Transmission of Streaming Data over an EGPRS Wireless Network

Kapil Chawla, Peter F. Driessen and Xiaoxin Qiu

AT&T Laboratories - Research

100 Schulz Drive, Red Bank, NJ 07701, USA

Email: {kapil,pfd,xqiu}@research.att.com

Abstract— **We investigate the feasibility of transmitting MPEG-2 AAC (Advanced Audio Coder) music over packet-switched bearers of an Enhanced General Packet Radio Services cellular network. Streaming data applications like AAC music differ from conversational voice in that they permit a larger transmission delay and delay jitter than the former. In addition, AAC music can tolerate a higher frame error rate than voice by using error concealment techniques at the receiver. This creates the opportunity to transmit AAC music over data-optimized, high-throughput and nominally high-error rate, wireless networks, by using appropriate performance enhancing techniques. In particular, we investigate the performance with the following techniques: (i) Link-layer retransmissions; (ii) Dynamic packet assignment; and (iii) Packet interleaving or shuffling. We find that all the above techniques are desirable in order to simulateneously achieve a high spectral efficiency and the required quality. While this paper is focused on one application, the conclusions and insights are applicable to most streaming audio and video applications.**

I. INTRODUCTION

With the emphasis on packet switched wireless data technology in proposed third generation (3G) wireless standards, it is important to understand the techniques that would enable offering the variety of applications envisioned for these networks. We investigate the feasibility of embedding *streaming* data applications over a wireless data network such as Enhanced General Packet Radio Services (EGPRS) [1], [2]. Examples of such applications include the transmission of encoded music, compressed video, and audio programming. Unlike interactive services (e.g. voice), which have very stringent delay requirements, streaming data applications can usually tolerate some amount of startup delay. Therefore such applications can benefit from link-level techniques, e.g. Automatic Repeat reQuest (ARQ), to achieve satisfactory performance while maintaining a high spectrum efficiency. In addition, streaming audio and video data, unlike file transfer, can also tolerate some packet loss, by using some error concealment techniques at the receiver to minimize the perceived quality degradation.

We focus on the transmission of streaming MPEG-2 Advanced Audio Coder (AAC) coded music over EGPRS [3], [4], [5]. MPEG-2 AAC offers the highest quality audio at data rates in the 32-128 Kbps range for two channel stereo. For a given perceived audio quality, AAC needs only one-half the bit rate of the popular MP3 (MPEG-1 Layer 3) audio coder. EGPRS is one of the proposals submitted to the IMT-2000 initiative of the International Telecommunications Union (ITU), for the 3G wireless services, which leads to a convergence of GSM and IS-136 (US-TDMA) based systems.

For acceptable Quality of Service (QoS), streaming mu-sic requires (i) a low average frame error rate ($\approx$ 5-10%); and (ii) frame error patterns that are "randomized", i.e. not bursty. Therefore, the challenge in transmitting streaming music over EGPRS is both the high error rates and the bursty error patterns that may be encountered. While the bursty error patterns are a characteristic of the wireless propagation channel and interference environment, the high error rates can result from the chosen "operating point" for EGPRS. This is the case in an EGPRS system that is optimized for the transmission of delay-tolerant data services, i.e. to maximize system throughput. Such a system would tend to have very tight frequency reuse, with consequent higher interference levels, and high (e.g. 10-50%) over-the-air Block Error Rates (BLER). Note that EGPRS does provide for circuit switched bearers, e.g. Enhanced Circuit-Switched Data (ECSD) [6], which are expected to operate at lower (e.g. 2%) error rates. However, for reasons of spectrum efficiency and system flexibility, it is still desirable to provide multimedia services, with streaming music as one such example, over wireless-data-optimized EGPRS packet-switched bearers.

In this paper, we investigate the performance enhancements that can be provided by the following techniques: (i) *Link layer retransmissions*, which help reduce the AAC frame error rate much below the raw EGPRS block error rate; (ii) *Dynamic packet assignment (DPA)*, which increases system capacity; and (iii) *Packet shuffling*, which is a new technique that we propose to randomize the AAC frame errors. We demonstrate through simulation that these techniques are effective, both in reducing the AAC frame error rate and in reducing the burstiness of the frame errors. We conclude that these techniques, along with the error concealment, are very useful for embedding streaming data in wireless networks optimized for best-effort data.

