Efficient Queuing Mechanism for QoS provisioning in Wireless IP Networks

Vijayakranthi chinthala[1], Manas Kumar yogi[2]
CSE Department[1], CSE Department[2], M.TECH[1], Sr.Assistant Professor[2]
Email:vijayakranthichinthala@gmail.com[1], manas.yogi@gmail.com[2]

Abstract-This paper presents the mechanism used in Class based queuing for enhancing the QoS in Wireless IP networks. Multiclass traffic increases the complexity in handling the data rate. Mobility of the user provides yet another challenge to this issue. This paper presents the design considerations while developing a wireless scheduling algorithm. Most of the algorithms used in real time are compensation based providing fairness between the same traffic classes. Wireless errors generally occur in bursts, because of the inability of signal propagation in a cellular network, as well as the inertia of users' movement in time intervals comparable to the time needed for processing of an individual IP packet. By using the compensation based algorithms, both issues: the location-dependence of wireless bit errors and the multiclass environment are discussed.

Index Terms- QoS;WFQ;WRR;SBFA;STFQ;WCBFQ.

1. INTRODUCTION

Time-varying transmission quality in wireless links precedes issues for QoS support to different traffic classes. The enquiry is how to provide the guaranteed data rate when there is a higher bit error ratio in the wireless channel (we use the notion of channel in the sense of a connection to a single user; it does not mean that it is a circuit-switched channel). On the other hand, it cannot predict the behavior of the wireless interface in a wider area because of the users' mobility. Also, one may expect wireless bit errors to occur in bursts.

To provide solvents to such issues, this paper defines a scheduler for wireless IP networks that should be used at wireless access points (i.e., base stations).Effort-limited scheduling for a wireless environment are already suggested. This is obtained by extension of WFQ via dynamic weight alteration. To provide fairness among the flows, the algorithm innovates factor coefficients that are used to adjust the throughputs of the flows at a higher bit error ratio. Through such factor coefficients, network operators are given the possibility of controlling the QoS level at error occurrence. But, this scheme does not provide QoS support in a multiclass environment, which is an expected in future wireless multimedia network.

WFQ can render service differentiation in cellular Internet only in specific network conditions. The performance of WFQ is acceptable only at higher traffic loads. Also, it is proven that propagation time, Existing TCP connections and user distribution have little influence on the performances of the WFQ scheme. Hence, appropriate alteration of WFQ may be helpful for packet scheduling in a wireless IP network.

Fair queuing of multiclass traffic for a hybrid wireless/wired network is proposed. In particular, scheduling is considered on a MAC layer in an ATM network. Dissimilar traffic classes are distinguished from one another by using priorities. In instance, real-time data uses a wireless mediocre queuing model. On the other hand, a weighted round robin scheduler processes non real-time data. Best-effort flows are serviced using the FIFO (i.e., FCFS) mechanism. The drawback of such a scheme, however, is the lack of a mechanism for support of real-time flow's throughput under location-dependent bit errors in the wireless channel.

In some overtures to the problem of bit errors in the wireless link, it uses a compensation method-that is, compensation of the flows that experienced bit errors using the bandwidth of the flows that received more bandwidth (i.e., higher QoS) during the error-state of other flows. This result can be found in different proposals .But, the query is whether the compensation method is applicable to real-time flows.

It is analyzes and reviews a scheduling algorithm for wireless IP networks that support multiclass traffic. The scheduler development is guided by the following petitions:

(1) When all flows are error-free, the throughput of the scheduler must be the same as with applied WFQ within every traffic class with QoS support (i.e., within class-A).
(2) The capacity loss of a specific flow in error-state should be dependent on traffic class.
(3) Flows within the same class experiencing equal error rates should experience equal capacity loss.
(4) Network decision maker should be involved only in setting the bounds for guaranteed services.

(5) Real-time flows should be adjusted to their error-free throughput in real time (if possible), and there should be no compensation on channel errors during error-free periods.
(6) Non real-time flows within class-A may be allowed to use some compensation model.
(7) Scheduling for best-effort traffic from class-B should be as simple as possible.

The scheduler is built in two steps:

- Differentiation between class-A (guaranteed) services and class-B (best effort) services, as well as between different subclasses within class-A is based on priority.
- Differentiation of the flows within a subclass of class-A is based on the modification of weights of the flows for real-time traffic, and wireless fair scheduling (e.g., compensation) for non real-time traffic.

It provides an overview of existing scheduling conditions for wireless networks. Then it suggests a scheduling mechanism for multiclass wireless IP networks.

