J. Indian Inst. Sci., Mar.–Apr. 2000, **80**, 99–112

A system architecture for mobile broadband access using wireless ATM

MANONEET SINGH[*], SUBRAT KAR[†] AND SURENDRA PRASAD
Department of Electrical Engineering, Indian Institute of Technology (IIT), Hauz Khas, New Delhi 110 016.
Phone. +91-11-6861977, Fax: +91-11-6966606
emails. manoneet@stanford.edu, {subrat, sprasad}@ee.iitd ernet in

Abstract

We present a system-level architecture for supporting mobility in an ATM-based wireless personal communication network (PCN). The proposed architecture uses a networking scheme based on intelligent, multicast-based trees with distributed mobility management and predictive resource allocation at the base stations The addressed functionalities include connection admission, location management and tracking and handoff procedure with QoS support. The associated protocols for implementing these functions are outlined and the architecture is compared vis-à-vis currently existing proposals in the literature. In doing so we discuss key design issues, and speculate on future directions in the development of broadband wireless packet communication systems

Keywords: Wireless ATM, mobility, QoS

1. Introduction

The emergence of portable multimedia computing platforms and growing demand for mobility among communication services has led to the need for extending ATM-like virtual connectivity from the wired to the wireless domain. The ATM cell relay paradigm, which forms the basis for the emerging B-ISDN, is now also being actively considered as a potential framework for next-generation wireless communication networks capable of supporting integrated, quality of service (QoS)-based multimedia services leading to a truly tetherless and seamless connectivity for mobile applications.[1]

The provisioning of wireless ATM involves several challenging and interesting issues, of which we mention just three here. *First*, in contrast to the reliable *physical* (wired) environment for which ATM was designed, the wireless medium suffers from high-bit error rates and limited bandwidth, necessitating the use of sophisticated strategies to mask the influence of phenomena like fading, multipath propagation, intersymbol interference (ISI) and shadowing on the QoS of an ATM VCC (virtual connection circuit). *Second*, the integration of W-ATM into the fixed B-ISDN requires that end-to-end connectivity be maintained through protocols that seamlessly *inter-network* the wired and wireline portions of the network. *Finally*, while current ATM standards proposed by the ITU are designed to provide connection-oriented sup-

*Now with the Department of Electrical Engineering, 350 Serra Mall, Stanford University, Stanford, CA 94305.
†To whom all correspondence should be addressed. The paper was presented, in part, at the IEEE *Int. Conf. on Personal Wireless Communications* ICPWC-99, Jaipur, India, Feb. 17–19, 1999.

port to wireline users at fixed locations, wireless users are *mobile*, and therefore require support for hand-off, location management and connection management functionalities. It is the last of these three issues, viz. mobility that forms the focus of this paper.

In an ATM network without support for mobile users, a (non-mobile) user is able to communicate with a host along a single pre-determined virtual path, and QoS can be maintained for the connection by simply reserving bandwidth during connection set-up at each router on the forwarding path.[2] In an environment with mobile users, however, there exists a need to dynamically reroute ongoing connections to/from mobile users as these users move among multiple base stations. Connection rerouting schemes must exhibit low handoff latency, maintain efficient routes, and limit disruption to continuous media traffic while minimizing updates to the network switches. The situation is complicated by the fact that, in order to provide broadband services, the frequency re-use factor must be increased, leading to a very small size of cells in the wireless access segment of the network (*picocells*) and a consequently higher frequency of handoffs as mobiles move relatively quickly between the small-sized cells.[3] Without some intervening mechanism, a mobile operating in such an environment would have to perform a connection set-up each time a handoff occurs. This would not only seriously degrade the QoS guaranteed to the connection, but would also impose a great computational bottleneck on the network's call-admission controller (CAC) performing the call processing and control functions in a centralized fashion.

Several alternatives for wireless ATM architectural design have been proposed in the literature to alleviate the mobility problem.[4–8] Among these is the virtual connection tree (VCT)[4], a construct that seeks to unburden the central network call processor by organizing a hierarchical grouping of backbone and wireless network resources, and making the handoff 'mobile-executed', thereby reducing the computational load on the CAC. The VCT and its variants[6] have been widely discussed; it has been pointed out that in several situations, the mobile-initiated handoff does not provide an optimal performance. Specifically, while providing a good starting point, the VCT algorithm in its native form suffers from the lack of inclusion of accurate knowledge of the mobile's trajectory. As such, there is a danger of underutilizing resources in base stations to which the mobile never connects, and a potential of overloading base stations when a large number of MTs attempt to connect at the same time. Furthermore, issues such as ordering in packet sequences and handovers in congested picocells cannot effectively be addressed by the mobiles themselves.[7, 8]

In this paper, we propose to augment the virtual tree topology with mechanisms that guarantee more reliable QoS during handoffs and provide local hot-spot relief besides reducing the computational burden on the mobiles, in effect simplifying their design. In particular, we envisage the use of three schemes where the *base stations* can be expected to play a significant role. The first is mobile tracking through sectored antennae. If the mobility patterns of a mobile can be made available to the base stations in the multicast cluster, the base stations can intelligently reserve resources for mobiles that may be anticipated to enter the pico-cell served by them (*predictive resource reservation*). The second idea is to borrow channels from adjacent cells if a particular base station is overloaded. Since this process may not be implementable in real time, base stations must use the advance warning provided by the motion information of the mobile to predict a hot spot and initialize channel borrowing measures, forcing a handoff of the overloading mobile (*base station-assisted handoffs*). The third idea is to use caching at the base stations and local retransmissions on the downlink to counter the effects of high BER experienced on the wireless link. We elaborate these ideas in the sequel.