The rest of the paper is organized as follows. In Section II we describe the system under consideration and some relevant aspects of AAC music. In Section III, we discuss the various techniques that are modeled. In Section IV, we discuss the resulting system performance. In Section V, we offer some concluding remarks and directions for future work.

II. WIRELESS NETWORK AND AAC MUSIC

A. System Architecture

Figure 1 shows a simplified architecture of the system under consideration. The figure shows AAC music servers, connected to a data network, e.g. the Internet, and mobile clients connected to an EGPRS network. For our current

purpose, the EGPRS network can be represented by a set of EGPRS serving nodes (E- GSNs)[1], which can be thought of as mobility aware routers, and base station systems (BSSs) connected to these serving nodes. Mobile clients connect wirelessly to the strongest BSS, through this BSS to the serving node, and onwards to the data network.

We assume that a mobile client requests a streaming music application from a server. We expect that a protocol like UDP/IP is used for the transport. We assume that the client can specify or choose a particular data rate at the beginning of the session.

B. EGPRS Network

EGPRS uses a TDMA-based, packet-switched radio technology with 200 KHz channels for the air interface, and an IP based packet switched core network. The air-interface, which is our focus, is also called EDGE (Enhanced Data rates for GSM Evolution). For transferring data between the IP serving node and the mobile client, EGPRS uses a set of protocols, which in effect, provide a secure logical pipe between the node and the client. EGPRS offers two types of bearers: packet-switched bearers, which are targeted for "connection-less" data transfers and circuit switched bearers, which are targeted for "connection-oriented" data transfers. The two types of bearers have different physical and link layer structures, and cannot be easily combined on the same channel. For smooth integration of the music and data services while achieving a high data throughput, we consider the transmission of music over the packet-switched bearers.

To improve throughput, EGPRS employs link adaptation, which adapts the information bit rate for each link, depending upon the link quality requirement, the link interference and noise conditions. We assume that for each link, this adaptation can occur once every 100 ms. The variable rate transmission is achieved by changing the modulation and coding format. As a result, the instantaneous information bit rate of a link can vary from around 11 kbps to 65 kbps, and different EGPRS radio blocks can contain a different number of information bits. In most cases, one AAC frame cannot fit into one radio block. Therefore, AAC frame segmentation at the base and frame re-assembly at the mobile is necessary.

C. Parameters of MPEG-2 Advanced Audio Coder

The basic idea behind perceptual audio coders such as MPEG-2 Advanced Audio Coder (AAC) is to hide quantization noise below the signal-dependent masking thresholds of the human auditory system [3], [4]. With AAC, the audio stream is divided into subbands using a filter bank, which uses a 1024 point Modified Discrete Cosine Transform (MDCT). Quantization noise is set separately in each subband to fall below the masking threshold. In addition, correlation between audio samples is used to remove redundancies.

The raw data produced by the AAC encoder is designed to be parsable so that it can be used with any data transport mechanism, including digital cellular air interfaces. The perceived audio quality is a function of this AAC raw data rate, which depends on the audio sampling rate and the compression rate, Typical AAC parameters are summarized in Table I, assuming a compression ratio of 16:1, i.e. from 16 to 1 bit per sample and two audio channels for stereo. The AAC frame lengths for each channel is variable from frame to frame, depending on the music source, with an average of 1024 bits per frame (1024 samples per frame times an average of 1 bit per sample).

TABLE I
AAC DATA RATE AND MUSIC QUALITY

Stereo data rate (Kbps)	Audio sample rate (KHz)	Highest audio freq. (KHz)	Quality	AAC frame time (msec)
96	48	24	CD	21.3
64	32	16	FM	32
48	24	12		42.7
32	16	8		64
24	12	6	AM	85.3

The AAC frames that use the 1024 point MDCT (so-called long frames) are used when the music is stationary over the frame transmission time. So-called short frames are used when the audio changes rapidly within the frame transmission time, in which 8 consecutive 128 point MDCTs are used, thus sacrificing frequency resolution to capture the time dynamics. These short frames cannot be predicted from surrounding long frames, and thus can be designated high priority AAC frames.

