RC 19980 (11/16/94)
Telecommunications

IBM Research Report

Issues in Wireless Access Broadband Networks

Mahmoud Naghshineh

IBM Research Division
T.J. Watson Research Center
Yorktown Heights, New York

Mischa Schwartz and Anthony S. Acampora

Department of Electrical Engineering &
Center for Telecommunications Research
Columbia University
New York, New York, 10027

IBM Research Division
Almaden · T.J. Watson · Tokyo · Zurich

Issues in Wireless Access Broadband Networks

Mahmoud Naghshineh* Mischa Schwartz† Anthony S. Acampora †

Abstract

In this paper, we discuss the effect of wireless link characteristics and mobility in wireless access broadband networks, and how major differences between wireline and wireless transmission links introduce a new networking paradigm with difficult and interesting challenges. We show that these differences have a direct impact on the packet-level quality-of-service defined by the packet loss, delay and jitter statistics. Moreover, we examine a number of alternatives such as forward error correction, antenna diversity and retransmission schemes to combat the time varying nature and high bit error rate of the wireless link, as well as how much and in what way these schemes can improve the packet-level quality-of-service. We also show that routing, call admission control, and capacity allocation functions of wireless access broadband networks need to take mobility into account. These functions can be invoked in a re-active manner at every hand-off event, or they can be invoked in a pro-active manner at the call setup time instead of being invoked at every hand-off event. In addition, we discuss the trade-offs between these two call control schemes, and show that each of these schemes is associated with a major cost. To find a balance between the cost of these two call control schemes, we use a connection tree as a basis for routing, call admission control, and capacity allocation in wireless access broadband networks. Finally, we discuss a number of issues in traffic integration for wireless access broadband networks and how they differ from wireline broadband networks.

*IBM Research Division, T. J. Watson Research Center, Yorktown Heights, NY 10598. Email: {mahmoud}@watson.ibm.com.

†Center for Telecommunications Research, and Department of Electrical Engineering, Columbia University, New York, NY 10027.

1 Introduction

The provisioning of wireless access broadband networks, carrying multimedia, packetized (cell-based) traffic introduces a set of challenging and interesting issues. All such issues are a direct result of two major differences between wireless and broadband wire-line networks: (1) link characteristics, and (2) mobility. The broadband network transmission links are characterized by high transmission rates (in the order of Gbps) and very low error rates. In contrast, wireless links have a much smaller transmission speed and a much higher bit error rate as compared to wire-line links of the broadband networks. In addition to the opposing characteristics of the wire-line and wireless links in terms of transmission speed and error rate, the wireless link is affected by shadowing and multipath fading which result in the fact that the bit error rate of a wireless link can be highly varying, with time and spatial dependencies. The other major difference between fixed, wire-line broadband networks and wireless access broadband networks is user mobility. In the wire-line broadband networks, the user-network-interface (UNI) remains fixed throughout the connection. However, in future wireless access broadband networks, the user access point to the wired network may change many times throughout the lifetime of a wireless/mobile connection (see Figure 1). Mobility of users connected through a wireless link to a wire-line broadband network necessitates revision and redesign of some essential attributes of wire-line broadband networks such as resource allocation, virtual path routing, call setup and admission control, and quality-of-service provisioning for integrated, multimedia connections. In this paper, we discuss a number of issues which are a direct result of the above mentioned differences between fixed and wireless-access broadband networks, and how these differences introduce a new paradigm with difficult and interesting challenges.

Broadband networks are designed based on the assumption that the packet (or cell) loss probability is in the order of 10^{-6}. With a typical bit-error-rate of 10^{-2} to 10^{-4}, the transmission quality of a wireless link is very different from a typical transmission link of today's broadband networks. In this paper, we show that the wireless link characteristics have a direct impact on the packet-level quality of service defined by the packet loss, delay and jitter statistics in broadband networks. We examine forward error correction, packet retransmission, and antenna diversity schemes to combat the time varying nature and high bit error rate of the wireless link, and to make the wireless link appear with bit-error-rate characteristics comparable to the transmission links of broadband networks. In Section 2, we discuss each of these techniques and compare their trade-offs.

Wired broadband networks are being designed with the assumption that call admission and resource allocation functions are invoked at the call setup time, and a set of virtual circuits along with their associated network resources are reserved for the duration of a connection. In wireless access broadband networks, the set of network resources assigned to a mobile user may change many times throughout the connection's lifetime. This is due to the frequent hand-off of mobile/wireless users which is a result of the use of small radio cells (micro/picocells) required for provisioning of the high capacity required to support bandwidth-intensive multimedia applications. We show that mobility (or hand-off) has a direct impact on the call-level quality-of-service as measured by the availability of network resources to support a mobile connection after a hand-off event. This means that existing call admission control policies of broadband networks need to be extended to guarantee not only a packet-level quality-of-service, but also to guarantee a call-level quality-of-service. In [1], it was suggested that the virtual connection tree can be used for routing, call admission and resource allocation in wireless access broadband networks. In Section 3, we discuss this issue in more detail and provide an analytical formulation of connection tree based resource allocation and call admission problem. Finally, we discuss the integration of multiple classes of traffic in wireless access broadband networks.

2 Wireless Link Issues

Integration of wireless networks into fixed broadband networks introduces a number of issues which arise from the inherent difference between wireless links and wired links of broadband networks in terms of transmission speed and bit-error-rate (BER). The immense difference between the transmission speed of these two types of links may result in buffer overflows at access points which are switches between the wired and wireless network. To overcome this, existing congestion control schemes proposed for integration of low speed links into high-speed broadband networks can be applied. The function of such schemes is to control the transmission of packets into the high-speed network in order to reduce the buffer overflow of switches connected to low speed links.

