Camera ready version of paper for

1995 Wireless Communication Systems Symposium

"Wireless Trends in 21ST Century"

November 28 & 29, 1995

Authors: Dietmar Petras

Title: **Medium Access Control Protocol for wireless, transparent ATM access**

Address: Aachen Univ. of Technology (RWTH)
Chair of Communication Networks (COMNETS)
Kopernikusstr. 16
52074 Aachen, Germany

Tel.: +49-241-80-7928

Fax.: +49-241-8888-242

E-Mail: petras@comnets.rwth-aachen.de

WWW: http://www.comnets.rwth-aachen.de/~petras

Medium Access Control Protocol for wireless, transparent ATM access

Dietmar Petras
Aachen Univ. of Technology (RWTH)
Chair of Communication Networks (COMNETS)
E-Mail: petras@comnets.rwth-aachen.de
WWW: http://www.comnets.rwth-aachen.de/~petras

***Abstract* — This paper describes a MAC protocol called DSA++ that allows the full integration of mobile ATM terminals into an ATM network by extending the statistical multiplexing from ATM multiplexers on the air interface, offering capacity on demand. The DSA++ protocol offers functions necessary for the implementation of a resource scanning algorithm and power saving mode in the mobile station as well as a dynamic frequency selection algorithm in the base station.**

I. Introduction

The Asynchronous-Transfer-Mode (ATM) turned out to be the transmission technique of future communication networks[1]. Following the general trend of extending services of fixed networks to mobile users, necessary network structures and protocols for a mobile extension of ATM networks are investigated within the project *Mobile Broadband System*[2] (MBS, R2067, supported by the European Commission under the RACE programme).

The goal of MBS is to allow full integration of mobile ATM terminals into a fixed ATM network. This integration should not affect the end-to-end relation of the protocols of the ATM adaption layer (AAL). The transmission of ATM cells over the air interface is to be transparent to the AAL. The resulting layering is shown in figure 1.

The transmission of ATM cells by means of virtual channels over the air interface between several mobile stations and a central base station leads to a medium access control (MAC) protocol, that extends the statistical multiplexing from ATM multiplexers to the specific situation characterized by multiple access of not easy to co-ordinate terminals. In contrast to 2nd generation cellular systems, where each mobile station accesses its own synchronous time division multiplex (STDM) channel, statistical multiplexing on the air interface requires several mobile stations sharing the same physical channel. The MAC protocol has to allocate the medium to mobile stations on demand, dependent on the instantaneous requirements also taking into account the Quality of Service (QoS) requirement of virtual channels of all other mobile stations. The critical QoS parameters are the *maximum cell delay* especially of ATM cells from real-time oriented services and the *maximum cell loss rate*.

The delays on the air interface are caused by competition of multiple mobile stations on the uplink on the one hand and on the other hand by the necessity of retransmissions caused by a higher bit error rate compared to fibres. The retransmissions have to be controlled by a special ARQ protocol in the LLC layer [3]. The uplink is critical for the MAC protocol because of contention and danger of collisions.

The outline of the rest of the paper is as follows. Section II. describes the requirements on the MAC protocol and the basic concept of the new DSA++ protocol. The following sections are dealing with specific aspects of the DSA++ protocol which are signalling capacity assignments, transmission of capacity request messages and scheduling of ATM cells. Section VI. introduces an extension of the protocol which allows to support the LLC layer in fast transmitting acknowledgement messages.

II. Dynamic Slot Assignment Protocol

Based on the requirements on the MAC protocol described above, the *Dynamic Slot Assignment* (DSA) protocol has been developed [4] and the performance of its ideal version has been validated by simulations in [5].

We now describe a more sophisticated version of the DSA protocol that is more suitable for the implementation in a cellular system with its constraints on physical transmission, signal processing and real-time protocol execution. Following a current trend in naming we have called this new protocol DSA++.

