**Source:** Ericsson, Nokia

**Title:** Additional information: AMR-WB+ performance at very-low  
bit rates

# Agenda item: 13.7.1

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Summary

This document gives additional information on the AMR-WB+ audio codec performance in various very-low-rate operation conditions. Very-low-rate audio codec performance is essential in PSS/MMS/MBMS applications with audio-visual content [1]. AMR-WB+ according to 3GPP TS 26.304 with flexible codec control is compared with Enhanced aacPlus (EAAC+) codec according to 3GPP TS 26.410 using MUSHRA testing methodology.

Results indicate that AMR-WB+ operated at around 10 kbps requires 5-6 kbps less bit rate than E-AAC+ to achieve an audio quality suitable for many applications that require very-low-rate audio content delivery, as e.g. PSS, MMS and MBMS of audio visual content via GPRS networks.

# Introduction

A couple of subjective listening tests have been conducted to characterize the performance of the AMR-WB+ at very low rates. Tests have been conducted using the same high quality procedures that have been used to perform the official selection tests. Two listening laboratories conducted tests independently. The results from Nokia listening test laboratory in Finland and Ericsson listening test laboratory in Sweden are presented separately, leading coinciding conclusions.

# Source material

The Ericsson laboratory performed listening on two sub-sets of the material (lowrate test sets A4a and A4b) while the Nokia lab performed testing only on test set A4a. Both sets consisted of 12 test and 4 practice items.

# Processing

Processing of the source material has been done according to the low rate audio selection test and processing plan [2]. Only the tested codecs were different. Main processing has been done using concatenated material. All the processing has been done using the very same tools (up/down sampling, file concatenation etc.) that were used in the processing phase of the official low rate test A4.

# Test conditions

|  |  |  |
| --- | --- | --- |
| **Main Codec Conditions** |  |  |
| Candidates | 3 | AMR-WB+@9.75kbps, 12kbps, 13.6 kbps, 15.2kbps E-AAC+ @12 kbps, 13.6 kbps 15.2 kbps |
| Use case | 1 | A (PSS) |
| Error Conditions | 1 | No errors |
| Mono/Stereo | 1 | Mono |
|  |  |  |
| **Codec references** |  |  |
| Codec references | 1 | AMR-WB+ @14 kbps use case A |
| Input sampling rate | 1 | 24 kHz |
| Input characteristics |  |  |
| Number of input channels | 1 | Mono |
| Number of output channels | 1 | Mono |
|  |  |  |
| **Other references** |  |  |
| Open Reference | 1 | Original signal |
| Hidden Reference | 1 | Original signal |
| Anchors | 2 | 3.5 kHz and 7 kHz low-pass filtered original signal |
|  |  |  |
| **Common Conditions** |  |  |
| Stimulus type |  | Sound item |
| Radio Channels | 0 | Clean |
| Number of test sets | 2 | Low-rate test sets A4a and A4b |
| Number of audio items per test set | 12 |  |
| Input sampling rate |  | 48 kHz |
| Number of input channels | 1 | Mono |
| Output sampling rate |  | Unspecified |
| Number of output channels | 1 | Mono |
| Listening Level | 1 | To be chosen by subject |
| Listening laboratories for test set A4a | 2 | Nokia and Ericsson listening test laboratories |
| Listening laboratories for test set A4b | 1 | Ericsson listening test laboratoriy |
| Listeners | 37 | Experienced listeners (Nokia 15, Ericsson 10/12) |
| Presentation randomizations | 46 | One for each listener |
| Rating Scale | 1 | Continuous quality scale |
| Listening System | 1 | Binaural high-quality headphones |
| Listening Environment |  | Room Noise: Hoth Spectrum at 30dBA (as defined by ITU-T, Recommendation P.800, Annex A, section A.1.1.2.2.1 Room Noise, with table A.1 and Figure A.1) |

Table 1: Overview of the test conditions for very-low-rate tests

# Listening sessions

## Nokia listening test laboratory

### Presentation sequences

All listeners listened their sound items and each trial of the item in unique order.

### Listeners

All the listeners were native finish speakers with prior experience in MUSHRA test methodology. They were all tested before listening with audiometer to have normal hearing (to fulfil ISO Standard 389 requirements).

### Listening environment

Listeners were placed in high quality, acoustically isolated booths. Six identical booths with internal dimensions of 1.4 x 1.1 x 2.1m were used. The background noise-rating curve of each booth fulfils the ISO NR15 requirement. The reverberation times within the booths are <300ms above 315Hz one-third octave bands. No discernible flutters are audible within the booths [3].

### Environmental noise

Environmental noise was fed into the booths with the required Hoth spectrum to represent typical room noise at the required 30dBA level (as defined by ITU-T, Recommendation P.800 [4]). Two loudspeaker units (type: Genelec 1029A) per booth were used. Speakers were positioned so that the sound pressure level was 30 dBA above the centre of the seat of subject's chair.

