Audio Engineering Society

Convention Paper 7394

Presented at the 124th Convention
2008 May 17–20 Amsterdam, The Netherlands

Cable-free Audio Delivery for Home Theater Entertainment Systems

Andreas Floros[1], Nicolas-Alexander Tatlas[2], John Mourjopoulos[2] and Dimitris Grimanis[2]

[1] Dept. of Audiovisual Arts, Ionian University, 49100 Corfu, Greece
floros@ionio.gr

[2] Audio and Acoustic Technology Group, Electrical Engineering and Computer Technology Department, University of Patras, 26500, Greece
mourjop@wcl.ee.upatras.gr

ABSTRACT

Real time, multichannel audio content delivery over the air is expected to significantly simplify the interconnection complexity required for setting up typical home theater applications. However, despite the technological advantages of wireless networking standards related to high transmission rates and Quality-of-Service support, a number of issues have to be additionally addressed, such as multiple loudspeaker synchronization and packet delay/losses containing compressed quality and multiplexed audio data. In this work, further developments in the area of wireless audio delivery are presented by considering in detail multichannel reproduction for wireless home theater applications. Using both subjective and objective performance evaluation criteria, it is shown that cable-free multichannel audio playback is feasible under specific networking and audio coding conditions.

1. INTRODUCTION

Packet-based delivery over computer data networks represents a very attractive format for distributing audio content in real-time within home and office environments [1], as well as over long distances [2]. Focusing on small scale applications, packet-based audio delivery allows the establishment of numerous audio playback setups that are otherwise impossible to realize using conventional audio systems sequential interconnections, such as music distribution from any digital audio source within a typical multi-room home environment to any networked speaker/playback device. Hence, any network-enabled audio device can be connected to any other, using the common network interface also used for computer communications and Internet access.

The evolving market requirements for wireless network interfaces can further enhance audio devices interconnection flexibility, by providing portable, wire-free and low cost means of packet-based audio content delivery. A number of wireless analog and digital audio transmission systems already exist in the market, operating in the S-Band ISM (2.40-2.48GHz, available worldwide) or in the C-Band ISM (5.725-5.875GHz, available in some countries). Typical implementations include wireless home networking setups [3], [4] and wireless headphones [5]. Additionally, integrated consumer home theater systems employing wireless surround channel reproduction via a single point–to–point link have been recently introduced [6]. In the above cases the network audio interface is usually realized on embedded hardware, while the transmission protocol is application-specific for reducing the implementation complexity and cost. However, the above approach raises interoperability issues between different vendor designs.

To overcome such compatibility issues, wireless networking standards should be employed which may eventually represent an attractive and cost-effective delivery scheme for high-quality digital audio distribution. For example, the Bluetooth protocol was employed in the past for delivering digital audio [7]. Due to its limited bandwidth, audio data should be psycho acoustically compressed prior to their transmission. On the other hand, the IEEE802.11 Wireless Local Area Networks (WLANs) family of standards extends the theoretical transmission rates currently up to 54Mbps, allowing non-compressed, high-quality audio transmissions. Moreover, new enhancements on the basic 802.11 standard has recently introduced a number of Quality-of-Service (QoS) mechanisms in the Medium Access Control (MAC) layer, which as it will be explained later represent an attractive framework for high-quality wireless audio transmissions.

A previous study [8] has assessed the performance of real-time CD-quality (uncompressed) audio transmission over WLANs and has provided a framework for defining optimal transmission and networking parameters. More specifically, it was found that the audibility of the distortions induced by the variable transmission delays and packet losses occurring during the wireless transmission of high-quality uncompressed audio traffic streams (TSs) strongly depends on the transmission parameters (such as the packet length and the transmission scheduler). The purpose for the current work is to introduce further developments in the area of wireless audio transmissions by considering multichannel reproduction for wireless home theater applications. This presents a novel and challenging topic, since audio coding for home theater applications incorporates psychoacoustic models for compressing and multiplexing all audio channels in a single data bitstream.