The nth consecutive AAC frame in a given channel has a limited useful lifetime, comprising the startup (playout) delay plus n times the AAC frame time. If it is delayed beyond its lifetime, the AAC frame has to be dropped. For a given data rate, the audio quality is determined by the *distribution* of the late or dropped frames, i.e. gaps in the AAC block sequence, at the receiver.

These gaps may be concealed using error mitigation techniques in which a prediction or interpolation algorithm is used to estimate the missing frame and thus "fill the gap". The details of the error mitigation algorithms will be presented elsewhere. The performance of the error mitigation schemes is best when only one or just a few consecutive AAC frames are missing. For most types of music it may be difficult to mitigate a gap of more than 3 AAC frames (192 msec at 32 Kbps) and still maintain audio quality. MPEG-4 AAC includes additional error robustness tools, which allow partially damaged AAC blocks to be reconstructed, as well as a more finely scalable data rate.

---

[1] Two types of serving nodes are defined: E-SGSN and E-GGSN.

III. DESCRIPTION OF TECHNIQUES

A. Link-Layer Retransmission

EGPRS provides an option of using link level Automatic Repeat reQuest (ARQ) to create a reliable pipe for the transfer over the radio-link. We assume that the network can choose a maximum number of retransmissions for each segment (radio block), and this number is set to 3 for most of the simulations.

B. Dynamic Packet Assignment

EGPRS facilitates dynamic packet assignment (DPA), in order to efficiently support bursty data traffic. A link can be reassigned different channels[2] based upon measurements of channel quality at the receiver. We assume that each mobile is assigned to the least-interfered channel at the beginning of a session. The same channel continues to be used unless M ($M = 2$ for most of the simulations) consecutive radio block errors occur, in which case the user is re-assigned to another channel. For our application, the primary benefit of DPA is to increase the system capacity and therefore also the "sustainable" link data rate for the music users. Therefore, a system with DPA can provide better coverage, i.e. service to a larger fraction of users, compared to a system without DPA.

C. Packet Shuffling

In order to increase the effectiveness of any error concealment techniques, the AAC frames may be shuffled or interleaved at the server or at an intermediate node, and then transmitted over the wireless channel. The re-assembled AAC frames are then re-ordered by the receiver and placed in the playout buffer. This shuffling and un-shuffling effectively spreads out bursts of AAC frame errors over the radio-link. Note, however, that as long as frame numbers (or a similar index) is included in the AAC frame, it is not necessary for the mobile to even be aware of the shuffling or interleaving. This is very desirable, as it implies that this feature can be activated without changing or re-programming the mobile.

The specific shuffling technique proposed and simulated is based upon a convolutional interleaver [7], [8]. Two integer parameters, N and B, specify this interleaver. The values that we use are $N = 6$ and $B = 1$. If $T(k)$ denotes the position at which frame k is transmitted, then we have:

$$T(k) = k + (k \bmod N)NB .\qquad(1)$$

Table II illustrates this interleaver. It follows that the or-

TABLE II
Convolutional interleaving for $N = 6$, $B = 1$.

k	0	1	2	3	4	5	6	7
$T(k)$	0	7	14	21	28	35	6	13

der of transmission of the frames is 0, 6, 1, 12, 7, 2, 18, 13, 8, 3, 24, 19, 14, 9, 4, 30, 25, 20, 15, 10, 5, ... It is easy to show that using this technique, two consecutive frames are separated by at least NB frames after the shuffling[3]. This is why this technique is very effective in shortening gaps in the received AAC frame sequence that result from a burst of over-the-air transmission errors. On the other hand, this shuffling introduces a delay in the transmission of some frames, and therefore comes at a price of reducing the effective playout buffer. Therefore, the choice of interleaver should provide a suitable trade-off between the frame error spreading and the reduced buffer length.