2. WIRELESS NETWORKS AND CHANNEL MODEL

The network consists of co-ordinate routers. The routers that are exploited as wireless access points are concerned to as base stations. It is pretended that every base station serves a unit cell in the network. A flow is said to be active if it has packets queued at the network nodes; otherwise, it is referred to as a passive flow. All active flows in a cell share the same wireless link. It is usually assumed that there is unity flow per active user.

Mobile hosts do not have data about the global state in the wireless link in terms of how many and which other mobile terminals have packets to transfer. Also, it is forced by battery power and processing power. Hence, base stations should perform scheduling in both the uplink and downlink.

Every mobile terminal in a cell communicates with a base station; thus, there is only one wireless hop in each direction. It is assumed that the scheduler in the base station views the traffic as a set of flows to the users. Users can be fast-moving mobile hosts that often make changes in the link state. Therefore, the wireless scheduler should be flexible sufficiency to follow the channel behavior. The error state is tied with one users (i.e., it is location-dependent). The flows of different users are acquired to be independent.

Let us make some main presumptions about the wireless channel model. A wireless channel refers to bandwidth allocated to a single connection, which may be fixed or varying in time. Due to different factors—such as shadowing, fading and multipath—the entire capacity of a wireless link, as well as the capacity of wireless channels, is dynamically variable. Due to the random position of mobile hosts inside a cell, errors are location-dependent. They are also bursty in nature due to different time scales on which changes occur in a user’s position and in packet transmission delay (for instance, a user with a velocity of 50 km/hr travels 0.29m during time intervals of 21 ms, which is the time needed to transmit 1,000-byte IP packet over a 384-Kbps link). All traffic with QoS support must go through the admission control phase, while for best-effort traffic multiple mobile terminals collide over the bandwidth.

It is assumed that the admission control module in the base station has included in all active flows by assigning traffic class and a bandwidth share. For specialization of the classes it uses ToS and DS bits in IPv4 and IPv6, respectively.

3. DESIGN OF WIRELESS SCHEDULING ALGORITHMS

Because of less wireless resources, the large user population, and burstiness of the traffic, it is necessary to apply aggressive admission control to fully utilize wireless resources. For future wireless networks will include multiple traffic types. In a multiclass environment different services have different QoS requirements. Also, within the same traffic class we should provide fairness between different flows because wireless media can exhibit high, variable error rates that affect network users.

For wire line networks, fluid fair queuing (i.e., WFQ) has long been a concept for providing bounded delay channel access and fairness among packets flows over a shared unidirectional link. WFQ provides full separation between flows. The minimum guarantees are unaffected by the behavior of other flows. Fluid-fair, however, assumes that the channel is error-free, or at least that errors are not location-dependent (i.e., all backlogged flows have the ability to transmit at a given time, or none of the flows can). Adapting fair queuing to the wireless environment is not a simple task because of the unique problems in the wireless channels, such as location-dependent bursty errors, channel contention as well as joint scheduling of uplink and downlink flows.

There are several existing proposals for wireless fair queuing. The main goal of wireless fair queuing is to emulate WFQ when all flows perceive error free channels, but to swap channel allocation between flows that perceive channel error and flows that perceive a clean channel. The main differences between different wireless fair queuing algorithms are:

- The process in which swapping takes place;
- Between which flows the swapping takes place;

- How the compensation model is designed.

3.1. *Wire line and Wireless Fluid Fair Queuing*

Let us consider the fair queuing in a case of no-channel errors, and consider why such an approach fails to provide fair service when the environment is error-prone. Weighted fair queuing model WFQ allows any flow i to be granted channel capacity over a given time interval $[t_1, t_2]$ so it minimizes. In WFQ each packet is associated with a start tag and a finish tag, which correspond to the virtual time at which the first bit of the packet and the last bit of the packet are served by that mechanism. Let B (t) denote the set of backlogged flows at time t. If it denotes with $A_{i,k}$ the arrival time of the k^{th} packet of the i^{th} flow, and $S_{i,k}$ and $F_{i,k}$ are start time and finish time for that packet, respectively, then it is represented as

$$S_{i,k} = \max\{V(A_{i,k});F_{i,k-1}\} \quad (1)$$

Where V(t) is the virtual time at time t, which denotes the current round of service. So, the packets are sorted according to the minimum eligible finish time. The finish time is computed from the start time by adding the time needed to send a packet of size L_p:

$$F_{i,k} = S_{i,k} + \frac{L_p}{r_i} \quad (2)$$

where r_i is the rate of the flow i. If it denotes with C(t) the link capacity at time t, which is dynamically varying, it can be obtained by the progression of the virtual time by using the following:

$$\frac{dV(t)}{dt} = \frac{C(t)}{\sum_{i \epsilon B(t)} r_i} \quad (3)$$

Often, approximations of WFQ are used, such as WRR and start-time fair queuing (STFQ) that do not need to compute dV/dt.

