H.264/AVC Video Transmission over MBMS in GERAN

Hrvoje Jenkač, Thomas Stockhammer, and Guenther Liebl
Institute for Communications Engineering (LNT)
Munich University of Technology
Munich, Germany

***Abstract*— In this work we investigate the integration of H.264/AVC based video into the Multimedia Broadcast/Multicast Service (MBMS). MBMS allows simultaneous distribution of live video with reasonable quality to several or even a large number of users within a serving area. We discuss several system design options and propose a simple, but nevertheless efficient cross-layer design concept, which helps to determine an optimal set of both application and transmission parameters for maximizing the quality at the user terminals for different link conditions. In order to validate our concept, we present detailed performance results using realistic test conditions provided by 3GPP.**

I. Introduction

Improvements in the ubiquitous framework of mobile communications are desired and necessary. This includes the addition of new services and applications, as well as enhancements and new concepts in the transport and physical layer. In this work, two of the most promising new additions to 3GPP Release 6 are combined and assessed, namely the transmission of H.264/AVC [1] based video within the Multimedia Broadcast/Multicast Service (MBMS) [2]. This combination enables the simultaneous distribution of live video with reasonable quality to several or even many receivers reusing the existing GERAN (GSM/EDGE Radio Access Network) infrastructure with just small modifications in the radio access network. A possible deployment scenario consists of several 10,000 spectators in a football stadium being able to have access to instantaneous replay of the most important scenes on their cell phone. However, due to the limited resources available on wireless links, a careful design of both the application and the radio network parameters is necessary, and cross-layer aspects are worth to be considered. In this paper we present performance results for AVC/H.264 video transmission over MBMS in GERAN for different radio bearer settings and provide parameter settings for appropriate video encoding.

II. MBMS for GERAN

A. General Framework

Via MBMS, simultaneous distribution of identical multimedia data to multiple receivers within one serving area shall be possible within the existing GERAN architecture. To accomplish this feature, an additional traffic channel is currently standardized by 3GPP: In addition to the already existing PDTCH (Packet Data Traffic CHannel) for point-to-point (p-t-p) transmission, the so-called MDTCH (Multicast Data Traffic CHannel) is planned to support point-to-multipoint (p-t-M) transmission allowing all receiving terminals to listen to the same time slot and frequency simultaneously.

Two different modes of operation are under consideration for MBMS in GERAN: The unacknowledged mode only utilizes forward error correction (FEC), whereas the acknowledged mode in addition to FEC exploits feedback information from multiple receivers. In the following subsections proposals for these two different modes are briefly presented, for mode details the interested reader is referred to [3]–[6].

Figure 1 shows the traditional (E)GPRS protocol stack which served as a starting point in the design of MBMS over GERAN: Incoming IP packets are pre-processed in the SNDCP and LLC layer (resulting in RLC-SDUs), before they are finally prepared for transmission over the air interface. At the RLC layer, the SDUs are segmented into equal-sized RLC segments with length adjusted to the coding scheme in use at the physical layer. Each segment is then mapped on the data part of an RLC/MAC block and forwarded to the physical layer. Here, a CRC is appended and channel encoding with one of the pre-defined coding schemes (CS-1. . .4 for GPRS and MCS-1. . .9 for EGPRS) is performed. The encoded block is then interleaved over four radio bursts, which are mapped on one out of eight time slots, spread over four successive TDMA frames. At the receiver channel decoding and error detection is performed on each of the received segments and any corrupt ones are discarded. Since the corresponding RLC-SDU cannot be reconstructed anymore, it is common practice that the loss of a single RLC segment results in the loss of the entire IP-packet. With RLC segments in the range of $20 - 50$ bytes, common loss rates are about 10%. Without the possibility of individual retransmissions a major QoS requirement of MBMS, namely to keep the IP packet loss rate below 10^{-2} for packet lengths up to 500 bytes, cannot be fulfilled with the plain GPRS mode. Therefore, different proposed modifications of the existing system are discussed in the following.