The difference between wireless and wire-line links in terms of bit-error-rate is another major factor to be considered. Here it is important to take the characteristics of a wireless link into account. The wireless environment is characterized by two major components: (1) Long-term fading and (2) Short-term fading. Long-term fading is the case in which the average or the local mean of the fading signal changes slowly with time. Short term or multipath fading represents the fast fluctuation of the fading signal due to multipath reflections and local scatterers [2]. Both of these components can result in the degradation of the received signal and a high bit-error-rate (BER) which results in a high probability that a packet is received in error. Packet-error-rate (PER) can have different consequences depending on the type of traffic. For example, data or non-realtime applications which require a reliable transport can tolerate delay but cannot tolerate error. This means that a data packet received in error needs to be corrected (retransmitted) by the link or transport layer until it is correctly received. On the other hand, realtime traffic is delay sensitive but loss tolerant. This means that error control for realtime traffic is only necessary if the packet-error-rate is unacceptably high. This is because packets can be dropped once in a while without major degradation of the video or voice quality. It is important to note that a realtime packet received in error can be re-transmitted if the strict delay requirements of its realtime application can be met. To clarify this point, let us assume that a realtime application running on a wireless terminal needs to receive one packet per time interval $[(i-1)T, iT]$. A packet received at time t ($(i-1)T \leq t \leq iT$) can be retransmitted if the retransmission takes place in the time period $T_r = iT - t$. If after the time period T_r the packet still contains errors, it must be dropped. We refer to this as the *limited retransmission* of realtime packets. Forward error correction (FEC) and antenna diversity techniques can also be used to reduce the bit- and packet-error-rate [2, 3]. These techniques increase the cost and complexity of the mobile terminal since FEC requires encoding/decoding circuits, and antenna diversity requires digital signal processing. One can also compare these schemes from the packet-level QOS and the throughput point of view. Retransmission results in higher delay and jitter and decreases the link throughput. FEC reduces the link throughput; however, it does not affect the packet delay and jitter. Antenna diversity neither reduces the link throughput nor increases the packet delay or jitter; however, its complexity increases exponentially with the number of antennas. Obviously, a combination of the above techniques can be used to tailor a system's cost/performance to a given radio environment. We provide the following example in order to clarify this point. We consider a radio environment with multipath fading (Rayleigh) in which fixed length ATM cells (53 Bytes) generated by a real time application are transmitted over the wireless link between the mobile and its access point to the broadband network. We assume that the realtime application (such as voice or video) can tolerate a 5% PER. Moreover, we assume that the average carrier-to-interference ratio (CIR) remains constant during a packet transmission, and that the bit errors due to multipath fading occur independently. In Table 1, we compare the packet-error-rate (PER) [1] of the system using FEC and/or antenna diversity with maximal ratio combining [2, 4]. To incorporate the effect of changing radio environment, we consider three different bit-error-rates (0.06, 0.01, and 0.0001). The first column shows the BER using Binary PSK modulation in a Rayleigh fading environment when no FEC or antenna diversity is used [5]. The second column represents the code characteristics (n, k, t) of a BCH code where n, k and t represent the block length, the number of information bits closest to an ATM cell (l=424 bits), and the error correcting capability [3]. The third column shows the number of antennas M deployed in the system where maximal-ratio-combining is used [2, 4]. Finally, PER shown in column four is calculated based on the BER of a system using a BCH code as shown in column two, and M antennas as shown in column three. The last column represents the link efficiency based on the FEC overhead where efficiency is defined as l/n. First, we consider a harsh radio environment with multipath fading giving rise to a high BER equal to 0.06. Here a PER of less than 5% can be achieved by using FEC with 60% overhead (40% efficiency) and a single antenna, by using FEC with 17% overhead (83% efficiency) and three antennas, or finally, by using six antennas and no error correcting code. All of these combinations have the same packet-level QOS in terms of packet loss or jitter. However, for this case, a solution based on only FEC seems to have a very low efficiency (40%). On the other hand, since the complexity of antenna diversity techniques increases exponentially with the number of antennas, a solution using six antennas seems to be very expensive/complex. Hence, a combination of FEC and antenna diversity seems to be a good design point for a high BER. Now, if we look at the system performance with a BER equal to 0.01, a PER less than 0.05 can be achieved by using FEC with 83% efficiency, or by using three antennas . For the case where BER is 0.0001, a single antenna would suffice without using

[1] Here PER represents the ATM cell-error-rate.

FEC. Since in the radio environment the carrier-to-interference ratio changes with time and spatial dependencies, an adaptive error correcting and diversity system can be used where the level of error correction and the number of activated antennas are selected based on the harshness of the radio environment. This is an important factor regarding the mobile terminal's battery or power consumption. Considering our example, in a radio environment where BER changes between 0.05 and 0.0001, a (551, 430) BCH coding/decoding logic along with three antennas can be applied adaptively to the fading environment. As the BER increases, the number of enabled antennas, encoding/decoding, and signal processing circuits is increased to provide the desired packet-level QOS.

Another important issue regarding the forward error control schemes for realtime traffic is the following. In the case that error detection with limited retransmission is used for realtime traffic, the error detection and the decision for a retransmission need to be made in a very short time due to the delay sensitivity of the realtime traffic. Hence, the error control and recovery need to be done at the time that a packet is received from the physical layer before passing it to the higher communication layers. This would increase the complexity of the MAC protocol. On the other hand, if by using forward error correcting codes and/or antenna diversity an acceptable PER (say 5%) can be achieved, the retransmission can be avoided and the MAC complexity can be reduced.

