Technical Specification Group Services and System Aspects ***TSGS#11(01)0095***
Meeting #11, Palm Springs, U.S.A, 19-22 March 2001 **Agenda Item: 7.4.3**

Source: **TSG_SA WG4**

Title: **Packet Switched Conversational Multimedia Applications; Default Codecs (Release 4)**

Document for: **Approval**

Presentation of Specification to TSG-SA

Presentation to:	**TSG SA Meeting #11**
Document for presentation:	**TS 26.235 Version 2.0.0**
Presented for:	**Approval**

Abstract of document:

This specification introduces the set of default codecs for packet switched conversational multimedia applications within 3GPP IP Multimedia Subsystem. Visual and sound communication are specifically addressed. The intended applications are assumed to require low-delay, real-time functionality.
The present document is applicable, but not limited, to PS video telephony.

Changes since last presentation to TSG Meeting #10:

Specification completed.

Outstanding Issues:

None.

Contentious Issues:

None.

3GPP TS 26.235 V2.0.0 (2001-03)

Technical Specification

3rd Generation Partnership Project;
Technical Specification Group Services and System Aspects;
Packet Switched Conversational Multimedia Applications;
Default Codecs
(Release 4)

Keywords

Conversational multimedia applications, codec, protocols, packet-switched

3GPP

Postal address

3GPP support office address

650 Route des Lucioles - Sophia Antipolis
Valbonne - FRANCE
Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Internet

http://www.3gpp.org

Contents

Foreword

This Technical Specification has been produced by the 3GPP.

The present document introduces the set of default codecs applying to 3G packet switched conversational multimedia applications within the 3GPP system.

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of this TS, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version 3.y.z

where:

x the first digit:

1 presented to TSG for information;

2 presented to TSG for approval;

3 Indicates TSG approved document under change control.

y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.

z the third digit is incremented when editorial only changes have been incorporated in the specification;

Introduction

This document contains a specification for default multimedia codecs to be used within 3GPP specified IP Multimedia Subsystem (IM Subsystem). IM Subsystem as a subsystem includes specifically the conversational IP multimedia services, whose service architecture, call control and media capability control procedures have been defined in 3GPP specifications TS 24.229 [15], and are based on the 3GPP adopted version of IETF Session Invitation Protocol (SIP).

The term codec is usually associated with a single media type. In case of packet switched transport domain, which IM Subsystem will depend on, the individual media types are independently encoded and packetised to appropriate separate Real Time Protocol (RTP) packets. These packets are then transported end-to-end inside UDP datagrams over real-time IP connections that have been negotiated and opened between the terminals during the SIP call as specified in 3GPP TS 24.229 [15].

From the codec definition viewpoint, the UEs operating within IM Subsystem need to provide encoding/decoding of the derived codecs, and perform corresponding packetisation/depacketisation functions. Logical bound between the media streams is handled in the SIP session layer, and inter-media synchronisation in the receiver is handled with the use of RTP time stamps.

Finally, since 3GPP networks are inherently error prone, error detection and/or correction must also be provided by the individual codecs within IM Subsystem, since they have a comprehensive view of the bit stream they produce and therefore can apply the most efficient form of error detection and/or correction.

1 Scope

This specification introduces the set of default codecs for packet switched conversational multimedia applications within 3GPP IP Multimedia Subsystem. Visual and sound communication are specifically addressed. The intended applications are assumed to require low-delay, real-time functionality.

The present document is applicable, but not limited, to PS video telephony.

2 References

The following documents contain provisions, which through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies.
- A non-specific reference to an ETS shall also be taken to refer to later versions published as an EN with the same number.