The basic concept of the DSA protocol is to handle a cell consisting of a central base station and several mobile stations as a distributed queueing system and to transmit ATM cells over the uplink according to their residual life time weighed by a parameter derived from the required maximum cell loss rate of the associated virtual channel. In this way it is possible to satisfy the QoS of real-time oriented services even with the low data rates on the air interface (e.g. in MBS: 34Mbit/s [6]) compared to ATM multiplexers of fixed ATM networks.

A base station operates a physical channel accessible by multiple mobile stations. As more mobile stations are accessing simultaneously the efficiency of statistical multiplexing improves. The physical channel offers a sequence of time slots each able to carry a burst with one ATM cell together with the necessary overhead for training sequence, synchronization, FEC and guard time. The allocation of capacity of the physical channel takes place slot–by–slot. It is controlled by the base station as the central instance.

The description of the DSA protocol can be separated into following subaspects:

- signalling of capacity (slot) assignments/reservations over the downlink

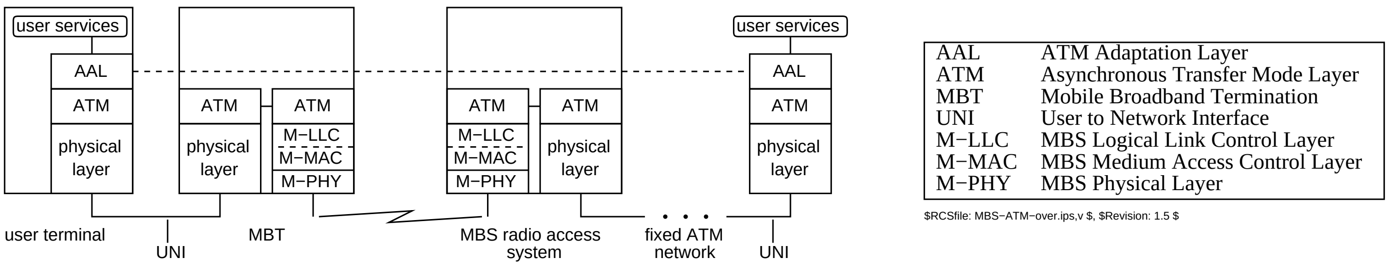

Figure 1: Protocol stacks for full integration of mobile ATM terminals into a fixed ATM network