B. Proposed MBMS Unacknowledged Mode

Adaptation of the coding schemes and the transmission power to varying channel conditions like in p-t-p communication is not feasible in a p-t-M scenario, as all terminals receive the same data from a single source. Therefore, coding scheme and transmit power are expected to be chosen in advance to satisfy a certain target user, which is for example specified by the worst or average class of users in the serving area. Obviously the required QoS for MBMS cannot be guaranteed to users with worse receiving conditions, though they still could be supported by p-t-p transmission with appropriate adaptation. For this reason, new coding and transmission schemes have been investigated and proposed for MBMS over GERAN (and also UTRAN) to provide satisfying and reliable services for most users within the serving area [3], [5], and to allow adaptivity to the user topology in the serving area. Different system concepts are still under consideration within 3GPP, but one of the most promising and efficient approach for an unacknowledged MDTCH operation mode relies on the introduction of outer Reed-Solomon (RS) coding at the RLC layer [4]. It is proposed that from k RLC data segments $(n-k)$ redundancy segments are generated by applying an (n,k) Reed-Solomon code byte-wise across segments. Both data and redundancy segments are then mapped on RLC/MAC blocks, which itself are encoded and transmitted in the usual manner. The loss of parts of the segments due to fading effects can thus be compensated for by applying simple erasure decoding of RS coded data. Due to the maximum-distance separability of RS codes, this outer coding achieves excellent performance, requires moderate complexity, provides flexibility as well as constant and moderate delay below 5 seconds, and allows the reuse of existing infrastructure and hardware components. For a detailed description and performance results of the MBMS system with outer RS coding we refer to [3] and [5].

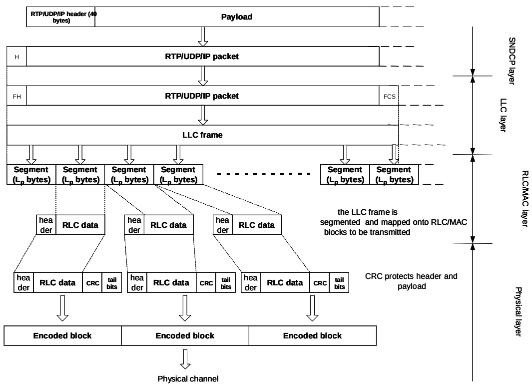

Fig. 1. Traditional GPRS protocol stack.

C. Proposed MBMS Acknowledged Mode

In various scenarios, like video streaming or file download, additional latency or data rate jitter on the MDTCH can be tolerated at least to some extent by applying appropriate receiver buffering. Thus, reliable transmission using an acknowledged mode is theoretically possible, if appropriate means are provided for receivers in broadcast or multicast scenarios to inform the transmitter about lost information. Within the GERAN standardization the introduction of the Common Feedback Channel (CFCH) has been considered [7], which allows that all receiving terminals inform the base station about lost RLC segments on one common frequency within the same time slot by sending only negative acknowledgments. For a detailed description of the CFCH we refer to [7]. Based on the syntax of the CFCH several proposals for retransmission strategies have been already discussed in, *eg* [6], [8], and performance comparison especially for a large number of receivers in the serving area has been provided. A combination of outer RS coding and retransmission strategy has been introduced in [9] and shows excellent performance. Thereby, based on outer RS coding, equivalent to the system approach for the unacknowledged mode described above, further redundancy packets are broadcasted to the receivers only if a common negative acknowledgment has been detected on the CFCH. The protocol is similar to well known incremental redundancy strategies proposed for p-t-p communication. Initial performance results [9] show that the throughput within the serving area only slightly depends on the number of receivers. Consequently, several hundreds of terminals can be served virtually error free, at the expense of only slightly increased delay. However, it should be mentioned that in contrast to the acknowledged mode used in p-t-p transmission, where retransmissions are performed fully-persistent, the MBMS acknowledged mode may fail to recover the segments once the redundancy in the RS code is fully exploited.

III. Packet Transmission of H.264/AVC Video

The H.264/AVC video codec represents one of the most advanced technologies in digital video: Besides offering good rate-distortion performance especially at low bit rates, it has been the first codec that has been designed with a particular focus on network and transmission issues. In addition, error resilience features especially for packet-based transmission have been introduced. This inherent flexibility, together with the excellent rate–distortion performance, encouraged us to assess H.264/AVC coded video within the MBMS framework.

