Designing the reliable transport layer for Satellite Digital Multimedia Broadcasting

M. Chipeta, M. Karaliopoulos, B.G. Evans, B. Garnier, and L. Roullet

***Abstract*—Multimedia Broadcast/Multicast Service (MBMS) is set to enrich the multimedia experience of terrestrial cellular mobile network users, raising at the same time many concerns regarding the required radio resources. The unfolding Satellite Digital Multimedia Broadcast (SDMB) system will play a key complementary role in the delivery of MBMS content. Nonetheless, the integrated nature of the system and the Land Mobile Satellite (LMS) channel pose challenging constraints upon the transport layer. This paper highlights these constraints and assesses proposed reliable transport mechanisms in the SDMB system context. The efficient delivery of each SDMB service may be achieved via a combination of several mechanisms, ranging from packet-level Forward Error Correction (FEC) and interleaving to data carousels and selective retransmissions, and the careful resolution of the tradeoff between transport reliability and system resource consumption.**

***Index Terms*— data carousels, FEC, interleaving, MBMS, reliable multicast, retransmissions, SDMB**

I. Introduction

THE delivery of rich multimedia content to mobile users in a resource-efficient manner has been the main motivation behind the introduction of Multimedia Broadcast/Multicast Service (MBMS) in 2.5G/3G terrestrial cellular mobile networks. MBMS, currently under standardization by the Third Generation Partnership Project (3GPP), promises data rates from 10kbps to 384kbps supporting applications such as file sharing, sports replay, music streaming and traffic information to name but a few [1], [2]. Transmission of MBMS content through 3G terrestrial networks, however, requires significant radio network resources, reducing the available capacity for more traditional services such as voice calls which generate the main revenues for mobile operators.

The Satellite Digital Multimedia Broadcasting (SDMB) concept [3] is a promising alternative for the delivery of MBMS content. The unidirectional satellite system is closely integrated with the terrestrial mobile cellular network; it could be viewed as an overlay distribution network that enhances the delivery of push-type applications, in particular when large, widely spread audiences are involved. Issues like mobility management that prove challenging in the context of mobile terrestrial networks are no longer relevant, while the terrestrial network capacity may be freed up for point-to-point (p-t-p) services.

On the other hand, the SDMB system has to face significant technical challenges, one of them being related to the reliable delivery of content. The hostility of the land mobile satellite (LMS) channel and the power limitations of the satellite are partly compensated by the use of terrestrial gap fillers, which are collocated with terrestrial base stations and retransmit the satellite signal to the mobile users. However, further requirements for the reliable transport layer stem from the hybrid nature of the system in combination with the energy and processing constraints on the handheld terminal side.

We discuss these, rather exceptional, requirements in section III of the paper, after reviewing the system and the requirements of the SDMB system services in section II. We present and assess the trade-offs related to mechanisms that can increase the transport reliability in section IV, identifying their combinations that are appropriate for each service. Finally, the conclusions thus far and the remaining tasks for the engineering of the SDMB reliable transport layer are discussed in section V.

II. SDMB System Architecture and Service Performance Requirements

A. SDMB System Architecture Outline

The SDMB system architecture is depicted in Fig. 1. The system is unidirectional, covering large parts of Europe with multiple geostationary satellite spot beams. The overall system is closely integrated into the architecture of 2.5G/3G mobile cellular networks, in a design that aims to maximize reuse of technology and infrastructure, and minimize system development cost.

M. Chipeta and B.G. Evans are with the Mobile Communications Research Group, Center for Communication Systems Research, University of Surrey, UK (phone: +441483683466; fax: +441483686011; email: {m.chipeta, b.evans}@surrey.ac.uk). M. Karaliopoulos was also with the University of Surrey. He is now with TELETEL SA, Athens, Greece (email: M.Karaliopoulos@teletel.gr).

B. Garnier is with Network & Services Division, Alcatel Space, Toulouse, France (e-mail: benoit.garnier@space.alcatel.fr).