However, WFQ provides two important guarantees: a bounded delay and associated minimum throughput of the flow. In WFQ the flow cannot reclaim time from another flow that used its empty channel time (when the first flow had no packets to transmit). In a wireless environment a flow may be backlogged, but unable to transmit due to channel errors.

It shows how the WFQ behaves in a wireless environment through a simple example. Let flows f_1 and f_2 be two flows that share a wireless channel, and let both have equal weights. So, when both flows are error-free, each of them should receive $W_1 = W_2 = 0.5$ channel allocation. Considering a time window [0,1]. It is assumed that flow f_1 is error-free over the entire time window. But, let us suppose that flow f_2 perceives channel error in the time interval [0, 0.5]. Then, in the interval [0, 0.5] WFQ will allocate all bandwidth to flow f_1, because f_2 perceives channel errors. In the interval [0.5, 1] both flows are error-free, and WFQ allocates half of channel capacity to each of them. Finally, over the considered time window, flow f_1 gets average channel allocation $W_1 = (1 + 0.5)/2 = 0.75$, while flow f_2 gets $W_2 = (0 + 0.5)/2 = 0.25$. So, the first flow receives 0.25 more channel allocation than the fair share of 0.5, while the second flow receives 0.25 less than its error-free channel share.

3.2. *WFQ Algorithms*

There are several different approaches for wireless fair queuing. One should note, however, that all of them are based on compensation (i.e., lead and lag model—or credit and debit model) and are created for non real-time communication such as best-effort traffic. Most of these algorithms are developed for wireless LANs (e.g., IEEE 802.11). All of them are modifications and adaptations of WFQ or its approximation algorithms (e.g., WRR) to wireless networks.

In this section we describe the most well-known wireless fair scheduling algorithms. At this point, it is convenient to define certain terms—such as lagging flow, leading flow, backlogged flow—that are used in the descriptions of the algorithms.

A flow is said to be leading if it has received channel allocation in excess of its error-free service. A flow is lagging if it has received less channel allocation than its error-free service. A flow is backlogged if it has packets to transmit over the channel.

3.2.1. Idealized Wireless Fair Queuing

Idealized wireless fair queuing (IWFQ) uses WFQ for the error-free service. Both start and finish tags are assigned according to the WFQ. The service tag for a flow is set to the finish tag of its head-of-line packet. IWFQ selects the flow with a least service tag among all backlogged flows that are error-free. The lead of the leading flow is the difference between its real service tag and its service tag in an error-free channel. However, the service tag is not allowed to increase/decrease by more/less than a predefined bound. IWFQ always allocates the slot (channel time) to the error-free flow with the lowest tag until it either perceives an error channel or its finish tag becomes greater than that of some other flow with an error-free channel. IWFQ was the first algorithm to propose adaptation of WFQ to a wireless environment. It provides long-term fairness and bounded delay channel access. The possible drawback is that lagging flows can capture the channel, and starve out other flows. Hence, IWFQ does not support graceful degradation of service.

3.2.2. Wireless Packet Scheduling

The wireless packet scheduling (WPS) packet scheduler involves WRR with spreading as its error-free service. WRR with spreading is identical to the schedule generated by WFQ if all flows are backlogged. WPS generates a frame of slot allocation from the WRR-spreading algorithm and provides fairness by swapping time allocations between mobile terminals experiencing error bursts and currently error-free terminals. The compensation is two-fold. WPS first tries to swap slots within a frame. If this fails, then it maintains the difference between the real service and the fair service for the flow by changing the effective weight in each frame based on the result of the previous frame. So, it attempts to provide graceful trading of the bandwidth between the leading and the lagging flows. This way it provides bounded delay channel access and long-term fairness, and at the same time it prevents the total channel capture by using the effective weights.

3.2.3 Channel-Condition Independent Packet Fair Queuing

In channel-condition independent packet fair queuing (CIF-Q), for error-free service STFQ is used. As already stated, STFQ is an approximation of WFQ that does not require dV/dt computation by setting the virtual time V (t) to the start tag of the transmitting packet. Each flow has a lag, which is defined as the difference between the error-free service and the real perceived service. If the lag is positive, than the flow is lagging; while in the opposite case it is a leading flow. This scheduling mechanism provides a graceful linear degradation for leading flows. For that purpose CIF-Q introduces a parameter α, which is a probability that a leading flow will retain its allocated slot, while $1 - \alpha$ is the probability that it will relinquish the slot to the lagging flows. CIF-Q can provide short-term and long-term fairness and bounded delay channel access.