The rest of the paper is organized as following: Section 2 describes the fundamental aspects of wireless audio transmission, while in Section 3, an overview of QoS mechanisms over WLANs is provided. Section 4 presents typical results obtained through computer simulations, which can be used for assessing the overall performance of the wireless audio distribution system in terms of the achieved audio playback quality. Finally, Section 5 concludes the work presented here.

2. WIRELESS AUDIO FUNDAMENTALS

According to [9], wireless installations can support two types of audio applications: (a) Simple point-to-point audio delivery where an audio server (possibly connected to the Internet) wirelessly transmits in real time the same or different audio streams to a number of wireless audio players/receivers and (b) wireless point-to-multipoint (or multichannel) transmission, where a number of loudspeakers (i.e. 5.1 for a typical home theater setup) are wirelessly connected to a digital audio source.

In the first case, no synchronization between the wireless receivers is required, while the maximum allowed number of remote players can be dynamically adjusted by bandwidth reservation algorithms and strongly depends on the amount of interference present in the wireless channel. In practice, using the wireless point-to-point setup, any compatible wireless-enabled audio device (including portable audio playback, laptop computers and consumer electronics equipment) can receive audio data from the audio source (e.g. a mass storage device, an internet access point etc) on user demand, provided that it is in reception range. Each of these audio devices can then reproduce the audio content via the usual wired audio links (e.g. headphones, speakers, etc).

On the contrary, in the multichannel transmission case, the digital audio source (e.g. a CD/DVD-player connected to an central wireless node) transmits audio data to the appropriate wireless loudspeaker which should perform simultaneous and synchronized (relative to all other receivers) playback in real-time. In practice, synchronization losses are very frequently raised because of the variable delay of packet delivery among the different audio receivers and the permanent packet losses imposed by the degraded channel conditions. Such delay or packet losses can cause (a) playback gaps due to the absence of the appropriate packet from the playout buffer and (b) relative channel phase shifting (one audio channel leading or lagging), which causes loss of the sound spatial information. As it is shown in [8], in both cases the playback distortion is audible. Following the above analysis it is clear that local (i.e. hardware) clock as well as packet playout synchronization is required for eliminating unpredicted channel shifts and phase distortions.

The audio coding type represents a parameter that may significantly affect the wireless delivery performance. For example, a recent study [10] has shown that when PCM and MPEG-1 Layer III stereo audio streams are wirelessly transmitted to a stereo pair of loudspeakers, the compressed audio delivery is more sensitive to both packet losses and transmission delays, although the total audio bitrate is much lower. Extending the above study, this work considers a 5.1 channel wireless streaming system that consists of a central node wirelessly transmitting digital audio to six loudspeakers.

3. QOS OVER WLANS

For efficient high-quality audio delivery, the wireless transmission protocol must provide sufficient QoS guarantees. In general, QoS refers to the capabilities of a network to provide service guarantees to data flows that are well described by specific traffic characteristics, such as priority, data rate and latency. In wired networking environments, QoS is a minor issue, since the error rate is low and the supported transmission rates are very high. However, this is not the case in wireless networking setups, where the limited channel bandwidth combined with the high packet error rate results into significant throughput and delay degradation [11]. This may render unacceptable the performance of time-bounded, high-rate applications such as audio/video streaming and Voice over IP (VoIP).

In order to define specific QoS mechanisms for the WLAN specification, the IEEE 802.11E MAC layer specification was recently ratified [12], which employs two wireless channel access methods: (a) the Enhanced Distributed Channel Access (EDCA), which provides distributed access using traffic differentiation and (b) a centralized access method termed as Hybrid Controlled Channel Access (HCCA). The suitability of these access methods especially for audio/video streaming applications was evaluated in [11] where it was found that, while the random access nature of EDCA couldn't provide strict service guarantees, HCCA fulfills this requirement by centrally scheduling the access for all associated receivers. However, the HCCA scheduler type that calculates and applies the final access schedule may dramatically affect the QoS performance of the network. The 802.11E specification defines the minimum requirements that must be met by any HCCA scheduling scheme and introduces as an example the Simple scheduler [12]. However, it is expected that several scheduling schemes will be developed, optimized for the targeted application.