As in the case of realtime traffic, transmission of data or non-realtime traffic over a wireless link with a high BER results in a degradation of the packet-level QOS in terms of packet delay and loss. [2] In the past, error control schemes for data traffic have been studied extensively, and compared in terms of complexity and performance (See references in [3] for example). Since data traffic is delay tolerant and loss sensitive, error control schemes designed for the delay sensitivity and loss tolerance characteristics of realtime traffic would not necessarily be efficient for data traffic. In wireless access broadband networks, realtime and non-realtime packets are multiplexed on the same wireless link, and hence, the design of efficient error control schemes for wireless links carrying mixed traffic types is an important issue. For example, error control schemes designed for realtime traffic along with packet retransmission can also be used for error free packet delivery of data traffic over a wireless link. Again referring to Table

[2] Here we assume that the non-realtime or data traffic is generated by an application which requires a reliable (error-free) stream transport.

1, in a radio environment with strong multipath fading giving rise to a BER of 0.06, FEC and antenna diversity can be used to provide a packet error-rate in the order of 10^{-5} which is acceptable for realtime traffic. In addition, packet retransmission can be used for error-free delivery of data traffic. [3]

Regarding the non-realtime traffic, an important issue in wireless access broadband networks is the placement of the error control function. To clarify this, in Figure 1, we show a wireless station (M) which is communicating over a wireless access broadband network with a fixed station (S) using a base station (B). Error control could be done at the transport layer of end stations M and S. This is shown in Figure 2(a). In this case, every packet received in error at the transport layer of M or S needs to be retransmitted. As a result, wired and wireless resources that comprise a route between the two end stations (M and S) will be utilized for every packet retransmission. For long-propagation paths in the wired network this could imply relatively long delays when retransmitting packets due to errors incurred on the relatively short-delay wireless-path segment. This potential problem could be overcome by terminating the transport layer at the access point B and providing a reliable link layer between the access point B and the mobile M. This is shown in Figure 2(b). This implies protocol conversion at the access point, leading to a tradeoff between delay (continuing the transport layer across to the mobile) and implementation complexity (terminating the protocol at the access point). Alternatively, a wireless link error control protocol could be implemented between the mobile M and the access point B to provide a PER comparable to the wired network, in addition to an end-to-end error control at the transport layer of end stations S and M. This is shown in Figure 2(c). In this case, the wireless/wire-line link characteristics would be hidden from the transport layer, and only the wireless resource would be used for the error control when the error is caused by the wireless link.

3 Mobility Issues

We can view wireless/mobile connections as consisting of (1) paths or routes through the broadband backbone network and (2) radio links between the mobile, wireless terminals and base stations (or access points) which provide the interface of mobile users to the fixed backbone network. When the quality of a radio link between a wireless

[3] This is reasonable since the packet error rate is very small.

terminal and its access point degrades, a new access point with acceptable quality must be found, leading to hand-off, and network control functions of both the fixed and wireless network need to be invoked. In the backbone network, hand-off requires that a new route needs to be established which transports the packets destined to (or originating from) the wireless terminal to (or from) the new access point. In this case network call processing functions need to be invoked to setup such a route and to ensure that the newly established route maintains acceptable quality-of-service (QOS) to both the wireless connection and to pre-existing calls sharing links of the new route [1]. In wireless access broadband networks, an issue of paramount importance is the fact that routing, call admission and resource allocation functions must take mobility into account. These call control functions can be performed in a re-active or pro-active manner. In the *re-active call control* methodology, network call control functions are invoked after every hand-off to establish a route, to allocate wired and wireless capacity, and to maintain the packet-level QOS of existing connections which are affected by the hand-off. In the *pro-active call control*, the routing, call admission control, and resource allocation functions of both the wired and wireless network take the mobility of end terminals into consideration. Namely, at the time that a mobile/wireless connection is admitted to the network, the routing, call admission, and capacity allocation functions ensure that the packet-level QOS of the admission seeking connection as well as the existing connections can be maintained throughout the duration of the connection which includes potentially a large number of hand-offs. Reducing the call processing load of a wireless access broadband network is an important advantage of the pro-active call control methodology. This is due to the fact that all network control functions are intensive tasks in terms of communication and/or computation, and their invocation at every hand-off event might result in a major bottleneck of future micro-picocellular networks [6] - [8]. [4]

3.1 Mobile Routing

Broadband networks are designed based on the virtual circuit/path routing concept where at call setup time a route is established which transports packets between end stations [9]. In wireless access broadband networks, new routing schemes are required

[4] Note that in order to support bandwidth intensive multimedia applications, wireless access broadband networks are likely to use micro/pico cellular architectures. This gives rise to a high frequency of hand-off events.

to accommodate mobility and hand-off. Moreover, these routing schemes must be based on the existing virtual circuit/path routing schemes of broadband networks. In the past, a number of mobile routing schemes were proposed for broadband, ATM-based networks [7], metropolitan area networks [6, 10], and IP-based networks [11] - [13]. In general, mobile routing schemes are tree-based, forwarding-based, or a combination of both. Forwarding-based mobile routing solutions define one or more mobile support stations (MSS) which act as forwarding agents. As a mobile hands off from one wireless subnet to another, the MSS in the old subnet will be notified about the new location of the mobile. As a result, packets destined to the mobile and received by the old MSS are *forwarded* to the wireless subnet where the mobile is residing. Basically, as the mobile roams in the network, the forwarding agents concatenate a set of routes in order to deliver its packets. This results in the fact that in strictly forwarding-based mobile routing schemes the number of network nodes and links that a packet destined to (or originating from) a mobile traverses could become extensive and result in an inefficient use of the network resources as well as an unacceptably high packet delay or jitter. In a sense, the forwarding scheme is a *re-active* mobile routing methodology. In the tree-based mobile routing scheme [7], a routing tree is defined at the time that a mobile/wireless user seeks admission to the wireless access broadband network. This routing (or connection) tree covers the mobile's radio cell and a number of adjacent radio cells and their base stations. As shown in Figure 3, the root of the tree is a fixed switching node of the wired network and the leaves of the tree are radio cells. For each mobile connection, the connection tree provides a set of virtual connections (in each direction), each providing a path from the root to one leaf. To complete a mobile connection in the backbone network, a fixed virtual connection is created for that connection from the tree's root node back to either a wired network port (if the connection is to a fixed port) or to the root of some other connection tree (if the connection is to another mobile user). After a hand-off, a new route to a mobile is enabled without the involvement of the network call processor and in a distributed manner. The connection tree-based routing scheme is a *pro-active* mobile routing methodology, where future hand-offs of a mobile are anticipated at the call setup time.

