A Survey of Wireless ATM Handover Issues

Chrysostomos Chrysostomou[1], Andreas Pitsillides[1], Fotini-Niovi Pavlidou[2]

[1]University of Cyprus, Department of Computer Science,
P.O. Box 20537, 1678 Nicosia, Cyprus
Tel: +357 2 892230, email: {cchrys, andreas.pitsillides}@ucy.ac.cy
[2]Aristotle University of Thessaloniki, School of Engineering,
Dept. of Electrical & Computer Engineering, Telecommunications Division,
540 06 Thessaloniki, Greece
Tel: +30 31 996285, email: niovi@vergina.eng.auth.gr

ABSTRACT

This paper reviews the requirements, characteristics and open issues of wireless ATM, particularly with regard to handover. It introduces key aspects of WATM and mobility extensions, which are added in the fixed ATM network. To support user mobility for a wireless network new mechanisms are needed and are fundamental, such as handover. Also an important key feature is mobile QoS offered by the WATM. A survey of the various schemes and types of network handover is provided. Several open issues for research have been identified. Failure to offer efficient solutions will result in handover severe delays and degradation of QoS offered to the applications.

I. INTRODUCTION

This paper gives an overview of wireless ATM (WATM) technology, in particular with regard to handover issues. It is appropriate to give first a short description of ATM technology and its basic principles in order to discuss the motivation for wireless ATM.

A. Asynchronous Transfer Mode (ATM): Basic Principles – Characteristics

The Asynchronous Transfer Mode (ATM) is a data transport technology that supports a single high-speed infrastructure for integrated broadband communication involving voice, video and data [1]. ATM technology combines some important features: short fixed-size packets or cells, virtual circuits, statistical multiplexing, and integrated services. All these concepts together provide a uniform framework that can carry multiple classes of traffic provided with a guaranteed quality of service (QoS).

ATM is a high-speed packet-based scheme, in which information is transferred in small packets (called cells) transported via transmission links at a bit rate of about 155 Mbit/sec or about 600 Mbit/sec. Each ATM cell is a small 53 bytes packet comprising of a 5 bytes header and a 48 bytes data field.

There are two ways to build a packet-switched network [2]. The first method is to have each packet to contain the full destination address (called a datagram). As addresses can be large, and if the average datagram length is small, bandwidth is wasted.

The second method of building a packet-switched network is to have packet headers to carry identifiers instead of addresses. In this way, each switch maintains a translation from the identifier to a destination. This saves header space, as identifiers are smaller than destination addresses. However, the mapping from an identifier to a destination must be set up at each switch along the path before data transmission begins. Hence, a call setup phase must precede a data transmission phase. This is called virtual circuit switching [2].

As stated above, for virtual circuit switching, although data is split into fixed size cells, those cells belonging to the same connection will travel along the same path. The path between the two parties is thus composed of virtual circuits (VCs) established between switching nodes. These VCs are set up on call establishment when various parameters such as acceptable QoS and demand profile (e.g., maximum and mean cell transmission rates, burst duration) can be negotiated.

This leads to a basic characteristic of ATM. ATM operates in a connection-oriented mode. That is, before any information is transferred from the terminal to the network, a logical/virtual connection set-up phase must first allow the network to do the reservation of the necessary resources, if these are available. If no sufficient resources are available the connection is refused to the requesting terminal. When the information transfer phase is finished, the resources are released. This connection-oriented mode of operation allows the network to guarantee in all cases a minimal packet loss ratio, and thus a maximal quality [3].

ATM scheme effectively offers bandwidth on demand. ATM accomplishes bandwidth efficiency through statistical multiplexing (sharing) of transmission bandwidth.

B. Wireless ATM

Wireless ATM (WATM) is mainly considered as an "access to an ATM network" issue. Depending on what kind of an ATM network is to be accessed, different aspects of wireless networking need to be addressed [4].

