An ATM based protocol for Wireless LANs

John Porter* and Andy Hopper*†

*Olivetti Research Limited
24a Trumpington Street
Cambridge, CB2 1QA, England

†University of Cambridge Computer Laboratory
New Museums Site, Pembroke Street
Cambridge, CB2 3QG, England

April 19, 1994

Abstract

This paper presents a solution to the problem of connectivity of portables to an ATM wired network in the local area environment. A compatible ATM approach is used to provide support for multi-media traffic. Spatial re-use of a single frequency with a large number of small cells is used to increase the aggregate throughput. An experimental system based on low-cost fixed ATM switches and software controlled base stations has been developed.

1 Introduction

There is a growing demand for mixed-media traffic support in wired networks and this is beginning to extend to the wireless domain. An example of a device requiring wireless multi-media communication is given by [She92]. Current and future local area WLANS either provide no support for multi-service data or use a hybrid scheme based on dividing the available bandwidth into reservation and collision intervals [DEB93] [Kru92]. ATM is considered to reduce the complexity of the network and improve the flexibility while providing end-end consideration of traffic performance.

Most current wireless LAN systems are low data-rate spread-spectrum systems either DS/CDMA or FH/CDMA, good for datagram services and compatible with Ethernet type local area networks. The low bit-rate is determined by the limited bandwidth and spreading factor for these systems. In general these systems are designed for a single access point in locations where it is difficult or expensive to install a wired network. With the advent of intelligent buildings which are provided with backbone networks these considerations are less important. Of more interest is short range, high aggregate bandwidth and compatibility with the wired network.

1.1 ATM

ATM (Asynchronous Transfer Mode) has been advocated as an important technology for the wide area interconnection of heterogeneous networks [bis90]. There is currently interest in using ATM in the local area to provide improved performance for applications end-to-end. The ATM-Forum are currently involved in defining the semantics of ATM in the local area in a manner broadly compatible with the CCITT recommendation.

Future (and some current) private ATM networks are likely to consist of a large number of small switches providing high aggregate bandwidth at low cost. The incremental cost of adding more interfaces and capacity is low in this environment and is thus ideal for the support of a flexible pico-cellular wireless network.

ATM adaption layers (AALs) provide mechanisms for supporting transport protocols over ATM cells. AAL1 and AAL2 have been defined by the CCITT for use in the wide area for support of constant and variable bit rate services respectively and AAL3/4 for connection-less data transport. However, AAL5 is being proposed by the ATM-Forum for *all* types of computer oriented multi-service traffic especially for the local area. AAL5 has lower pay-load overhead per cell and relies on quality of service and statistical mechanisms to provide multi-service capability.

1.2 Compatibility with ATM

The architecture presented is compatible with the emerging standards for private ATM LANs as defined by the ATM Forum which in turn are designed to be compatible with the wider area ATM networks to be provided by the PTTs. In this context, compatibility means that no modification to the wired part of the network is necessary to support the wireless extension and the network provides a standard ATM interface to both mobile and fixed hosts. The currently favoured AAL5 adaption layer will be supported and the Virtual Circuit connection based data model will be the same. In the future, it is expected that many buildings will be wired with ATM networks and the ability to connect a wireless base-station directly to the existing network, without adding new wiring, will be a major advantage.

A cell format standard is required for wired ATM so that equipment from different vendors can communicate over a physical link. The precise details of the header format for the wired ATM network standard are currently being formulated but the following will certainly be supported: cell specific quality of service parameters (cell loss priority and generic flow control), virtual circuit identifier, and some limited pay-load type information. In the wireless data-link a different cell format can be used as both ends of the link are specific to the radio network. Although the contents of the cell header can be different only information that can be forwarded through the wired network can be used outside the wireless link. In particular there is no provision for cell sequence numbering within the standard cell header.

One of the main attractions of ATM is that out-of-band control and the connection oriented approach allow a Quality of Service (QoS) to be associated with each connection. Certain guarantees can be made about the performance of individual circuits because the aggregate load from all circuits can be calculated [FRV92]. Ideally the wireless network should provide the same QoS interface as the wired network, but in practice some details of the semantics are likely to be different. In a mobile environment, more flexibility may well be desirable such as providing a range of acceptable performance values which a circuit can have before the application is notified or the circuit is cancelled.

In the mobile environment network re-configuration happens much more frequently than in the wired network and consequently the problem of mis-ordering of cells is important. Although the higher layer protocols will recover from this mis-ordering the overall performance of the network will be affected unless the hand-over mechanisms designed to minimise the probability of mis-ordering. The ATM standards state that the underlying protocols will not re-order cells which is difficult to guarantee on a link by link basis. In practice mis-ordered cells are detected by the various ATM Adaption Layers and affected PDUs are discarded.

2 RATM Architecture

2.1 Architectural Principles

The mobile network consists of a large number of small transmission cells, termed pico-cells, each served by a base-station. All base-stations operate on the same frequency so there is no hard boundary between pico-cells. The base-stations are interconnected via the wired ATM network. The unit of transmission over the wireless link is also the ATM cell, this provides ease of interconnection with the wired network and simplifies the base-station hardware requirement. Base-stations are simple cell-relays which translate the header formats from the Radio ATM network to the wired ATM network.

Reducing the size of the pico-cells has major advantages in mitigating some of the major problems associated with in-building wireless LANS [ARR87]. The main difficulties encountered are the delay spread due multi-path effects and the lack of a line-of-sight path resulting in high attenuation. The RMS delay spread [SV87] is a good measure of the multi-path dispersion and [BMZ93] shows that there is some correspondence between the receive and transmit antenna separation (pico-cell size here) and the RMS delay spread. There have been a large number of large-scale attenuation models proposed for in-building radio propagation [Has93] and attenuation in most of these cases is considerably greater than line-of-sight.

Small cells have some drawbacks compared to larger cells. There are on average a small number of mobiles within range of any base-station so base-station cost and connectivity is critical. Hand-over rates also increase as the cell size is reduced (see Appendix A). The RATM network operates on a single frequency across all pico-cells, therefore there is no hand-over required at the physical layer. Also, because the pico-cells overlap there is no hard boundary at which a mobile must hand-over and a form of soft hand-off is possible similar to that in the CDMA digital cellular system [Vit92].

In this scheme there is no explicit colouring of pico-cells and an access mechanism is used which is resilient to co-cell interference. In a traditional mobile network, such as the cellular telephone networks, transmission-cells are "coloured" using frequency division multiplexing or code division multiplexing to prevent interference between cells [Ber87]. Within a building it is notoriously difficult to predict the propagation characteristics of micro-waves and this colouring of cells is a difficult problem [FFG+93]. Colouring is also wasteful of bandwidth since in order for it be successful there must be areas between re-use of the colour in which it is idle. These inactive areas could potentially be used for transmission.

Extension of the network by adding a new pico-cell is done by placement of a new base-station connected to the wired network. There is no re-configuration necessary except to register the base-station with the management entity allocating base station identifiers.

Although colouring is not required at the pico-cell level at some stage it is necessary for non-cooperating domains to co-exist without interference. By choosing a simple multi-access protocol and assuming that adjacent domains run the same protocol the interference between domains is the same as that between adjacent pico-cells. However, it is desirable to restrict the scope of management of a domain and a form of colouring is used to partition the address space in order to achieve this.