The remainder of this paper is structured as follows. In Section 2, we describe the components and salient features of the proposed wireless ATM network architecture, elaborating the construct of the *intelligent multicast tree* (IMT) which forms the essence of this proposal. In Section 3, the network protocols to be implemented in the present scheme are briefly outlined. Section 4 offers a performance evaluation of the architecture using discrete-event simulation. Section 5 summarizes the major notions contained in this paper.

2. Proposed architecture: Salient features

In this section, we describe the components, organization and key features of the proposed IMT-based architecture. It may be pointed out that the 'intelligence' in the IMT really derives itself from the buffering and switching capability embedded in the base stations, which, when coupled with information about mobility patterns of the MHs (mobile hosts), helps to reduce the processing burden on both the CAC and the mobiles, and enables a more distributed sharing of computational load among the various entities in the network (Table I).

2.1. *Components*

The wireless access structure based on IMT is shown in Fig. 1. The radio access region consists of the union of many IMTs, each defined by a cluster of contiguous picocells. An access point/ base station manages and regulates traffic in one picocell. Each base station (BS) is connected to the root switch of that IMT by a virtual path through a network of fiber links and is assumed to be equipped with directional, sectored antennae. Sectoring is important in this scheme, not only to increase the effective channel reuse factor and assist in directional channel

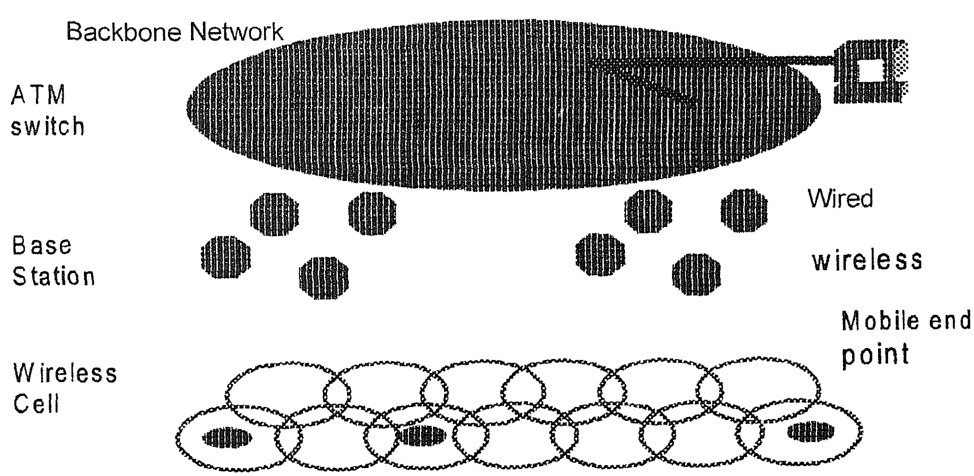

FIG. 1. Schematic of the proposed architecture.

borrowing during hot spots,[9] but also to facilitate the estimation of velocity and position of mobile terminals.

An MH is admitted once into an IMT, based on a statistically pre-computed look up. When it hands off between access points inside the IMT, the network CAC is not invoked; the mobile's movement inside the IMT is thus transparent to the rest of the network. Only when it moves out of the cluster of the IMT is the existing connection replaced with another to the root switch of the second tree. By making the IMTs large and overlapping, frequent handoffs across clusters can be avoided.

2.2. *Estimating mobility*

A key feature of the intelligent multicast topology is the ability of the base stations to process and use the mobility information of the MTs to proactively track users and allocate resources to them. The motivation behind creating a local mobility model is based on the observation that the seemingly random choice of inter-cell movement is actually a logical function of the user's position, speed, direction and cell geometry.[10, 11] Thus, in order to predict the time-varying movement patterns of the MTs, we need to mathematically model the mobility patterns in terms of dynamical and measurement equations. We present such a preliminary mathematical model next.

In existing cellular systems, the distance between the mobile and its closest base stations is practically observable. Such information is inherent on the forward link received signal strength indicator (RSSI) of a reachable base station. Let $z_i \in \mathbb{R}^2$ denote the position of a base station i and $d_i = ||x - z_i||$ the Euclidean distance of $x \in \mathbb{R}^2$ from z_i. If the average signal strength follows a propagation law of the type $c.\ d^{-\eta}$, d denoting the distance from the transmitter and $\eta > 2$ the attenuation exponent, then the average signal strength $s_i(x)$ of base station i measured at position x is obtained as

$$s_i(x) = c\,[d_i(x)]^{-\eta} \tag{1}$$

Let $\gamma_i(t)$ denote the measured signal strength of BS i, $i = 1, 2, \ldots, n$, at a certain position $\boldsymbol{x}$. Obviously, $\gamma_i(t)$ is subject to random fluctuations due to short-term Rayleigh and Rician fading. The transformed values $\gamma_i(t)^{\eta}/c = \delta_i(t)$ correspond to the distance from BS i, such that the solution ξ of

$$\text{minimize}\, f(x) = \sum_{i=1}^{n} \left[d_i(x) - \delta_i(t)\right]^2 \text{over}\; x \in \mathbb{R}^2 \tag{2}$$

is a least-squares estimator of the actual position of the MH at time t. A local minimum of (2) can be calculated by a Newton-type iteration