The following sections introduce key aspects of WATM and mobility extensions, which are added in the fixed

ATM network. To support user mobility for a wireless network new mechanisms are needed and are fundamental, such as handover, routing, and location management. Also an important key feature is mobile QoS offered by the WATM as opposed with that of other technologies.
It is worth mentioning that WATM has not been standardized yet. There are many proposals, and discussions by various researchers and companies contributing to the WATM working group of the ATM Forum. Research laboratories have developed several WATM prototypes. They include WATMnet, BAHAMA/MII, Magic WAND, AWACS, and others.

II. THE NEED FOR WIRELESS ATM

The area of wireless transmission systems has been increasing rapidly. Mobility raises a new set of questions, techniques, and solutions. This growth will occur in an environment characterized by rapid development of end-user applications and services towards the Internet and broadband multimedia delivery over the evolving fixed-wired infrastructure [1].
Therefore, new developments of wireless networks are needed to enable wireless technologies to interwork with existing wired networks.
A short look back into the history shows that all the current popular wired networks have their own wireless extension [5]. Therefore, in order for ATM to be successful, it must offer a wireless extension. Otherwise it cannot participate in the rapidly growing field of mobile communications [6].
As ATM networks scale well from local area networks (LANs) to wide area networks (WANs), and there is a need for mobility in local and wide area applications, a mobile extension of ATM is required in order to have wireless access in local and wide environments.
Many other wireless technologies, such as IEEE 802.11, typically only offer best-effort services or to some extend time-bounded services. However, these services do not provide as many QoS parameters as ATM networks do. WATM could offer QoS for adequate support of multimedia data streams [6].
Also, other wireless technologies are implemented in specific environments. For example, IEEE 802.11 only covers local area access methods, Bluetooth only builds up piconets, and Mobile IP only works on the network layer. WATM, on the other hand, tries to build up a comprehensive system covering many different networking scenarios, such as private and public, local and global, mobility and wireless access [7].
Considering all these aspects of WATM, it is apparent that this wireless system will be more complex than most of the other wireless technologies. Many open issues remain to be addressed and resolved. In particular, ATM was designed for media whose bit error rates are very low (about 10^{-10}). This performance benchmark is difficult to match with highly noisy wireless communication links.
As nodes do not have permanent access points to the fixed network while moving in a wireless environment, the need to accommodate mobility while satisfying established QoS presents a serious problem.
It is interesting to note that the performance of conventional ATM networks based on bandwidth optical fibers are limited by the switching capacity of the ATM switches. However, with WATM, the performance bottleneck has now shifted from the switching capacity of the switches to the transmission bandwidth of the wireless link [1].

III. WATM CHARACTERISTICS

Some of the critical requirements and characteristics of wireless ATM are outlined. Key aspects of WATM and mobility extensions, which are added in the fixed ATM network, are introduced.

A. *Cellular Architecture*

WATM networks covering reasonable distances must be built in a group of small geographical coverage zones. Since bandwidth is shared and spatially reused by many nodes, it is possible to give rise to co-channel interference. Reducing the size of the coverage area - to accommodate greater capacity per unit area- increments handover rate (due to increased crossings between coverage areas per unit of time). Consequently, the probability of dropped connections is increased. Also, routing becomes more dynamic because routes may need to be re-established whenever a handover occurs. This places great demands on the switching architecture [1].

B. *Resource Allocation*

The base stations of cellular WATM networks will need to provide assurance that QoS requirements will be met. This can be achieved by explicit resource allocation using a combination of admission, traffic shaping, and policing mechanisms. Requests for new connections are blocked if the anticipated traffic load presented by a new connection causes unacceptable congestion to build for existing connections.
The connection admission mechanism must also insure a low rate of dropped connections as users roam among different wireless coverage areas. The admission decision is usually based on several criteria [1] such as: traffic and handover characteristics; call holding time statistics; desired QoS of each class of traffic; and amount of radio resource available.