2.2 System Components

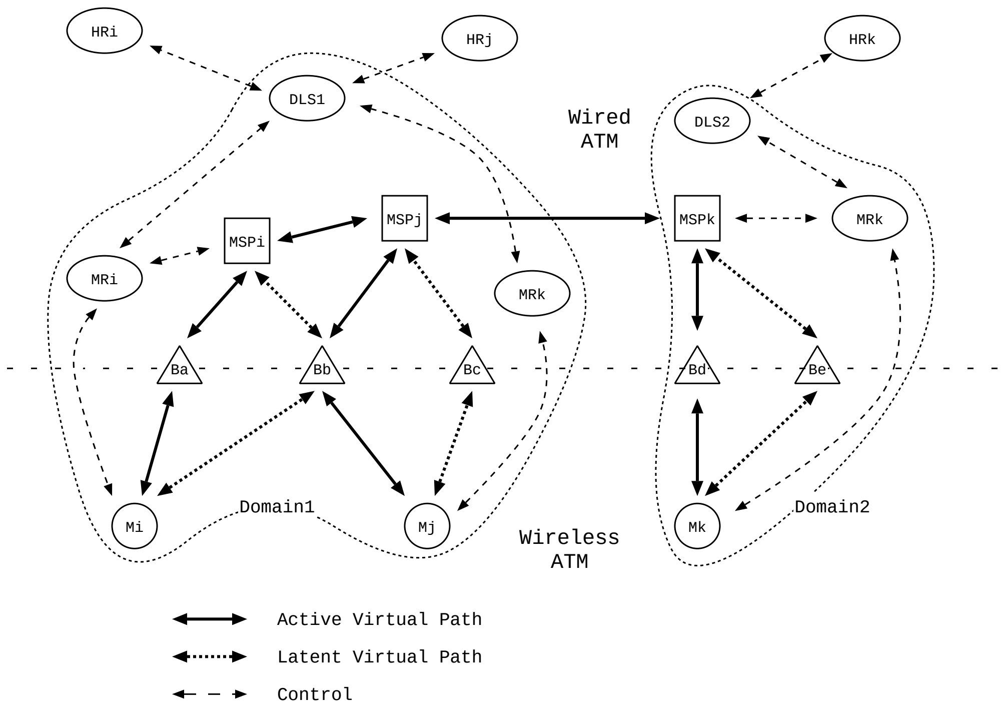


Figure 1: System Components

Mobile (M_i, M_j, M_k in Figure 1) A mobile terminal which is interacting with the network. In this architecture no explicit provision is made for mobile to mobile communication.

Base station (B_a,B_b,B_c,B_d) The access point for a set of mobiles. Acts as a gateway or bridge between the wired and wireless part.

Mobile Representative (MR_i, MR_j, MR_k) A Software entity handling control and management functions for a particular mobile.

Mobile Switching Point (MSP_i, MSP_j, MSP_k) A Switching point within the network which is used to decouple the small-scale movement of a mobile from connection establishment within the wider area wired network.

Domain Location Service (DLS_1, DLS_2) A management entity which maintains tables of the various dynamic entities within a particular domain.

Home Register (HR_i, HR_j, HR_k) A database which is used to store the current location of a mobile.

The domain location service and home register will be co-located and the mobile representative may be located at a base-station or mobile switching point.

2.3 Network Layer Issues

2.3.1 Mobility Management

Mobiles must have enough intelligence to perform the normal ATM connection control and management functionality [1] but other functionality is not necessarily available. The wired network also has a much higher bandwidth than the wireless network. For these reasons control and management of mobility is initiated from the wired network. Base station functionality and mobility management are also decoupled to reduce the complexity required of the base-station.

Each mobile is allocated a Mobile Representative(MR) when it is registered within a domain. The MR is a software entity which is initiated when the allocation occurs and runs logically close to the location of the mobile in the wired network. The use of a MR per mobile reduces the protocol overhead that would be associated with a single management entity for all mobiles in the domain. It provides the signalling service and maintains the small scale location information for the mobile. The Mobile Representative maintains a signalling channel to the mobile over which all control is performed and communicates with both the Domain Location Service and the base stations within range of that mobile.

Virtual Connection establishment, while being reasonably fast in the local area, is likely to take longer in the wider area. Consequently, it is not practical to re-establish all virtual circuits whenever a mobile moves between pico-cells. To isolate the small scale mobility of the mobile from the rest of the wired network latent virtual paths and the Mobile Switching Point (MSP) are used. The MSP provides a routeing point through which all virtual circuits to the mobile are routed. From the mobile switching point there may be a number of potential base-stations which can be used to contact the mobile. Each of these routes for virtual circuits is termed a virtual path which can be manipulated by a single signalling command at the MSP. Whenever a virtual path is active all of the associated virtual circuits will also be active, other virtual paths are latent – the virtual circuits have all been established but no traffic is flowing.

Routeing virtual circuits between the end-point and the MSP is a traditional wired ATM routeing problem. The Mobile Switching Point is simply an ATM switch in which it is possible to associate a set of virtual circuits to form a virtual path.

2.3.2 Inter-pico-cell Hand-over

All the virtual circuits for a mobile can be controlled together reducing the protocol overhead for the management functions. In order to hand-over between pico-cells, the incoming virtual circuits are mapped between the previously active virtual circuits to the latent virtual circuits simply by changing the virtual path.

Although a base-station may not be relaying cells for a mobile, it may be receiving cells from the mobile and can monitor the quality of the link. This information is forwarded periodically to the Mobile Representative which also has knowledge of the physical relationship between base-stations and uses a heuristic to make decisions about when hand-over is wise. Hand-over becomes necessary when the quality of the wireless link falls below a usable level. However, the normal case is that a link to a new base station will be better than the current one before the throughput falls,

[1] Some systems propose a proxy protocol mechanism in order to implement simple hosts, but this is probably unreasonable over an unreliable wireless link

so hand-over will take place before there is any data loss. The physical topology of the network, the actual bit-error rates at the receiver and the observed received signal strength are used to make the hand-over decision. This is more robust than relying on signal strength only as in [VH93] and [Ber93].

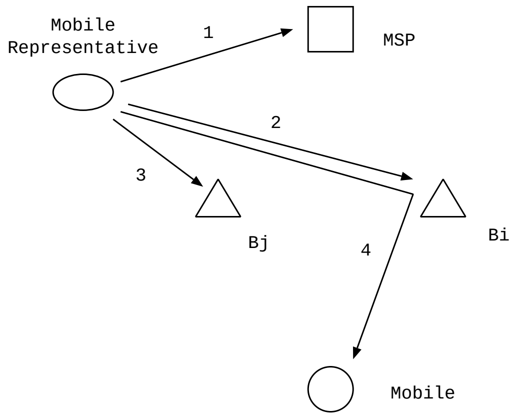


Figure 2: Hand-over Control Messages

The control messages from the MR for hand-over between base stations Bi and Bj using signalling over pre-established circuits are shown in figure 2 . The sequence of messages is:

1. Switch Virtual Paths at Mobile Switching Point.
2. Disable transmission from Bi to mobile, discard any queued cells.
3. Enable transmission between Bj and mobile.
4. Inform Mobile of hand-over.
5. Cancel receptions at Bi from mobile.

The complexity of the hand-over mechanism is considerably reduced, compared to say [KMS+93] due to the decoupling of the various control, management and data-transfer functions within the network. In [KMS+93] a single hand-over (using incremental re-establishment) requires at-least 10 separate stages, some of which are in themselves complex interactions.

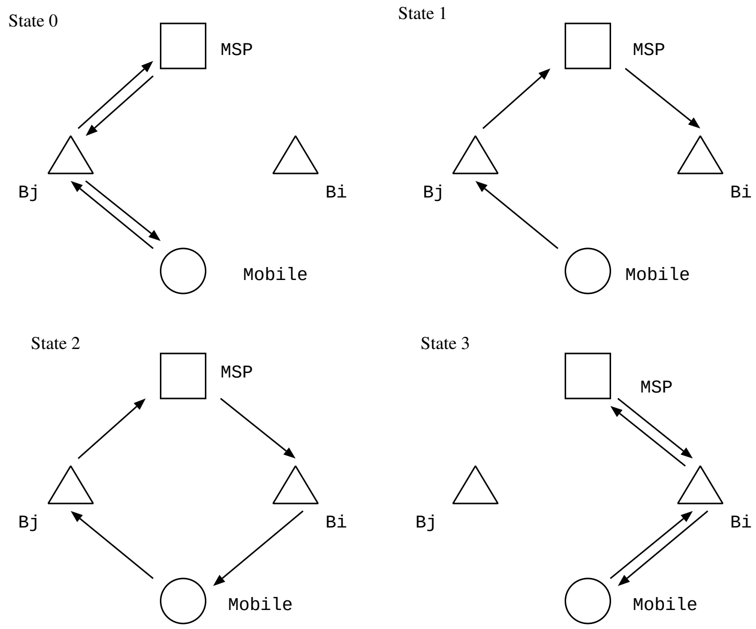


Figure 3: Active data paths during hand-over

The hand-over sequence is designed to reduce the probability of mis-ordering of cells. In the fixed/mobile direction mis-ordering is a potentially serious problem. If one base-station has a number of cells queued for transmission to a particular mobile but is then switched out of the active communication path these cells would continue to be transmitted. The solution is to contact the currently active base-station first which discards any cells which are buffered or subsequently arrive for the particular mobile.

The bottleneck, due to bandwidth mismatch, in the combined wireless/wired network is the access to the wireless network. In the mobile/fixed direction the hand-over switching point is after the buffering. Thus there is little chance of mis-ordering due to hand-over, in this direction, unless the fixed network is suffering severe congestion.