4. RESULTS

4.1. Simulation Parameters

For the purposes of this work, two multichannel 5.1 audio streams (Theme 1 and Theme 2) where considered, encoded using the AC-3 [13] coding scheme, using different coding bitrates. The signal and coding characteristics of these multichannel streams are shown in Table 1. The complete AC-3 stream was wirelessly transmitted from the digital audio source to all 6 audio receiver-loudspeakers. The positions of audio receivers was pre-determined as Left (L), Right (R), Center (C), Left Surround (LS), Right Surround (RS) and Low Frequency Effect (LFE). Hence, each of these receivers was responsible for decoding the multiplexed audio stream and for selecting the appropriate channel for playback.

	Theme 1	Theme 2
Duration (sec)	54	57
Coding bitrate (kbps)	384	448
Sampling rate (kHz)	48	48
Frame size (bytes)	1536	1792
Total transmitted bitrate (kbps)	2376	2760

Table 1 Multichannel audio stream parameters

In all test cases, the lower cost 802.11b protocol was considered with a physical transmission rate equal to 11Mbps. Additionally, the HCCA 802.11e polling-based enhancements were considered, with three different centralized scheduling schemes: the Simple scheduler mentioned previously, the WAVES adaptive scheduler [14] as well as the SETT-EDD scheduling scheme [15]. The transmitted packet size was equal to L_p=882bytes, which as was found in [10], represents the optimal choice for uncompressed audio real-time delivery. It should be also noted that in order to match the AC-3 frame size appeared in Table 1 and the MAC layer packet size L_p, data-appropriate packet fragmentation is performed.

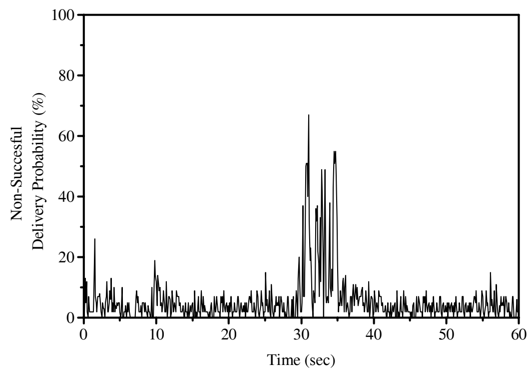

Figure 1 Non-successful transmission probability vs time

During the wireless transmission simulations, two different channel models were considered: (a) the ideal model, where no packet losses are present and (b) the noisy model. In the latter case, the instantaneous packet loss probability was obtained through measurements of real-world 802.11b-based transmission patterns, with two wireless stations transmitting at 2m distances, under heavy interference induced by two neighboring stations in the 2.4GHz frequency band. More specifically the number of the total transmissions R_{tot} (including retransmissions) and the total transmissions that resulted into successful delivery (R_s) were measured within every beacon period (T_b) and the successful delivery probability p was calculated as a function of time (expressed in multiples of the beacon period) using the equation:

$$p(kT_b) = \frac{R_s(kT_b)}{R_{tot}(kT_b)} \qquad (1)$$

In Figure 1, the measured non-successful transmission probability $1\text{-}p(kT_b)$ is illustrated as a function of time.

4.2. Objective measurements

In order to objectively evaluate the effect of the wireless system on the transmitted digital streams, the mean time deviation between channels during reproduction was calculated. These deviations are caused by either lost or excessively delayed packets that are not equivalent for all streams being transmitted. Depending on the time shift, the resulting loss of synchronization may be perceived as phase mismatch, shifts in the acoustic image or even as separate sources. The time deviations were calculated in comparison with one channel. The reference channel was arbitrary selected as the (L) channel.

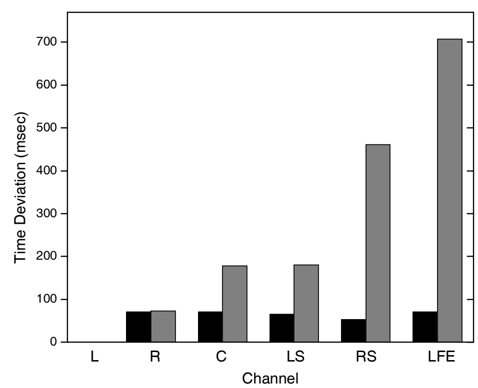

Figure 2. Measured time deviation for the Simple scheduler (Black: Ideal channel, Gray: Noisy channel)