C. *Mobility Management*

Mobility is added to the current ATM protocols. Mobility management refers to roaming issues such as handover signaling, location management, and connection control. Location management is responsible for finding the mobile node. Handover refers to the process of changing frequency channels so that uninterrupted service can be maintained when nodes move across wireless coverage areas. This process also helps to track the mobile node dynamically. Connection control deals with connection routing and QoS maintenance [1].

Since a mobile node is usually located in a wireless coverage area on a temporary basis, all of these functions must be performed more frequently and at higher speeds.

In a mobile environment, management of the VC with QoS is not easy since the end-to-end path has to be continually modified as terminals move during the lifetime of a connection.

Because wireless base stations and mobile routing nodes are normally less capable than the wired network counterparts, WATM networks may potentially suffer from excessive delay and latency. In addition, the allocation of resources has to be re-evaluated each time a node moves to a new location. Moreover, mobile routing protocols need to operate in both wired and wireless environments if they are to be usefully integrated into future networks. Hence, the routing of ATM cells to mobile terminals requires new mechanisms.

Developing solutions that ensure QoS resources keep pace with continually changing network states resulting from user mobility, without consuming large amounts of overhead in the process, is a major subject for WATM research [8].

IV. HANDOVER

The system is responsible to route the traffic through the wireless network to the access point (AP), which is currently responsible for the wireless terminal. As the wireless terminal moves to a new position (AP), the system must reroute traffic. Therefore, the network must apply mechanisms responsible for searching new APs, and setting up new connections between intermediate systems.

One of the most important issues in a WATM network is handover. Handover involves rerouting of connections, as well as reserving resources in switches, testing of availability of radio bandwidth, tracking of terminals to perform look-ahead reservations etc [6]. The main consideration during handover in a WATM environment is to maintain connection quality.

The requirements for the handover procedure are expanded and detailed in ATM Forum/97-0153 [9]. Some of the key points of the requirements defined are explained [9]:

- The handover process should be fast enough so that the handover decision is still valid for the new position of wireless terminal after the handover process is complete.
- The switching of the active VCs from the old data path to new data path should be as efficient as possible in order to minimise the interruption to cell transport.
- The handover procedure should support handover between:
  - APs in a private network
  - APs in different private networks connected by a public network
  - APs in public networks.
- The handover procedure should aim to preserve the requested QoS of all VCs at handover. This may not always be possible and some form of QoS renegotiation and/or dropping of certain VCs on a priority basis may be required.
- Minimise cell loss but avoid cell duplication or cell reordering.

The purpose of the handover procedure is to ensure user mobility among the APs of the mobile network with minimal degradation on their QoS. Ensuring the completion of this procedure with very low delay has always been of primary importance [10].

These requirements are used to determine the suitability and performance of different handover schemes proposed by various researchers. Various solutions are developed performing efficient connection management in the case of handover.

A. Path Rerouting Scheme

Path rerouting involves changing the route of some portion of a connection from a suitable switch called a Crossover Switch (COS) to the new AP. Depending on the COS selection, the new route of the connection can be close to optimal [9].

For example, in Fig.1, let a mobile terminal move from its current coverage area (AP1) to a new area under AP2. Then the portion of the existing connection from the "handover switch" (COS) to AP1 is removed, and a new sub-path, from the point of detachment, is added with the remainder of the original connection path being extended to AP2. Details of this scheme can be found in [11].