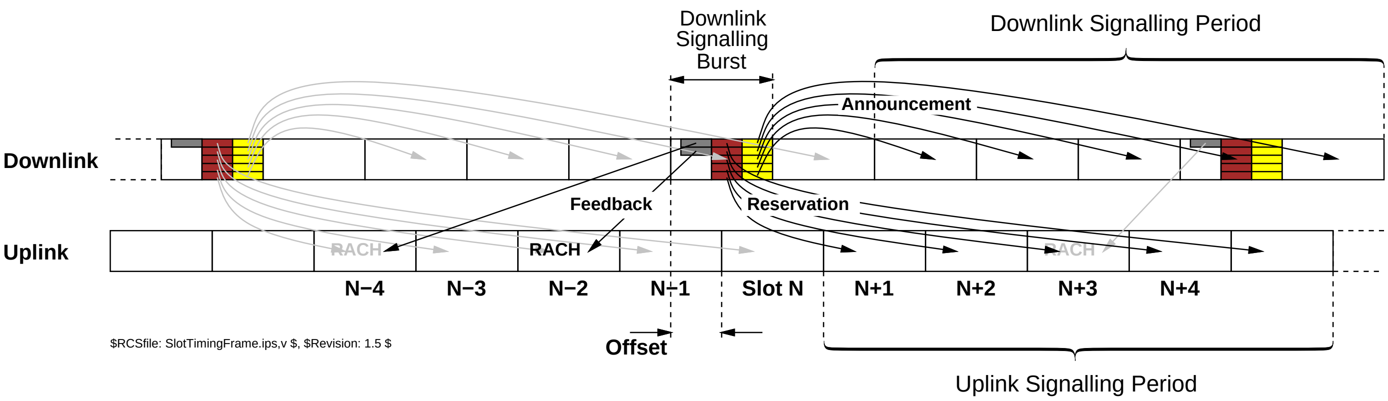

Figure 2: Downlink signalling scheme of DSA++ protocol

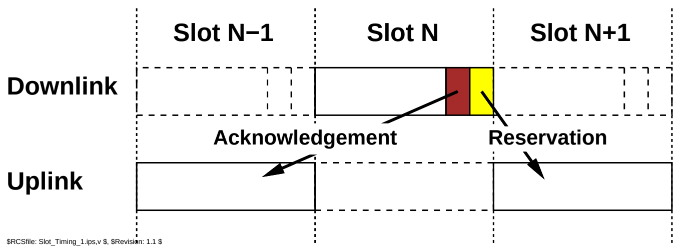

Figure 3: Downlink signalling scheme of DSA protocol

- transmission of capacity requirement messages over the uplink (by inband signalling, random access, polling)
- algorithm in the base station for determining the order of ATM cell transmission on uplink and downlink
- fast collision resolution algorithm and stability of random access

III. Signalling on Downlink

In [5], a simple scheme for signalling of slot reservation and acknowledgement messages over the downlink is described and evaluated (figure 3). Because of its short round trip delay between requests and reservations it allows very dynamic reactions on changing capacity requirements which is why we call it *ideal* DSA. A disadvantage of this simple scheme is, that all downlink bursts have to be sent in broadcast mode, since the receivers of reservation and acknowledgement messages as well as payload (ATM cell) are not the same in general. This prevents the usage of power control mechanisms. Furthermore the mobile stations have to receive each downlink burst in order not to miss messages a station is supposed to receive. That is why it is difficult to implement a power saving mode as well as a resource scanning algorithm to find alternate physical channels (operated by other base stations) necessary for executing fast and seamless handovers.

To avoid the previous described problems, the DSA++ protocol groups the signalling messages for several consecutive slots in a *downlink signalling burst* as shown in figure 2. Thus each downlink signalling burst opens a signalling period of a specific length. They have to be transmitted in broadcast mode. But all other bursts are directed only to one mobile station so that the usage of power control is possible. A downlink signalling burst contains the following messages:

- a reservation message for each uplink slot of the signalling period
- an announcement message for each downlink slot of the signalling period
- a feedback message for each slot with random access of the previous signalling period
- a field for other system signalling messages (paging channel, info channel, etc.)

The number of messages within a downlink signalling burst and therefore the length of a signalling period depends on the available number of bits in a burst. Because of their importance for the correct protocol execution it is useful to spend additional redundancy for error correction for downlink signalling bursts. The usage of the various signalling messages of a signalling period is illustrated in figure 2. The slots on the uplink are shifted against the slots on the downlink so that the resulting period τ_{offset} can be used for decoding a received burst.

Transmission of announcement messages allows mobile stations to leave a physical channel for short time (e.g. to execute resource scanning or switch to power saving mode) without loosing synchronity or missing messages.

The feedback messages are necessary to enable fast collision resolution for transmission in contention mode. For simplicity the random access illustrated in figure 2 occurs in normal slots. In section IV. it is explained that random access is always executed in subslots. Thus a downlink signalling burst has to contain one feedback entry for each subslot with random access of the previous signalling period.

IV. Transmission of Capacity Request Messages over the Uplink

Uplink slot reservations are determined by the base station. Therefore it has to have a mostly actual knowledge of the capacity requirements of the mobile stations. The instantaneous requirements of mobile stations are described by their *dynamic parameters.* In order to inform the base station about the newest requirements, each uplink burst contains the dynamic parameters of the sending mobile station in its header. Examples of dynamic parameters are:

- residual lifetime of most critical queued ATM cell
- mean residual lifetime of n most critical queued ATM cells
- number of queued ATM cells

But often situations occur, in which the mobile station's necessity for transmission is increased without being able to inform the base station. Two cases are possible:

1. If a transmitted ATM cell is the last one in its queue, this mobile station will not request further slots. Since it only makes sense to assign a slot to a mobile station if a transmission in that slot can be guaranteed or at least is very likely, this mobile station cannot inform the base station on an arrival of a new ATM cell.
2. The residual lifetime of the most critical queued ATM cell in a mobile station can decrease very fast if, e.g. a mobile station has several parallel VCs with different maximal delays and a new ATM cell with a very low maximum delay is generated. In this case the base station will handle the mobile station with too low priority.