3.2.4 Server-Based Fairness Approach

Server-based fairness approach (SBFA) reserves part of the bandwidth for compensation of the lagging flows via so-called virtual compensation flow. It conceptually differs from other wireless fair scheduling algorithms. When a backlogged flow is allocated channel time, but it cannot transmit due to channel errors, then it requests service time (e.g., a slot) in the compensation flow. When a compensation flow is allocated a slot, it gives the slot to the flow to which its head-of-line request belongs. If there are no slots for compensation, then the bandwidth of the compensation flow is shared among all flows. SBFA does not monitor the lead of the leading flows. Subsequently, leading flows do not give up their lead. This algorithm provides long-term fairness, but not short-term fairness or worst-case delay bounds. A lagging flow would request compensation slots till it receives its error-free fair service. However, SBFA is bounded by the reserved bandwidth for the virtual compensation flow. If this portion of the link bandwidth is less than the lags of all backlogged flows over some time interval, then long-term fairness cannot be guaranteed.

3.2.5 Wireless Fair Service

The wireless fair service (WFS) scheduling algorithm uses WFQ scheduling for error-free wireless link. It allocates to each flow two parameters: a rate weight r_i and delay weight φ_i for a flow i. The start tag is computed using the rate weight $S_{i,k} = \{v(A_{i,k}), S_{i,k-1} + \frac{L_{i,k-1}}{r_i}\}$The finish tag is computed using the delay tag: $F_{i,k} = S_{i,k} + L_{i,k/\varphi i}$. Using the delay and bandwidth weights allows for delay-bandwidth decoupling. If a backlogged flow perceives channel errors, its lag is increased only if there is a backlogged error-free flow that increases its lead. Each flow is bounded by per-flow parameters—that is, a lead bound l_i^{max} and a lag bound b_i^{max}. A leading flow with a current lead l_i relinquishes l_i/l_i^{max} of its allocated service time. A lagging flow with a current lag bi receives a fraction $b_i / \Sigma_{j \epsilon B} b_j$ of all relinquished slots by leading flows, where B is the set of backlogged flows. This way, WFS provides fair compensation among the lagging flows. Degradation of leading flows is graceful, and a fraction of the bandwidth relinquished by the leading flows decreases exponentially.

3.2.6 Channel State Dependent Packet scheduling

Channel state dependent packet scheduling (CSDPS) uses a WFQ-like scheduling discipline for error-free service (e.g., WFQ and WRR). This algorithm does not provide compensation between lagging and leading flows. CSPDS does not measure lead and lag of flows, and therefore it is simple for implementation. When service time is allocated to a flow that perceives channel error, then that flow is skipped and the service time is given to the next eligible flow in the WRR cycle. Thus, it may happen that a leading flow increases its lead. Because there is no compensation, this mechanism does not provide short-term and long-term fairness. However, it provides throughput guarantees to error-free channels. Also, if all flows are backlogged with equal probability, lagging flows can reduce their lag over the long term.

3.3. *Service Differentiation Applied to Existing Systems*

In this section we give examples of particular proposals for service differentiation in existing or standardized mobile packet-based networks, such as IEEE 802.11 wireless LAN and 3G mobile networks.

3.3.1 Service Differentiation in IEEE 802.11 Wireless LAN

Wireless LANs provide superior bandwidth compared to any existing cellular technology. The state-of-the-art standard in this area is IEEE 802.11b, which provides data rates up to 11 Mbps using the 2.4-GHz frequency band (there are also higher speed alternatives, such as IEEE 802.11a and IEEE 802.11g). However, it lacks QoS support—that is, it does not have implemented mechanisms for service differentiation.

For example, service differentiation may be based on modification of function of the IEEE 802.11 network, which was initially created to support best effort traffic. IEEE 802.11 networks have two basic functions on the MAC layer: point coordination function (PCF) and distributed coordination function (DCF). PCF is intended to support real-time services by polling mobile terminals in its service area. DCF is created for best-effort traffic by using the CSMA/CA protocol. In the DCF mode, a terminal must sense the medium before sending a packet. The sensing time must be long enough to avoid collision between different mobile terminals, and this time is referred to as distributed interface space (DIFS). If a mobile terminal detects a signal, it backs off a QoS Provisioning in Wireless IP Networks Through Class-Based Queuing 331 random time interval within a specified contention window (CW). The 802.11 standard specifies alternation between PCF and DCF intervals, although PCF may be not supported by some wireless card interfaces. Support of both PCF and DCF may lead to inefficient usage of wireless resource. Therefore, some authors propose an extension of DCF to provide service differentiation. One way to accomplish such a task is to create a Diff Serv-enabled MAC, where packets are differentiated by DS field in the IP packet's header. Specifying different CW sizes for different services provides support to different classes in this algorithm. Packets with a smaller CW value are more likely to be transmitted first; that is, high-class service can get better service than lower-class service. To provide absolute QoS guarantees, one needs an accurate estimation of traffic parameters in the cell. For such purposes, one may find it suitable to use a virtual MAC (VMAC) that simulates real MAC behavior and thus provides, in advance, traffic information needed for admission control.