Regarding mobile routing schemes for wireless access broadband networks, two important issues need to be considered: (1) re-routing overhead measured in terms of frequency of the network routing functions' involvement for every hand-off and (2) network efficiency measured in terms of the average network resources used by mobile connections in the network. The re-routing overhead is important since mobile routing schemes which require the involvement of network control functions at every hand-off event would impose a major limit on the number of connections that can be admitted to a network. In addition, the amount of time required for changing a route plays an important role in the ability of a mobile routing scheme to support delay sensitive realtime applications. The efficiency of the re-routing scheme is another important factor, and represents its optimality in terms of usage of network resources.

3.2 Call Admission Control and Capacity Allocation

Call admission control and capacity allocation in wireless access broadband networks must take mobility and the wireless link characteristics into account. As far as the wireless spectrum is concerned, overload conditions may occur if the communication needs of a number of wireless terminals populating a small area exceed the total wireless capacity of all access points within their reach. We refer to this situation as a radio congestion state. Depending on the load of the wireless network and the manner in which the radio resource is shared among mobile users (fixed channel allocation, dynamic channel allocation, etc.), a radio congestion state might be encountered at hand-off, an event which would result in either the termination of the connection, large delays, and/or packet loss. In the past, several wireless call admission schemes were proposed to keep the probability of a congested radio state suitably low [1], [14] - [17]. In [17], it was found that hand-off dropping can be improved by taking into consideration the number of calls in adjacent cells in addition to the number of calls in the cell where a new call request is made, in order to make a wireless call admission decision. These solutions reserve some resource (e.g. radio spectrum) in the wireless network to limit the forced call termination (or congestion) probability of admitted calls in the network. Since these call admission control functions take only the wireless resources into account, they are not sufficient to provide quality-of-service in wireless access broadband networks which requires consideration of both wired and wireless resources in the call admission and capacity allocation functions.

As we discussed earlier, the capacity allocation and call admission control in wireless access broadband networks can be performed in a re-active, or a pro-active manner. The re-active scheme has two disadvantages. First of all, it requires the call control functions to be invoked at every hand-off, hence leading to a potential bottleneck on the network capacity. Second, a re-active call admission control and capacity allocation scheme does not guarantee the availability of wireless and wired network resources to support the desired packet-level QOS of a mobile connection after a hand-off. This is due to the fact that the availability of a wireless access broadband network's resources to provide a connection from a mobile terminal to its destination(s) at the call setup time does not necessarily mean that network resources are available to provide a new connection after a hand-off. This is a major departure from the framework of call admission control and capacity (or resource) allocation in the fixed broadband networks. To guarantee that the packet-level QOS of a wireless/mobile connection can be maintained throughout its lifetime, a pro-active call control scheme can be used. Namely, at the time that a wireless/mobile connection seeks admission to a wireless access broadband network, call control functions of both wired and wireless networks must be invoked to find a set of network resources that can provide the desired packet-level QOS of the admission seeking connection throughout its connection lifetime. This is an important factor from the point of view of subscribers of wireless access broadband networks who presumably want a guarantee on the availability of wireless and wired resources during their connection to support their desired packet-level QOS. The pro-active call control methodology overcomes the two disadvantages of the re-active call control methodology. The cost for achieving this is the fact that the pro-active call control scheme requires over-allocation of the network resources. To clarify this point, let us define $C_{P_{QOS},r}$ as the amount of capacity required using the resource r of the wired or wireless network in order to meet the packet-level QOS of a connection based on its traffic parameters (for example, see [18, 19, 20]) and the link characteristics. [5] Also, let us denote $\mathbf{R}$ as the set of all wired and wireless network resources that the mobile uses during its connection time which might include a number of hand-offs. To provide an absolute guarantee that the packet-level QOS of a mobile connection is maintained throughout its lifetime, an amount of capacity $C_{P_{QOS},r}$ must be reserved on every network resource $r \in \mathbf{R}$ at call setup time. There are two important issues to be considered here. First, as suggested in [7], $\mathbf{R}$ can be arranged in a connection tree in order to maximize the reuse of network resources and to limit the total capacity reserved for the connection. Second, this

[5] As we discussed in Section 2, for the wireless resource, the characteristic of the radio link and the amount of capacity required to support a realtime or data connection need to be taken into account.