In these cases an alternate way of transmitting dynamic parameters is necessary. It is done by cyclic insertion of special slots, which are intended for random access. For simplification the sequence of these contention slots is called *random access channel* (RACH). To prevent unnecessary collisions, only those mobile stations may access the RACH, whose state corresponds to one of the two cases stated above. Such a mobile station is said to be in *contention mode.* Likewise, a mobile station, which has sent its dynamic parameters to the base station and according to its need of transmission can expect a reserved slot, is said to be in *reservation mode.* With this combination of reservation and contention it is not necessary for mobile stations with bursty services in times of low data rates to request reservation slots.

To reduce the amount of data lost in a collision, RACH slots are divided into several subslots resulting in increased density of contention periods. In these subslots short bursts are transmitted, which only contain the dynamic parameters. Because of more contention subslots and constant input the collision probability decreases. Therefore the base station earlier gets notice of changing capacity requirements leading to shorter delays. Because of the overhead for guard time, training sequence etc. in each subslot, a contention slot cannot be subdivided indefinitely. Based on the burst structure of MBS [6] four subslots seem to be realistic.

For stability reasons and for achieving short reaction times on changing capacity requirements, a fast collision resolution algorithm is necessary. Stability can be guaranteed with each well known collision resolution algorithms with maximum stable throughput $\lambda_{max} > 0$. A new *Priority Splitting Algorithm*, a modification of the First-Come First-Serve (FCFS) Splitting Algorithm [7], seems to be the best candidate when weighing efficiency against implementation/execution effort. The concept of all splitting algorithms is to divide the set of colliding packets into two (or more) subsets (e.g. with the FCFS algorithm splitting is done by arrival time). With the Priority Splitting Algorithm collided mobile stations are split according to their priority, which is calculated based on their dynamic parameters. The algorithm for priority calculation has to produce values that are mostly equally distributed over a discrete range (e.g. $[0, 1 \ldots 127]$), so that the splitting of a collision set produces two subsets of approximately equal size.

An efficient collision resolution algorithm is a prerequisite for fast transmission of dynamic parameters in random access mode. The DSA protocol offers an additional way in order to reduce the length of a collision resolution period: Because of broadcasting uplink slot reservations inside downlink signalling bursts it is possible to schedule new RACH slots very fast and flexible. Thus after occurrence of collisions additional RACH slots for fast collision resolution can be inserted in the next signalling period. A problem is the delay between a random access slot and reception of the associated feedback, transmitted with the next downlink signalling burst. This delay can be further reduced by shortening the length of signalling periods. Such a shorted period has to be announced by the downlink signalling burst opening the period. A collision resolution period consisting of several shorted signalling periods is illustrated in figure 4.

With the introduction of subslots for random access it is necessary to coordinate multiple RACH slots in one signalling period. But splitting a set of collided stations requires the knowledge of the associated feedback, which will not be available before the start of the next signalling period. Therefore multiple random access of a mobile station within the same signalling period is not useful. But these

