

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] C. Chou and K. G. Shin, "Analysis of Combined Adaptive Bandwidth Allocation and Admission Control in Wireless Networks", *Proceedings of Annual Joint Conference of the IEEE Computer Societies (INFOCOM)*, New York, U.S.A, vol. 2, pp. 676 -684, June 2002.

[4] R. Hutchens and S. Singh, "Bandwidth Reservation Strategies for Mobility Support of Wireless Connections with QoS Guarantees", *Proceedings of the Australasian Computer Science Conference (ACSC)*, Melbourne, Australia, vol. 4, pp. 119-128, January 2002.

[5] M. R. Sherif, I. W. Habib, M. Nagshineh, and P. Kermani, "Adaptive Allocation of Resources and Call Admission Control for Wireless ATM Using Genetic Algorithms", *IEEE Journal of Selected Areas in Communications*, vol. 18,no. 2, pp. 268-282, February 2000

[6] N. Nasser, H. Hassanein; "Multi-class Bandwidth Allocation Policy for 3G Wireless Networks", *Proceedings of th IEEE International Conference on Local Computer Networks* (LCN), Bonn, Germany, pp. 203-209, October 2003.

[7] C. Oliviera, J. Kim, and T. Suda, "An Adaptive Bandwidth Reservation Scheme for High Speed Multimedia Wireless Networks," *IEEE Journal of Selected Areas in Communications*, vol. 16, no. 6, pp. 858-874, August 1998.

[8] B. Kraimeche and M. Schwartz, "Bandwidth Allocation Strategies in Wide-Band Integrated Networks", *IEEE Journal on Selected Areas in Communications*, vol.4 , no.6, pp. 869-878, September 1986.

[9] Y. Afek, M. Cohen, E. Hallman and Y. Mansour, "Dynamic Bandwidth Allocation Policies", *Proceedings of the Annual Joint Conference of the IEEE Computer Societies (INFOCOM),* San Francisco, U.S.A, vol.2, pp. 880-887, March 1996.

[10] X. Xu, D. Liang, H. Jiang and X. Lin,, "Dynamic Bandwidth Allocation in Fixed BWA Systems", *Proceedings of the International Conference on Communication Technolog (ICCT)*, Beijing, China, vol.2, pp.1000-1003, April 2003.

[11] A. *Duel*-Hallen, S. Hu, and H. Hallen, "Long-range prediction of fading signals", *IEEE Signal Processing Magazine*, vol. 17, pp. 62–75, May 2000.

[12] R. Vaughan, P. Teal, and R. Raich, "Short-term mobile channel prediction using discrete scatterer propagation model and subspace signal processing algorithms", *Proceedings. of the IEEE Vehicular Technology Conference (VTC)*, Boston, MA, USA, September 2000.

[13] T. Ekman, Prediction of mobile radio channels, modeling and design, Ph.D. thesis, Uppsala University, Sweden, 2002.

[14] H. Varian, "Intermediate Microeconomics: A Modern Approach", 6th edition, W.W. Norton & Company, 2003.

[15] 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.

TABLE II
AVERAGE UTILITIES

w_1	w_2	w_3	w_4	U_1	U_2	U_3	U_4
9	5	3	1	0.68	0.43	0.19	0.05
7	4	2	1	0.561	0.39	0.263	0.19
4	3	2	1	0.442	0.41	0.331	0.295

TABLE III
AVERAGE UTILITIES WITH DIFFERENT OPPORTUNITY COST VALUES

w_1	w_2	w_3	w_4	U_1	U_2	U_3	U_4	ζ
1	1	1	1	0. 47	0.38	0.248	0.133	0.33
1	1	1	1	0.65	0.37	0.127	0.09	0.66
1	1	1	1	0.91	0.294	0	0	0

TABLE IV
PROPORTION OF ASSIGNED FRAMES

w_1	w_2	w_3	w_4	$\overline{P_1}$	$\overline{P_2}$	$\overline{P_3}$	$\overline{P_4}$	ζ
9	5	3	1	57%	23%	15.4%	4.6%	1
7	4	2	1	45.2%	21.3%	18.7%	14.8%	1
4	3	2	1	32.2%	26.3%	22.3%	19.2%	1

TABLE V
PROPORTION OF ASSIGNED FRAMES WITH DIFFERENT OPPORTUNITY COST VALUES

w_1	w_2	w_3	w_4	$\overline{P_1}$	$\overline{P_2}$	$\overline{P_3}$	$\overline{P_4}$	ζ
1	1	1	1	39.3%	28.7%	20.2%	11.8%	0.33
1	1	1	1	54.9%	26.1%	12.6%	6.4%	0.66
1	1	1	1	78.5%	21.5%	0%	0%	0