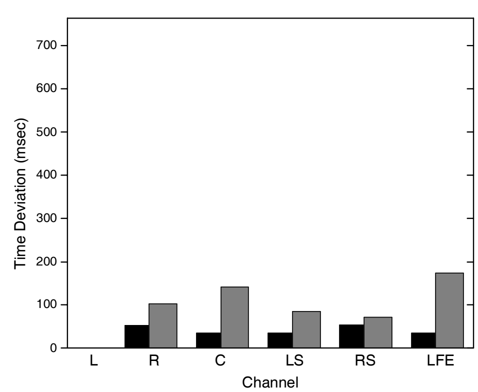

Figure 3. Measured time deviation for the WAVES scheduler (Black: Ideal channel, Gray: Noisy channel)

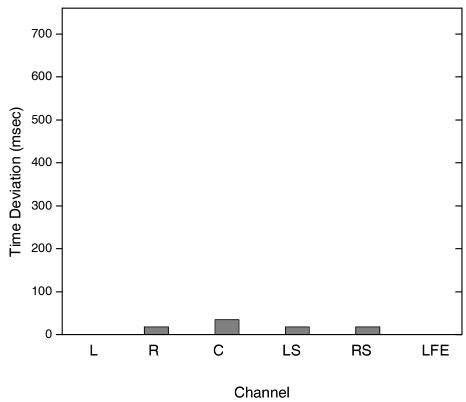

Figure 4. Measured time deviation for the SETT-EDD scheduler (Black: Ideal channel, Gray: Noisy channel)

Figures 2, 3 and 4 show the time deviation values obtained for each audio channel when using the Simple, WAVES and SETT-EDD schedulers respectively. It should be noted that the illustrated values are the mean absolute values obtained from both audio themes considered here. The following conclusions can be drawn from the above figures:

(a) The simple scheduler causes significant time deviations for both channel models. For the case of the ideal channel, the deviations are limited, while for the noisy channel they approach 0,7sec.
(b) Accordingly, the Waves scheduler leads to the introduction of time deviations for both channel models. Compared to the simple scheduler these deviations are almost equivalent for the case of the ideal channel, while quite restricted for the noisy channel.
(c) Finally, no time deviations are measured using the SETT-EDD scheduler over the ideal channel. However, even the SETT-EDD scheduler introduces some limited delays for the case of the noisy channel model.

4.3. Subjective measurements

In order to perceptually assess the effect of the audible distortions raised during the wireless multichannel audio AC-3 streams transmission, the Noise-to-Mask (NMR) criterion [16] was employed. The usefulness and validity of NMR has been established in many engineering fields [17] as it utilizes masking pattern and internal auditory representations introduced by a reference signal to assess the audibility of the distortions imposed by any audio system. In this work, for NMR estimation, the original AC-3 decoded audio (prior to wireless transmission) was used as reference and the quality assessment of the multichannel audio streams was based on the averaged (segmental) NMR values, using the equation

$$avgNMR = \frac{1}{K}\sum_{i=1}^{K} NMR(i) \tag{2}$$

where NMR(i) is the NMR value of the i-th transmitted audio frame, for a total number of K frames. It should be also noted that NMR values above 0dB indicate the presence of audible distortions, while NMR values below -10dB indicate an audio signal free of audible distortions.

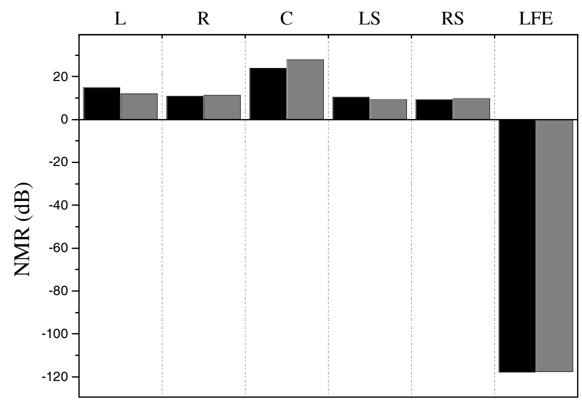

Figure 5 Measured NMR for the Simple scheduler (Black: Ideal channel, Gray: Noisy channel)