approach can be very wasteful since only a small subset of $\mathbf{R}$ is used at any time by the mobile connection. By sacrificing the absolute guarantee on the packet-level QOS, we can reduce the total amount of capacity allocated to the mobile connection. This can be done by allocating capacity in the wired and wireless network such that after a hand-off a set of network resources are available with the probability $P_{C_{QOS}}$ to provide the packet-level QOS desired by the mobile connection. We refer to $P_{C_{QOS}}$ as the connection-level QOS. As suggested in [1], and [8], this can be done by using a connection tree for joint routing, call admission, and capacity assignment in wireless access broadband networks. Namely, we break the QOS provisioning in mobile/wireless networks up into three sub-problems: (1) we identify a connection tree (used for mobile routing) which covers possible access points that a mobile might use during its connection based on its mobility characteristics (rate or pattern); (2) we identify the amount of capacity $C_{P_{QOS},r}$ required in order to meet the connection's packet-level QOS when using wired or wireless resource r belonging to the connection tree; and (3) we allocate enough capacity for the mobile in the connection tree such that a capacity $C_{P_{QOS},r}$ is available to the mobile connection with a probability $P_{C_{QOS}}$ when using resource r of the connection tree. The connection tree can be viewed as a single collective resource of the wireless access broadband network which can be used for mobile routing, resource allocation, and call admission control in the wired and wireless network. As in statistical sharing of a virtual path in broadband networks, the connection tree resources are statistically shared among its mobile connections. Statistical sharing of resources on a virtual path affects the packet-level QOS, and statistical sharing of the connection tree resources affects the connection-level QOS. In addition, since routing, call admission, and and capacity assignment functions are based on the connection tree, they do not need to be invoked as long as a mobile station remains within the area covered by a connection tree. Whenever a mobile reaches the boundary of a connection tree, it seeks admission to a new connection tree. This is referred to as a connection tree hand-off [7]. At this point, network call control functions need to be invoked to admit the call to a new connection tree, and to allocate capacity in the new connection tree. As we decrease the size of a connection tree (which is defined as the number of radio cells that the connection tree covers), the call control scheme becomes more re-active and less pro-active. It is important to note that there is a trade-off between, on the one hand the level of QOS that can be guaranteed to mobile users and the call processing load, and on the other hand the amount of capacity allocated to a mobile user. In the following we provide a model to quantify this trade-off and to gain more insight into the connection tree-based call admission control and capacity assignment.

Let us consider a homogeneous system in which a mobile connection requires C_p units of bandwidth to meet its packet-level QOS. Furthermore, let us assume that the connection tree shown in Figure 4 is used for mobile routing and call admission in the wireless-access broadband network. Finally, let us assume that there are N mobiles admitted to the connection tree and that any test mobile can be found within any base station of the connection tree with the probability $1/B$. Here B represents the number of base stations in the connection tree. Referring to Figure 4, where we have numbered the links of the connection tree, C_k is the capacity of link k in terms of number of connections. It is important to note that the capacity of a leaf link (6, ..., 17 in Figure 2) represents also its radio cell capacity (in terms of number of calls). Let us define p_k as the probability that a test mobile's route uses link k. Then, due to our assumptions, $p_1 = 1$, $p_3 = 2/3$, $p_2 = p_4 = p_5 = 1/3$, and $p_i = 1/12$ for $i \geq 6$. The probability of having i calls with link k is given by a binomial distribution which we approximate by a Gaussian distribution $G(m, \sigma)$ with $m = N\ p_k$ representing the mean and $\sigma^2 = N\ p_k\ (1 - p_k)$ representing the variance of G. We define the overload probability P_O as the probability that link k has more than C_k mobiles. [6] P_O can then be approximated by $Q((C_k - Np_k)/\sqrt{N\ p_k\ (1 - p_k)})$ with $Q(.)$ defined as the integral over the tail of the Gaussian or normal distribution. This function can in turn be calculated in terms of the error function [21]. For a given overload probability P_O, there exists a value α such that $P_O = Q(\alpha)$. As a result, we have the following relation:

$$\alpha = (C_k - Np_k)/\sqrt{N\ p_k\ (1 - p_k)} \tag{1}$$

For the capacity assignment problem, we can solve for C_k in Equation (1), resulting in

$$C_k = \alpha\sqrt{N\ p_k\ (1 - p_k)} + Np_k. \tag{2}$$

As we can see, the capacity of the link k depends on the desired limit on the overload probability (expressed in terms of α), the size and shape of the tree (expressed in

[6] Here we assume that a Fixed Channel Assignment (FCA) scheme is used in the cellular network.

terms of p_k), and the number of calls (N) admitted to the connection tree. Using the capacity allocation scheme described above, we calculate the capacity required by each link of the connection tree described by Table 2. In that table k represents the link number, p_k represents the probability that any test mobile uses link k, C_k represents the required link capacity in terms of number of connections such that the link overload probability is limited to 1%, and ρ_k represents the utilization efficiency measured in terms of $(Np_k)/(C_k)$. As can be seen, the link utilization efficiency increases as we move toward the root of the connection tree. This is an important feature of connection tree-based call admission control and capacity assignment. It suggests that the design of a connection tree to maximize the overall utilization efficiency of network resources is an important issue. Moreover, using Equation (2), it can be easily shown that as the size of the connection tree increases, its overall utilization efficiency decreases. This shows the above mentioned trade-off between the frequency of invocation of call control functions, and QOS provisioning, and the amount of over-allocated capacity in the connection tree

Regarding the call admission problem, we need to find the maximum number of mobiles that can be admitted to a connection tree such that a desired packet-level quality of service can be maintained with the probability P_{CQOS}. First, using Equation (1), we find the number of mobiles N_k that can be admitted to the connection tree such that link k with capacity C_k of the connection tree can provide the desired QOS.

$$N_k = \frac{\alpha^2(1-p_k) + 2C_k - \alpha\sqrt{\alpha^2(1-p_k)^2 + 4C_k(1-p_k)}}{2p_k} \tag{3}$$

Since the desired QOS must be provided by all radio cells and connections of the connection tree, $N = min_{all\ k}\ (N_k)$ represents the number of calls that can be admitted to the connection tree providing the desired QOS. Using Equation (2), one can show that, as the number of mobiles per connection tree increases, the link capacity C_k approaches $N\,p_k$, or equivalently utilization efficiency ρ_k approaches 1. This suggests that, as we move toward higher capacity radio cells, connection tree-based call admission control and resource allocation can provide a packet- and connection level QOS at a very slight cost as measured by overall utilization efficiency of the connection tree wireless and wired resources. In the above model, we consider a fixed mobile population that remains in the system for a long period of time. The same call admission control and capacity assignment methodology can be used in a model with dynamic user population that incorporates the arrival and departure of mobile users. In [7] and [8], it was shown that as the arrival rate of new call admission requests increases, the overload or congestion probability of connection-tree-based call admission scheme approaches an asymptote. This asymptote depends on the system and call admission parameters.