IV. PERFORMANCE RESULTS

A. Simulation Model

We simulate a standard macrocellular system with three-sectored base stations on a hexagonal grid. A nominal propagation model is used: path loss exponent equal to 3.5 and Lognormal shadow fading with dB standard deviation of 6 dB. Multipath fading is not modeled explicitly in the system level study, but is modeled in the link level results that are used. At the physical layer, a GSM-like structure is assumed. Time is divided into 20 ms logical frames, which consist of eight blocks. A block is the basic transmission unit of EDGE (EGPRS). An interference limited environment is considered: the median SNR at the cell boundary is set to 30 dB, and a 1/3 frequency reuse is assumed. The spectrum allocated is 600 kHz, so that each sector is assigned one 200 kHz carrier (eight channels). We focus on the downlink performance, as this is expected to be the limiting link due to traffic asymmetry.

A combined snap-shot and time-driven simulation is implemented. The system consists of background data users and music clients. Data users are uniformly distributed in the system. To represent different loading scenarios, the user density is set to 12 or 24 data users per sector. No admission control policy is used for data. Data traffic is modeled using a simple on-off model: for each user, a message arrives once every 3.3 seconds on average, and has an average message length of 1.6 kB (12.8 kbits). Both the interarrival time and the message length follow exponential distributions. Therefore, the average data rate of each data user is 3.9 kb/s. We assume that this data traffic is "best-effort data", i.e. it can tolerate transmission delay and does not have an explicit data rate requirement. Upon arrival, the message is stored in an infinite queue and segmented appropriately for transmission. As we are mainly concerned with the music users, we do not present any performance results for the data users. There are several suitable references (e.g. [9]) that discuss the impact of various techniques, e.g. DPA, on data throughput and message delay performance for EGPRS-like systems.

A number of music clients (users) are placed in the system, assuming a uniform distribution with a user density

---

[2] A channel is usually specified as a carrier frequency and a specific time-slot.

[3] This is not true for a few initial frames, i.e. when the interleaver begins operation. However, this should have a minimal impact on performance for reasonable sized files.

of one per sector. The music sequence is around 60 s long, and a playout delay of 5 s is permitted. In order not to suddenly load the system, the start time for each music file is generated randomly. For most of the simulations, we assume that the music is transmitted at 32 kb/s. For the simulation with variable rates, the system has a choice of two rates: 32 kb/s and 16 kb/s. In this case, for each client, an appropriate rate is chosen at the start of the session, and cannot then be changed in the session. As discussed earlier, each AAC frame has a useful lifetime, and a frame is lost if any segment either arrives too late, or is received in error. The perceived music quality is characterized by (i) the AAC frame error rate (FER); and (ii) the AAC frame error run lengths (number of consecutive AAC frames in an error burst). The latter performance measure is important since it determines the effectiveness of error concealment techniques.

B. Numerical Results

Figure 2 shows the impact of link level ARQ on the AAC frame error performance for the case of 12 data users (on average) per sector and random packet assignment (RPA). The figure plots the cumulative distribution function (CDF) of the AAC FER over the set of music users for (i) no ARQ; (ii) ARQ with 2 re-transmissions; and (iii) ARQ with 9 re-transmissions. Assuming an acceptable FER of 0.1, the figure shows that without ARQ, only around 17% of the users can be supported, whereas with ARQ, this fraction increases to around 75%. As long as ARQ is used, the actual number of retransmissions used does not appear to have a large impact on the performance. For the rest of the simulation, we assume ARQ with 2 retransmissions. The figure also shows that with ARQ, the system tends to separate the users into two categories: users with excellent FER and users with unacceptable FER. To investigate this further, we consider the distribution of the EGPRS throughput.

Figure 3 plots the CDF of the EGPRS throughput over the set of music users for (i) RPA and (ii) DPA. In order to transmit 60 sec of 32 kbps streaming music in 65 s (60 s + startup delay) with FER better than 0.1, the EGPRS throughput should be at least $(32 * 0.9 * 60/65) \approx 27$ kbps. The figure shows that around 25% of the users have a throughput lower than 27 kbps. This agrees with the conclusion of Figure 2, where around 25% of the users had an AAC FER higher than 0.1. Using the same logic, the possible gain using DPA is also clear from the correponding CDF in Figure 3. In fact, it follows that with DPA, more than 90% of the users should have an AAC FER better than 0.1. We have verified that this is the case through simulations.