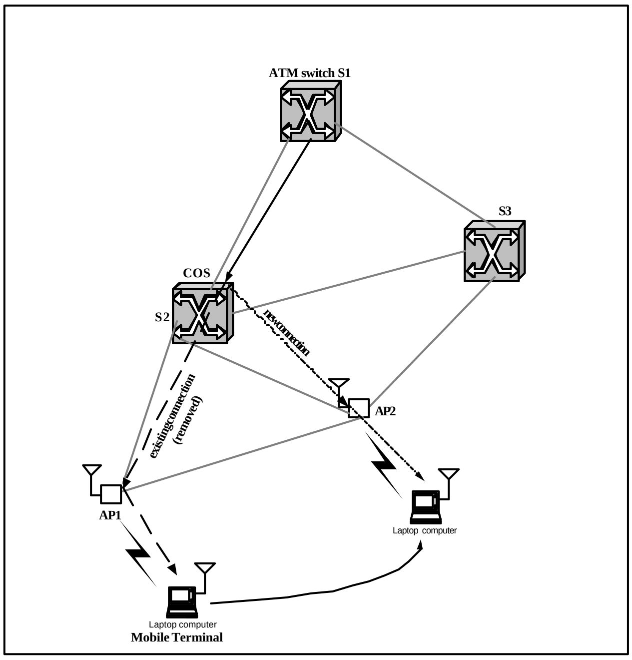

Fig. 1: Path Rerouting Scheme

Behind the simplicity of this method, some issues arise like the selection/discovery of the COS that need to be considered and further investigated. Different methods on selecting the COS give different performance for handover control in terms of latency, data loss and resource utilization [13]. One method is to iteratively probe each switch on the existing connection path such that the rerouted path through the switch satisfies the QoS of the original connection path [8]. However this method has the disadvantage of not preventing cell loss or reordering.

B. Path Extension Scheme

Another simple approach to handover consists of extending the route of a connection from the old AP to the new AP. The key issue behind this scheme is that after handover, the new connection consists of the existing connection from the source to the old AP followed by an additional sub-path, called the "extension", from the old AP to the new AP [8].

For example, in Fig. 2, let two users A and B, initialising a communication between them, being located under the coverage area of AP1 and AP2, respectively. Suppose that user A moves from AP1 to AP3. The traffic transmitted from the user A to user B will be transmitted from its current AP, AP3, to its source home AP, AP1, and from there to the current location of user B through AP2. Details of such an approach can be found in [16,17].

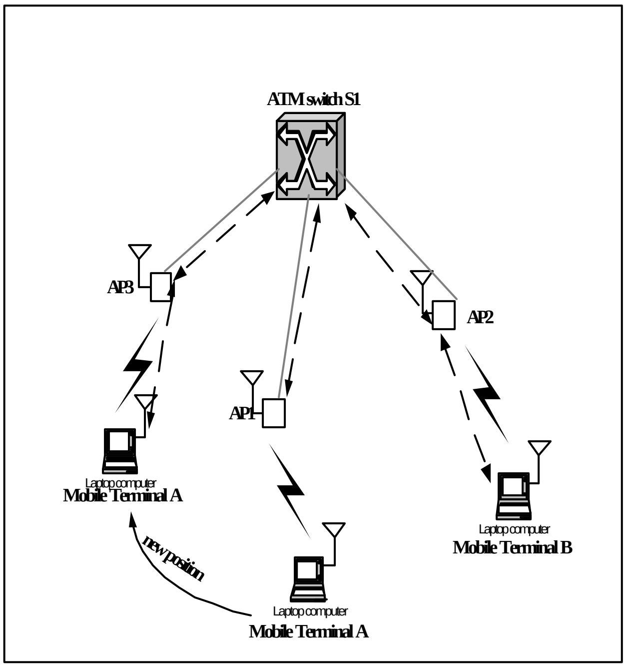

Fig. 2: Path Extension Scheme

The first obvious advantage of this kind of scheme is that no COS discovery phase is required, and the existing path is maximally reused [8]. Moreover, this approach makes it easier to implement connection handover without affecting data integrity [9], that is, maintaining the transmission order of the ATM cells during the handover procedure.

However, the extended path increases the end-to-end delay and also reduces network utilisation due to the creation of loops – since the extended path may traverse the same link, more than once [8]. Some form of route optimisation has been proposed, so as to counter these inefficiencies. A simple optimisation that is used to improve network efficiency during the path extension process is to detect and eliminate loops [12].