3.3.2 Service Differentiation in 3G CDMA-Based Mobile Networks

Several 3G mobile standards are CDMA-based, such as UMTS and CDMA 2000. Therefore, we consider an example of service differentiation in a CDMA network. In such networks, resource allocation to users is mainly controlled by SIR and spreading control. One approach is to use adaptive power control based on fixed target SIR, in conjunction with variable spreading control to adjust bandwidth offered to a user in a particular frame. In such an environment, class-based scheduling can be provided by introducing additional parameter elasticity (besides the bandwidth requirements), which refers to how the rate will decrease in a period of congestion. In the uplink, the mobiles can reduce its rate upon congestion according to the elasticity. In the downlink, the limiting factors are path loss and total base station transmitted power to users. Therefore, in the downlink case elasticity must be considered together with the path loss the corresponding mobile terminal sees from base station. To provide multiclass communication from a single mobile terminal, each class should be assigned a different code. Also, base stations control the scheduling in the wireless channel. While downlink scheduling is trivial because the base station has a complete knowledge about the traffic, uplink scheduling requires signaling information from mobile terminals to base stations.

The above approach in CDMA mobile networks can be extended by allocation of resources proportionally to weights, thus leading to fair allocation .With such an approach, naturally one should take into account the difference in resource scarcity for the uplink and downlink. First, let us consider service differentiation in the uplink. It is assumed that each mobile user has associated weight that corresponds to its service class. In 3G UMTS's WCDMA, transmission occurs in fixed-frame sizes with minimal duration of 10 ms, and the rate may change only between frames (it is fixed within a single frame). Let us denote with $r_i = R_i\ \nu_i$ the transmission rate of the user i (R_i is the bit rate, and ν_i is the activity factor), and with $SIR_i = (E_b/N_0)_i$ the signal-to-interference ratio of user i. If we assume a large number of users in a cell (e.g., low-rate service), then the assumption $(W/r_iSIR_i)>>1$ is valid. In this case, using (7.86) it is obtained

$$\sum_{i=1}^{N} r_i SIR_i = \frac{\eta_{UL}}{(1+i)}\, W = \eta_{UL}\ 'w \qquad (4)$$

Where W is the chip rate (e.g., W = 3.84 Mcps for WCDMA) and η_{UL} is the uplink load factor. With the aim of achieving fair resource allocation, wireless channels should be allocated in proportional weights, as given by

$$r_i SIR_i = \frac{w_i}{\sum_j w_j} \eta_{UL}{}' w \qquad (5)$$

Assuming that the user can potentially control both the transmission rate in the uplink and the SIR, we can use the above relation to calculate the needed SIR_i for fixed rate requirements r_i (e.g., CBR service), or to provide a given frame error ratio (FER) for user i (i.e., fixed SIR_i) by applying rate adaptation (i.e., by varying r_i).

In the downlink the limiting factors are the base station's total transmission power and multipath fading. Because of multipath fading, the received signal quality at mobile terminals will fluctuate. Therefore, it is convenient to use average power levels in the downlink and then calculate the transmission rate. The average power for user i can be written as

$$\bar{P}_i = \frac{w_i}{\sum_j w_j} \eta_{DL}\ P \qquad (6)$$

where η_{DL} is the downlink load factor, and P is the total transmission power of the base station. Because of the multipath, users at different locations in the cell experience different path loss and interference. Therefore, one may find it suitable to use average values on these parameters with the aim of avoiding dependence of service differentiation upon the mobile's location. Then transmission rates in the downlink can be calculated by

$$r_i = \frac{w_i}{\sum_j w_j} \frac{W}{SIR_i \bar{I} \bar{L}} \eta_{DL} P \qquad (7)$$

Where $\bar{I}$ and $\bar{L}$ are average values of the interference and the path loss in the cell, respectively.