[1] IETF RFC 2543 "SIP: Session Initiation Protocol"

[2] IETF RFC 2327 "SDP: Session Description Protocol"

[3] IETF RFC 2429 "RTP Payload Format for the 1998 Version of ITU-T Rec. H.263 Video (H.263+)"

[4] IETF RFC 1889 "RTP: A Transport Protocol for Real-Time Applications"

[5] IETF RFC 3016 "RTP Payload Format for the MPEG-4 visual"

[6] ITU-T Recommendation H.263: "Video coding for low bitrate communication"

[7] 3GPP Technical Specification 3GPP TS 26.110: “Codec for Circuit Switched Multimedia Telephony Service; General Description”

[8] 3GPP Technical Specification 3GPP TS 26.111: “Codec for Circuit Switched Multimedia Telephony Service; Modifications to H.324”

[9] 3GPP Technical Specification 3GPP TS 26.071: “Mandatory Speech Codec; General Description”

[10] 3GPP Technical Specification 3GPP TS 26.090: “Mandatory Speech Codec; Speech Transcoding Functions”

[11] 3GPP Technical Specification 3GPP TS 26.073: “Mandatory Speech Codec; ANSI C-Code”

[12] 3GPP Technical Specification 3GPP TS 26.104: "AMR speech Codec; Floating point C-Code"

[13] International Standard ISO/IEC 14494-2: "Information technology - Generic coding of audio-visual object - Part 2: Visual, 1999".

[14] 3GPP Technical Specification 3GPP TS 24.228: “Signalling flows for the IP multimedia call control based on SIP and SDP”

[15] 3GPP Technical Specification 3GPP TS 24.229: “IP Multimedia Call Control Protocol based on SIP and SDP”

[16] 3GPP Technical Specification 3GPP TS 26.171: "AMR Wideband Speech Codec; General description".

[17] 3GPP Technical Specification 3GPP TS 26.190: "AMR Wideband Speech Codec; Transcoding Functions".

[18] 3GPP Technical Specification 3GPP TS 26.201: "AMR Wideband Speech Codec; Frame Structure".

[19] ITU-T Recommendation H.263: "Annex X, Profiles and Level Definition"

[20] 3GPP Technical Specification 3GPP TS 23.228: "IP Multimedia (IM) Subsystem - Stage 2"

[21] 3GPP Technical Specification 3GPP TS 23.107: "UMTS QoS and architecture"

[22] 3GPP Technical Specification 3GPP TS 23.207: "End-to-End QoS Concept and Architecture"

[23] 3GPP Technical Specification 3GPP TS 23.060: "General Packet Radio Service (GPRS); Service Description; Stage 2"

[24] IETF RFC 2793 “RTP Payload format for Text Conversation”.

[25] ITU-T T.140 “Text conversation presentation protocol” (1998 with amendment 2000)

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the following terms and definitions apply.

3G PS multimedia terminal: A terminal based on IETF SIP/SDP internet standards modified by 3GPP for purposes of 3GPP packet switched network based multimedia telephony

3.2 Abbreviations

For the purposes of the present document, the following abbreviations apply:

AMR	Adaptive Multirate Codec
IETF	Internet Engineering Task Force
IM Subsystem	Internet Protocol Multimedia Subsystem
ITU-T	International Telecommunications Union - Telecommunications
RFC	IETF Request For Comments
RTP	Real-time Transport Protocol
RTCP	RTP Control Protocol
SDP	Session Description Protocol
SIP	Session Initiation Protocol

4 General

3G PS multimedia terminals provide real-time video, audio, or data, in any combination, including none, over 3GPP IM Subsystem. Terminals are based on IETF defined multimedia protocols SIP, SDP, RTP and RTCP. Communication may be either 1-way or 2-way. Such terminals may be part of a portable device or integrated into an automobile or other non-fixed location device. They may also be fixed, stand-alone devices; for example, a video telephone or kiosk. Multimedia terminals may also be integrated into PCs and workstations.

In addition, interoperation with other types of multimedia telephone terminals, such as 3G-324M may be possible, however in such case a media gateway functionality supporting 3G-324M - IM Subsystem interworking will be required within or outside the IM subsystem.