L. Roullet is with UDcast, Sophia Antipolis, France (e-mail: roullet@udcast.com).

This work has been supported by the European Union in the context of the Framework Program 6 Integrated Project MAESTRO (Mobile Applications & sErvices based on Satellite and Terrestrial inteRwOrking). MAESTRO started in January 2004 and its first phase is planned to finish in December 2005. Website: http://www.ist-maestro.dyndns.org

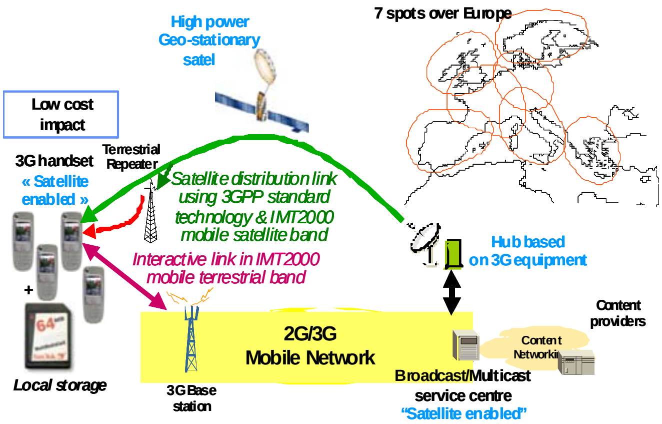

Fig. 1. SDMB system architecture

The SDMB-enabled Broadcast/Multicast Service Center (BMSC), hereafter called SDMB service center, is the standard 3GPP MBMS BMSC [1], enhanced with SDMB-specific functions. The SDMB radio interface is an adaptation of the Wideband Code Division Multiple Access (WCDMA) [4], with the satellite gateway hosting both the Radio Network Controller and the Node B functional entities of the UMTS Radio Access Network. The user equipment (UE) is a standard 3G terminal enriched with SDMB-enabling functions, which, given the unidirectional system nature, are significantly limited.

Terrestrial gap-fillers are installed in the physical locations of terrestrial base stations to enhance signal reception quality in urban, built-up areas and provide indoor coverage. A more detailed description of the SDMB concept and architecture can be found in [3], while [5] provides additional information on main functions of the satellite radio interface.

B. SDMB Service Requirements

Three types of user services are envisaged in the SDMB system: streaming, hot download and cold download services.

1) Streaming: this is a sequence of moving images and sound that are sent in compressed form and played by the UE on the fly. The UE may join the stream any moment.

Streaming is mainly oriented towards real-time services such as live audio/video broadcasts. Therefore, an upper limit of 1 minute is set for the overall delay, from the moment the content is produced at the source and queued at the SDMB service center till its presentation to the user at the terminal. Furthermore, the user should be able to "change channel", namely tune to the reception of different content, within 2 seconds.

The service has therefore both delay and delay variation (jitter) requirements, whereas it is more loss-tolerant in comparison with the hot/cold download services.

2) Hot download: this service supports the delivery of content that, contrary to the streaming paradigm, is stored at the terminal to be accessed later offline. Broadcasting of emergency information, which may be updated regularly by the application, and messages refreshing frequently changing data such as stock values, are examples of applications that could use the hot download service.

The size of the applicable content is small: the overall time spent on the scheduling and queuing of the content at the network nodes and the UE processing plus the content transmission time should not exceed 1 min. Note that at 64kbps, accounting for headers at lower layers, the transmission of 100kBytes requires around 15s.

The content is less loss-tolerant than streaming content. Considering its limited size (e.g., fixed images, sound, text), it may be acceptable to drop it if the UE has not completely or correctly received it after one transmission.

3) Cold download: the service is oriented to type of content that does not the have timeliness constraints of the hot download. Much like the hot download service, the tolerance to information loss depends on the individual item. Cold download supports applications disseminating content, which by nature does not change frequently, such as audio/video clips, images, software.