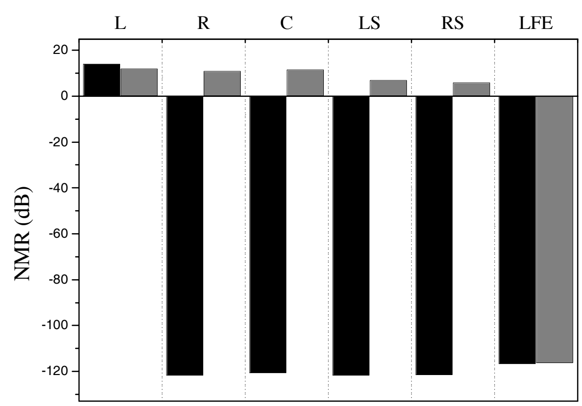

Figure 6 Measured NMR for the WAVES scheduler (Black: Ideal channel, Gray: Noisy channel)

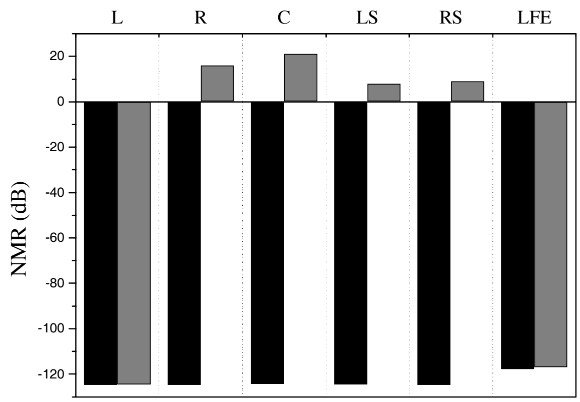

Figure 7. Measured NMR for the SETT-EDD scheduler (Black: Ideal channel, Gray: Noisy channel)

Figures 5, 6 and 7 show the NMR values obtained for each audio channel when using the Simple, WAVES and SETT-EDD schedulers respectively. It should be noted that the illustrated NMR values are the mean values obtained from both audio themes considered here.

As can be deduced from the above figures, the simple scheduler causes the introduction of significant distortions to five streams, regardless to the channel model being used. Positive NMR values for all channels were measured, except for the LFE. On the other hand, employing the Waves scheduler leads to audible distortion for only one channel for the ideal wireless channel case, while for the noisy channel the results are almost equivalent to the simple scheduler. Finally, bit-accurate reproduction for all six channels is achieved using the SETT-EDD scheduler over an ideal wireless channel. However, even the SETT-EDD scheduler can not satisfy the service requirements for all six-streams for the case of the noisy channel model, leading to significant audible distortions for four channels.

5. CONCLUSIONS

In this paper, digital multichannel wireless audio transmission was investigated, using the 802.11b WLAN protocol in addition with 8011.e QoS provisions. A number of different scenarios were considered, employing three scheduling algorithms while using two wireless channel models, obtained by real-life measurements.

The results presented in the previous Section make evident that the selection of the QoS scheduler is crucial to ensure successful servicing for all six streams. More specifically, it was found that the simple scheduler introduces significant reproduction distortions in all test cases, while for the case of the noisy channel, the overall playback quality is further reduced. The employment of the WAVES scheduler leads to improved however not satisfactory reproduction fidelity. On the other hand, the error-free reproduction achieved for the case of the SETT-EDD scheduler for the ideal wireless channel, shows that the employment of an adaptive scheduler which dynamically adjusts the service schedule based on the networking conditions, represents a fundamental requirement for real-time multichannel audio streaming applications at the expense of the increased implementation complexity and processing power. However, channel interference introduces additional transmission errors, due to the bandwidth limitations. Increasing the available bandwidth can reduce the effect of the channel interference on the final playback quality. Hence, higher rate wireless protocols (e.g. 802.11g/n) must be preferred for high-quality multichannel applications.