5 System overview

This specification describes the required codec related elements for 3G PS multimedia terminal:

- Mandatory and optional codecs for 3G PS multimedia terminal
- Media encapsulation and decapsulation rules for each mandatory and optional codec

6 Functional requirements

SIP protocol itself does not mandate any codecs. Standardisation of mandatory codecs does not prevent the use of other codecs that can be signalled using the SDP protocol. 3G PS multimedia terminals shall be able to use the same audio and video codecs applied in 3G-324M [8]. This will ensure the interoperability with 3G circuit switched multimedia telephony.

6.1 Audio

3G PS multimedia terminals offering audio communication shall support AMR narrowband speech codec [9, 12]. This is the mandatory speech codec.

The AMR wideband speech codec shall be supported when the 3G PS multimedia terminal supports wideband speech working at 16 kHz sampling frequency [16].

6.2 Video

3G PS multimedia terminals offering video communication shall support ITU-T recommendation H.263 baseline [6]. This is the mandatory video codec.

H.263 Version 2 Interactive and Streaming Wireless Profile (Profile 3) Level 10 should be supported [19]. This is an optional video codec.

ISO/IEC 14496-2 (MPEG-4 Visual) Simple Profile at Level 0 should be supported [13]. This is an optional video codec.

6.3 Real time text

3G PS multimedia terminals offering real time text conversation should support ITU-T T.140 Text Conversation presentation coding [25].

6.4 Interactive and background data

SIP signalling offers initialisation of packet switched interactive or background class reliable data services as well. However specification of such data services are outside the scope of this specification.

7 Call control

Functional requirements for call control are specified in 3GPP TS 23.228 [20].

The required signalling functions are specified in 3GPP TS 24.228 [14] and call control protocols in 3GPP TS 24.229 [15].

8 Bearer control

The media control is based on declaration of terminal media capability sets in SDP part of appropriate SIP messages.

Relation of application level SDP signalling and radio access bearer assignment is defined outside this specification. The QoS architecture and concept for WCDMA and GERAN is specified in 3GPP TS 23.107 [21]. The end-to-end QoS framework involving GPRS and UMTS is specified in 3GPP TS 23.207 [22]. The applicable general QoS mechanism and service description for the GPRS in GSM and UMTS is specified in 3GPP TS 23.060 [23].

9 Multimedia stream encapsulation

9.1 MIME media types

The terminal shall declare the mandatory and any optional media streams using the codec specific MIME media types in the associated SDP syntax. The MIME media types for the mandatory and optional codecs shall be according to the corresponding types registered by IANA.

- AMR narrowband speech codec MIME media type as specified in Annex D.
- AMR wideband speech codec MIME media type is specified in Annex B.
- H.263 video codec MIME media type is specified in Annex C.
- MPEG-4 visual simple profile level 0 MIME media type as specified in RFC 3016 [5].
- T.140 Text Conversation MIME media type as specified by RFC 2793 [24].

9.2 RTP payload

RTP payload formats specified by IETF shall be used for real time media streams.

RTP payload format for the AMR narrowband speech codec is specified in Annex D.

RTP payload format for the AMR wideband speech codec is specified in Annex B.

RTP payload format for the ITU-T H.263 video codec is specified in IETF RFC 2429 [3].

RTP payload format for the MPEG-4 visual simple profile level 0 is specified in IETF RFC 3016 [5].

RTP payload format for the ITU-T T.140 text conversation coding is specified in IETF RFC 2793 [24].

Annex A (informative): Information on optional enhancements

The section is intended for informational purposes only. This is not an integral part of this specification.

A.1 Video

This section gives recommendations for the video codec implementations within 3G PS multimedia terminals.

Regardless of which specific video codec standard is used, all video decoder implementations should include basic error concealment techniques. These techniques may include replacing erroneous parts of the decoded video frame with interpolated picture material from previous decoded frames or from spatially different locations of the erroneous frame. The decoder should aim to prevent the display of substantially corrupted parts of the picture. In any case, it is recommended that the terminal should tolerate *every* possible bitstream without catastrophic behaviour (such as the need for a user-initiated reset of the terminal).