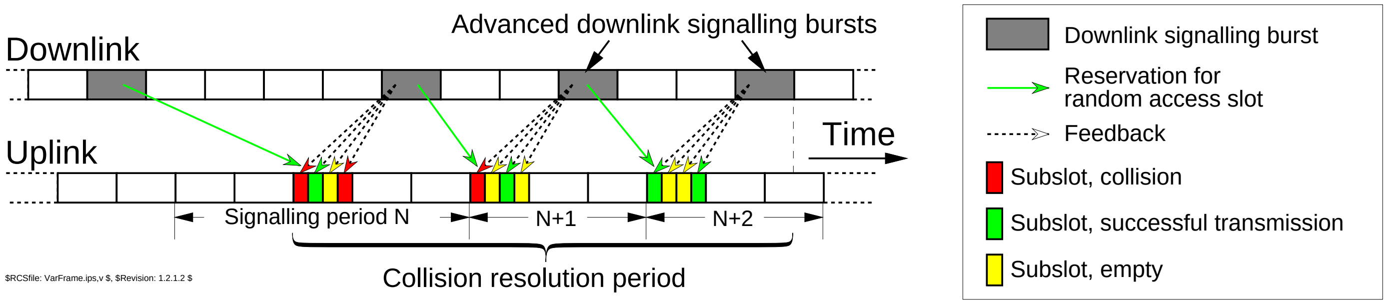

Figure 4: Collision resolution with shorted signalling periods

quasi parallel RACH slots can efficiently be used with multiple simultaneously running instances of the splitting algorithm. One instance can use one or several RACH slots of a signalling period. During the generation of the next downlink signalling burst the number of necessary RACH slots resp. subslots is determined dependent on the number of splitting algorithm instances and their internal states. New instances are generated if there are mobile stations which want to access to RACH slots but are currently not associated with an instance. The size of this *waiting set* can be estimated taking into account the QoS parameters of a mobile station. After the resolution of a collision set an instance will finish execution.

V. Determining the Order of ATM Cell Transmissions

For each signalling period the base station has to determine reservations for all uplink slots of this period. It can be assumed, that the transmission of dynamic parameters by the two possible ways described above leads to a sufficient knowledge of the capacity requirements inside the base station. Thus the system consisting of the central base station and the associated mobile stations can be modelled as a distributed queueing system with the server being determined by the channel with its deterministic service time.

The order of transmissions has to be determined in such a way, that the required values for maximum delay are observed for each virtual channel and that furthermore the rate of violations on the maximum delay, which lead to cell losses because of discarding delayed cells, does not exceed the required maximum cell loss rate of a virtual channel.

The requirements on the transmission of ATM cells are described by the previous mentioned dynamic parameters. Special attention has to be payed to the residual life time of a cell resp. the latest acceptable time for transmission. This leads to a queueing strategy based on dynamic priorities, e.g. G/D/1/FCFS/RU-NONPRE, relative urgency.

In order to assign multiple uplink slots of a signalling period to the same mobile station, the dynamic parameters of this station have to contain information on as many ATM cells as reservations which have to be determined. The information has to be extracted from highest prioritized ATM cells buffered in the mobile station.

Because of infavourable conditions when transmitting dynamic parameters, it may happen that the base station does not receive any information about the dynamic parameters of a mobile station for some period of time. In this case the base station has to work with the latest values of the dynamic parameters but can execute a prediction based on the QoS parameters of the mobile station (e.g. mean cell rate). This results in polling the mobile station in order to transmit the dynamic parameters in special subslots which are not declared for random access but reserved for this station.

Another adaptation of the dynamic parameters has to be executed to consider erroneous transmission of ATM cells in reserved slots. Whenever the base station schedules an uplink slot for transmission of an ATM cell, the dynamic parameters of this cell are removed from the queueing system. But if the transmission in the reserved slot was not successful, the base station can reestablish the corresponding dynamic parameters in order to schedule another slot. For stability reasons this procedure has to be abandoned after n tries.

The queueing algorithm has to offer a special procedure for inserting slots dedicated for transmission of dynamic parameters in random access. This can easily be done by generating dynamic parameters for dummy ATM cells which, when selected, will result in the insertion of a RACH slot divided into subslots. The number of dummy cells depends on the number of instances of the collision resolution algorithm and their requirements. The dynamic parameters (priority) of a dummy cell depends on the internal state of the associated instance. The priority calculation has to be adjusted to the algorithm for calculating priority of normal ATM cells.