III. Transport Layer Constraints

Special requirements for the SDMB reliable transport layer stem from the hybrid nature of the system. The key system hypothesis affecting the transport layer is that there is a single reception chain in the UE to receive both the terrestrial network (UMTS/GSM) and the SDMB signals. Depending on the individual UE architecture, simultaneous reception of both the satellite and the terrestrial network signals may be possible. However, the discussion that follows is based on the conservative hypothesis that it is not, which is more realistic for low-end, low-cost handheld devices.

In this case, the terrestrial network signal has preemptive priority over the SDMB signal, namely the SDMB reception is interrupted during the reception of the former. Hence, the SDMB signal reception quality is not only affected by the satellite path propagation impairments, but also by the unavailability of the SDMB signal reception chain at the terminal (equivalent to 100% propagation loss for that duration) whenever terrestrial reception occurs. The terminal switches to terrestrial network reception every time critical network signaling information is transmitted and during calls. In the following, we look more closely into these types of SDMB signal impairments and the requirements they pose on the reliable transport layer. Although the discussion evolves around UMTS, similar arguments apply for the GSM/GPRS network.

A. SDMB data loss due to terrestrial network calls

The main source of call-related SDMB signal interruptions are inbound/outbound calls and short message services (SMS). Data coming from a study over a set of seven European countries (Germany, Spain, Finland, France, Italy, UK, Sweden) suggest that the mean traffic voice per subscriber is 82mins outbound voice calls, 48mins inbound voice calls and 30 SMS messages per month, with an average of 1min48s per voice call [6].

B. SDMB data loss due to critical terrestrial network signaling

1) Idle mode tasks: While the UE is attached to the UMTS network, it is most likely in the idle state, since this state has the longest sleep periods and is the most energy-efficient one. But the "idle" mode is actually not idle at all. The following tasks need to be executed by the terminal in order to use the mobile network services when required:

- Neighbor Cell measurements
- Paging Channel (PCH) and Paging Indication Channel (PICH) reception
- Broadcast Channel (BCH) reception
- Cell (re-) selection
- Location and Routing Area updates

Discontinuous reception (DRx) is used to reduce the UE power consumption related to some of these tasks. The amount of times the user has to wake up to listen to the PCH, PICH and BCH and perform measurements is dependent on the DRx cycle. This cycle is chosen by the terrestrial mobile network and may range from 8 up to 512 radio frames (one WCDMA frame=10 ms); the most common values are between 0.5 and 1 second.

2) Computation of interruption period due to idle mode tasks: Overall, the tasks performed and the time it takes to complete these tasks within an idle mode duty cycle are affected by a number of parameters, which are dependent on the network – given that the DRx value is fixed by the Public Land Mobile Network–, the movement of the UE that dictates how often the cell reselection process is executed and the network topology, which defines the number of neighbor cell measurements. Users attached to different mobile networks could have different duty cycles. The following hypotheses were considered:

- All types of measurements require 7 slots, including set-up and hold time of the reception and transmission circuitry and the acquisition time. The duration of one slot is 0.667ms.
- Every DRx cycle, there are 2 serving cell measurements.
- Every DRx cycle, during the cell reselection procedure, there are 2 intra-frequency measurements and 1 inter-frequency measurement per UMTS neighbor cell.
- Every 1, 1.5 or 2.5 DRx cycle(s), depending on the DRx length, there are 4 GSM measurements per GSM neighbor cell
- The system information has to be decoded (8 radio frames) only when the PICH indicates a change of this information or when the UE changes cell
- Each DRx, the UE will devote 1 radio frame to the monitoring of PICH
- All measurements are considered to be grouped together to form a single interruption of the SDMB signal reception within the DRx cycle.

With these hypotheses in mind, and based on typical user mobility models [7], we can estimate the duration of outages required for terrestrial network signaling and derive the number of lost radio frames and transport blocks at the physical layer (Layer 1).

However, the main driver for transport layer technique selection and the optimization of parameters is the estimate of information loss statistics (percentage of loss and distribution of burst loss) at the UDP packet level. Thus, we consider the impact of the protocol stack layer below the transport layer, the overhead due to the headers and the effect of fragmentation and compression. The loss of Layer 1 (L1) blocks is mapped to UDP datagram loss taking into account the fragmentation of the datagram along the protocol stack.