The mobile will receive cells from any base station provided the MId and DId are set correctly in the RATM cell header (see 2.4.1). The mobile is not involved in the selection of a new base-station during the switch of the down-link. The mobile interface supports selection of a new up-link without loss of data, under control of the MR. The switch is made by selecting a new BId under software control and can be altered without flushing the contents of the transmit buffers see 3.3.2.

2.3.3 Network Location and Registration

The mechanism outlined above for virtual circuit hand-over between pico-cells is transparent above the network layer. The end-points in the communication see a normal ATM link (with the possible exception of the QoS re-negotiation mechanism). However, in order to establish connections between the end-points the mobile must be located. Two extremes are possible for the mobile location problem — searching and registration.

- Searching. Involves a form of broadcast in which the whole network is queried. This is practical for small systems, but inappropriate for a general system.
- Registration. Objects are responsible for their own registration at a well known registration point. Subsequent enquiries about the object are directed to this register using a static routeing mechanism. This is a more scalable approach, but requires a database system for registering and recording information.

The RATM architecture uses a hierarchical registration scheme. When a mobile is within a domain, it is registered at the appropriate Domain Location Server (DLS) and this registers the mobile at its Home Register (HR) which merely keeps a record of the mobiles current DLS location. Each mobile has a statically bound home address which is mapped to the HR address. This address is equivalent to the Internet notion of the IP address of a fixed host and the mapping between the mobile's name and address, when the mobile is referenced, uses a traditional name server.

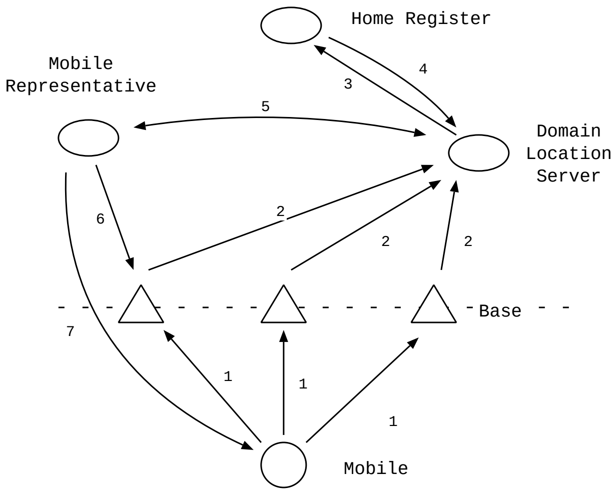


Figure 4: Mobile Registration

Figure 4 shows the control flow when a previously inactive mobile becomes active. Initially, the mobile, which is not registered at any domain, transmits a broadcast message identifying itself (1 in Figure 4). Any base-station receiving this message informs the DLS (2). The DLS may receive a number of these relayed copies of the broadcast message and will later forward information about which base-stations were involved to the MR. The DLS has no record of the mobile so it contacts the mobile's HR and registers the mobile (3,4). The DLS then allocates a domain specific identifier for the mobile (see 2.4.1) and initiates an MR and chooses an MSP for the mobile (5). The MR sets up a signalling channel to the new mobile via the chosen base-station (6,7). This signalling channel is then used for subsequent communication between the mobile and MR.[2]

When a mobile, which has been registered and already has a number of virtual circuits active, comes within range of a new base-station the base-station will begin to receive transmissions from this previously unknown mobile. The derived information about the new link quality is forwarded

[2] Hand-over of the signalling channel is treated in the same way as a data connection to the mobile

(quickly) to the MR which decides whether the new base-station is a likely candidate for future communication. The latent virtual circuits are established between the base and the MSP and hand-over can then take place if it becomes necessary.

There are some occasions when it is desirable to change the MSP assigned to a mobile. It would be possible to do this by tearing down all the old connections, establishing the new MSP and re-establishing all the old connections. However, the cost involved here means that moving an MSP would have to be a very rare event. Instead the new MSP is treated in a similar manner to a new base-station and a new virtual path is established between the old MSP and the new one. The virtual paths between the base-stations and the new MSP are established and then the inter-MSP path is made active. The latent virtual paths involving the old MSP are then re-claimed. The effect is that all new virtual circuits are established directly through the new MSP while the previously active circuits continue to operate but are relayed between the old and new MSP.

2.3.4 Connection Establishment

A signalling connection request is generated by a host which specifies the network addresses of the two end-points. If the destination address is that of a mobile then the local DLS is consulted. If the local DLS has no knowledge of the mobile then the routeing request is forwarded to the HR of the mobile. The HR returns the address of the remote DLS which in turn returns the address of the mobile's MR. Once the address of the MR is known, the connection request and all subsequent requests are directed to the MR.

When the connection request arrives at the MR it first consults the Mobile, any host must decide whether to accept a particular connection. If the mobile accepts the connection it allocates a virtual circuit number (see 2.4.1), and returns it to the MR.

The MR then creates the virtual circuits between the MSP and the remote end-point and adds new virtual circuit to the active and inactive virtual paths between the MSP and the base stations close to the mobile.

2.4 MAC Layer

2.4.1 Cell Format

The radio ATM network has a compatible pay-load size and addressing scheme but there are a number of reasons why the Radio ATM header must be different from ATM. The base-station can translate the headers of cells as they are relayed between the wireless and wired part of the network.

Mobility should be as transparent as possible to the end-points and therefore the VCIs used by the end-points should not change during hand-over. This requires that allocation of the VCI space should remain valid as the mobile moves through different pico-cells within a domain. It is also desirable that translation of the VCI space due to movement between domains should be as simple as possible. A practical means of doing this is to split the VCI space hierarchically into a number of fields: DId, MId and VCn.

DId Domain Identifier - Identifies the domain in which the subsequent fields have been allocated. As mentioned earlier this field is the "colour" which is used to distinguish adjacent management domains.

MId Mobile Identifier - Uniquely identifies the mobile within the domain. It is allocated dynamically within the domain when a mobile is first registered. It is used by a base station to select a particular mobile from those within range.

BId Base station identifier - allocated statically when a new base station is added within the domain and used by mobiles to select a particular base-station for reception.

VCn Virtual Circuit number - combined with the Mobile Identifier uniquely identifies a particular virtual circuit within the domain.

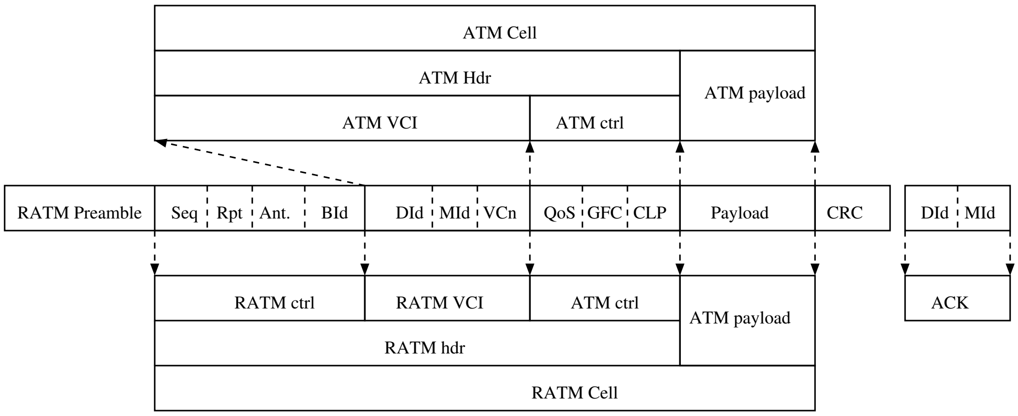


Figure 5: RATM Cell format

Figure 5 shows the correspondence between the RATM cell format and the ATM cell format. The RATM header has a number of extra fields, but the VCI/VPI field is more condensed.

Media Access Control (MAC) acknowledgements are generated when a bit error is detected in the cell, a 16 bit CRC is used to detect bit errors. No distinction is made between errors due to signal/noise induced bit-errors or collisions; this must be taken into account in the design of the back-off algorithm.

The short-cell and MAC-layer acknowledgement gives tolerance to independent bit-errors, a bit-error rate of 10^-3 gives a load of $1.5 * S$ with reliable communication. At a BER of 10^-3 and a repeat threshold of 10 repeats, the cell-error rate is 10^-5 equivalent to an uncorrected bit-error-rate of $3 * 10^-8$.

The sequence number is used to detect missing acknowledgements – when a cell is correctly received but for some reason the acknowledgement is lost, a duplicate cell will be generated by the source. If the cell were to be forwarded onto the wired network it would only be detected by the ATM adaption layer and would likely result in the loss of the associated block. The transmitting interface inserts a sequence number into the RATM cell header which is incremented on every successful transmission, this sequence number is checked in software when the cell is relayed.