4. WIRELESS CLASS-BASED FLEXIBLE QUEUING

The wireless class-based flexible queuing (WCBFQ) algorithm is a scheduling scheme created to support multiple traffic classes in wireless IP networks [i.e., real-time flows, CBR, VBR, as well as best-effort traffic (Web, FTP, and so forth)]. It should be applied at wireless access points. Our tendency in creating this scheduling algorithm was to take into consideration the high BER in the wireless environment. BER is flow-specific due to the different location of single users and the different states of the air interface. Location-dependent errors are more likely to be expected than uniformly distributed errors over the whole bandwidth of the cell. In such conditions it is to be satisfy guaranteed services when they are experiencing high error rate by increasing their share of the bandwidth. On the other hand, it is not desirable to allow flows in the error state to decrease significantly the performances of the entire wireless link.

4.1. *Class Differentiation*

The base station assigns the traffic flow a channel according to a hierarchy of priorities. The first differentiation of the traffic is into two main classes: class-A with bandwidth guarantees, and class-B for best-effort traffic. A class selector separates arriving packets into different queues for every class. Class-A is divided into CBR subclass, VBR subclass, and BEmin. CBR subclass should be used for real-time applications that have strict demands on network delay, such as voice over IP. This is high-priority class. The flows belonging to the CBR subclass will be first served until the buffer for this class is emptied. VBR is intended for real-time applications with time-varying rate, such as video streams. Because video usually has higher bandwidth demands than voice, it is given lower priority to this subclass compared with CBR. That is a consequence of the characteristics of video information, where information is referred to a limited number of video frames per second that are less deterministic than traffic such as voice. Also, video flows require many times greater bandwidth than voice-oriented services. Video communication is usually one-way (e.g., video streaming), although it can be bidirectional (e.g., video telephony). In the latter case one may decide to apply CBR subclass instead of VBR. Due to such characteristics of VBR sources, we give lower priority to VBR subclass than to CBR. But, to avoid monopolization of the bandwidth by the CBR flows, we should limit the maximal capacity that can be allocated to them. This can be accomplished by an admission control mechanism. The last subclass of class-A is dedicated to users who want to have some QoS guarantees (they should pay more for their services than class-B users).

Let us use B for a bandwidth of the wireless link. The weights assigned to flows in a subclass j are w_{ji}, i = 1, …, N, where N is the number of active flows on the link. It is defined that the throughput of each flow, normalized on the link bandwidth admitted for that subclass (RT: relative throughput)

$$RT_{ji}(t) = \frac{w_{ji}(t)}{\sum_{j=1}^{N_c} \sum_{i=1}^{N_f} w_{ji}\ (t)} \qquad (8)$$

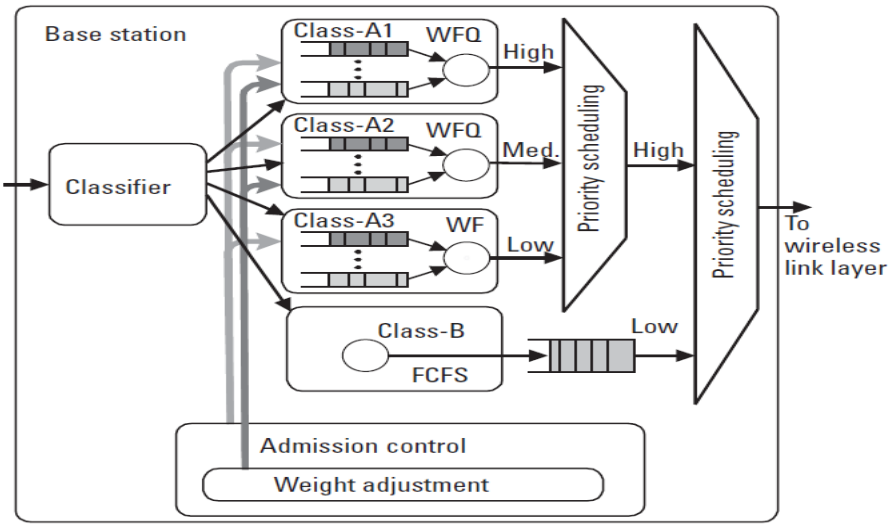

Figure 1. Model of WCBFQ scheduler

When the wireless path is error-free, the flow should get bandwidth share $b_{ji}(t)$:

$$b_{ji}(t) = RT_{ji}(t) * B = \frac{w_{ji}(t)}{\sum_{j=1}^{N_c}\sum_{i=1}^{N_f} w_{ji}(t)} B \qquad (9)$$

The above relations refer to a situation when we are using absolute weights for all flows from all classes over the entire bandwidth of the wireless link. However, we may also apply weights relatively within each class that uses fair-like queuing.