3.3 Support of Multiple Classes of Traffic

With recent developments in the area of multimedia applications, support of multiple classes of traffic is an essential requirement in the wireless access broadband networks. Multimedia applications necessitate support of different classes of traffic with diverse quality-of-service requirements [22]. Networks supporting multimedia wireless applications require dynamic resource allocation and admission control which is a much more complex problem in wireless networks as compared to wire-line networks [23]. Hierarchical decomposition techniques have been proposed in the past for dynamic control of integrated networks with multiple classes of traffic [24]. Similarly, the dynamic control and resource allocation in integrated wireless networks can be broken up into two independent sub-problems: (1) class-based wireless call admission control; (2) wireless bandwidth allocation to all admitted connections according to some scheduling or resource sharing algorithm directly related to the class of traffic and its quality-of-service requirements.

In [25], an adaptive call admission control mechanism for wireless/mobile networks supporting multimedia traffic was introduced, and a number of resource sharing schemes which can be used to allocate wireless bandwidth to different classes of traffic were discussed. The adaptive call admission control reacts to changing new call arrival rates, and the resource sharing mechanism reacts to rapidly changing traffic conditions in every radio cell due to mobility of mobile users. In a sense, results shown in [25] identify an admission region for the wireless access broadband networks supporting multiple classes of traffic with call-level QOS guarantees. This is an extension of the admissible load region defined for packet-level QOS in broadband packet-switched networks [20].

4 Conclusion

Integration of wireless links into broadband networks necessitates provisioning of reliable radio links and new call control methodologies which take mobility into account. The quality of the wireless link (in terms of packet-error-rate) can be improved by applying schemes such as forward error correction, antenna diversity, and retransmission. A key factor to consider in such error control schemes is that they must be designed based on the traffic characteristics, and application requirements in terms of packet-error-rate. Moreover, a combination of these error control techniques can be used in an adaptive manner to mitigate the highly time varying bit-error-rate of the radio environment. The provisioning of mobile communication in wireless access broadband networks requires new routing, capacity allocation, and call admission control schemes in order to provide a packet-level quality of service. In order to reduce the call processing load and to maintain a packet-level quality-of-service, these call control functions need to anticipate future locations of a mobile user and reserve network resources for such locations. Finally, these call control schemes need to consider different classes of traffic which are differentiated based on the capacity requirements, the packet-level QOS, and the mobility pattern. This will be an important element for the support of multimedia applications in wireless access broadband networks.

References

[1] A. S. Acampora and M. Naghshineh, *Control and Quality-of-Service Provisioning in High-Speed Microcellular Networks*, IEEE Personal Communications, Vol. 1, No. 2, 2nd Quarter 1994.

[2] W. Jakes, *Microwave Mobile Communications*, IEEE Press, 1994.

[3] S. Lin, and D. J. Costello, Error Control Coding: Fundamentals and Applications, Prentice-Hall, 1993.

[4] W. C. Lee, *Mobile Communications Design Fundamentals*, John Wiley & Sons, Inc., New York, 1993.

[5] J. G. Proakis, *Digital Communications*, McGraw-Hill, 1989.

[6] D. Goodman, *Cellular Packet Communications*, IEEE Transactions on Communications, vol. 38, No. 8, August 1990.

[7] A. S. Acampora and M. Naghshineh, *An Architecture and Methodology for Mobile-Executed Hand-off in Cellular ATM Networks*, IEEE Journal on Selected Areas in Communications, Vol. 12, No. 8, October 1994. Also in the Proceedings

of the 1994 International Zurich Seminar on Digital Communications, Zurich Switzerland, March 1994.

[8] M. Naghshineh, *Distributed Control of Wireless/Mobile Networks*, Doctoral Thesis, Columbia University, 1994.

[9] J.-Y. Le Boudec, *The asynchronous transfer mode: A tutorial*, Comput. Net. ISDN Syst. 24, 1992.

[10] G. Pollini, K. Meier Hellerstern, and D. Goodman, *Handover Protocols Between Metropolitan Area Networks*, GLOBECOM'92, Orlando, FL., 1992

[11] J. Ioannidis, D. Duchamp, and G. Maguire, *IP-based protocols for Mobile Internetworking*, Proc. ACM SIGCOMM, 1991.

[12] F. Teraoka, M. Tokoro, *Host Migration Transparency in IP Networks*, Comp. Commun. Rev., Jan 1993.

[13] C. Perkins, *Providing Continuous Network Access to Mobile Hosts Using TCP/IP*, Computer Networks and ISDN Systems, Nov. 1993.

[14] E.C. Posner and R. Guerin, *Traffic policies in cellular radio that minimize blocking of handoff calls*, Proc. 11th Teletraffic Cong. (ITC 11), Kyoto, Japan, Sept. 1985.

[15] S. Rappaport, *The Multi-Call Hand-off Problem in High Capacity Cellular Communications Systems*, IEEE Trans. on Veh. Tech., Vol. 40, No. 3, August 1991.

[16] M. Naghshineh, A. S. Acampora, *Design and Control of Micro-Cellular Networks with Quality-of-Service Provisioning for Real-Time Traffic*, Intl. Conference on Universal Personal Communications (ICUPC'94), San Diego, CA., September 1994. Also, to appear in the Journal of High-Speed networks' special issue on PCN, 1995.