There are a number of types of wireless cell transmission:

- DId Broadcast - wireless broadcast, limited to domain DId - not acknowledged .

- Full Broadcast - wireless broadcast to any base station within range - not acknowledged.
- BId selected - DId, no MId, only Base matching BId responds.
- MId selected - DId, (BId?), only mobile matching MId responds.

Broadcasts are used before a mobile has established a signalling channel to it's representative see 2.3.3. Broadcasts might also be used for non-critical broadcast data from a base to multiple mobiles eg. a multi-cast video system.

2.4.2 MAC protocol

The protocol used for the MAC is slotted ALOHA with exponential back-off. The ALOHA slot structure is maintained by re-synchronising the mobile to the base-station on each reception, thus each base-station is the source of synchronisation for the mobiles which are using it for communication.

Slotted ALOHA is a natural choice for the RATM system with the slot size being equivalent to an ATM cell. Slotted ALOHA has considerably better delay performance at low utilisation than a fixed allocation scheme and fits in well with the statistical multiplexing of ATM. It also allows hand-over or re-configuration at the higher network layers without requiring any re-allocation of resources at the MAC layer.

The multiple overlapping pico-cell approach results in potential interference between adjacent cells. ALOHA is robust in the presence of this - the state of the receiver is all that is important. CSMA on the other hand monitors the state of the carrier in the vicinity of the transmitter. In an indoor environment with small cells this is a poor measure of the state at the receiver - the hidden terminal problem means that active sources will not be detected and the interference between adjacent pico-cells causes false carrier detection [TK86], [NK85].

Small scale antenna diversity is used to improve the probability of reception. Repeated transmissions are sent from different antennas which are separated by less than a wavelength. There is very low correlation between the multi-path fading of spatially separated antennas even when the separation is around $\lambda/4$ (4cm at 2.5GHz) [TETM92]

The information indicating which antenna was used and the number of repeat attempts for a cell is available in the control information within the header and is used by the MR to select the initial antenna for each transmission attempt by the mobile.

2.4.3 Performance considerations

Theoretically pure slotted ALOHA has a maximum throughput of $1/e$, in a practical implementation there is the possibility of "capture" in which a collision between two sources is not necessarily destructive [Nam84], [CS85]. For perfect capture, that is where one of any colliding cells will be received correctly, the throughput approaches the capacity of the base station and capture improves the performance of all sources in the system, not just those close to the base [GS87]. Simulation of a realistic propagation model suggest that throughput of approximately half the bit-rate are achievable (see Appendix B). The combination of capture and large-scale diversity also has a large impact on the overall throughput. Increasing the number of base-stations increases the overall capacity even when the characteristics are the same at each base station [Sak92].

As is well known ALOHA is inherently unstable, when high load is applied the network falls into a low throughput mode in which almost all transmission attempts are repeats. A number of techniques are known for stabilising slotted ALOHA, (pseudo-Bayesian, FCFS) [BG87], but in general these rely on determining the outcome of each transmission attempt. All terminals communicating via a base-station can monitor acknowledgements with a high degree of success, but it is not necessarily possible to detect collisions or idle slots.

A number of different back-off strategies have been proposed which stabilise ALOHA by decreasing the probability of repeating a cell as the number of repeats decreases [RJ90]. These mechanisms are much easier to implement in a diverse environment as only the local success of a transmission is used as a metric. However, each station is only aware that it is failing to succeed so it is possible for successful sources to be unaware that they are swamping other transmissions. This means that if a number of sources are saturated there is no guarantee of evenness of access to the medium. [CH92] present a mechanism for implementing the pseudo-Bayesian algorithm in the presence of diversity and capture - but global information is still required regarding the success of a transmission.

Slotted ALOHA with capture is inherently unfair, [Abr77] shows that outside a particular range, the “Sisyphus distance”, users achieve no throughput. This can be mitigated through the use of ATM QoS by regulating the load generated by each source.

2.5 QoS management

As already discussed, one of the major reasons that ATM is proposed as a multi-media transmission mechanism is the ability to make Quality of Service (QoS) reservations for particular Virtual Circuits. The precise semantics of QoS have yet to be defined, but there will certainly be provision for supporting guaranteed QoS and best-effort QoS.

Guaranteed QoS These are services which demand connections with particular characteristics and the connection may be rejected if this cannot be guaranteed. Once the connection has been established the service must be informed if the service can no longer be supported. For example this might be where the application explicitly determines the data-rate and must change some aspect of the user interface if the required data-rate is changed such as a video phone application.

Best effort QoS These are services which are generally bursty and will tolerate back-pressure from the network interface to control the overall load on the network eg. traditional data services such as file transfer.

Within the combined wired/wireless network there are particular areas for QoS control which have different characteristics:

- Mobile-Base wireless link - QoS cannot be guaranteed across this link. Performance cannot be guaranteed when a mobile moves with respect to this base-station. The available link throughput can fall due to bit-errors, but hand-over and diversity are used to try and maintain performance
- Base capacity - The aggregate capacity of the pico-cell shared between all mobiles using the cell. This is likely to be the limiting capacity for the wireless network, whenever an attempt is made to over-subscribe any capacity this is probably the bottleneck.

- Wired network capacity - QoS guarantees are available in the wired ATM network and the capacity is of the order of an order of magnitude bigger than wireless network.
- Terminal-capacity - determined by application, interface and operating system performance.

In the wired network QoS is allocated in the normal way to active virtual circuits — but latent circuits are also allocated, this is because the wired network has excess capacity. If the latent virtual circuits did not have QoS allocated, then before a hand-over was completed it would be necessary to allocate QoS in all the intermediate switches which would impose a large overhead on the hand-over.

The maximum QoS allocation at a base-station is limited by the throughput of the wireless network this is more than an order of magnitude less than the throughput of a single ATM link. So the average over-allocation of service within the wired network will be a small proportion of the total.

Each virtual circuit has its own QoS requirement. An aggregate QoS requirement for a set of VCs can be determined by combining the individual requirements. At the base-station, where the throughput bottleneck exists, the QoS allocation for latent circuits is *not* included in the total QoS allocation. The base-station has a measure of the current QoS allocation and the cost of realising the VC set from a particular mobile. If the QoS demanded by the new circuits exceeds the total available then hand-over to that base-station is deferred if possible. If it becomes necessary to hand over to the new base then hand-over happens and the QoS allocation to each circuit needs to be re-negotiated.

When the total capacity has been exceeded due to hand-over it is necessary to decide which VCs to degrade. A parameter to the QoS allocation is the priority within a particular host for that circuit to maintain it's quality of service. For rate controlled systems the application does not need to be informed if the available data-rate is reduced. Back pressure from the network interface will just reduce the throughput. The fixed rate-services typically need application intervention if the available QoS is changed.

When a mobile moves into a new pico-cell it may have a number of active circuits which it has been given QoS guarantees for in a different pico-cell. There is relative priority of VC QoS within the set and absolute priority of QoS. A new set being made active can potentially cause renegotiation of VCs for other mobiles.

In previous multi-media experiments it has been shown that application layer feedback is an important factor in coping with varying service in an interactive environment [JH93].

2.6 Alternative Approaches

If the requirement for strict compatibility with wired ATM is relaxed and a private switching network is provided between base stations an interesting possible modification to the system is possible. The addition of a sequence number within the ATM cell header would allow the use of base-station diversity on a per-cell basis. A cell could be received by any base-station within range and forwarded to a filtering point which would filter out duplicates (and possibly error-correct the transmission) a delayed acknowledgement would then be generated. Unfortunately this increases the load on the backbone network because all the duplicate cells must be transmitted. The filtering of duplicate cells also requires some additional hardware within network. As previously discussed the advantage of using a standard switching network between base-stations outweighs the diversity advantage obtainable.

3 Experimental Wireless ATM Network

3.1 ATM Framework

An experimental radio system has been built with the aim of investigating the performance of the ATM protocol. The system is based around a range of low-end ATM switches which use the Advance Risc Machines ARM processor as a central switching unit. The switching is either directly through shared memory or by using broadcast backplane fabrics. Where a backplane is used, the ARM CPU sets up DMA transfers between input and output ports on the switch. There is a control interface to the switch which can be used to communicate to another system which evaluates VCI routes and performs signalling and other control and management functions. If performance permits these two functions can be combined on a single CPU.