In conclusion, the 802.11e HCCA access method was found to be marginally suitable for 5.1 channel real-time audio playback using the 802.11b transmission protocol, taking into account the limitations raised during the tests. What has become clearly evident is that uninterrupted streaming for wireless audio applications requires the use of adaptive, dynamic HCCA schedulers as well as higher physical transmission rates. While SETT-EDD is suitable for some test cases investigated here, a scheduler optimized for audio would ensure in-phase multichannel playback with no audible distortions under any channel and network conditions.

Future research efforts should focus on the perceptual evaluation of such distortions, leading to a set of practical rules for schedulers optimized for audio applications. Moreover, the development of higher layer protocols would be advantageous for minimizing the audibility of the distortions induced by the wireless transmission and for synchronizing the remote playback devices.

6. REFERENCES

[1] T. Blank, B. Atkinson, M. Isard, J. Johnston and K. Olynyk, “An Internet Protocol (IP) Sound System”,

presented at the 117th AES Convention, San Francisco, October 28-31 2004 (preprint 6211).

[2] A. Xu, W. Woszczyk, Z. Settel, B. Pennycook, R. Rowe, P. Galanter, J. Bary, "Real-Time Streaming of Multichannel Audio Data over Internet", JAES Volume 48 Issue 7/8 pp. 627-641; July 2000

[3] http://www.streamium.com

[4] http://www.elevenengineering.com

[5] http://www.amphony.com/products/h2500.htm

[6] http://www.pioneerelectronics.com

[7] Floros, A., Tatlas, N. A. and Mourjopoulos J., BlueBox: A cable-free Digital Jukebox for compressed-quality Audio Delivery. IEEE Trans. on Consumer Electronics, 51, 2, (May 2005), 534 – 539.

[8] Tatlas, N.A., Floros, A., and Mourjopoulos, J., Wireless Digital Audio Delivery Analysis and Evaluation. In Proceedings of the IEEE 31st International Conference on Acoustics, Speech, and Signal Processing (ICASSP'06), Toulouse, May 2006.

[9] Floros, A. and Karoubalis, T. Delivering high-quality Audio over WLANs. In Proceedings of the 116th AES Convention, Berlin, May 2004, (preprint 5996).

[10] N.-A. Tatlas, A. Floros, T. Zarouchas, and J. Mourjopoulos, "WLAN technologies for audio delivery", published in the Advances in Multimedia (Hindawi Publishing Corporation) Journal, Volume 2007, Article ID 12308.

[11] Floros, A., Karoubalis, T. and Koutroubinas S. Bringing Quality in the 802.11 Wireless Arena. Broadband Wireless and WiMax IEC Comprehensive Report, ISBN: 978-1-931695-30-x, January 2005, 443 – 454.

[12] IEEE802.11 WG, IEEE802.11e. IEEE Standard for Information Technology - Telecommunications and information exchange between systems – Local and metropolitan area networks – Specific requirements - Part 11: Wireless Medium Access Control (MAC) and Physical Layer (PHY) specifications: Amendment: Medium Access Control (MAC) Quality of Service Enhancements. November 2005.

[13] M. Davis, "AC-3 operation, bitstream syntax and Features", presented at the AES 97th convention, November 1994, San Francisco (preprint 3910).

[14] A. Floros and N. Kanellopoulos, "Efficient Wireless Delivery for Digital Audio/Video Applications", to be presented at the 3rd International Symposium on Communications, Control and Signal Processing (ISCCSP 2008) to be held in Malta, 2008.

[15] A. Grillo, M. Macedo, and M. Nunes, "A Scheduling Algorithm for QoS Support in IEEE802.11E Networks", IEEE Wireless Communications, 2003, Vol. 10, No 3, pp. 36 – 43.

[16] K. Brandenburg and T. Sporer, "NMRand masking flag: evaluation of quality using perceptual criteria," in Proceedings of the 11th International AES Conference: Audio Test &Measurement, pp. 169–179, Portland, Oregon, May 1992.

[17] J.Herre, E. Eberlein, H. Schott, and K. Brandeburg, "Advanced audio measurement system using psychoacoustic properties," in Proceedings of the 92nd AES Convention of the Audio Engineering Society, New York, NY, USA, March 1992, (preprint 3321).