C. Handover Classes

Handover, which is implemented by the network, allows the users to roam freely within a limited wireless coverage area, while they are communicating. This has as a consequence to have handovers of virtual connections frequently.

Handover requires radio resource management functions for the air interface as well as network signalling and control functions for handover control, QoS management and rerouting of the connection to the new point of access to the ATM network [14].

WATM supports the following classes of handover: hard and soft, backward and forward handover.

- *Hard handover:* In a hard handover, a wireless mobile terminal has a radio connection with only one AP at any time.
- *Soft handover:* Soft handover supports simultaneous communication of a wireless mobile terminal with more than one AP during the handover.
- *Backward handover:* In a backward handover, the wireless mobile terminal notices, for example, a fading signal and initialises the handover to a new AP. The terminal continues to maintain the radio link while the handover is in process and switches over to a new AP after radio resources have been reserved and all entities involved are prepared for the handover [6]. Hence the handover execution can be initiated via the old AP.
- *Forward handover:* A forward handover is characterised by a wireless mobile terminal arriving at a new AP suddenly. The handover can only be initiated after the terminal has associated itself with the new AP. In this case, the new AP has to initiate and control the handover from there after. This happens when the terminal suddenly loses its connection to the old AP (due to interference or a fast-moving terminal), so there is no time to perform a backward handover.

These types of handover can be flexibly chosen, depending on radio conditions and QoS requirements, to enhance handover performance and robustness.

The two basic handover mechanisms are hard and soft handover. In the case of hard handover, the handover control flow can be directed either across the current AP's air interface (backward handover) or across the target AP's one (forward handover) [14]. A number of handover protocols have already been proposed based on hard handover [8, 15, 16, 17, 18].

In the case of having the radio link deteriorating rapidly, what is important is to have reliability and robustness of the handover signalling protocol. However, under these circumstances the regular backward handover mechanism will not perform reliably any more. In practice, the quality of the radio link prevents any communication on the dying radio link before the handover procedure can be completed. The signalling control flow will most probably be severely impacted [14].

Therefore the handover protocol has to be supplemented by a forward handover procedure, so as to avoid having handover failure and subsequent call dropping.

The forward handover aim is to maintain a connection despite the fact of having unexpected link failures during a handover situation. However, this is achieved at the expense of losing data on the old mobile connection segment and a higher cost of resynchronisation.

Summarising, the advantage of having a backward handover mechanism is that it gives you the possibility to choose the best AP for the wireless mobile terminal to connect to. In case of forward handover, the terminal suddenly interrupts the old connection and tries to connect to a new AP. However, this new AP cannot reserve resources in advance and therefore may not be such a good solution. But if the current connection is interrupted by radio interference, this is the solution of reconnecting fast enough. The terminal has to optimise its AP locally [6].

In [19-21] a handover scheme is designed in such a way as to support both backward and forward handover in a flexible way. It enables the wireless mobile terminal to instantaneously detach from its current AP and hand over its connection to the target AP at any time instance. As during backward handover the handover signalling is exchanged with the old AP, this instantaneous detach facility of the proposed scheme is of no particular use in this case. However, it can be very successfully exploited to provide zero cell loss forward handover [14].

An enhance performance of the proposed handover scheme in [19-21] through soft handover mechanism is proposed in [14].

However, [15] excludes the support for soft handover, arguing that soft handover is most often related to CDMA, which has been judged less likely for high bandwidth systems such as WATM.

Coming back to [14], it states that soft handover will be a basic feature for high-tier Mobile ATM systems, as air interface technologies for next generation wideband mobile services will enable soft handover, like for example 3rd generation CDMA technology. Soft handover poses the highest requirements on the radio technology, because it requires the wireless mobile terminal to be able to communicate concurrently with two APs. Therefore, in the overlapping boundary region, it enables dynamic selection of the best radio path. Provided that the overlapping region is sufficiently large and both APs can maintain a sufficiently strong signal in this region, this ensures enhanced QoS for the connection as well as handover reliability.