We assume that the base station has knowledge of the channel state (e.g., by monitoring or prediction), as well as which mobiles attend to send uplink data. Since location-dependent error is a specific of the wireless interface, it suggests queuing the packets during the error period. But this is not appropriate for traffic with strict delay requirements, such as voice traffic. In this scheduler there is no queuing of the packets during error state, but also there is no compensation on errors for real-time flows because it is redundant.

Maximum delay for a CBR flow i without errors is denoted as D_{CBR}^{max}, and it is given by

$$D_{CBR,i}^{max} = \frac{L_{p,max}}{B} + \frac{L_{p,max}}{B}\frac{\sum_{j\in F_{CBR}}^{N_{CBR}} w_j}{w_i} + \Delta t_p \qquad (10)$$

Where N_{CBR} is number of CBR flows, maximum packet length is $L_{p,max}$, and F_{CBR} is the set of all CBR flows. The last term Δ^{t}_{p} includes all delays due to processing, such as framing, segmentation, encoding, spreading, rate matching, and multiplexing. Usually, however, queuing delay in packet networks is higher than processing delay in order of magnitude, due to the statistical multiplexing of data.

Because the CBR subclass has the highest priority, CBR packets use all of link bandwidth B until they are all served. The maximum delay corresponds to the situation when the packet of a flow is the last on the list of the active CBR flows. Total buffer space for CBR flows can be calculated, where LCBR is the maximum length of CBR packets and NCBR is the number of CBR flows:

$$Q_{CBR} = L_{CBR} N_{CBR} \qquad (11)$$

When all CBR queues are emptied, the scheduler will start serving VBR flows. The bandwidth that is left for VBR flows can be calculated.

$$B_{VBR} = B - \sum_{i\in CBR} b_i \qquad (12)$$

Considering "Eq.(11)", the buffer requirement for the flows of the VBR subclass of class-A is calculated as follows:

$$Q_{VBR} = q_{burst} + \frac{L_{p\,max} N_{CBR}}{B} r_{VBR} \qquad (13)$$

In the calculation of buffer space for VBR flows, the bursty nature of the VBR traffic (e.g., video) should be taken into account. The additional length of the VBR queue, which is aimed to capture burstiness of VBR flow, is denoted as q_{burst}. If maximum burst duration is t_{burst} with peak rate of the flow r_{peak} and admitted rate r_{VBR}, then it can be calculated using

$$q_{burst} = t_{burst}\left(r_{peak} - r_{VBR}\right) \quad (14)$$

Because VBR flows are serviced with a lower priority than CBR traffic, the additional delay due to higher-level traffic must be considered. The worst case delay of VBR flow includes delay due to serving higher-level A1 packets, and delay for serving packets from other VBR flows. Using the effective throughput of VBR traffic, we may calculate the worst-case delay by the following equation:

$$D_{VBR,i}{}^{max} = \frac{N_{CBR} L_{p,max}}{B_{VBR}} + \frac{L_{p,max}}{B_{VBR}} + \frac{\sum_{j \in F_{VBR}} w_j}{w_i} \frac{L_{p,max}}{B_{VBR}} + \Delta t_p \quad (15)$$

The third subclass, called best-effort with minimum guarantees (BE min), is targeted to non real-time traffic with minimal QoS guarantees. Therefore, we use a fair scheduling mechanism for this subclass, such as WFQ or WRR, together with admission control to provide the minimal QoS support. These flows are serviced with lowest priority from all subclasses within class-A. Therefore, the packets of this subclass have to wait until CBR and VBR queues are drained out. Also, a packet might wait for all other BE min flows to be served. Therefore, the A3 traffic subclass requires the following buffer space:

$$Q_{BE\,min} = \frac{L_{p\,max} N_{CBR}}{B} r_{BE\,min} + \frac{\sum_{i \in F_{VBR}} Q_{VBR_i}}{\sum_{j \in F_{VBR}} r_{VBR_j}} r_{BEmin} \quad (16)$$

Each of the classes, class-A and class-B, are scheduled in different queues. Modification of the WFQ is applied for class-A traffic. Class-B flows get the remaining part of the bandwidth after class-A flows are serviced. Most class-B flows are based on the TCP protocol. TCP adjusts to the available bandwidth by managing its congestion window, and in longer time intervals TCP flows get equal bandwidth shares of the link. However, some application may start several simultaneous TCP connections to get a larger share of the bandwidth. Hence, TCP gets as it can, but best-effort can suffer from some other aggressive flows that are established between peers based on some other protocol or agent module. Therefore, if one needs minimal QoS guarantees, then the A3 subclass for best-effort traffic should be used. Otherwise, the option is class-B, which does not offer any QoS guarantees. All class-B packets are serviced according to the FCFS principle.