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \boldsymbol{H}_f(\boldsymbol{x}_k)\, \nabla f(\boldsymbol{x}_k)\; k \in \angle_0. \tag{3}$$

The gradient $\nabla f(x_k)$ at a differentiable point $\boldsymbol{x}$ is given by

$$\nabla f(x) = 2\sum_{i=1}^{n} \frac{d_i(x) - \delta_i(t)}{d_i(x)}(x - z_i) \tag{4}$$

and with $\boldsymbol{x} = (x, y)$ and $z_\iota = (\alpha_\iota, \beta_\iota)$, the matrix $\boldsymbol{H}$ is determined as:

$$\boldsymbol{H}_f(\boldsymbol{x}) = \begin{bmatrix} f_{11}(\boldsymbol{x}) & f_{12}(\boldsymbol{x}) \\ f_{21}(\boldsymbol{x}) & f_{22}(\boldsymbol{x}) \end{bmatrix} \tag{5}$$

where

$$f_{11}(\boldsymbol{x}) = 2\sum_{\iota=1}^{n} \frac{\delta_\iota(t)(x-\alpha_\iota)^2}{d_\iota^3(\boldsymbol{x})} + \frac{d_\iota(\boldsymbol{x}) - \delta_\iota(t)}{d_\iota(\boldsymbol{x})}, \tag{6}$$

$$f_{22}(\boldsymbol{x}) = 2\sum_{\iota=1}^{n} \frac{\delta_\iota(t)(y-\beta_\iota)^2}{d_\iota^3(\boldsymbol{x})} + \frac{d_\iota(\boldsymbol{x}) - \delta_\iota(t)}{d_\iota(\boldsymbol{x})}, \tag{7}$$

$$f_{12}(\boldsymbol{x}) = f_{21}(\boldsymbol{x}) = 2\sum_{\iota=1}^{n} \frac{\delta_\iota(t)}{d_\iota^3(\boldsymbol{x})}(x-\alpha_\iota)(y-\beta_\iota). \tag{8}$$

Having obtained $\boldsymbol{x} = (x, y)$ using eqns (3) through (8) above, we can then proceed to estimate the velocity of the MH in the following manner. It is known that in two-dimensional Cartesian coordinates, the movement can be described by a first-order vector differential equation with the dynamic state vector $\boldsymbol{u}(t) = [x(t), x'(t), y(t), y'(t)]^{\mathrm{T}}$, where x and y represent the position at time t, and their first-order derivatives of $x'(t)$ and $y'(t)$ represent the relative speed along the x and y directions, so that $v = (x', y')$. Furthermore, let $\boldsymbol{r}(t) = [r_x(t), r_y(t)]^{\mathrm{T}}$ denote the two-dimensional random acceleration vector. Then we have

$$\boldsymbol{u}'(t) = F\,\boldsymbol{u}(t) + G\,\boldsymbol{r}(t), \tag{9}$$

where

$$F = \begin{bmatrix} \Re & 0 \\ 0 & \Re \end{bmatrix}, G = \begin{bmatrix} \varpi & 0 \\ 0 & \varpi \end{bmatrix}, \quad \Re = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \varpi = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{10}$$

Random acceleration $\boldsymbol{r}(t)$ is correlated in time, i.e. if a moving object is accelerating at time t, it is likely to continue accelerating at time $t + \tau$ for sufficiently small τ. A typical representative model of the correlation function is

$$R_r(\tau) = E\,[\boldsymbol{r}(t)\,\boldsymbol{r}(t+\tau)] = \sigma_{\mathrm{m}}^2 \mathrm{e}^{-\alpha|\tau|} I \tag{11}$$

where I is a 2×2 identity matrix, σ_{m}^2 is the variance of the random acceleration of a single dimension, and α is the reciprocal of the random acceleration time constant. Such a random process can be obtained by passing the white gaussian signals $\boldsymbol{w}(t) = [w_x(t), w_y(t)]^{\mathrm{T}}$ though a one-pole shaping filter, where

$$r'(t) = -\alpha\, r(t) + w(t), \text{ with } R_w^2 = 2\,\alpha\,\sigma_m^2\,\delta(t)I. \tag{12}$$

Equations (9) and (12) can be combined and discretized using the state-space method[11] in order to obtain the velocity estimate v of the MH as well. It is not our purpose here to discuss the algorithms that may be used to solve the above equations (9 through 12). Rather, we note that once the mobility information of MHs is available in the form of velocity and position estimates, these can be multicast to all BSs in the IMT. Next, we describe ways in which this mobility information may be used to greatly enhance system efficiency and improve the quality of service received by the mobiles admitted to the IMT.