[17] M. Naghshineh, M. Schwartz, *Distributed Call Admission Control in Mobile/Wireless Networks*, IBM Research Report Feb. 1995. Also, submitted to PIMRC'95

[18] J. M. Hyman, A. A. Lazar, G. Pacifici, *Real-Time Scheduling with Quality of Service Constraints*, IEEE Sel. Areas in Communications, No. ISAC-7, Sep. 1991.

[19] R. Guerin, H. Ahmadi, and M. Naghshineh, *Equivalent Capacity and its Applications to Bandwidth Allocation in High-Speed Networks*, IEEE Sel. Areas in Communications, No. ISAC-7, Sep. 1991.

[20] J. M. Hyman, A. A. Lazar, and G. Pacifici, *A Separation Principle Between Scheduling and Admission Control for Broadband Switching*, IEEE Journal on Selected Areas in Communications, Vol. 11, N0. 4, May 1993.

[21] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, 1991.

[22] A. Campbell, G. Coulson, F. Garcia, D. Hutchison, and H. Leopold, *Integrated Quality of Service for Multimedia Communications*, Proceedings IEEE Infocom'93.

[23] R. Schaefer, *European R & D Activities in Mobile Multimedia Communications and Expectations on Future Services*, IEICE Trans. Commun. Vol. E77-B, N0. 9, Dept. 1994.

[24] R. Bolla, F. Danovaro, F. Davoli, and M. Marchese, *An integrated dynamic resource allocation scheme for ATM networks*, Proceedings of IEEE INFOCOM '94.

[25] M. Naghshineh, and A. S. Acampora, *QOS Provisioning in Micro-Cellular Networks Supporting Multimedia Traffic*, INFOCOM'95, Boston, MA, April 1995.

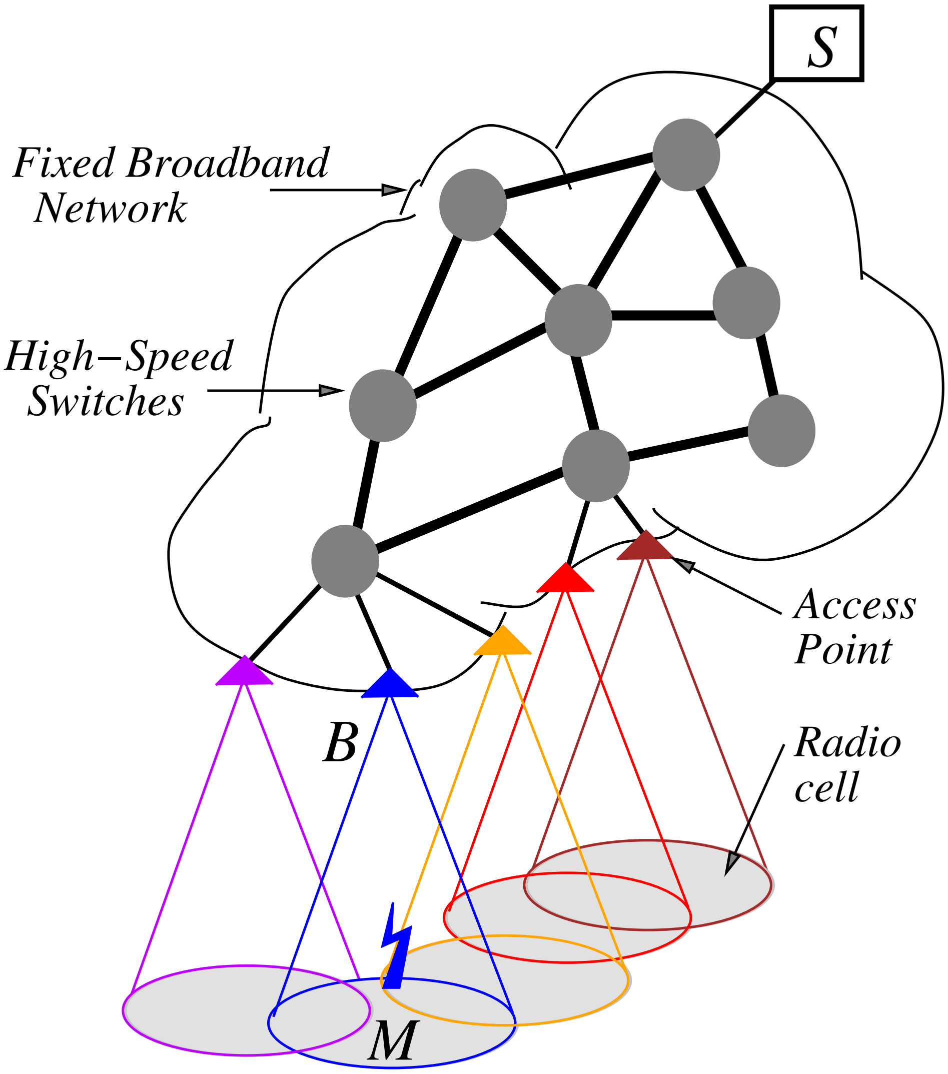

Figure 1: A Wireless Access Broadband Network

BER BPSK	**FEC (n, k, t)**	**M**	**PER**	**FEC Eff.**
0.06	none	1	$\simeq 1$	-
	(1023,433,74)	1	0.05	40%
	none	2	0.99	-
	none	3	0.77	-
	(511,430,9)	3	0.00002	83%
	none	6	0.02	-
0.01	none	1	0.98	-
	(511,430,9)	1	0.03	83%
	none	2	0.12	-
	none	3	0.005	-
0.0001	none	1	0.05	-
	none	2	0.00002	-