The previous description was focused on the transmission of ATM cells over the uplink. The algorithm can also be used for the transmission over the downlink. But because the ATM cells waiting on transmission over the downlink are buffered inside the base station, it has the full knowledge about the capacity requirements and thus the problem of transmitting dynamic parameters does not exist.

VI. Delivering Feedback Values to the LLC Layer to be used as fast Acknowledgement

The delays on the air interface are not only caused by the queueing delay because of statistical multiplexing, but also by the necessary retransmission because of erroneous transmission. Therefore the transmission of positive as well as negative acknowledgements especially for ATM cells of real-time oriented services has to be handled with high priority[3]. The acknowledgements and retransmissions are controlled by the LLC layer. But the MAC layer can support the LLC layer by delivering inside a downlink signalling burst not only feedback messages for RACH slots but for every uplink slot of the previous signalling period which has been reserved for the transmission of an ATM cell. The MAC protocol of mobile stations can combine these feedback values with a reference (e.g. handle ID) to the ATM cell belonging to the feedback and deliver it to the LLC layer, where it can be used by an ARQ protocol as fast acknowledgement.

With some additional effort it is possible also to transmit feedback messages over the uplink in order to acknowledge the previous transmission of an ATM cell on the downlink. For this purpose special subslots can be inserted in the stream of slots on the uplink. Three aspects have to be considered:

- timing relation between uplink and downlink and the order of normal slots and subslots
- necessary signalling for indicating and reserving feedback subslots
- algorithm for determining the density and position of feedback subslots

The simplest method for solving the timing and signalling problem is by using a fixed scheme for the sequence of normal slots and subslots on the uplink. This can be done by inserting a feedback subslot for each slot of the downlink in which transmission of ATM cells has been scheduled. By that way there is a fixed timing relation between receiving an ATM cell burst on the downlink and sending a feedback burst over the uplink, resulting in an immanent signalling scheme which makes it unnecessary to send any further signalling message in downlink signalling bursts.

In figure 5 two variants of a possible timing scheme are shown. With the first variant all feedback subslots are grouped together in order to keep the synchrony between slots on uplink and downlink. With the second variant always one feedback subslot is inserted between two normal slots. This leads to a very early transmission of feedback messages resulting in reduced delay till a necessary retransmission arranged by the LLC layer. Asynchronities between uplink and downlink because of shorted signalling periods can be compensated by inserting additional subslots e.g. for transmission of dynamic parameters.

With some more signalling effort it is possible to insert any number of feedback subslots. Furthermore this can be performed asynchronous to the transmission of ATM cell burst on the downlink and it is not necessary, that a uplink part and the downlink part of a signalling period (figure 2) are of equal length but can vary by a multiple of the length of one subslot. In this way the transmission of bursts in feedback subslots can be signaled to the LLC layer of the mobile stations which will generate spe-

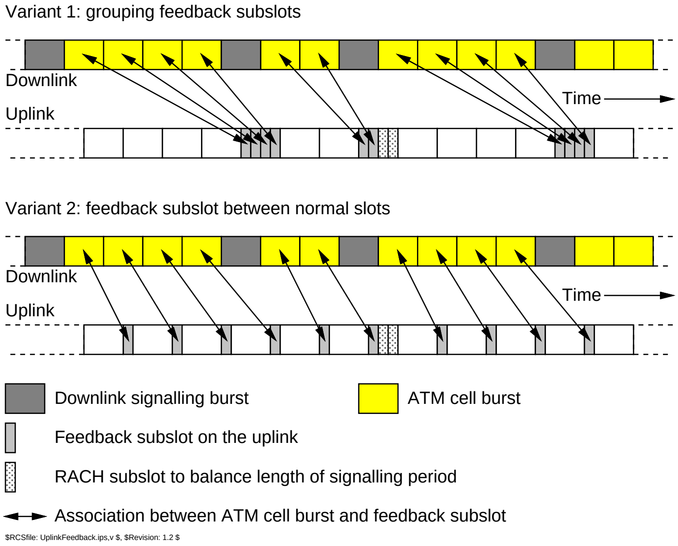

Figure 5: Timing schemes for immanent signalling of feedback subslots on the uplink

cial short frames containing acknowledgement messages of the ARQ protocol.