2.3. *Mobility-based predictive resource reservation: blind vs intelligent multicast*

The goal of any multicast-based mobility approach is to minimize the impact of disruption in loss and delay sensitive continuous media traffic by grouping the neighbors of the cell in which the mobile user is currently resident, and perform advance multicasting of data to the base stations in this group. The mobile can then freely roam in the area covered by the multicast, without invoking the network call acceptance capabilities during each handover. Although the approach is fast and statistically guarantees QoS, the overhead of buffering data for each mobile host in multiple base stations in the cluster could form a potential implementation bottleneck in the future. The use of mobility information based on motion estimates of the mobiles can alleviate this problem. By processing the mobility information of the mobiles available to them along with the multicast data, the base stations can effectively allocate their resources, reducing the net buffering overhead per admitted mobile. Note that since this processing is done by the base stations themselves, the CAC as well as the mobiles are not burdened computationally.

In order to quantify the benefit of including mobility-based information on buffer allocation in the present architecture, we follow the notation introduced in Section 2.2 and obtain the amount of resources that must be reserved by a base station i at any given instant as

$$\beta_i = \sum_{j=1}^{M} \sum_{k=1}^{N_j} r_{jk} t_{jk} p_{ji}(x, v, z) \tag{13}$$

where N_j is the number of virtual connections set up by the jth mobile, r_{kj} is the peak cell rate of kth vc set up by the jth MH, t_{kj} is the time-out interval of kth vc set up by the jth mobile, M is the total number of MHs admitted into the multicast cluster, and p_{ji} (x, v, z) is a function which represents the likelihood of the jth mobile moving into the test base station i within any specified interval of time.

In general, p_{ji} would depend on the Euclidean distance $d_i(x)$ between the mobile's current position x and the BS i; the rate of approach v of the mobile along $|x - z_i|$; and, perhaps most importantly, the actual topology under consideration†, so that:

†Most built structures have fixed paths along which movement may take place, e.g. corridors, stairs, elevators, etc. The knowledge of this topology should be aadvantageously exploited in predicting the trajectories of the mobiles and the time they would take to move from one BS to another in the cellular array.

$$p_{ji}(x,v,z)=\begin{cases}1 \text{ if } z=z_i \\ \omega_1 d_i(x)+\omega_2 \frac{|x-z_i|}{\|x-z_i\|}\cdot v+\omega_3\varphi(z,z_i) \text{ otherwise.}\end{cases} \tag{14}$$

Here $\varphi(z, z_i)$ is a function determined by the specific architectural constraints of the considered topology and $\omega_1,\ldots,\omega_3$ are appropriate weights that could be set either heuristically or analytically. Of course, if $p_{ji}(x, v, z)$ exceeds t_{jk} (plus some tolerance), it may be set to zero—the BS need not reserve any resources for MHs which are not expected to move into its domain within the next predetermined interval of time. Furthermore, once it is obtained, the same likelihood function p can also be used as a trigger for initiating the borrowing of channels from neighboring cells, in order to avert an impending hot spot situation (Section 2.4).

A simple numerical example of the above analysis is obtained by considering a network consisting of 10 picocells, serving a total of 30 admitted videoconference calls at 512 kbps, each with a time out interval of 500 milliseconds.[1] Assuming the mobiles to be uniformly distributed in the cellular array at all times, the buffer requirement per BS as per the present scheme would be a little below 2 MB. According to the naïve (blind) multicast approach, however, each BS would have to reserve resources for each of the admitted mobiles, leading to a total buffer space requirement of 7.5 MB per cell! Clearly, the advantage of including mobility-based information on buffer allocation in a picocellular network is substantial, and this gain only increases as the size of the multicast cluster is increased.

2.4. *Hot spot alleviation: Using the velocity and position estimates*

As far as the wireless bandwidth is concerned, overload conditions might occur if the communication requirement of a number of wireless terminals populating a picocell (or cluster of adjoining picocells) exceeds the total capacity of all access points within their reach. In the virtual tree[4], such a 'hot spot' radio congestion state would simply cause the incoming mobiles to be dropped, thereby increasing the call-dropping probability. The problem can be alleviated by the use of channel borrowing, which, at least in a semi-dynamic way, is known to be a viable method to alleviate local radio congestion in cellular systems.[12] The protocol that an overloaded access point must follow to borrow channels is necessarily lengthy[+] and this could cause a break in the service received by the mobile. Using motion prediction, however, a picocell can predict a hot spot and initialize hot-spot relief measures so as to be able to deploy them quickly when a hot spot actually occurs. In specific, the base station can pre-empt the borrowing of spare channels in neighboring cells to relieve local congestion. Sectoring in this set-up enables the borrowed channels to be used in such a way that interference between co-channel cells is minimized.[9, 12]

2.5. *Caching for error control*

The final application of base station 'intelligence' in the IMT architecture is in caching ATM packets and performing local retransmissions of lost or corrupted packets, alleviating the prob-