Two basic features must be taken into account during a soft handover process: synchronisation of the two communication paths over the different APs and dynamic path selection.

D. Handover QoS

A main key issue that must be taken into account during the process of handover is that of minimising the effects of QoS disruption [8]. For example, handover blocking due to limited resources at target APs, cell loss during handover, or the speed of the whole handover process are some of the critical factors for QoS.

One way to minimise QoS disruption during handover is to ensure a lossless handover [15]. In a network consisting of very small radio cells, the number of handovers could be high. As argued in [15], if even a single ATM cell is lost at each handover, the overall effect of the cell loss on the system throughput could be significant with the experience of delay too. Therefore, a key design aim of handover is that it is lossless.

To ensure a lossless handover, all cells in transit during the handover procedure are buffered within the network in order to maintain in-sequence cell delivery, without loss, to the wireless mobile terminal [8].

However, ensuring lossless operation is typically done at the expense of introducing additional buffering and therefore delay. This delay needs to be minimised [15]. Hence, an open issue to be further investigated is the planning of a lossless handover mechanism that also has low delay and delay variation.

In [15] the presented handover procedure can support both forward and backward hard handover, showing that lossless operation is important in order to achieve high radio link utilisation for loss sensitive connections subject to buffering in the AP.

E. Handover Research Issues

The main complexity of WATM arises from the functions and protocols for handover. In this paper, several schemes and types for handover procedure have been presented. The complexity of WATM is due to its ability to maintain QoS parameters for connections during handover and the connection-oriented paradigm of ATM [6]. As WATM has these critical characteristics, a main consequence is the need for resource reservation, checking for available resources at APs, as well as rerouting of connections.

An important issue in WATM that needs further investigation is the planning of an optimal handover procedure that enables the network with a guaranteed level of QoS being protected against cell loss, cell duplication, and loss of cell sequence.

An optimal design of handover should give a lossless mechanism that also has low delay and delay variation. What is important in the design of handover is to have in mind some critical factors that influence handover QoS. Such factors [8] are: handover blocking due to limited resources at target APs, cell loss during handover, and the speed of the whole handover process.

The way of implementing handover by means of choosing the right scheme and type for handover must ensure enhanced QoS for the connection as well as handover reliability. The main consideration during handover is to maintain connection quality. Ensuring the completion of handover procedure by preventing any cell loss and avoiding cell duplication or cell reordering with very low delay is of primary importance.

Hence a major subject for WATM research is the development of solutions ensuring that QoS resources keep pace with continually changing network states - resulting from user mobility – without consuming large amounts of overhead in the process.

V. CONCLUSIONS

ATM is a connection-oriented technology using cell-based switching/multiplexing techniques. It allows voice, video, and data services to be carried transparently through a single integrated network. WATM exploits these capabilities with the added benefits of mobility. However, due to the differences in characteristics between wired and wireless links, the design of WATM networks is aimed primarily at resolving issues related to resource allocation, mobility management and maintaining QoS in the presence of intermittent connectivity. Handover is a key function of a WATM network. Various types and schemes used for the network handover are addressed. In this paper several open issues for research have been identified, as for example in the field of giving to the network a guaranteed level of QoS, being protected against cell loss, cell duplication, cell reordering, handover blocking, and the speed of the whole handover process. Failure to offer efficient solutions in above will result in increased handover delays and thus influences the ability to offer QoS to the applications.

REFERENCES

[1] B. Bing, *High-Speed Wireless ATM And LANs*, Norwood: Artech House, ch. 6, pp. 167-214.

[2] S. Keshav, *An Engineering Approach to Computer Networking: ATN Networks, The Internet, and The Telephone Network*, Massachusetts: Addison-Wesley Publishing Company, ch. 4, pp. 47-64.