When executing an ARQ protocol, the information of acknowledgement messages is important for the sender of information frames because it has to arrange retransmissions. Therefore the sender has to be asked for the amount of capacity that has to be offered for transmission of acknowledgement messages. Based on this considerations the base station has to execute an extra scheduling algorithm for insertion of feedback subslots which is comparable to that one coordinating the transmission of ATM cells described above. Based on the requirements of the LLC protocol this algorithm has to schedule feedback subslots and generate the necessary signalling messages contained in the downlink signalling burst of each signalling period.

VII. Performance Evaluation by Simulation

For performance evaluation of the *ideal* DSA protocol presented in [5] an event driven simulation tool has been developed, which is currently being extended by the DSA++ protocol. The simulator is based on the SIMCO3++ tool developed at COMNETS[8]. It contains models for propagation, mobility and traffic generation enabling the validation of protocols under realistic environmental conditions.

The simulator is intended to allow the combination of all protocols and components of a mobile ATM system. Therefore the LLC protocol and higher layer functionalities have been implemented[3]. Furthermore algorithms for resource scanning in the mobile station and dynamic frequency selection in the base station are under investigation. The integration of this functions will allow to produce a detailed evaluation of the performance of the single protocols as well as of the complete system.

VIII. Conclusions

We have described an improvement of the DSA protocol called DSA++ which is able to extend the statistical multiplexing from ATM multiplexers on the air interface by offering capacity on demand. The protocol offers functions necessary for the implementation of a resource scanning algorithm and power saving mode in the mobile station as well as a dynamic frequency selection algorithm in the base station. This is an prerequisite for the usage of the protocol in a cellular system offering a full integration of mobile ATM terminals into a fixed ATM network.

Further studies are dealing with the analysis and optimization of algorithms for collision resolution and scheduling of ATM cell transmissions.

IX. Acknowledgements

The author would like to thank the members of the MBS project for their suggestions. Furthermore this work would not have been possible without the hard work and enthusiasm of the students working on their diploma thesis.

X. References

[1] R. Händel, M.-N. Huber, and S. Schröder, *ATM Networks: Concepts, Protocols, Applications.* Addison–Wesley, 1994.

[2] L. Fernandes, "Developing a system concept and technologies for mobile broadband communications," *IEEE Personal Communications*, vol. 2, pp. 54–59, Feb. 1995.

[3] D. Petras and A. Hettich, "Performance evaluation of the ASR-ARQ Protocol for wireless ATM," in *IEEE Wireless Communication Systems Symposium*, (Long Island, NY), pp. 71–77, Nov. 1995. available at http://www.comnets.rwth-aachen.de/~petras.

[4] D. Petras, "Performance evaluation of medium access control schemes for mobile broadband systems," in *DMR VI*, (Stockholm, S), pp. 255 – 261, June 1994. available at http://www.comnets.rwth-aachen.de/~petras.

[5] D. Petras, "Medium Access Control Protocol for transparent ATM access in MBS," in *RACE Mobile Telecommunications Summit*, (Cascais, P), pp. 218–224, Nov. 1995. available at http://www.comnets.rwth-aachen.de/~petras.

[6] M. Prögler, "MBS Air Interface Principles," in *RACE Mobile Telecommunications Summit*, (Cascais, P), Nov. 1995.

[7] D. Bertsekas and R. Gallager, *Data Networks.* Englewood Cliffs, NJ: Prentice-Hall, 1987.

[8] C.-H. Rokitanski, A. Guntsch, B. Bjelajac, and C. Wietfeld, "Simulation of mobile communication for performance evaluation of cellular networks, short-range and satellite communications," in *RACE Mobile Workshop*, (Amsterdam, NL), pp. 578 – 584, May 1994.