[+]This is because the borrowing of a 'free' channel from a neighbouring cell leads to the locking of this channel in all interfering co-channel neighbors (see Fig. 2). Schemes that do not use such locking also exist in the literature, but are not considered here.[13]

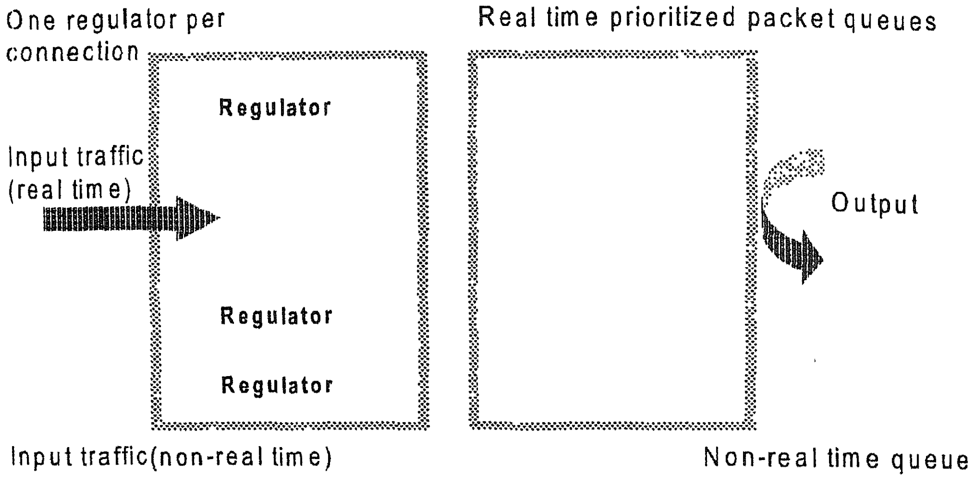

FIG. 2. A borrowing strategy. Channel a4 is borrowed and now locked to cells marked N. Cells marked X were already prohibited from using a4.

lems caused by high BERs on the wireless link. This local retransmission not only reduces the effective latency, but also circumvents the need for repeated transmissions back from the source terminal, preserving the end-to-end semantics of the wired ATM protocol. By detecting missing packets and generating corresponding negative acknowledgements, the base stations can shield the sender from transient situations of very low communication quality and temporary disconnectivities on the wireless link. In doing so, a substantial increase in effective throughput can be expected.

3. Network protocols

In this section, we briefly delineate the major protocols that manage mobility in the given architecture (Figs 2–4).

3.1. *Connection set-up protocol*

When an MH powers up or roams into an IMT, it is required to register itself with the network. The registration procedure allows the network to authenticate the device and allocate resources to it. To register, an MH meta signals its global MAC identifier to the nearest BS; the CAC (and in turn, the BS) responds by allocating an M_ID (mobile identifier) for the MH. A mobile terminal's M_ID is used to identify an admitted the terminal uniquely within the multicast cluster, and is part of the extended ATM header used for each cell transmitted over the wireless link.

Once a mobile has been registered, it can request a connection set-up in the IMT cluster by randomly accessing the call admission request segment of the local base station's signaling field with its local identification number (M_ID) and traffic parameters required for the desired QoS. The set-up and connect functions to the backbone network are handled by the root switch, and successful establishment of an ATM VCI/VPI and decision to admit are transmit-

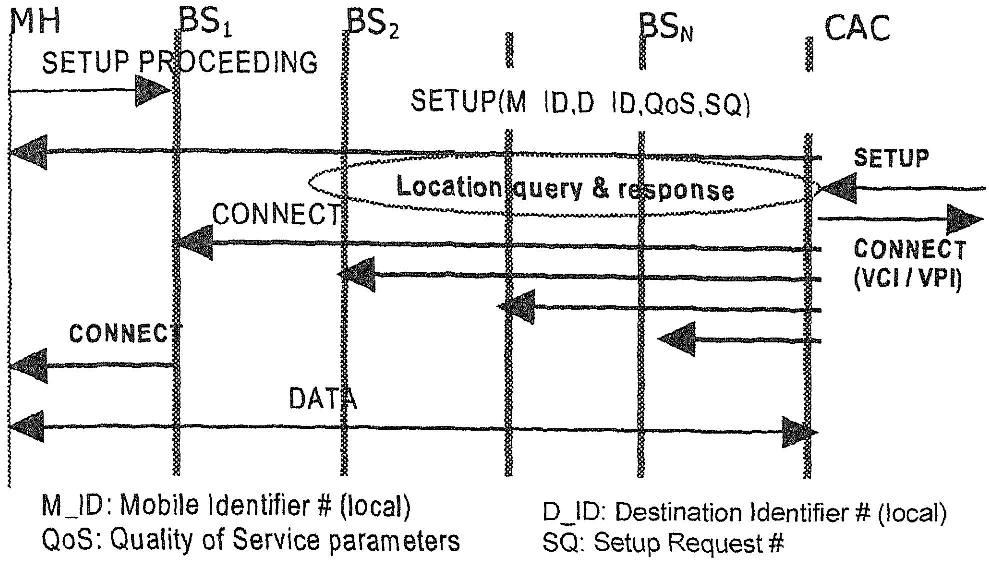

FIG. 3. Connection set-up procedure.

ted to the mobile in the connection acknowledgement field of a later frame. Once the connection has been admitted, data exchange can proceed following the mobile's ID (M_ID) being successively multicast to all base stations in the IMT. The above sequence of steps is enumerated in Fig. 3.