4.2. *Characteristics of WCBFQ*

The choice of the limits for weight adjustment of CBR flows is left to network administrators. Typical values of the limits Li should be 2 or higher for flows that occupy the smaller part of the bandwidth, and less for flows that highly utilize the link resources. In every situation, guaranteed service means an error
free and providing the minimum guaranteed data rate.

A CBR flow carrying voice will not cause high degradation of the wireless link performance, but this is not the case with video content. Video streams usually occupy a larger amount of the bandwidth and they may produce higher performance oscillation in the wireless link. For best-effort flows we may apply any of the existing schedulers created for a wireless LAN environment.

When does a flow enter an error state? The scheduler at the base station with TDD access technology services packets in both the uplink and downlink. In a multiple access technology, different schedulers may be applied in different directions. The flow transits into an error state if the average number of time slots or frames with detected errors divided by the total number of allocated time slot/frames to that flow is over the predefined error threshold Compensation methods refer only to the location-dependence of bit errors in the wireless link, but they do not capture the requirements from real time flows. Wireless errors usually occur in bursts, because of the inertia of signal propagation in a cellular network, as well as the inertia of users' movement in time intervals comparable to the time needed for processing of an individual IP packet (e.g., several milliseconds). By using the WCBFQ algorithm, we address both issues: the location-dependence of wireless bit errors and the multiclass environment.

5. Conclusion

It is concluded that Future generation mobile systems are expected to include heterogeneous wireless access networks (3G, WLAN, WPAN) with multiple traffic classes. Such a scenario requires traffic classifications, appropriate dimensioning, admission control, efficient mobility, and location management. In Wireless networks side, the key characteristics are Mobility of the users, Bit errors in the wireless channels, Scarce wireless resources and In IP network side, the key problems are Lack of QoS support, Lack of data synchronization. A scheduling algorithm is proposed for multiclass wireless IP networks called wireless class-based flexible queuing, which is flexible to different traffic demands from different traffic classes. It provides real-time compensation for A1 and A2 flows, Where A1 traffic is given higher priority for compensation than A2. Because subclass-A3 is targeted to non real-time traffic, servicing these packets with a lower priority than subclasses A1 and A2, but minimal bandwidth guarantees are provided by some of the wireless fair algorithms (e.g., CSDPS and WFS), which are adaptations of WFQ to the wireless environment. In, class-B traffic is serviced using the FCFS scheduler because this traffic class is defined for traffic without any QoS guarantees (identical to today's best-effort traffic in the Internet). Finally this paper provides traffic dimensioning, analysis, and optimization, as well as for the design of wireless IP networks.

REFERENCES

[1] Eckhardi, D. A., and P. Steenkiste, *Effort-Limited Fair (ELF) Scheduling for Wireless Networks,* INFOCOM 2000, Tel Aviv, Israel, March 2000.

[2] Jiang, Z., L. F. Chang, and N. K. Shankara narayanan, *Providing Multiple Service Classes for Bursty Data Traffic in Cellular Network*, INFOCOM 2000, Tel Aviv, Israel, March 2000.

[3] Moorman, J., and J. Lockwood, *Multiclass Priority Fair Queuing for Hybrid Wired/Wireless Quality of Service Support,* IEEE Mobicom/Wow Mom, Seattle, WA, August 1999.

[4] Eugene, T. S., I. Stoica, and H. Zhang, *Packet Fair Queuing Algorithms for Wireless Networks with Location-Dependent Errors,* INFOCOM 1998.

[5] Lu, S., V. Bharghavan, and R. Srikant, *Fair Scheduling in Wireless Packet Networks,* ACM Sigcomm '97, Cannes, France, September 1997.

[6] Bharghavan, V., S. Lu, and T. Nandagopal, *Fair Queueing in Wireless Networks: Issues and Approaches,*IEEE Personal Communications Magazine, Vol. 6, No. 1, February 1999.

[7] Ramanathan, P., and P. Agrawal, *Adapting Packet Fair Queuing Algorithms to Wireless Networks,*ACM Mobicom'98, Dallas, TX, October 1998.

[8] Lu, S., T. Nandagopal, and V. Bharghavan, *A Wireless Fair Service Algorithm for Packet Cellular Networks,* ACM Mobicom'98, Dallas, TX, October 1998.

[9] Siris, V. A., B. Briscoe, and D. Songhurst, *Service Differentiation in Third Generation Mobile Networks,* 3rd International Workshop on Quality on Future Internet Services (QofIS'02), Zurich, Switzerland, October 16–18, 2002.