3G PS terminal video encoders and decoders are recommended to support the 1:1 pixel format (square format).

A.1.1 H.263 video codec

H.263 was approved as a standard in 1996. Since then, version 2 and version 3 enhancing version 1 have been approved in 1998 and 2000 respectively. As of today, H.263 contains an extensive set of mandatory and optional coding tools. H.263 Annex X (going to be approved in 2001) defines codec profiles for various target environments.

The Baseline Profile (Profile 0) stands for H.263 with no optional modes of operation. It includes the basic coding tool set common in modern video coding standards. It provides simple means to insert resynchronisation points within the video bitstream, and, therefore, it enables recovery from erroneous or lost data.

The Version 2 Interactive and Streaming Wireless Profile (Profile 3) provides enhanced compression efficiency when compared to the Baseline Profile. Moreover, it provides enhanced error resilience for delivery to wireless devices. Specifically, Profile 3 includes the following optional coding modes:

1. Advanced INTRA Coding (Annex I). Use of this mode improves the compression efficiency for INTRA macroblocks (whether within INTRA pictures or predictively-coded pictures).

2. Deblocking Filter (Annex J). A deblocking filter improves image quality by reducing blocking artifacts. When compared to deblocking filtering performed as a postprocessing operation, the Deblocking Filter Mode reduces the amount of required memory, as no additional picture memory is needed for the filtered images. This mode also includes the four-motion-vector-per-macroblock feature and picture boundary extrapolation for motion compensation, both of which can further improve compression efficiency.

3. Slice Structured Mode (Annex K). This mode provides a flexible mechanism to insert resynchronisation points within the video bitstream for recovery from erroneous or lost data.

4. Modified Quantisation (Annex T). This mode enables flexible quantiser control that can be used in sophisticated bit-rate control algorithms. In addition, it improves chrominance fidelity.

[FFS]

A.2 Audio

[FFS]

A.3 Text

Use of the redundancy coding variant specified in RFC 2793 [24] is recommended for error resilience.

Annex B (normative): AMR-WB RTP payload and MIME type registration

This section specifies the AMR-WB speech codec RTP payload and MIME type registration.

Note: The intention is to replace this normative annex with the IETF RFC defining the AMR Wideband RTP payload and MIME media type registration when the RFC is available.

B.1 AMR-WB RTP payload

The AMR-WB payload format supports transmission of multiple frames per payload, the use of fast codec mode adaptation, and robustness against packet losses and bit errors.

The AMR-WB payload format consists of one payload header, a table of content, optionally one CRC per payload frame, and zero or more AMR-WB payload frames. The payload format is made as bandwidth efficient as possible by not using octet alignment for the payload header, table of content or the payload frames. However, the full payload is octet aligned.

B.1.1 Payload header

The length of the payload header is either 7 or 15 bits, depending on whether the interleaving is used or not. Figures B.1a and B.1b illustrate the header structure. Header bits are specified in following two subclauses.

B.1.1.1 Required fields of the payload header

S (1 bit): Indicates, if set, that the bits in the payload is robust sorted. If not set, simple payload sorting is employed. Note that this bit can be set only if the receiver has signalled support for the OPTIONAL robust payload sorting.

C (1 bit): Indicates the existence of OPTIONAL CRC fields in the payload table of content. Note that this bit can be set only if the receiver has signalled support for the OPTIONAL CRC.

I (1 bit): Indicates, if set, that frames in this payload are interleaved, and that ILL and ILP fields are present in the payload header. If not set, frames in this payload are successive frames and ILL and ILP fields are not present in the payload header. Note that this bit can be set only if the receiver has signalled support for interleaving.

CMR (4 bits): Indicates Codec Mode Requested for the other communication direction. It is only allowed to request one of the AMR-WB speech modes (frame type index 0...8, see [18]). CMR value 15 indicates that no mode request is present, other values are for future use.

B.1.1.2. Optional fields of the payload header

ILL (4 bits): OPTIONAL field that is present only if I=1. The value of this field specifies the interleaving length used for frames in this payload.