The Network Abstraction Layer (NAL) provides a common interface to adapt the output of the media codec to a format suitable for transmission over different types of networks, including bandwidth constrained and error-prone wireless links [10]. NAL units can be directly encapsulated into RTP packets according to the respective RTP payload specification [11]. Therefore, the length of the NAL unit directly influences the length of the IP packet, which is a crucial factor in a wireless scenario: The longer the IP packets, the higher the probability that at least one segment is lost, and thus that the whole packet is lost. Hence, shorter IP packets are definitely beneficial in lossy wireless environments. In H.264, this can be achieved by using slice structured coding such that each primary NAL unit generated by the video codec contains a single slice representing a sequence of macroblocks. Slices within one video frame are independently coded and therefore provide spatially distinct resynchronization points within the video data. Advanced error concealment can be applied in this case [10]. The number of macroblocks encoded within one frame is arbitrary and therefore, this mode allows to generate NAL units of similar length. In this way, drastic changes in the IP packet size due to the traditional output sequence of, for example, large intra- and small predictive-coded frames are avoided.

Despite the adaptation of the packet length in some of our considered broadcast scenarios IP packet loss cannot be completely avoided. Then the application experiences spatio-temporal error propagation in the video reconstruction process. To some extent this disturbing effect can be compensated for by the usage of completely intra-coded frames, which are nevertheless to be used sparsely for random access, but not too frequently in order not to harm the overall compression efficiency. For this reason, we decided to use the possibility in H.264/AVC to encode only a single macroblock in intra mode, which allows early recovery in case of losses. The placement and frequency of these intra macroblocks is best done in a rate-distortion optimized way where the distortion should reflect an expectation of the decoder distortion including the channel characteristics [10], [12]. Hence, we propose not only to adjust the packet length for a given scenario, but also the macroblock intra update ratio in case of infrequent, but still present packet loss. The encoder is configured such that for error resilience and network adaptation purpose the maximum NAL unit length in bytes, N_{NALU}, and the macroblock intra update ratio specified by the percentage of expected NAL unit losses, P_{NALU}, can be selected [10].

By applying this simple, but nevertheless effective type of cross-layer design between application and transmission parameters, we achieve good performance at each point of operation in the system even if the underlying network is not optimized and lossy. Tradeoffs between residual error rates, throughput, video coding parameters, availability of cross-layer information, as well as system and implementation complexity based on some experimental results will be discussed in the following.

IV. Experimental Setup and Simulation Results

In order to evaluate the performance of H.264/AVC video within the MBMS/GERAN framework, we performed various simulations. For all of them we assumed that MBMS exploits six of the eight available GSM time slots resulting in a maximum throughput of 48 kbit/s for coding scheme CS-1 without any outer RS coding. The transmitted video is the baseline H.264/AVC encoded QCIF sequence foreman at 7.5 fps. Intra frames are inserted every 10 s. Flexibility in the video encoding is provided by allowing to adapt the bit rate R_v including packetization overhead for NAL and IP headers, the maximum NAL unit length in bytes, N_{NALU} and the macroblock intra update ratio specified by P_{NALU}. The detailed parameter settings will be discussed below for each experiment. Moreover, we assumed the mobiles to move in a typical urban environment with 3 km/h.

In a first set of experiments the system parameters (GPRS coding scheme, RS code parameters (n, k)) are chosen such that users operating at C/I=7.5 dB are supported with IP packet loss rates of at most 10^{-2} for packet lengths equal to 500 bytes. Users with lower C/I are not considered to be supported with satisfying QoS from a cell planing point of view. For all simulated combinations coding scheme CS-1 has been selected as it is the only one which yields satisfying RLC segment loss rates at C/I=7.5 dB. We compare three different transmission strategies: MBMS without any modifications (*ie* using the traditional GPRS stack), and the two proposed enhancements MBMS unacknowledged mode and MBMS acknowledged mode. For traditional GPRS, an optimized parameter set of N_{NALU}=50 bytes and P_{NALU}=40% for an RTP/UDP/IP header size of H=40 bytes and P_{NALU}=25% for H=5 bytes has been selected. The total video bit rate R_v including packetization overhead matches the maximum throughput of 48 kbit/s. The unacknowledged mode is investigated with the two RS code parameter settings (n, k)=(64,48) and (n, k)=(128,101), both resulting in IP packet loss rates lower than 10^{-2} for packet lengths equal to 500 bytes. For a detailed description on how RS code parameters can be optimized to achieve a desired IP packet loss rate, we refer to [3]. To be consistent with the requirements, video is encoded with N_{NALU}=500 bytes and P_{NALU}=1% and the bit rate R_v is adapted to the supported bit rate of the link. In case of the acknowledged mode, we have also selected and evaluated two different RS mother codes (n, k)=(255,64) and (n, k)=(255,128). At this point we want to stress that the effective code length n' depends on the number of negative feedbacks from the receivers, since the number of redundancy symbols is not fixed, but varies on demand [6]. No specific error resilience is applied in this case, *ie* the length of the NAL units corresponds to the size of a video frame and P_{NALU}=0%.