3.2. *Intra-cluster handoff protocol*

When a mobile terminal moves from one picocell to another within the same IMT, all the connections originating or terminating at the terminal must be re-routed from the old base station (BS-1 here) to a new base station (BS-2). An ATM mobile host may have several virtual connections open at the same time, belonging to either different calls (i.e. with different remote users) or to the same call but supporting different applications. The handover of each VC must be carried out separately (possibly at the same time, but with different protocol instances). Highly correlated connections can also require a coordinated handover procedure. For the sake of simplicity here we describe only the handover procedure for a single connection.

The message flow for the handoff of an MH between two picocells that are part of the same IMT is indicated in Fig. 4. The handoff is initiated when, based on its beacon measurements and some hysteresis, the MH transmits a handoff request to the new base station BS-2. While the request is being delivered and processed, packets continue to be transmitted to/from the MH through BS-1. The control message to activate forwarding on BS-2 includes a sequence of identifiers that identify the sequence of packets last received by the MH from BS-1. Once the handoff is accomplished, the old base station automatically tears down the routing entry for MH and de-allocates its resources accordingly. Since only one path through the network is active at any one time, and by ensuring that all the remaining cells at the old BS are flushed to

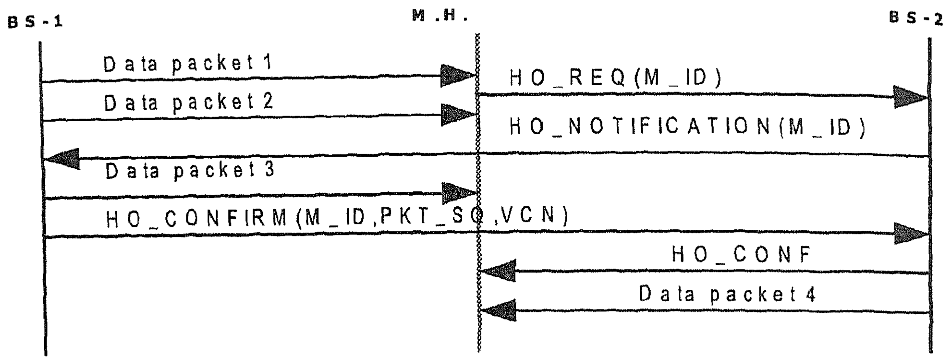

FIG 4 Handoff protocol message flow.

the MH before the radio handover occurs, the indicated protocol effectively maintains per-VC cell sequence integrity during the handover process.*

Another desirable trait of a handover protocol is that it should minimize the length of time that an MH is out of touch with the wireline network (*minimal service interruption*). As such, there are two separate but overlapping procedures involved in the handover of an MH: the *network* handover, where the virtual circuits of the MH are transferred from the old BS to the new one; and the *radio* handover, where the MH relinquishes its radio channel to the old BS, and establishes a new one with the new BS. It has been shown[14] that while the network handover takes in the order of tens of milliseconds (using current implementations of the ATM-F-UNI signaling), the radio handoff can be completed in the range of a few hundred microseconds. Therefore, minimal service interruption requires that the *network* handover is performed as quickly as possible. Our scheme accomplishes this task explicitly: for, being a multicast-based approach, no time is spent in setting up a new connection from the root switch to the BS, every time a handoff takes place. Furthermore, by the use of in-band mobility-enhanced signaling based on the AAL-5,[15] it is possible to achieve a deterministic performance of the above described handoff protocol.[16] Finally, we note that none of the above descriptions is applicable for the case of handoffs forced by base station during hot-spot conditions.[17]

3.3. *Medium access control*

To provide class-based quality of service to mobile applications in the cluster, a novel medium access scheme, MASCOT (*medium access support for multiple classes of traffic*)[18] is proposed in conjunction with a weighted fluid fair queuing-based service discipline at the base station (Fig. 5).

4. Simulation results

Discrete event simulation is used to characterize the performance of the proposed architecture under various operating conditions. For our simulation we consider a multicast cluster consist-

*One of the fundamental principles of ATM is that the relay of cells within a single VC must maintain cell sequence integrity.[2] Thus it is essential (in the wireless case) that cells are not reordered or duplicated during the handover.