### Testing facility

The listening test was controlled by remote PCs with a keyboard, mouse and an LCD screen in the booths. Six machines were used to play the samples to the listeners and to collect their answers. Each one is furnished with a high quality digital sound card (type: RME DIGI 96/8 PRO), providing 44.1kHz or 48kHz output at a resolution of 24 bits. The digital audio output signals were subsequently fed to a Studer D19 24bit multi-channel digital to analogue converter employing an AES/EBU bus. A Symmetrix 304 headphone amplifier was used. Samples were presented binaurally to the listeners over high quality Sennheiser HD580 headphones.

## Ericsson listening test laboratory

### Presentation sequences

All listeners listened their sound items and each trial of the item in unique order.

### Listeners

The listening panel was selected from experienced listeners inside Ericsson. A pre-screening procedure was used were previous performance in intermediate quality audio listening tests served as an indication of the listeners’ ability to judge anchors and references in a correct way, as well as the ability to repeatedly grade in a consistent manner. The listeners, both male and female, were between 21 to 44 years of age and had all had previous experience of audio listening tests using the MUSHRA methodology.

### Listening environment

The listening environments were two listening labs, which both conformed to the standard requirements.

### Environmental noise

Environmental noise was fed into the booths with the required Hoth spectrum to represent typical room noise at the required 30dBA level (as defined by ITU-T, Recommendation P.800 [4]).

### Testing facility

Open-back circum-aural headphones were used (Sennheiser HD600) and the listeners could individually adjust the listening level to their preference. The audio was fed from the computer to the listener using M-audio USB Duo sound cards.

The MUSHRA software has been developed in-house. It has a similar GUI as the CRC-SEAQ software shown in the test plan although there is a possibility to show the waveform of the current test item. The waveform is rendered from the open reference clip, thus showing no information about the encoded clips. The software performs both inter-item and intra-item randomization of the test sequence, and provides a raw output of the test results into individual listener output files.

# AMR-WB+ codec

The flexible control of AMR-WB+ is utilised in the experiment as the instrument for trading intrinsic codec bit rate against bandwidth. The codec mono core rates and internal sampling frequency (ISF) are set according to Table 2. To compare the performance of the flexible codec control the configuration utilised in PSS/MMS codec selection (14 kbps PSS) is included, which was constrained to a sampling rate of 24 kHz.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Codec** | **Core** | **Stereo** | **ISF** |
| **1** | AMR-WB+ 9.75 kbps | 10.4 | 0 | 24.0 |
| **2** | AMR-WB+ 12 kbps | 12.0 | 0 | 25.6 |
| **3** | AMR-WB+ 13.6 kbps | 13.6 | 0 | 25.6 |
| **4** | AMR-WB+ 15.2 kbps | 15.2 | 0 | 25.6 |
| **5** | AMR-WB+ 14 kbps mono (PSS) | Configuration utilised in PSS/MMS selection tests | | |

Table 2: AMR-WB+ codec configurations

# E-AAC+ codec

The E-AAC+ codec is operated at the same bit rates as AMR-WB+, except for the lowest AMR-WB+ bit rate of 9.75, as the lowest mono bit rate supported by E-AAC+ is 12 kbps.

|  |  |  |
| --- | --- | --- |
|  | **Codec** | **Stereo** |
| **1** | E-AAC+ 12 kbps | 0 |
| **2** | E-AAC+ 13.6 kbps | 0 |
| **3** | E-AAC+ 15.2 kbps | 0 |

Table 3: E-AAC+ codec configurations

# Test results

The test results by laboratory and by test set are given below in numerical formats as well as the total scores in graphical format.



Table 4: Results for test set A4a, Ericsson



Figure 1: Total results for test set A4a, Ericsson



Table 5: Results for test set A4b, Ericsson



Figure 2: Total results for test set A4b, Ericsson



Table 6: Results for test set A4a, Nokia

Figure 3: Total results for test set A4a, Nokia



# Combined results

The individual test results have been combined to quality-rate curves for the two codecs. Figure 2 displays curves for the total quality and for speech and music content separately.

Figure 4: Quality - rate curves

# Conclusion

As can be concluded from figure 2, AMR-WB+ is particularly strong at very low bit rates. At about 10 kbps a quality level is reached which is suitable for many applications that require very-low-rate audio content delivery, as e.g. PSS, MMS and MBMS of audio visual content via GPRS networks.

In comparison of the two 3GPP audio codecs it can be stated that in the studied very-low-rate bit rate range:

* E-AAC+ requires about 5-6 kbps more bit rate for achieving the same quality level as AMR-WB+ at 9.75 kbps
* Comparing the codecs at a rate of 15.2 kbps, AMR-WB+ provides about one Mushra category higher audio quality than E-AAC+, which at this rate is still slightly below the quality level of AMR-WB+ at 9.75 kbps
* Both codecs exhibit a certain quality dependence on the content. Generally, the quality gap between speech and music content is smaller for AMR-WB+. I.e. AMR-WB+ provides a more consistent quality across various content types.

# References

[1] S4-040484: “Considerations for MBMS audio codec selection”

[2] S4-030824 “Low-Rate Audio Selection Test and Processing Plan”

[3] M. Kylliäinen et al.; “Compact high performance listening spaces.” Euronoise, Naples, 2003.

[4] ITU-T, Recommendation P.800