Fig. 2a shows the simulation results considering an RTP/IP/UDP overhead of H=40 bytes. We measure the achieved video quality at the receiver in terms of average Peak-Signal-to-Noise Ratio (PSNR). Although the maximum throughput of MBMS with traditional GPRS is high (48 kbit/s for CS-1), many of the received RLC/MAC blocks are erroneous (11% for C/I=7.5 dB). This, together with the overhead required for error resilience, results in poor video quality at 7.5 dB. However, even for receivers with higher C/I the PSNR is not satisfying as the performance of the video is restricted by the overhead for error resilience adapted to the transmission conditions at 7.5 dB. If outer RS coding is applied with (n, k)=(64,48), at C/I=7.5 dB a performance gain of 9 dB in terms of PSNR is achieved. This is due to the RS code reducing the RLC/MAC block loss rate from 11% to 0.03%. Although in the RS case the video was encoded at a target bit rate of only 36 kbit/s, the number of received video packets is much higher in this case and less overhead for error resilience is introduced in the video bitstream. Since the RS code is already optimized to provide low error rates at C/I=7.5 dB, only a small gain in quality is achieved for users with higher C/I in the serving area. Moreover, the PSNR is limited due to the fixed low bit rate encoding. By selecting a stronger RS code (n, k)=(128,101), a slight performance gain can be achieved. However, the processing delay will also slightly increase [3].

In [6] it was shown that the system throughput of the MBMS acknowledged mode depends on the selected RS mother code and only slightly on the number of served receivers. Therefore, we investigated the performance of the acknowledged mode for different mother codes (255,12) and (255,128) and a different allowed number of receivers U=10 and U=50. Generally, compared to the unacknowledged mode slight performance gains in the range of 0.5 dB to 1 dB can be achieved due to the reliable service. Similarly, the PSNR

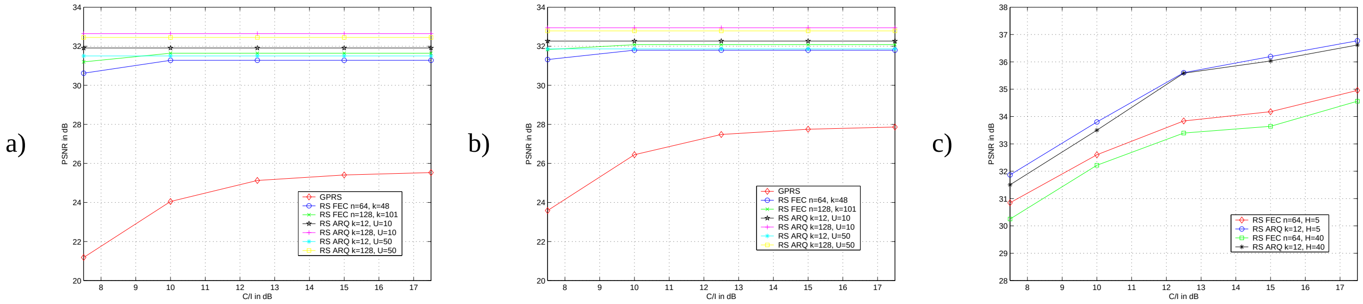

Fig. 2. a),b) Results for fixed $C/I = 7.5$ dB optimization and $H = 40$ bytes and $H = 5$ bytes, respectively. c) Results for variable C/I optimization point.

is limited due to the fixed low rate video encoding. Again, by applying a stronger code, the performance results improve. Moreover, with increasing allowed number of receivers a slight performance degradation is observed, whereas when applying the unacknowledged mode the performance does not depend on the number of active receivers. The same experiments have also been performed considering header compression, where RTP/UDP/IP headers are assumed to be compressed resulting in an average overhead of H=5 bytes. Fig. 2b shows the simulation results: Since in this case more bit rate is available for the video, we get a slight performance gain for both unacknowledged and acknowledged mode, while the performance of MBMS with traditional GPRS is increased by several dBs, but is still not sufficient.