The first segment of Table I shows the decomposition of a UDP packet as it progress down the protocol stack; the loss of an L1 block results in 1 lost UDP datagram in the best case, and 2 in the worst case for the 800- and 1500-byte datagrams. An L1 block error rate (BLER) of 1% corresponds to a minimum of 3.2% 1500-byte UDP datagram loss or 1.7% 800-byte UDP datagram loss, considering a uniform and independent error spread.

The percentage of lost UDP packets has been computed under different assumptions for the duration of the DRx period, the number of neighbor cells, the UDP packet size and the radio bearer data rates for both the static and mobile cases in [7]; a sample of the results is presented in the second and third segments of Table I.

All in all, the figures obtained suggest that most of the time, the resulting data loss at the transport layer due to the idle mode tasks will be less than 3% when attached to a UMTS network. There will be many cases where the transport layer will have to cope with 10% loss, while loss of 20% should be seen as the target conditions that the transport layer should be capable to handle. Results are typically less stringent if a GPRS system with pure GSM idle mode activity is considered [7].

C. Propagation losses

Transmission outages usually result from the complete blockage of the satellite line-of-sight path, or other events such as lightning, electrical interference from sparking plugs. In addition, significantly poor signal reception conditions cause bit level misinterpretation resulting in perturbations. Normally, such losses range in duration from some microseconds to milliseconds. Due to user mobility, longer outages can occur:

- travelling under a bridge (typically 1 second)
- travelling through a tunnel (typically 1 minute)

The SDMB link is dimensioned for a mean Block Error Rate (BLER) of 1%, but that figure can vary significantly over time depending on user mobility.

D. Duration-based classification of SDMB signal reception interruptions and countermeasures

In summary, the interruptions of SDMB signal reception, which have to be addressed by the reliable transport layer, may be categorized based on their duration and call for different mitigation mechanisms (Fig. 2): we discuss these mechanisms in the section that follows (section IV).

TABLE I. PACKET DECOMPOSITION AND LOSS RATES FOR 384 KBPS DATA RATE WITH DRX = 1.28 SECONDS

Data in UDP datagrams (bytes)		459	800	1500
UDP header (bytes)		8	8	8
IP header w/o PDCP compression (bytes)		20	20	20
PDCP (IP header) compression in (bytes)		3	3	3
Total data at RLC SDU level (bytes)		470	811	1511
Number of MAC blocks per RLC SDU		1.00	1.73	3.21
RLC and MAC header (bytes)		10	10	10
Fragmentation data value (bytes)		470	470	470
Layer 1 block (bytes)		480	480	480
Layer 1 data information in bits per 10ms		3840	3840	3840
Number of UDP packets per second		100	57.95	31.11
UDP packets lost for 1% BLER		1.00%	1.73%	3.21%
[redacted]				
Static case: percentage of UDP packets lost as a function of DRx time & idle mode PICH monitoring				
Serving cell only	1.5% (interruption period, % of DRx)	1.6%	2.7%	2.5%
4 UMTS + 4 GSM	8.1%	8.6%	9.4%	10.0%
8 UMTS + 8 GSM	14.0%	14.1%	14.8%	15.1%
16 UMTS + 16 GSM	25.6%	25.8%	27.0%	27.6
24 UMTS + 24 GSM	37.3%	37.5%	37.7%	37.7%
[redacted]				
Mobile case: percentage of UDP packets lost as a function of DRx time, idle mode PICH monitoring & SI decoding				
Serving cell only	7.8%	7.8%	8.1%	10.0%
4 UMTS + 4 GSM	14.4%	14.8%	16.2%	15.1%
8 UMTS + 8 GSM	20.2%	20.3%	21.6%	22.6%
16 UMTS + 16 GSM	31.9%	32.0%	32.4%	32.7%
24 UMTS + 24 GSM	43.5%	43.8%	44.5%	45.2%