Table 1: Effect of antenna diversity, and FEC on Packet-level QOS in a multi-path fading environment

k	p_k	C_k	ρ_k
6, ...,17	1/12	15	56%
2, 4, 5	1/3	45	74%
3	2/3	78	85%
1	1	100	100%

Table 2: Capacity assignment in a connection tree with $N = 100$ mobiles and overload probability ≤ 0.01

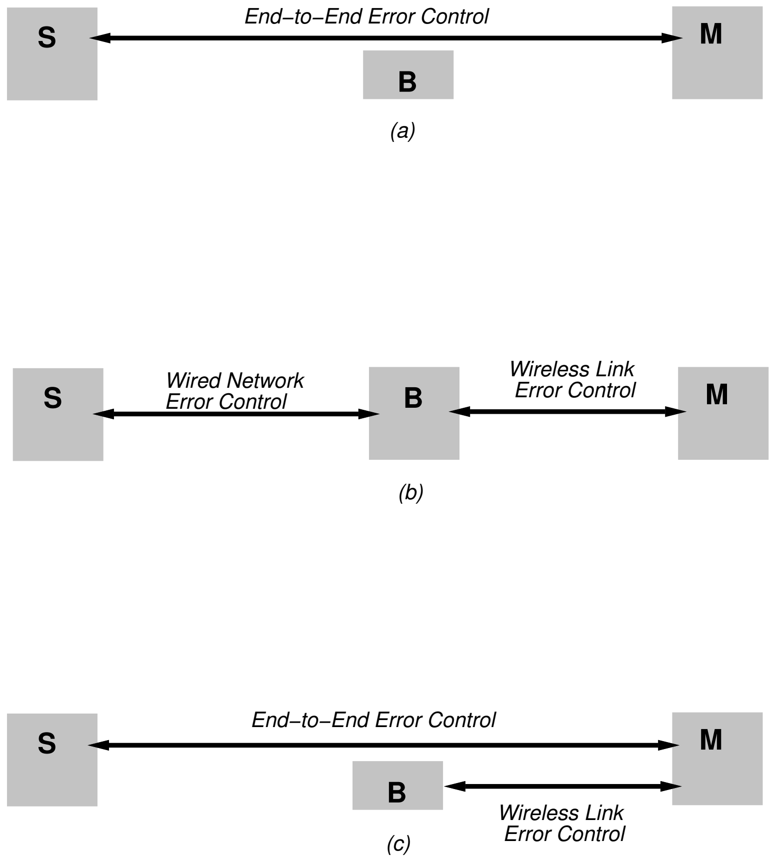

Figure 2: Placement of error control function

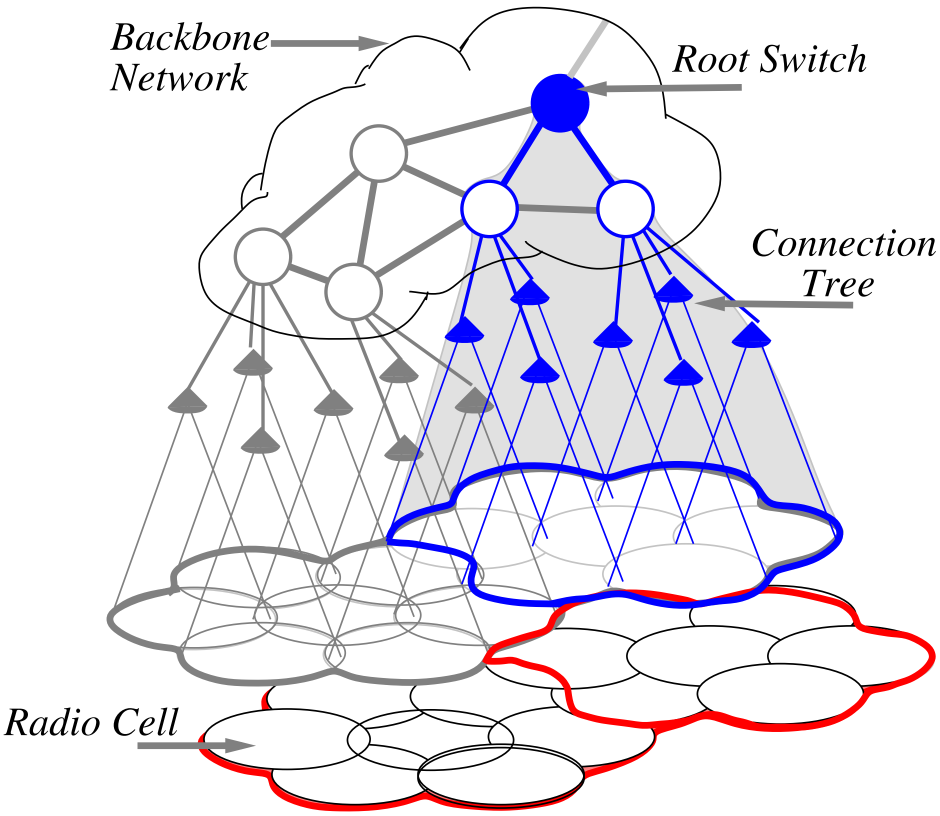

Figure 3: A Connection Tree in a Wireless Access Broadband Network

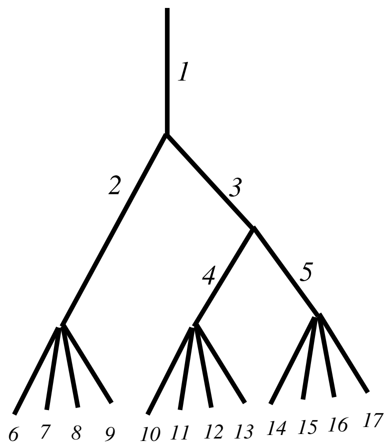

Figure 4: An Example Connection Tree