This approach makes possible a flexible strategy for handling control and data in the switch. At some cost in performance it is possible to derail the data pay-load and execute priority or other QoS mechanisms in software. The range of switches includes 4x4, 8x8 and 16x16 designs and typical performance for the backplane is the switching of 1M cells per second (where no additional control functions are being performed). As the speed of the ARM CPU increases we would expect a performance increase of a factor of two. At the current speed the performance of the ATM switches is sufficiently good that the assumptions made about the ratio of RATM communication speeds and fixed network communication speeds are valid. The fixed ATM network is sufficiently fast that a number of communication state changes can be made in a single RATM cell transmission time.

3.2 ATM Direct Peripherals

The basic hardware design of the switch is based on a building block which can also be used for attaching peripherals directly to the ATM network. The design consists of line modules on the network side typically operating at 100Mbps and based on TAXIchip[3] technology. On the data side the hardware presents a general purpose data interface to which peripherals (or switch fabrics) can be attached. Thus the radio module becomes a direct peripheral which in this role performs the functions of a bridge between the fixed and wireless networks.

A software kernel, called ATMos, runs on the standard platform and is designed to facilitate the writing of control programs for small to medium size applications. The kernel provides no protection for direct access hardware. It consists of a boot ROM, applications libraries and initial core modules. There is a UNIX[4] based development and debugging environment as well as interfaces to POSIX threads and general purpose distributed platforms based on the CORBA specification [GX91]. A typical embedded ATMos application will consist of the scheduler, a number of device handlers, as well as the application related processes. Running as a RATM base station, with an ARM processor, about 75 percent of CPU power is still available after the fixed ATM network has been serviced.

[3] TAXIchip is a trademark of Advanced Micro Devices

[4] UNIX is a registered trademark of UNIX System Laboratories

3.3 Radio Interface

Initial experiments with a prototype data-link showed that the proposed data-rate could be achieved with tolerable error-rates. The prototype system was based around an 8-bit ISA interface, while the system under development interfaces to the 32-bit ATMos platform described.

The system is designed to minimise the complexity of the radio frequency interface while providing as much flexibility as possible for experimentation with prototype access protocols. One of the major design goals is to minimise the overhead per cell. This requires that the radio turn-around time must be minimised, currently of the order of 20 bit-times.

The system designed uses a carrier frequency of 2.45GHz and a bandwidth of 10MHz. This is currently an unlicenced band in the UK available for short-range data links in buildings. Transmission power is less than 10dBm and this gives sufficient power for pico-cells of a radius of around 10 metres. The modulation scheme is QPSK with a bit-rate of 10Mbit/s – QPSK being used to reduce the symbol rate and allow us to keep within the 10MHz frequency bandwidth. Figure 6 shows the radio front end architecture.

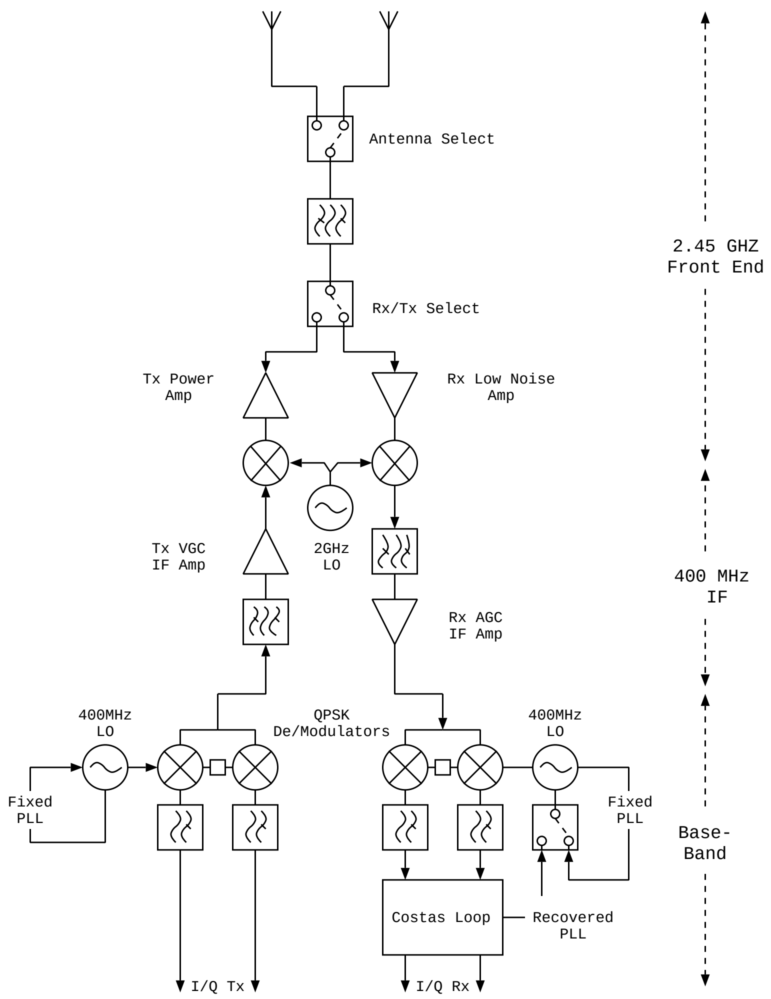


Figure 6: Radio Front End

3.3.1 Receive RF

The receive LNA amplifier is a GAAS FET and a silicon MIMIC. An active down-converter is used to convert to an Intermediate Frequency (IF) of 400MHz with a 2.05 GHz fixed frequency oscillator. The 400MHz IF uses an AGC with a fast attack and very slow decay. The fast attack gives a quick response to the beginning of the cell and when the end of the cell is detected by the MAC layer, the gain of the amplifier is forced low.

QPSK phase recovery uses an analogue Costas Loop. When there is no detected signal the passive QPSK demodulator is driven from a fixed frequency phase-locked 400MHz oscillator. When a received signal is detected the phase error signal from the Costas Loop is used to lock the oscillator. The recovery oscillator, therefore, only has to track the total frequency error between transmitter and receiver and is already centred at close to the correct frequency.

3.3.2 Transmit side

The front end-antenna switch selects a particular antenna before transmission of each cell. Each time a cell is repeated the selected antenna is changed.

A variable gain amplifier stage is used at the IF to provide isolation when receiving and variable power output for optimising the pico-cell coverage.

The transmit side uses a separate fixed frequency 400MHz oscillator so that the switch between reception and transmission can be as fast as possible.

The total switching time between reception and transmission is of the order of 2 micro-seconds. Given that the total cell transmission time is around 50 micro-seconds, this is considered to be a reasonable overhead.

3.3.3 Host Interface

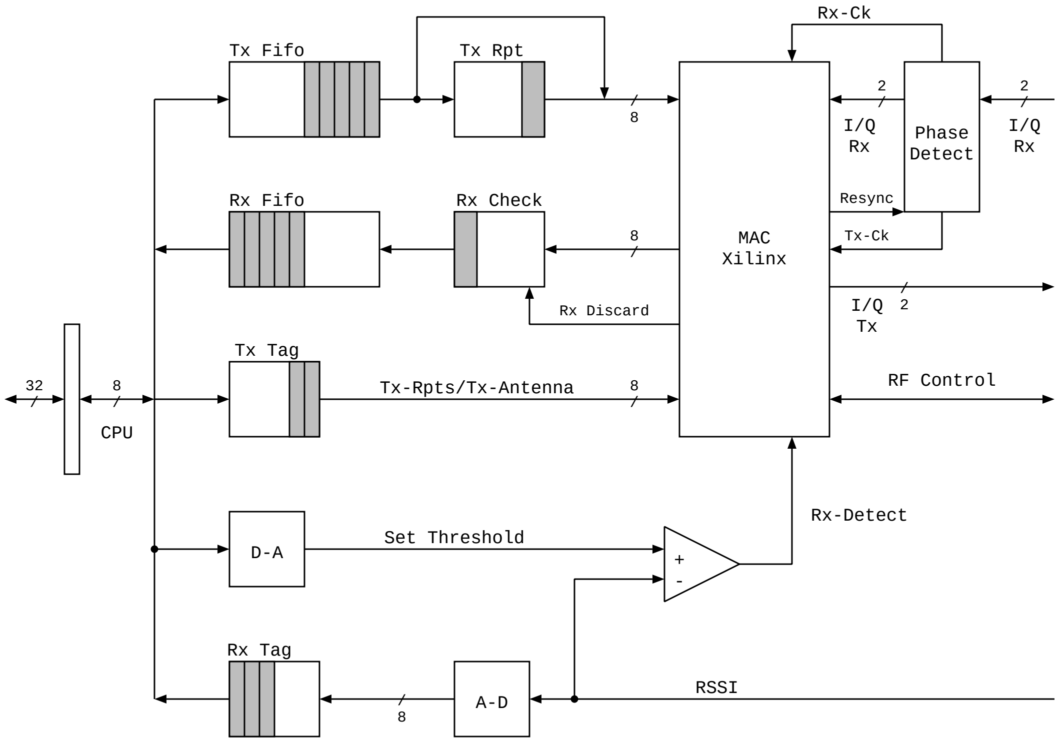


Figure 7: 8 bit ISA Interface

Figure 7 shows the prototype host interface to the radio network which interfaces to the 32-bit ARM host.