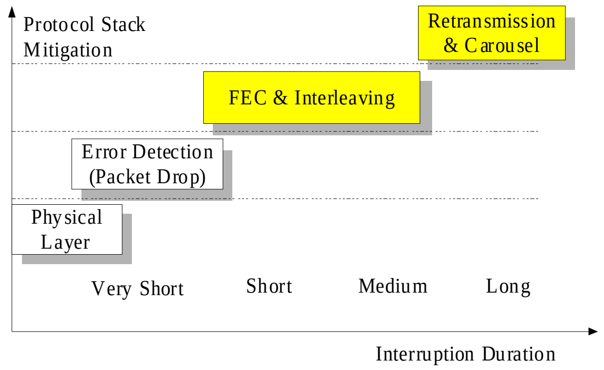

Fig. 2. Applicability of transport layer mitigation to system interruptions

IV. RELIABLE MULTICAST TRANSPORT MECHANISMS AND THEIR APPLICABILITY TO SDMB SERVICES

A. Packet-level FEC

FEC has traditionally been used at the physical layer to combat bit errors, but it has been emerging for some time now at packet level to mitigate packet erasures [8].

The SDMB streaming service shall adopt systematic small [9], [10] block codes, such as Reed-Solomon (RS) codes, since the requirement for quick UE tuning to content channels implies that the FEC block size *n*, derived from *k* original packets and leading to a code rate $R = k/n$, should be small enough to span a time interval that does not exceed the 2 seconds. As a minimum, a small systematic FEC decoder requires any *k* packets out of the *n* transmitted to recover the original *k* packets.

Small codes are also relevant to the SDMB hot download services. For the limited file size of individual contents and UE capabilities resembling those of an Intel Pentium II processor scheduled at 300 MHz with 1 MB RAM [7], the bottleneck on the hot download content's path to the UE is the network transmission capacity rather than the FEC codec.

The alternative to small block codes is large block codes such as the Low Density Generator Matrix (LDGM) codes [10]. A disadvantage of these codes is their *reception overhead* r_o; the decoder requires $(1+r_o)\cdot k$ packets to recover the original *k*. The advantage of large codes is that their codec encoding (decoding) throughput decreases much more slowly with the file size [9], [10] than that of the small codes. Therefore, large codes can support large FEC block sizes, which result in higher file recovery probabilities [11]. The advantage of large codes with respect to the achievable encoding/decoding throughput is partially reduced due to the relatively low data rates supported in the SDMB system (in the order of 384 kbps), which imply that, as already mentioned, the system bottleneck will be the transmission capacity rather than the coder/decoder speed. Nevertheless, significant benefits are expected in the case of low-end handheld devices with limited processing resources.

In SDMB, both small and large codes can be used for cold download services. The choice of code depends on the file size [10]: small codes are more suited to small files as the whole file can be encoded in one FEC block or in a very few blocks whereas large codes exhibit high r_o values; on the other hand, large codes are more suited to large files as r_o is now significantly lower whereas the several FEC blocks required for small codes lead to the coupon collector problem [9] where a decoder can fail to recover certain FEC blocks despite receiving an aggregate number of packets greater than the aggregate number of original packets.

B. Interleaving

The rationale behind interleaving is to randomize burst errors. In general, there are two types of interleavers, block and convolutional (see, for example, [11]).

It must be noted that interleaving does not correct errors. However, when used in streaming services, the user may experience a distinctly higher audio/video quality due to the help of error concealment techniques [12]. A critical parameter is the interleaving depth, which is defined as the number of blocks spanned by packets from a certain block. In general the optimum interleaving depth is a function of the second-order statistics of the packet error process [13].

In SDMB, interleaving is applicable to all three services. In the case of the SDMB streaming service, the interleaving period should be smaller than 2 seconds, in order to satisfy the requirement for quick scanning of content channels by the user. Interleaving may not be applicable at all for low transmission rates. In general, media-specific error concealment techniques at the application layer make up for the reduced protection offered by the transport layer.