[3] M. D. Prycker, *Asynchronous Transfer Mode: Solution For Broadband ISDN*, Hertfordshire: Prentice Hall International (UK), ch. 2, pp. 61-96.

[4] E. Ayanoglou, K. Y. Eng, M. J. Karol, “Wireless ATM: Limits, Challenges, and Proposals”, IEEE Personal Communications, vol. 3, no. 4, 1996, pp. 18-33.

[5] J. Mikkonen, “Wireless ATM Overview”, Mobile Communications International, issue 28, 1996, pp. 59-62.

[6] J. Schiller, *Mobile Communications*, London: Pearson Education Limited, ch. 8, pp. 215-253.

[7] D. Raychaudhuri, “Wireless ATM Networks: Architecture, System Design and Prototyping”, IEEE Personal Communications, vol. 3, no. 4, 1996, pp. 42-49.

[8] A. Acharya, J. Li, B. Rajagopalan, D. Raychaudhuri, “Mobility Management in Wireless ATM Networks”, IEEE Communications Magazine, vol. 35, no. 11, 1997, pp. 100-109.

[9] J. Porter, D. Gilmurray, A. Massarella, J. Naylon, “Wireless ATM Handover Requirements and Issues”, ATM Forum/97-0153.

[10] V. Mocci, Y. F. Hu, “Virtual Tree Based Multicast with a Distributed Numbering Algorithm for WM-ATM Handover”, Mobile Networks and Applications, vol. 2, no. 5, 2000, pp. 111-118.

[11] R. Yuan, S. K. Biswas, L. J. French, J. Li, D. Raychaudhuri, “A Signaling and Control Architecture for Mobility Support in Wireless ATM Networks”, MONET – Mobile Networks and Applications, vol. 1, no. 3, 1996, pp. 287-298.

[12] P. Agrawal, E. Hyden, P. Krzyzanowski, P. Mishra, M. B. Srivastava, J. A. Trotter, “SWAN: A Mobile Multimedia Wireless Network”, IEEE Personal Communications Magazine, vol. 3, no. 2, 1996, pp. 18-33.

[13] A. Acharya, J. Li, A. Bakre, D. Raychaudhuri, “Design and Prototyping of Location Management and Handoff Protocols for Wireless ATM networks”, in Proceedings of ICUPC, 1997, pp. 213-217.

[14] H.-J. Vogel, “Robust and Soft: Handover Design for High-tier Mobile ATM systems”, in Proceedings of Wireless’99, 1999.

[15] H. Mitts, H. Hansen, J. Immonen, S. Veikkolainen, “Lossless Handover for Wireless ATM”, MONET – Mobile Networks and Applications, vol. 1, no. 3, 1996, pp. 299-312.

[16] B. A. Akyol, D. C. Cox, “Rerouting for Handoff in a Wireless ATM Network”, IEEE Personal Communications Magazine, vol. 3, no. 5, 1996, pp. 26-33.

[17] C.-K. Toh, “A Hybrid Handover Protocol for Local Area Wireless ATM Networks”, MONET – Mobile Networks and Applications, vol. 1, no.3, 1996, pp. 313-334.

[18] O. T. W. Yu, V. C. M. Leung, “Connection Architecture and Protocols to Support Efficient Handoffs Over an ATM/B-ISDN Personal Communications Network”, MONET – Mobile Networks and Applications, vol. 1, no. 2, 1996, pp. 123-139.

[19] H.-J. Vogel, “Handover Switching in Mobile ATM Networks”, in Proceedings of European Personal Mobile Communications Conference EPMCC’97, 1997.

[20] H.-J. Vogel, “Virtual Connection Handover in ATM-based Cellular Networks”, in Proceedings of European Conference on Networks and Optical Communications NOC’98, 1998, pp. 298-305.

[21] H.-J. Vogel, “A Networking Concept for Wide Area Mobility and Fast Handover in ATM Networks", in Proceedings of IEEE Globecom’97, 1997, pp. 1124-1128.