In the investigations discussed before, the system and video parameters have been adapted to worst case users operating at C/I=7.5 dB. However, in some cells a higher worst case C/I value might be possible. If a fixed scheme is deployed, users with better C/I can not yield significant quality gains. However, RS coding provides the flexibility to adjust the transmission parameters [3]. Hence, in a second set of experiments we investigated the video performance at different operation points (target C/I). For MBMS unacknowledged mode the video is encoded with N_{NALU}=500 bytes and P_{NALU}=1%, and bit rate R_v adapted to the operation point. In case of the acknowledged mode we applied no specific error resilience. Fig. 2c shows the resulting quality: As already observed before, again the acknowledged mode provides higher quality due to the reliable service. For increasing target C/I the gap between both modes mode is even several dBs. Moreover, due to the adaptation the received quality is increasing with higher target C/I. As expected, header compression yields further gain, but only in the range of less than 0.5 dB.

V. Conclusions

In this work we investigated the integration of H.264/AVC based video into the Multimedia Broadcast/Multicast Service (MBMS) in order to enable simultaneous distribution of live video to a large number of users within a serving area. We compared several transmission strategies which are under consideration for MBMS in GERAN. We showed that the traditional GPRS system without extensions and enhancements is not suitable for video delivery in broadcast scenarios even if the video is adapted to the harsh transmission conditions. Enormous gains in terms of video quality are obtained by introducing additional outer RS coding at the RLC layer in combination with channel aware error resilient video coding. The video has been adapted to the target operation point of 1% error rate of packets with length 500 bytes. By additionally exploiting feedback from multiple receivers, further improvements in the resulting PSNR can be achieved. However, the gains compared to the unacknowledged mode are less impressive (in the range of 0.5 dB - 1 dB). Considering the increase of the overall system complexity with the use of a feedback channel, the use of feedback is not recommended for live video in this broadcast scenario. The situation might change for reliable data distribution, *eg*file download, where reliability is much more important than throughput and delay jitter. Furthermore, we showed that by appropriate adaptation to the worst expected user in the cell, additional performance gains are possible. Finally, header compression yields some gain in all investigated cases, but is only essential for the plain GPRS transmission.

References

[1] ITU–T and ISO/IEC JTC 1, *Advanced Video Coding for Generic Audiovisual Services*, ITU-T Recommendation H.264 - ISO/IEC 14496-10(AVC), 2003.

[2] 3GPP TR 25.992-140, *Multimedia Broadcast/Multicast Service (MBMS); UTRAN/GERAN requirements*, ETSI, 2003.

[3] H. Jenkac, G. Liebl, T. Stockhammer, and W. Xu, "Flexible Outer Reed-Solomon Coding on RLC Layer for MBMS over GERAN," in *Proc. VTC Spring 2004*, Milan, Italy, May 2004.

[4] TSG GERAN Tdoc GP-032100, *Outer Reed-Solomon Coding on RLC Layer for MBMS over GERAN*, Siemens, TSG GERAN # 16, New York (USA), Aug. 2003.

[5] TSG GERAN Tdoc GP-032614, *Outer coding in the RLC for MBMS*, Siemens, TSG GERAN # 17, Budapest (Hungary), Nov. 2003.

[6] TSG GERAN Tdoc GP-032647, *Retransmission Strategies for MBMS*, Siemens, TSG GERAN # 17, Budapest (Hungary), Nov. 2003.

[7] TSG GERAN Tdoc GP-0400301, *Common Feedback Channel for MBMS delivery*, Siemens, TSG GERAN # 18, Reykjavik (Iceland), Feb. 2004.

[8] J. Metzner, "An improved broadcast retransmission protocol," *IEEE Trans. Commun.*, vol. COM-32, pp. 679–683, June 1984.

[9] TSG GERAN Tdoc GP-0400356, *Performance of ptm retransmission strategies*, Siemens, TSG GERAN # 18, Reykjavik (Iceland), Feb. 2004.

[10] T. Stockhammer, M. Hannuksela, and T. Wiegand, "H.264/AVC in Wireless Environments," *IEEE Trans. on Circuits and Systems*, vol. 13, no. 7, pp. 657–673, July 2003.

[11] S. Wenger, T. Stockhammer, M. Hannuksela, M. Westerlund, and D. Singer, "RTP payload format for H.264 video," Internet Engineering Task Force (IETF)," draft-ietf-avt-rtp-h264-04.txt, Feb. 2004.

[12] R. Zhang, S. Regunthan, and K. Rose, "Video Coding with Optimal Inter/Intra-Mode Switching for Packet Loss Resilience," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 6, pp. 966–976, June 2000.