The flexibility with respect to the interleaver configuration is higher for the hot download service and even higher for the cold download service. As mentioned earlier the choice of the interleaver depth has to match the packet loss dynamics, while respecting the memory and processing requirements at the receivers.

C. Data carousels

Data carousels are a common way of data scheduling and delivery in a push broadcast system such as the SDMB system. Content is transmitted multiple times, periodically or not, to SDMB receivers, which can complement missing blocks of content items of interest after each data carousel pass. The main user-centric performance metric for data carousels is the response time, namely the time that elapses between the time instant the user expresses his desire to access a certain item until the time the item is retrieved from its schedule and stored in his terminal. The response time is sensitive to the nature of the transmissions (cyclic or non-cyclic), network capacity, number of items on the carousel and packet losses [7], [14], [15].

The combined use of FEC and data carousels introduces a trade-off between the system capacity and the responsiveness of the system [16]. The addition of more parity packets accelerates the download time of the content (assuming that the terminal saves and caches correctly received portions of the content during successive repetitions). However, it reduces the effective system capacity that is available for the content transfer, or, in other words, the content, which can be transferred for given capacity.

In light of the delay requirements of SDMB services (section II.B), data carousels are only considered for the support of the cold download SDMB service.

D. Retransmissions

Retransmissions are another means of loss recovery, which can be activated after data carousels for the SDMB cold download service. Two types of retransmissions are envisaged, as shown in Fig. 3:

1) Satellite selective retransmissions: The content server gathers user feedback on the return link via the terrestrial mobile network regarding which blocks of content are still missing and retransmits them via satellite. The retransmissions are multicast; hence, they scale well with the number of receivers.

2) Terrestrial selective retransmissions: This is the ultimate loss recovery phase, which may be initiated when the satellite selective retransmission phase terminates. Missing content is transmitted via unicast connections over the terrestrial mobile network. This mechanism is the only one that can guarantee full reliability of content delivery but does not scale well and, hence, it is invoked in exceptional cases. The duration of these two loss recovery phases depends on the content nature and its tolerance to loss.

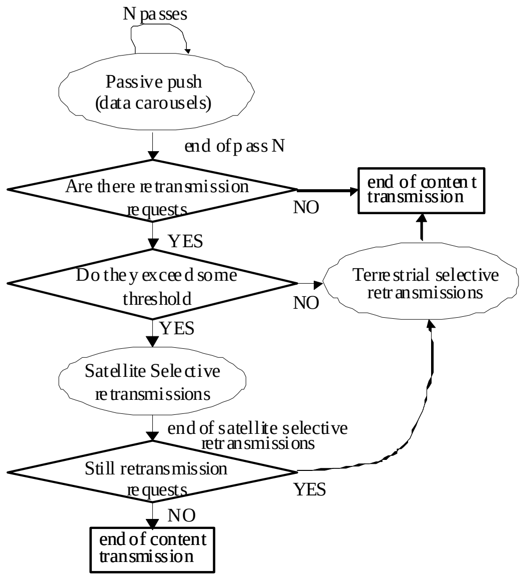

Fig. 3. Transition amongst reliability mechanisms and loss recovery phases for the SDMB cold download services

V. CONCLUSIONS

We have described the reliable transport layer of the SDMB system, a unidirectional satellite system acting as an overlay distribution network over the third generation terrestrial mobile cellular networks for the delivery of push-type point-to-multipoint services. The close integration of the system with the terrestrial networks, in conjunction with energy and computational constraints on the terminal side, generates quite unique signal reception conditions. The combination of several data loss mitigation techniques is deemed mandatory in order to address the system service requirements in the light of the widely different types of SDMB signal interruptions.

The use of small packet-level FEC codes with or without interleaving is envisaged for the support of the loss tolerant, delay-constrained streaming services, where the reduced protection provided by the transport layer will be complemented by application-level, media-specific data loss mitigation techniques. The same combination will serve the more loss-sensitive hot download services, both the length of the FEC scheme and the depth of interleaving being constrained by the requirement for timely delivery of the content to the subscribers. Finally, far more options are available for cold download services, making possible to address SDMB signal interruptions of all time-scales. Data carousels and selective retransmissions, either in unicast or in multicast mode, can significantly increase the achieved reliability in the delivery of this type of content, which is not updated frequently and, hence, does not pose strict requirements on the timing of its delivery.