The MAC protocol is implemented in a re-programmable Xilinx gate array allowing experimentation with the details of the protocol. The details of the repeat strategy are important for the performance of slotted ALOHA and can be altered on a per-cell basis.

Up to 64 cells can be buffered in the *Tx Fifo* for transmission. When a complete cell is written into this fifo a single word is written into the *Tx Tag* fifo. This acts as an asynchronous up-down counter with an entry for each cell in the *Tx Fifo*. The tag fifo also contains per-cell MAC information such as the initial antenna to use and the repeat configuration for that cell. The first time a cell is transmitted it is copied into the single cell repeat fifo *Tx Rpt*; subsequent repetitions of the cell are copied from this fifo rather than the *Tx Fifo*. Reception of the cell begins when the received signal strength indicator reaches a threshold value set from the threshold digital to analogue converter.

The receive side is similar to the transmit side. An incoming cell is copied into the $RxCheck$ fifo and if the CRC check is valid is then copied into the cell buffer fifo $RxFifo$. When the copy is complete the signal strength of the reception is written into the receive tag fifo.

A receive interrupt is generated when the $RxFifo$ is half-full or an end-of-block marker is detected in a cell. The host can then read out the completed block from the fifo in one go.

A digital phase locked loop is used to recover the incoming symbol phase and this is also used to generate the transmission clock. When a synchronisation cell arrives at the interface the transmit phase can be forced to that of the receive phase. This is used to maintain the slotted nature of the ALOHA protocol.

4 Conclusions

The RATM system is being designed to extend the multi-media communications environment at Olivetti Research Ltd. This environment is based on the notions of plentiful network connectivity and a range of direct peripherals attached to the ATM network. This approach takes advantage of the network to make possible a wide range of applications — in particular ones which use a large number of network streams. The hardware modules that have been implemented include storage servers, video sources, audio sinks and sources, and simple display devices such as ATM LCD tiles. These direct peripherals have been used in applications ranging from video-conference and video-mail to permanent video links between laboratories. The RATM system is designed to be used within this context and make possible experiments with mobile versions of these applications. The distributed system context is based on typing of streams and negotiation of QoS of service attributes at all levels. It treats the physical location of a module as a first class parameter for help with the QoS negotiation. We hope to learn both the style and flexibility of types required for support of RATM and also the way location information (something which is natural to mobile systems) can be used to simplify and automate choices.

5 Acknowledgements

D.J.Greaves for discussion and clarification of ATM issues.

T.Blackie, O.Mason and S.Biswas for help with the protoype implementations.

C.A.Nicolaou and D.P.Gilmurray for help with preparation of this paper.

Professor M.V.Wilkes for many pertinent comments regarding earlier versions of the paper.

References

[Abr77] N. Abramson. The throughput of packet broadcasting channels. *IEEE Trans. Comm.*, COM-25:117–128, January 1977.

[amd]

[ARR87] A.A.M.Saleh, J. Rustako, and R.S.Roman. Distributed antennas for indoor radio communications. *IEEE Trans. Comm.*, COM-35:1245–1521, December 1987.

[Ber87] R.C. Bernhardt. Macroscopic diversity in frequency reuse radio systems. *JSAC*, SAC-5(5):862–870, June 1987.

[Ber93] R.C. Bernhardt. The effect of multipath fading on automatic link transfer for a random walk trajectory. *IEEE Trans. Veh. Tech.*, 42(4), November 1993.

[BG87] D. Bertsekas and G.R. Gallager. *Data networks*. Prentice-Hall, Eaglwood Cliffs, NJ, 1987.

[bis90] B-isdn draft recommendation i.150, 1990.

[BMZ93] R. Bultitiude, P. Melancon, and H. Zaghoul. The dependence of indoor radio channel multipath characteristics on transmit/receive ranges. *JSAC*, SAC-11(7), 1993.

[CH92] R. Clark Robertson and Tri T. Ha. A model for local mobile radio communications with correct packet capture. *IEEE Trans. Comm.*, COM-40(3):577–583, March 1992.

[CS85] I. Cidon and M. Sidi. The effect of capture on collision resolution algorithms. *IEEE Trans. Comm.*, 33(4):317–324, April 1985.

[DBMR90] D.M.J. Devasirvatham, C. Banarjee, R.R. Murray, and D.A. Rappaport. Multi-frequency radiowave propagation measurements in the portable radio environment. In *Proc. IEEE Int. Conf. Comms. ICC'90*, pages 1334–1340. IEEE, April 1990.

[DEB93] W. Diepstraten, G. Ennis, and P. Belanger. Wireless access method and physical specification, doc: Ieee p802.11-93/190, November 1993.

[FFG$^+$93] M. Frullone, G. Falciasecca, P. Grazioso, G. Riva, and A.M. Serra. On the performance of packet reservation multiple access with fixed and dynamic channel allocation. *IEEE Trans. Veh. Tech.*, 42(1), Feb 1993.

[FRV92] D. Ferrari, J. Ramaekers, and G. Ventre. Client-network interactions in quality of service communication environments. Technical report, Tenet Group, University of California, Berkeley, 1992.

[GS87] D.J. Goodman and A.A.M. Saleh. The near-far effect in local aloha radio communication. *TVeh*, 36(1):137–, February 1987.

[GX91] Object Management Group and X/Open. *Common Object Request Broker: Architecture and Specification*, Dec 1991.

[Has93] H. Hashemi. The indoor radio propagation channel. *Proceedings of the IEEE*, 81(7), July 1993.

[JH93] A. Jones and A. Hopper. Handling audio and video streams in a distributed environment. In *The 14th ACM Symposium on Operating System Principles*. ACM, December 1993.

[KMS$^+$93] K. Keeton, B.A. Mah, S. Seshan, R.H. Katz, and D. Ferrari. Providing connection-oriented network services to mobile hosts. In *USENIX Symposium on Mobile and Location-Indepenent Computing*, August 1993.

[Kru92] J. Kruys. Hiperlan, applications and requirements. In *PIMRC 9́2*. IEEE, October 1992.

[Nam84] C. Namislo. An anlysis of mobile radio slotted aloha networks. *JSAC*, SAC-2:583–588, July 1984.

[NK85] R. Nelson and L. Kleinrock. Rude-csma: A multihop channel access protocol. *IEEE Trans. Comm.*, COM-33(8):785–791, August 1985.

[RJ90] D. Raychaudhuri and K. Joseph. Performance evaluation of slotted aloha with generalized retransmission backoff. *IEEE Trans. Comm.*, COM-38(1), January 1990.

[Sak92] K. Sakakibara. Performance approximation of a multi-base station slotted aloha for wireless lans. *IEEE Trans. on Veh. Tech.*, 41(4), November 1992.

[She92] S. Sheng. A portable multimedia terminal. *IEEE Communications Magazine*, 30(12):64–76, December 1992.

[SV87] A.A.M Saleh and R. Valenzuela. A statistical model for indoor multipath propagation. *JSAC*, SAC-5(2):138–146, Feb 1987.

[TETM92] S.R. Todd, M.S. El-Tannay, and S.A. Mahmoud. Space and frequenc diversity measurements of the 1.7ghz indoor radio channel using a four-branch receiver. *IEEE Trans. on Veh. Technology*, 41(3), August 1992.

[TK86] H. Takagi and L. Kleinrock. Approximate output processes in hidden-user packet radio. *IEEE Transactions on Communications*, COM-34:685–693, 1986.

[VH93] R. Vijayan and J.M. Holtzman. Sensitiviy of handoff algorithms to variations in the propagation environment. In *CUPC '93 2nd Int.Conf on Universal Personal Cummunications*, pages 158–161, October 1993.

[Vit92] A.J. Viterbi. Spread spectrum cdma for digital cellular and beyond: Field test data. In *The third IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Boston, Massachusetts, Oct 1992. IEEE.