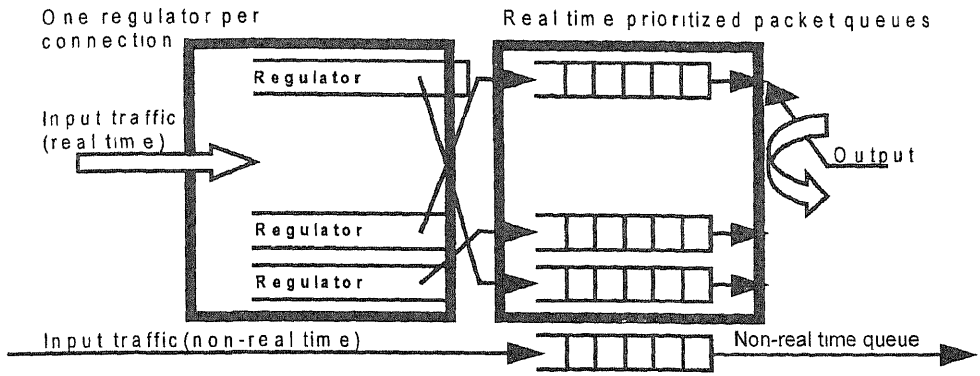

FIG. 5 Base-station scheduling for various traffic classes.

ing of 49 cells each having a capacity equivalent to 7 voice calls. New call arrivals are assumed to be Poisson distributed with mean arrival rate λ. Call-holding time is exponentially distributed with mean service time of $1/\mu_d$, and the cell sojourn time, i.e. time during which the mobile resides in a cell is exponentially distributed[†] with a mean sojourn time of $1/\mu_s$. Under these assumptions, each BS in the IMT behaves like a M/M/m/m queue with m channels[20] (Fig. 6). Baseline values for various system parameters are indicated in Table II.

Table I
Intelligence distribution among MH, BS and root switch in handover procedures

Task	MH	BS	Root switch
Radio channel measurements	Make periodic measurements on current and neighboring broadcast channels, send results to BS using common channel signaling	Monitor reverse link, give measurement command to MH	—
Issue handover request	—	Send measurement results to RS, request borrowing	Evaluate handover request. Inform new BS, relevant neighbors
Confirm/disconfirm handover	—	Accept/block/delay handover request	Permit/drop/ delay handover

Table II
Baseline values of simulation parameters

Parameter	Performed variation
No. of cells in the IMT	49
Frequency of reuse factor	K=7
No. of channels per cell	7–10
Mean call-holding period	3 min
Mean cell sojourn time	20 s
Aggregate arrival rate (1/s) per cell	0.01–0.07

†This assumption is particularly appropriate in a microcell/picocell environment, see, for example, 19

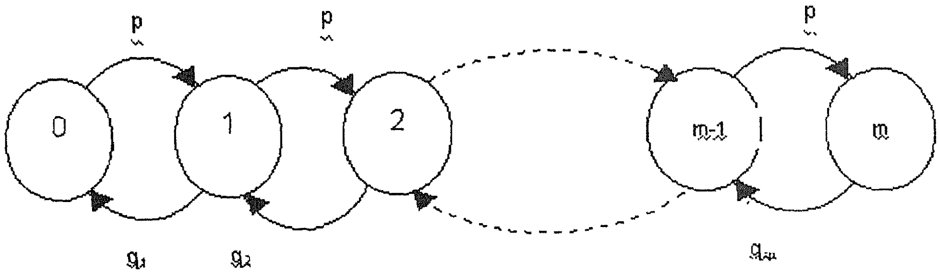

FIG. 6 Discrete time Markov chain for M/M/m/m system.

User mobility is modeled in this simulation according to the *random walk* approach, whereby the mobile resides in a cell for an interval equal to its sojourn time and then moves on to one of the adjacent neighbors (adjacent cells being defined as those having common edges) with equal probability, i.e. 1/6 for the hexagonal cell configuration assumed here. The mobility of the MHs is therefore a random walk in a 2D hexagonal plane.[21] In case it handoffs into a picocell that is already serving its maximal capacity, the mobile must be dropped. Figure 7 plots this dropping probability against the net offered load (in Erlangs) on the network. From the observations it is evident that, as the offered load per cell is increased, the call dropping rates increase for both the uncontrolled and the predictive system; however, the increase in call-dropping rate for the proposed scheme is substantially smaller. Figure 8 repeats this observation taking the capacity of each cell as equivalent to 10 voice calls. The call-dropping rate in this case is lower than that in Fig. 7, and once more, the proposed scheme is seen to outperform the original one. Of course, this improved performance is at the cost of a somewhat increased implementational complexity of the proposed system, and the greater signaling overheads that the proposed predictive reservation protocol would involve. Also note that in this analysis, we do not distinguish between newly generated calls and calls arising out of handoffs between

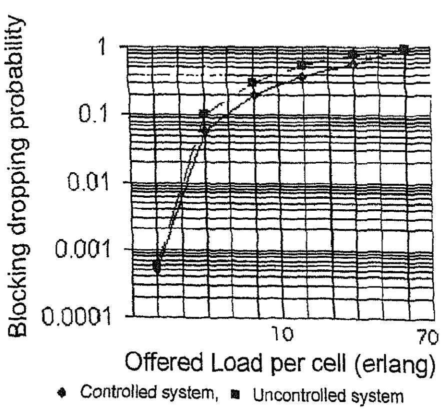

FIG. 7. Call-dropping probability vs. Erlang load per cell (Case-I).