ILP (4 bits): OPTIONAL field that is present only if I=1. The value of this field indicates the interleaving index for frames in this payload. The value of ILP MUST be smaller than or equal to the value of ILL. Erroneous value of ILP SHOULD cause the payload to be discarded.

The value of the ILL field defines the length of an interleave group: ILL=L implies that frames in (L+1)-frame intervals are picked into the same interleaved payload, and the interleave group consists of L+1 payloads. The value of ILP=p in payloads belonging to the same group runs from 0 to L. The interleaving is meaningful only when number of frames per payload N is greater than or equal to 2. Thus, when N frames are transmitted in each payload of a group, the interleave group consists of payloads with sequence numbers s...s+L, and frames encapsulated into these payloads are f...f+N*(L+1)-1.

To put this in a form of an equation, let's assume that the first frame of an interleave group is n, the first payload of the group is s, number of frames per payload is N, ILL=L and ILP=p (p in range 0...L), the frames contained by the payload s+p are n + p + k*(L+1), where k runs from 0 to N-1.

Interleaved frames MUST be stored in the payload in timestamp-increasing order. Furthermore, the interleaved payloads within an interleave group MUST be sent according to increasing order of ILP field, and each payload of an interleave group MUST contain equal number of frames. It is RECOMMENDED that ILL remains constant throughout the session. If ILL are to be changed, the change MUST be done between interleaving groups, i.e. the ILP of the previous packet was L. Furthermore, because of the inter-frame dependent nature of AMR-WB coding, it is RECOMMENDED that ILL values greater than or equal to 2 are used to enable better error recovery in the decoder in case of lost interleaved payload. Note also that using value ILL=0 or using interleaving for payload carrying only one frame is not meaningful.

```
 0
 0 1 2 3 4 5 6
+-+-+-+-+-+-+-+
|S|C|I|  CMR  |
+-+-+-+-+-+-+-+
```

Figure B.1a: AMR-WB payload header, I=0.

```
 0                   1
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|S|C|I|  CMR  |  ILL  |  ILP  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure B.1b: AMR-WB payload header, I=1.

B.1.2. The payload table of content and CRCs

The table of content (ToC) consists of one table of content entry for each speech frame in the payload. A table of content entry includes several specified fields as follows:

F (1 bit): Indicates if this frame is followed by further frames in this payload. F=1 further frames follow, F=0 last frame.

FT (4 bits): Frame type indicator, indicating the AMR-WB speech coding mode or comfort noise (CN) mode. The mapping of AMR-WB modes to FT is given in Table 1a in [18]. If FT=14 (lost frame) or FT=15 (no transmission/no reception), no CRC or payload frame is present.

Q (1 bit): The frame quality bit indicates, if not set, that the payload is corrupted and the receiver should set the RX_TYPE (see [18]) to SPEECH_BAD or SID_BAD depending on the frame type (FT).

```
 0
 0 1 2 3 4 5
+-+-+-+-+-+-+
|F|   FT  |Q|
+-+-+-+-+-+-+
```

Figure B.2: Table of content (ToC) entry field.

CRC (8 bits): OPTIONAL field, exists if the payload header bit C is set (C=1). The 8 bit CRC is used for error detection. These 8 parity bits are generated according to [18].

```
 0
 0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|      CRC      |
+-+-+-+-+-+-+-+-+
```

Figure B.3: CRC field.

The ToC and CRCs are arranged with all table of content entries fields first followed by all CRC fields. The ToC starts with the frame data belonging to the oldest speech frame in the payload.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|F|  FT   |Q|F|  FT   |Q|F|  FT   |Q|      CRC      |      CRC  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   |      CRC      |
+-+-+-+-+-+-+-+-+-+-+
```

Figure B.4: The ToC and CRCs for a payload with three speech frames.