[WH92] R. Want and A. Hopper. Active badges and personal interactive computing objects. *IEEE Trans. Consumer Electronics*, February 1992.

Appendix A Mobility results

A necessary effect of reducing the coverage of each transmission cell is to increase the rate of hand-overs. The impact of this effect on system performance depends on the precise characteristics of mobility in the real world. The mobility of people within buildings is a good model for the mobility of wireless terminals within the network as all mobility in the indoor wireless networks we are considering is associated with the movement of people. This at-least gives an upper-bound for the expected number of hand-overs.

The measurement of mobility was made possible by an infra-red active tagging system which is used for the location of people within buildings [WH92].

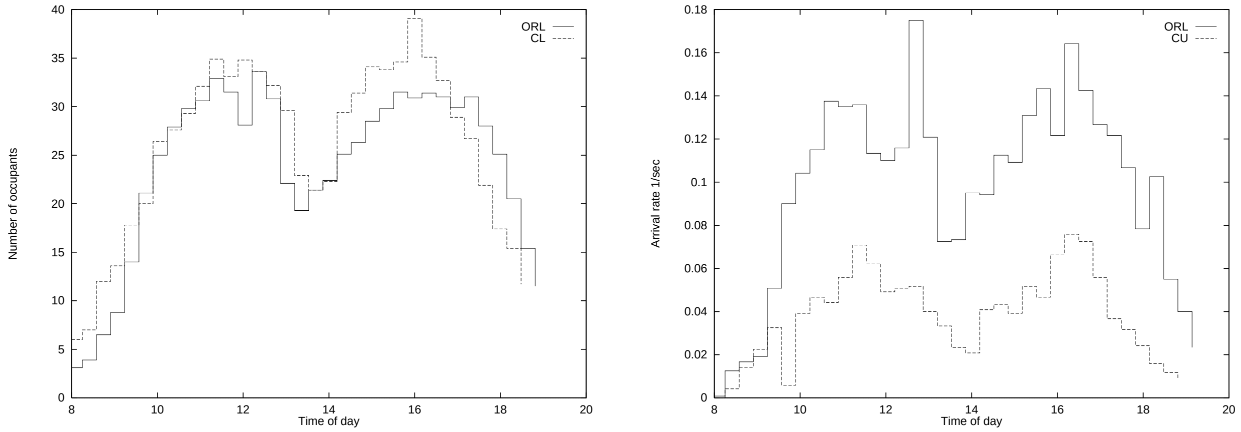
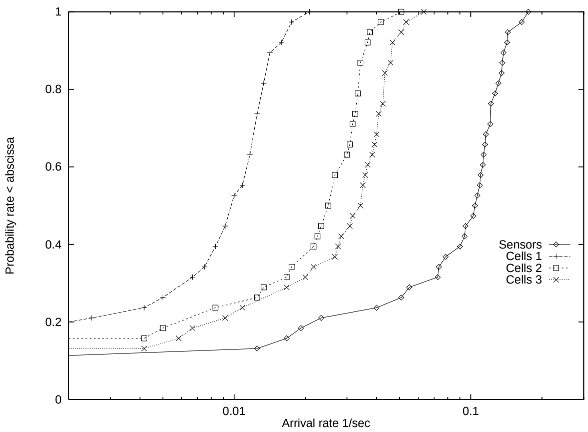


Figure 8: Mobility for occupants of ORL and CUCL

The sightings from all badges where recorded for a week, only those which where actually mobile are displayed, no equipment is included in the results. Studies were made of mobility in two buildings, ORL and Cambridge University Computer Lab. The ORL building has 100 rooms and 3 floors connected by a single stairwell, the CUCL building has 264 sensors on a number of floors with no sensors between them.The total number of sightings recorded was over one hundred thousand. The average number of people within the system during the day is shown in Figure 8a

Transitions between regions are of the most interest in the context of a wireless network. Figure 9 shows the number of transitions between rooms against time for the two locations. Of interest is that although both locations have almost identical numbers of occupants, the number of transitions is very different. This can be explained to some extent by the different types of organisation but also by the difference in coverage of sensor zones.

The peak number of transitions remains under 0.2/second and this gives an indication of the maximum expected load on the wireless network.



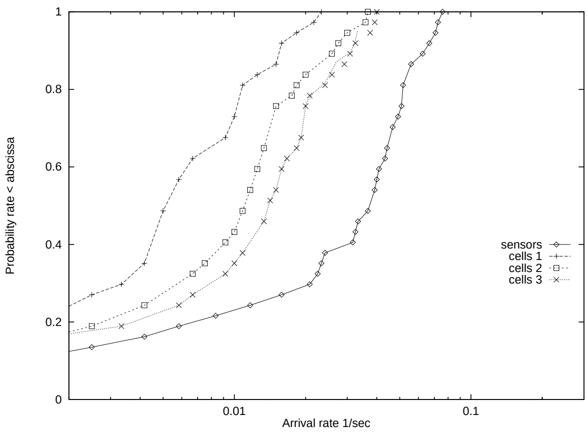


Figure 9: Arrival rate distribution in both buildings

Appendix B Performance Results

Appendix B.1 Propagation model

Choosing a particular attenuation model is difficult. There is a vast literature discussing propagation characteristics within buildings. For this series of simulations the attenuation between pairs of stations is determined using the free-space and linear path attenuation model based on that in [DBMR90]. The mean large scale attenuation is given by: $A(r) = 20Log10(4\pi r/\lambda) + Br$

λ is the wavelength of the carrier, r is the range in metres and B is the attenuation coefficient. A frequency of 2.5GHz with isotropic antennas and attenuation coefficient B of $0.6dB/metre$ is assumed.

The small-scale attenuation is log-normal around the mean value with a variance of 8dB, this seems to be generally accepted as a reasonable approximation for in-building propagation at the frequency of interest.

Appendix B.2 Diversity

This section analyses the affect of the propagation model in an overlapping pico-cell environment.

A grid of square pico-cells is considered where the separation between pico-cells is S metres. The radius of coverage of a particular base-station R is the range at which the mean level falls to a value of -90dBm. This reception threshold gives a signal to noise ratio of around 20dB.

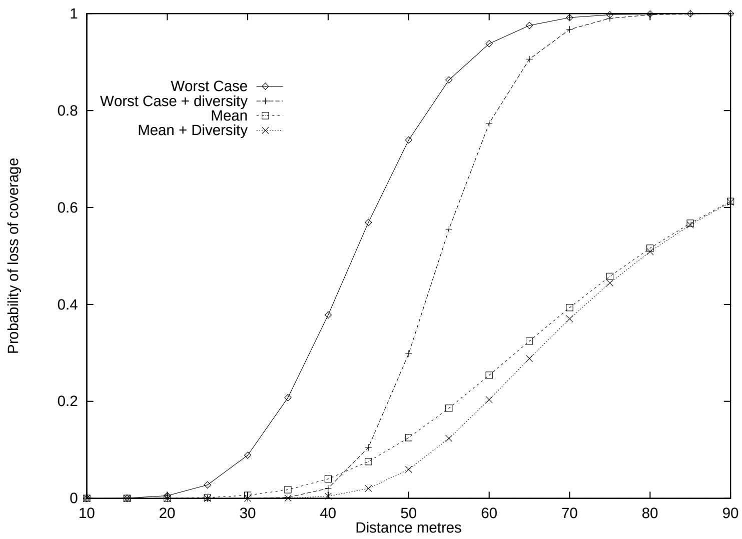


Figure 10: Probability of loss of contact against cell size

The probability that a mobile is not received at any base-station is taken to be the probability that its received power is less than the reception threshold. Figure 10 shows the probability of loss of contact for a mobile at the worst case position and the mean probability for all mobiles within the cell. The results are shown with and without diversity. Loss of coverage for a system with diversity requires that the mobile is not received at any base-station.

A measure of the diversity of a system is the average number of base-stations a mobile has in range. This diversity measure is given by the integral over the area of the pico-cell of the coverage from each base-station. Only the adjacent 8 cells are considered.

The diversity measure is: $C = 1/(S^2) * \int_0^S \int_0^S \sum_{i=1}^9 PB_i(x, y) dx dy$

Where S is the cell seperation and $PB_i(x, y)$ is the probability that the power from neighbouring base-station B_i is greater than the threshold at position x, y in the cell.