Figures 2 and 3 also demonstrate the benefits of a system that can successfully adapt the streaming data rate on a per session basis. In particular, it follows from Figure 3 that with DPA, around 50% of the users can be transmitted streaming music at 48 kbps, and that close to complete coverage can be achieved if the rate can be set to 16 kbps for some of the users. This highlights the need

for the system to be able to predict the performance of the user, and to provide an appropriate level of service. This problem is even more challenging if the users are mobile, or if longer session durations are considered. In both these cases, the user environment can change substantially over a session, so that predicting an appropriate service level may need to be very conservative.

Figure 4 shows the benefit of the packet shuffling. The figure plots the histogram of the AAC frame error run lengths for the 90% of the users with AAC FER below 0.1, with and without the shuffling. DPA and ARQ with 2 retransmissions in assumed in both cases, and an interleaver with $N = 6, B = 1$ is used for the shuffling. The figure shows that the error patterns (gaps in the AAC frame sequence) are much more randomized with the shuffling. In particular, we find that with the shuffling, there are essentially no error patterns of length larger than 3, whereas this is not the case without the shuffling. Further, we find that most of the error patterns with shuffling are of length 1, whereas those without shuffling are of length 2. This is due to the segmentation of the AAC frames into EGPRS radio blocks. As a result of the segmentation, a large fraction of the radio blocks contain pieces of two AAC frames. Therefore, if such a radio block is dropped, two AAC frames are damaged. In the case without shuffling, the two damaged frames are necessarily adjacent frames, whereas in the case with shuffling, the two damaged frames are separated by at least NB frames in the playout buffer.

V. Conclusions

In this paper, we investigate the possibility of embedding streaming data services in an EGPRS network. In particular, we consider the following techniques and conclude that they provide the listed benefits: (i) *Link layer retransmissions*, improve the AAC frame error rate for a large fraction of the users, and therefore improve coverage; (ii) *Dynamic packet assignment (DPA)*, which increases system capacity and therefore also improves coverage; and (iii) *Packet shuffling*, which is a new technique that we propose to randomize the AAC frame errors and therefore assist in the error concealment. We conclude that these techniques, along with error concealment techniques, significantly improve the feasibility of embedding streaming data services in data-optimized, packet-switched bearers.

Our study assumes a limited interaction between the server and the mobile client. The benefits that result from a continuous monitoring and prediction of user performence, and a more complex, ongoing interaction between the client and the server, are subjects for future study. Further, MPEG-4 AAC includes error robustness tools to reconstruct partially damaged AAC frames. The benefits of this are also interesting topics for future work.

VI. Acknowledgments

References

[1] Special Issue on "The Evolution of TDMA to 3G," *IEEE Personal Communications Magazine*, Vol. 6, No. 3, June 1999.

[2] A. Furuskar et al, "EDGE: Enhanced Data Rates for GSM and TDMA/136 Evolution," *IEEE Personal Communications Magazine*, Vol. 6, No. 3, pp. 56-66, June 1999.

[3] M. Bosi et al, "ISO/IEC MPEG-2 Advanced Audio Coding", *J. Audio Eng. Soc.*, vol. 45, No. 10, October 1997, pp. 789-812.

[4] http://www.mpeg.org/MPEG/aac.html

[5] C. Perkins, O. Hodson, V. Hardman, "A survey of packet loss recovery techniques for streaming audio", *IEEE Network*, September/October 1998, pp. 40-48.

[6] S. Hamati et al, "Enhanced Circuit Switched Data for Real Time Services over GSM," *Proc. IEEE Vehic. Tech. Conf.*, Amsterdam, 1999.