B.1.3. AMR-WB speech frame

An AMR-WB speech frame represents one encoded speech frame encoded using the mode according to the FT field in ToC entry corresponding to this frame. The length of this field is implicitly defined by the AMR-WB mode in the FT field. The AMR-WB speech bits SHALL be sorted according to [18].

B.1.4. Compound AMR-WB payload

The compound AMR-WB payload consists of one AMR-WB payload header, the table of content, and one or more AMR-WB payload frames. These can be combined either by using robust or simple payload sorting. The S-bit in the AMR-WB payload header indicates which method is used.

Definitions for describing the compound AMR-WB payload:

b(m) - bit m of the compound AMR-WB payload
t(n,m) - bit m in the table of content entry for speech frame n
p(n,m) - bit m in the CRC for speech frame n
f(n,m) - bit m in speech frame n
F(n) - number of bits in speech frame n, defined by FT
h(m) - bit m of payload header
H - number of bits in payload header, 7 or 15 bits
C - number of CRC bits , 0 or 8 bits
N - number of payload frames in the payload
S - number of unused bits in the last octet of the payload

Payload frames f(n,m) are ordered in the order they are delivered by the AMR-WB speech encoder, i.e. frame n is preceding frame n+1. All frames between the oldest one and the most recent one must be present in the payload, the only exception is interleaving, when the frame order is defined in B.1.1.2. If some of the frames are not available because of a frame loss or they are not transmitted due to DTX, those MUST be replaced by lost speech or by no transmission/no reception type frames, respectively.

B.1.4.1. Robust payload sorting

A bit error in a more sensitive bit is subjectively more annoying than in a less sensitive bit. Therefore, to enable protection of only the most sensitive bits of a payload with a forward error detection code, e.g. a CRC outside RTP, the bits inside a payload can be ordered into sensitivity order. The protection SHOULD cover an appropriate number of octets from the beginning of the payload, covering at least the AMR-WB payload header, ToC, and class A bits. Exactly how many octets that needs protection depends on the network and application. To maintain sensitivity ordering inside the AMR payload, when more than one speech frame is transmitted in one payload, reordering of the bits in the payload is needed.

The AMR-WB payload header, ToC and CRCs SHALL still be placed unchanged in the beginning of the robust sorted payload. Thereafter, the payload frames are sorted with one bit alternating from each AMR-WB payload frame.

The robust payload sorting algorithm is defined in C-style as:

```
/* payload header */
```

```
k=0;
for (i = 0; i < H; i++){
  b(k++) = h(i);
}
/* table of content */
for (j = 0; j < N; j++){
  for (i = 0; i < 6; i++){
    b(k++) = t(j,i);
  }
}
/* CRCs */
for (j = 0; j < N; j++){
  for (i = 0; i < C; i++){
    b(k++) = p(j,i);
  }
}
/* payload frames */
max = max(F(0),..,F(N-1));
for (i = 0; i < max; i++){
  for (j = 0; j < N; j++){
    if (i < F(j)){
      b(k++) = f(j,i);
    }
  }
}
/* padding */
S = 8 - k%8;
if (S < 8){
  for (i = 0; i < S; i++){
    b(k++) = 0;
  }
}
```

B.1.4.2 Simple payload sorting

If multiple frames are encapsulated into the payload and robust payload sorting is not used, the payload is formed as concatenation of the AMR-WB payload header, ToC, possibly optional CRC fields, and the AMR-WB speech frames. However, the bits inside each AMR-WB payload frame are ordered into sensitivity order as defined in [18].

The simple payload sorting algorithm is defined in C-style as:

```
/* payload header */
k=0;
for (i = 0; i < H; i++){
  b(k++) = h(i);
}
/* table of content */
for (j = 0; j < N; j++){
  for (i = 0; i < 6; i++){
    b(k++) = t(j,i);
  }
}
/* CRCs */
for (j = 0; j < N; j++){
  for (i = 0; i < C; i++){
    b(k++) = p(j,i);
  }
}
/* payload frames */
```

```
for (j = 0; j < N; j++){
  for (i = 0; i < F(j); i++){
      b(k++) = f(j,i);
    }
  }
}
/* padding */
S = 8 - k%8;
if (S < 8){
  for (i = 0; i < S; i++){
    b(k++) = 0;
  }
}
```