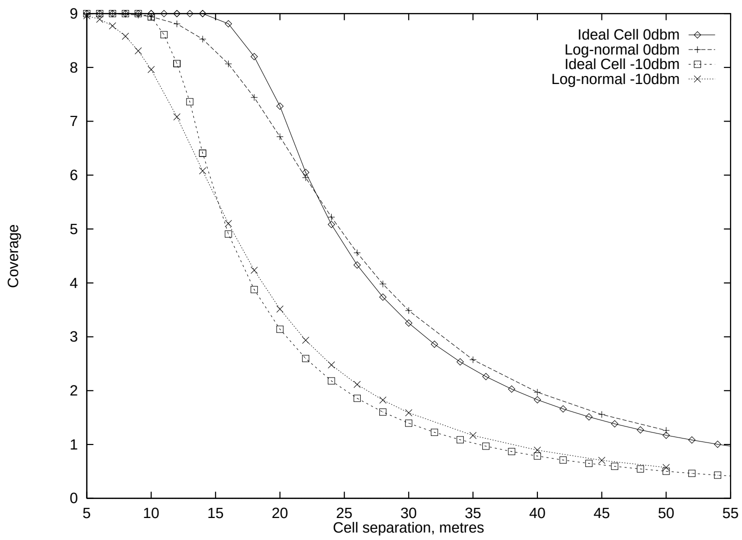


Figure 11: Mobile to base-station diversity

Figure 11 shows the diversity of the system as the pico-cell separation is increased for both a sharp cutoff at the threshold value ie. with no variation around the mean and when the log-normal distribution of signal strengths is considered. Cell sizes for transmission powers of 0dBm and -10dBm are considered.

Appendix B.3 Simulation Model

The physical interface to the radio medium is simulated in a manner which models the effect of interference and noise generated bit errors over time. Each medium interface is associated with a position and a transmission power. The attenuation between each pair of stations, and hence received signal strength, is determined statically at the beginning of each simulation run for a particular topology.

When a block of data arrives a start event begins a transmission which is propagated to all stations within range at the appropriate signal strengths. The receiving stations then compares the signal with that currently being received and if the incoming signal is greater "capture" is assumed. The previous signal is then assumed to be equivalent to Gaussian noise and adds to the current noise level at the station. This new signal and noise is used to generate the new signal/noise ratio.

If the incoming signal has a lower power than that currently being received it is added to the apparent noise at the receiving station.

At each change in received signal or noise the new BER is calculated and the previous BER rate is used to calculate the probability of error in the number of bits received at that level. This signal to noise ratio is used to determine the bit-error-rate using a look up table with values generated from the standard QPSK results.

When a transmitter finishes transmission the receiving station is informed. If there has been no bit-error during the period of transmission the data is passed up to the destination - otherwise it is discarded.

The interaction between transmitter and receiver is modelled by swamping the receiver when a transmission occurs.

Appendix B.4   Slotted ALOHA

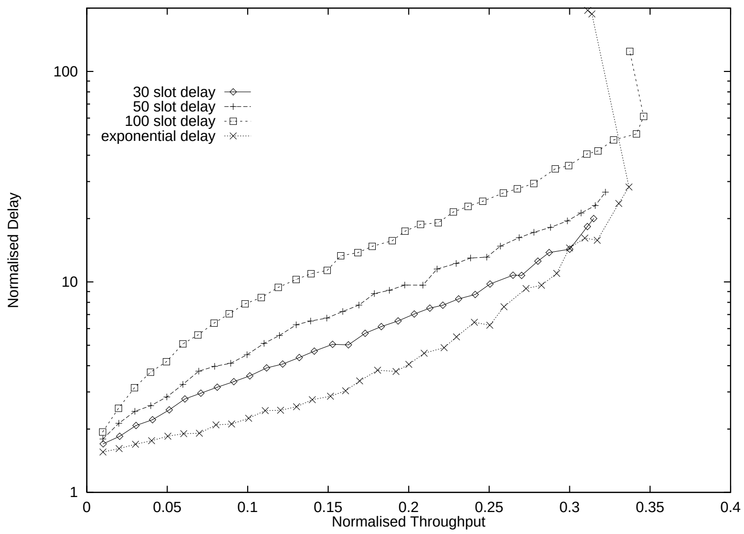
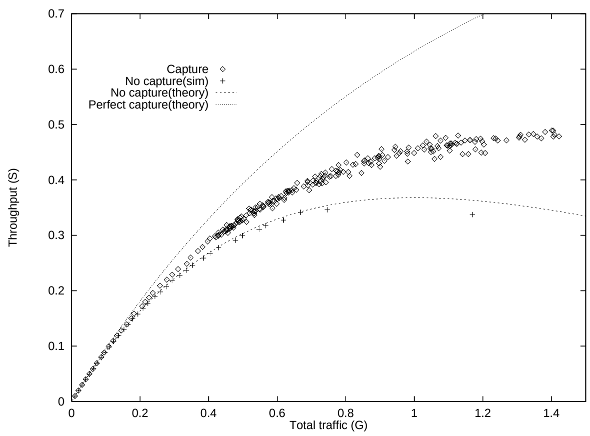


Figure 12: Slotted ALOHA with varying repeat probability

Figure 12 shows the stabilising effect of decreasing the repetition probability. This simulation has 100 sources transmitting to a base station with no diversity. The effect of exponential back-off is shown. At low load, a large static back-off delay increases the average delay even when there are few repetitions. The exponential back-off case does not suffer excess delays until there are significantly more repeats.



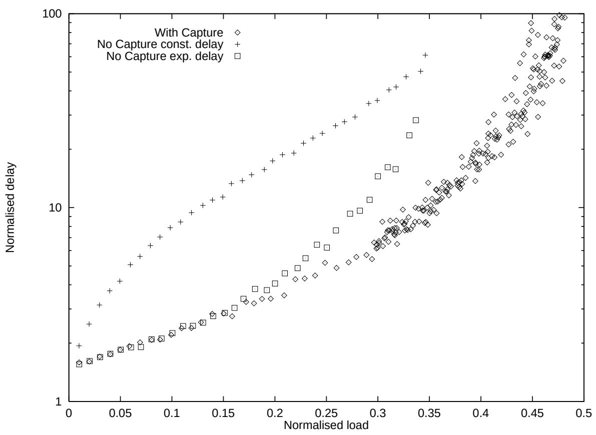


Figure 13: Effect of capture on single cell system

Appendix B.5 Overlapping Pico-Cells

The aim of the system is to support multiple overlapping pico-cells. There are diversity advantages in increasing throughput, but the utilisation of each base-station is reduced.

The simulation is based on a square 3x3 grid of micro-cells. The mobiles are distributed uniformly across the total grid and the performance of the mobiles nearest the central base-station is measured.

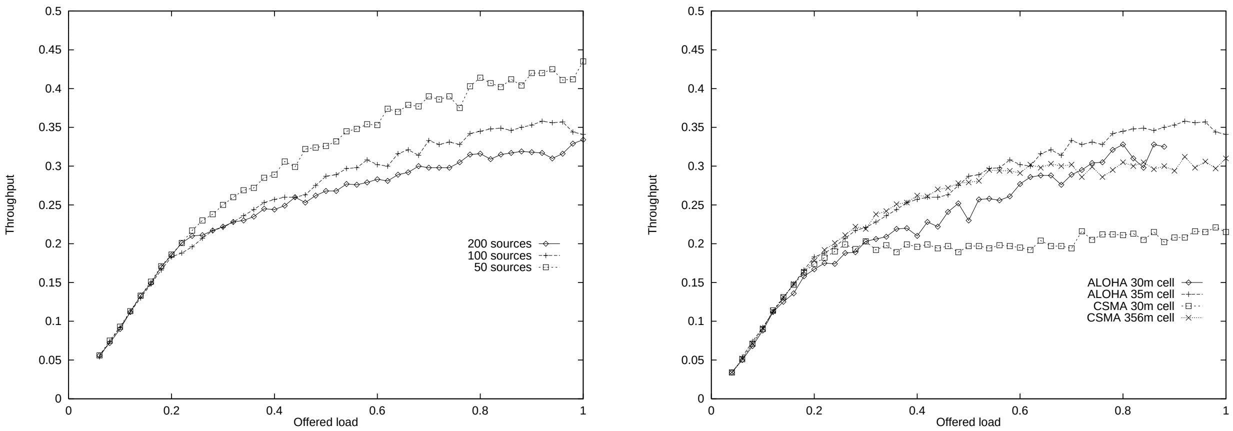


Figure 14: Multiple pico-cell system

Figure 14a shows the throughput of the system as the number of sources is increased.

Figure 14b shows the throughput of the system compared to a CSMA based approach. CSMA performs poorly due to interference between adjacent pico-cells causing unreliability of carrier sensing due to the hidden-terminal problem and false detection of activity in neighbouring cells.