[7] J. L. Ramsey, "Realization of Optimum Interleavers," *IEEE Trans. Inform. Theory*, Vol. IT-6, No. 8, pp 338-345, May 1970

[8] E. K. Hall and S. J. Wilson, " Convolutional Interleavers for Stream-Oriented Parallel Concatenated Convolutional Codes," *Proc. of ISIT*, page 33, August 1998

[9] J. Chuang, X. Qiu and J. Whitehead, "Data Throughput Enhancement in Wireless Packet Systems by Improved Link Adaptation with application to the EDGE System," *IEEE VTC99-Fall, Sept. 1999*.
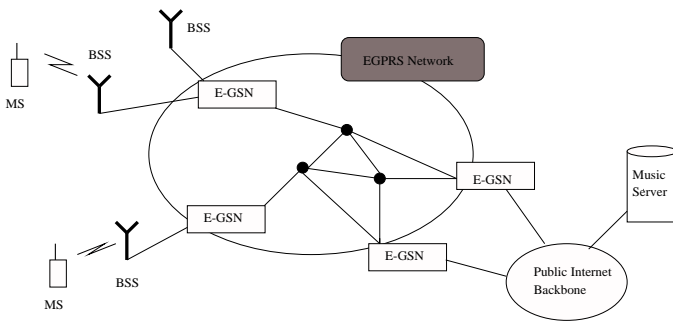
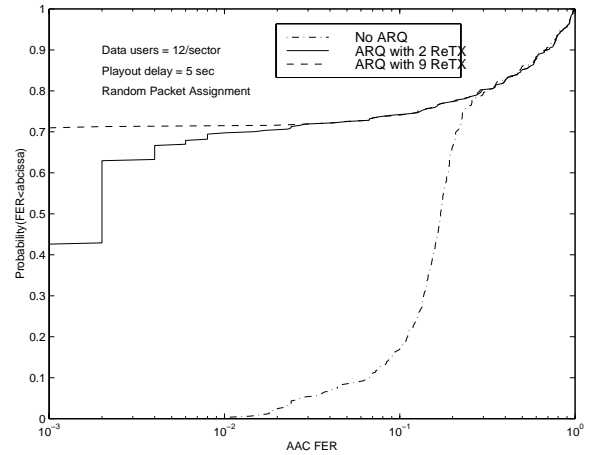Fig. 1. The system architecture.

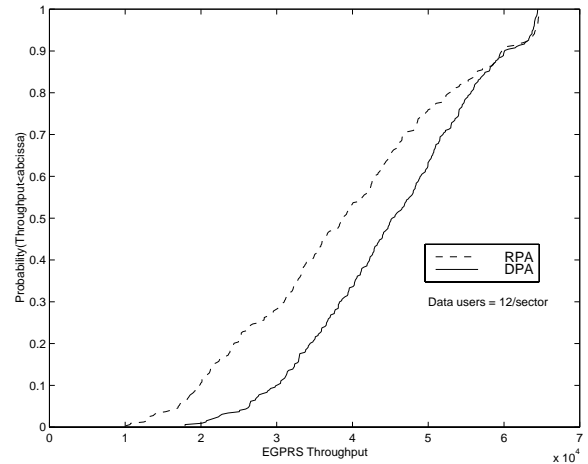Fig. 2. Impact of ARQ on the AAC frame error rate.

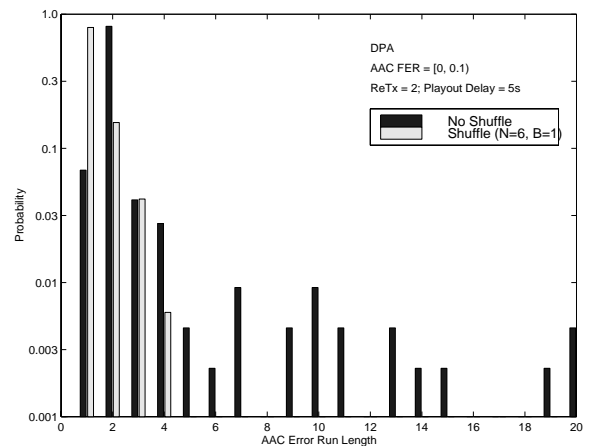Fig. 3. EGPRS throughput with random and dynamic packet assignment.

Fig. 4. Impact of shuffling on the AAC frame error pattern lengths.