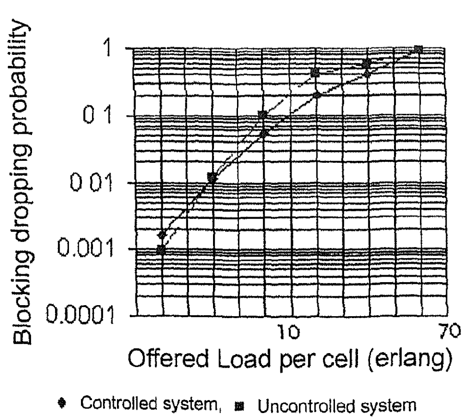

FIG. 8. Call-dropping probability vs. Erlang load per cell (Case-II).

adjacent cells. In general, the latter must be given preference over the former so that the QoS of already admitted mobile hosts is not disadvantageously compromised.

5. Conclusions

This paper proposes a system architecture to provide mobility support in a futuristic wireless ATM-based personal communication network (PCN). The architecture augments the static virtual connection tree concept presented in Acampora and Nagshineh[4] by introducing the notion of *intelligent multicast*, whereby the mobility management functionality is distributed among the various entities in the network *(mobile hosts, base stations* and *wireline switches*). In doing so, system resources (*buffers, wireless bandwidth*) can be utilized more efficiently, hot-spot conditions can be alleviated and, in effect, the call-blocking rates can be reduced.

The intelligent multicast concept introduces some processing and signaling overheads. However, these overheads involve only the base stations and the underlying high-capacity wireline network, and not the scarce wireless resources. In this paper, we sketch the network protocols to be used for the proposed scheme, and characterize the performance of the architecture using discrete event simulation. A more rigorous and comparative evaluation of performance forms the subject of future work.

References

1. RAYCHAUDHURI, D *et al.* WATMnet. A prototype wireless ATM system for multimedia personal communications, *IEEE J.*, 1997, **SAC-15**, 83–95.
2. ONVURAL, R. O. *Asynchronous transfer mode· performance issues*, Artech House, 1994
3. LEE, W. C Y. Smaller cells for greater performance, *IEEE Commun. Mag.*, 1991, **29**, 19–23.
4. ACAMPORA, A. S AND NAGSHINEH, M. An architecture and methodology for mobile-executed handoff in cellular ATM networks, *IEEE J.*, 1994, **SAC-12**, 1365–1374.
5. VEERARAGHAVAN, M., KAROL, M J. AND ENG, K. Y. Mobility and connection management in a wireless ATM LAN, *IEEE J.*, 1997, **SAC-15**, 50–68.
6. YU, O. AND LEUNG, V. Extending B-ISDN to support user terminal mobility over an ATM based PCN, *Proc. GLOBECOM '95*, 1995, pp 2289–2293.
7. WALKE, B., PETRAS, D. AND PLASSMAN, D. Wireless ATM: Air interface and network protocols of the mobile broadband system, *IEEE Personal Commun* , 1996, **3**, 50–56.
8. LEVINE, D. *et al.* A resource estimation and call admission algorithm for wireless multimedia networks using the shadow cluster concept, *IEEE/ACM Trans Networking*, 1997, **5**, 1–12
9. TAK-SHING, YUNG, P. AND WONG, W. S. Hot spot traffic relief in cellular systems, *IEEE J.*, 1993, **SAC-11**, 934–939
10. HELLEBRANDT, M., MATHAR, R. AND SCHIEBENBOGEN, M. Estimating position and velocity of mobiles in a cellular radio network, *IEEE Trans.*, 1997, **VT-46**, 65–71.
11. LUI, T., BAHL, P. AND CHLAMTAC, I. Mobility modeling, location tracking, trajectory prediction in wireless ATM networks, *IEEE J.*, 1998, **SAC-16**, 922–935.

12. COX, D. C AND REUDNIK, D. O. Increasing channel occupancy in large scale mobile radio systems: Dynamic channel assignments, *IEEE Trans.*, 1973, **VT-22**, 218–222

13. JIANG, H. AND RAPPAPORT, S S. Channel borrowing without locking for sectorized cellular communications, *IEEE Trans.*, 1994, **VT-43**.

14. NIEHAUS, D. *et al.* Performance benchmarking of signaling in ATM networks, *IEEE Commun Mag*, 1997, **35**, 134–143.

15. MARSAN, M. A. *et al.* Local and global handovers for mobility management in wireless ATM networks, *IEEE Personal Commun*, 1997, **3**(5), 16–24

16. NAYLON, J. *et al.* Low latency handover in a wireless ATM LAN, *IEEE J.*, 1998, **SAC-16**, 909–921.

17. SINGH, M. AND ARVIND, D *A system architecture for mobile broadband access*, Internal Report, Dept. of Electrical Engineering, IIT, New Delhi, Dec 1998.

18. SINGH, M., PRASAD, S. AND KAR, S MASCOT: A novel medium access control protocol for wireless ATM networks, *Globecomm 2000* (submitted).

19. GUERIN, R. *et al.* Channel occupancy time distribution in a cellular radio system, *IEEE Trans.*, 1988, **VT-36**, 89–99.

20. BERTSEKAS, D AND GALLAGER, R. *Data networks*, 2nd edn, Prentice Hall, 1997

21. AKYILDIZ, I *et al.* Movement-based location update and selective paging for PCS networks, *IEEE/ACM Trans. Networking*, 1996, 4, 629–638.