The trade-off resolution at the level of individual reliable transport mechanisms but also between mechanisms is the next step, which will be pursued by means of both analytical work and extensive simulation campaigns. The first results in this direction are reported in [16]. In the longer term, lab and field trials will evaluate the SDMB reliable transport layer under more realistic test conditions, gaining also better insight to its implications upon the energy and computational resources of the SDMB terminals.

ACKNOWLEDGMENT

We would like to thank the anonymous reviewers for their thoughtful comments and suggestions.

REFERENCES

[1] 3GPP TS 23.246, "Multimedia Broadcast/Multicast Service; Architecture and functional description," version 6.3.0, June 2004.

[2] 3GPP TS 22.246, "Multimedia Broadcast/Multicast Service (MBMS) user services; Stage 1," version 6.1.0, March 2004.

[3] N. Chuberre *et al.*, "Satellite Digital Multimedia Broadcasting for 3G and beyond 3G systems," in *Proc. 13th IST Mobile and Wireless Communications Summit*, June 2004.

[4] H. Holma and A. Toskala, *WCDMA for UMTS: Radio access for third generation mobile communications*, 2nd ed., J. Wiley & Sons Ltd., 2002.

[5] M. Karaliopoulos *et al.*, "Satellite radio interface and radio resource management strategy for the delivery of multicast/broadcast services via an integrated satellite terrestrial system", *IEEE Comm. Magazine*, Sept. 2004.

[6] OMSYC study, "Comparative measurements of concurrent intensity indices of European mobile markets in 2002", February 2004. Available: http://www.art-telekom/publications/etudes/omsyc-comobeurope.pdf

[7] MAESTRO project Deliverable 5.1, "Reliable Transport Layer Technical Requirement," July 2004. Available: http://www.ist-maestro.dyndns.org

[8] C. Huitema, "The case for packet level FEC," in *Proc. 5th IFIP Intl. Workshop on Protocols for High Speed Networks*, Oct. 1996.

[9] J.W. Byers, M. Luby, M. Mitzenmacher, and A. Rege, "A digital fountain approach to reliable distribution of bulk data," in *Proc. ACM SIGCOMM*, Sept. 1998, pp. 56-67.

[10] V. Roca and C. Neumann, "Design, Evaluation and Comparison of Four Large Block FEC Codecs, LDPC, LDGM, LDGM Staircase and LDGM Triangle, plus a Reed-Solomon Small Block FEC Codec," INRIA Research Report No 5225, June 2004. Available: http://www.inrialpes.fr/planete/people/roca/mcl/ldpc_infos.html

[11] S.B. Wicker, *Error Control Systems for Digital Communication and Storage*, Prentice Hall, 1995.

[12] M. Claypool and Y. Zhu, "Using Interleaving to Ameliorate the Effects of Packet Loss in a Video Stream" in *Proc. Intl. Workshop on Multimedia Network Systems and Applications (MNSA)*, May 2003.

[13] K.I. Chang and J.C.I. Chuang, "Required Interleaving Depth in Rayleigh Fading Channels," in *Proc. IEEE GLOBECOM*, November 1996.

[14] S. Acharya, M. Franklin, and S. Zdonik, "Dissemination-based data delivery using broadcast disks," *IEEE Personal Comm. Magazine*, vol. 2, no. 6, Dec. 1995.

[15] N.H. Vaidya and S. Hameed, "Scheduling data broadcast in asymmetric communication environments," *ACM Wireless Networks*, vol.5, no. 3, May 1999, pp. 171-182.

[16] M. Chipeta, M. Karaliopoulos, B.G. Evans and R. Tafazolli, "On the use of packet-level FEC and data carousels for the delivery of broadcast/multicast services to mobile terminals", in *Proc. 61st IEEE VTC May/June 2005*, to be published.