B.1.5. Simple example

In the simple example one AMR-WB frame is encapsulated into the payload. Simple payload sorting is used (S=0), no CRC fields are present (C=0), and interleaving is not used (I=0). A 23.05 kbps mode is requested for the reverse link (CMR=7), and the payload was not damaged at IP origin (Q=1). The AMR-WB mode is the 12.65 kbps mode (FT=2). The speech encoded bits are put into f(0...252) in descending sensitivity order according to [18].

```
   |                           Bit no.                             |
Oct|   0       1       2       3       4       5       6       7   |
---+-------+-------+-------+-------+-------+-------+-------+-------+
 0 |  S=0  |  C=0  |  I=0  |   0   |   1   |   1   |   1   |  F=0  |
---+-------+-------+-------+-------+-------+-------+-------+-------+
 1 |   0   |   0   |   1   |   0   |  Q=1  | f(0)  | f(1)  |  ...  |
---+-------+-------+-------+-------+-------+-------+-------+-------+
32 |  ...  |  ...  |  ...  |  ...  |  ...  |  ...  |f1(249)|f1(250)|
---+-------+-------+-------+-------+-------+-------+-------+-------+
33 | f(251)| f(252)|   0   |   0   |   0   |   0   |   0   |   0   |
---+-------+-------+-------+-------+-------+-------+-------+-------+
```

Figure B.5: One AMR-WB frame per payload example.

B.2 The AMR-WB MIME type registration

This chapter defines the MIME type for the Adaptive Multi-Rate Wideband (AMR-WB) speech codec. AMR-WB implementations according to [17] MUST support all nine coding modes. The fast mode adaptation is supported by transmitting the mode information in-band together with encoded speech data to allow mode change without any additional signaling. Furthermore, fast mode adaptation requires transmission of codec mode request inside payload.

In addition to the speech codec, AMR-WB specifications also include Discontinuous Transmission / comfort noise (DTX/CN) functionality. The DTX/CN switches the transmission off during silent periods of the speech and only SID frames containing CN parameter updates are sent at regular intervals. Also the AMR-WB DTX/CN MUST be supported.

It is possible that the receiver may only want to receive a certain AMR-WB mode or a subset of AMR-WB modes, due to link limitations in some cellular systems, e.g. the GSM/GERAN radio link can require that only a subset of AMR-WB modes is used. Therefore, it is possible to request a specific set of AMR-WB modes in capability description and the encoder MUST abide this request. If the request for mode set is not given, any mode may be used or requested.

The AMR-WB codec can in principle perform a mode change at any time between any two modes. To support interoperability with GSM through a gateway it is possible to set limitations for mode changes. The decoder has possibility to define the minimum number of frames between mode changes and to limit the mode change to happen into neighboring modes only.

The receiver can limit the number of AMR-WB frames encapsulated into one RTP packet, and if maximum number of frames per packet is given in capability description, the transmitter MUST comply with this limitation. This is an

OPTIONAL feature and if no parameter is given in capability description, the transmitter can encapsulate any number of AMR-WB speech frames into one RTP packet.

The payload CRC UED MUST only be used if the receiver has signalled support for this functionality in the capability description.

To enable unequal error protection and/or detection outside RTP, the payload format supports robust payload sorting. The robust payload sorting is an optional feature and MUST only be used if the receiver has signalled support for this functionality in the capability description.

The speech quality in case of packet losses when transmitting several AMR-WB frames per packet can be improved by using OPTIONAL frame interleaving. The interleaving improves perceived speech quality since it introduces series of single frame errors instead of several consecutive frame errors. Interleaving MUST only be applied if the receiver has signalled support for it, and if used, the interleaving length MUST NOT exceed the limitation given in capability description. Note that the receiver can use the MIME parameters to limit increased buffering requirements caused by the interleaving. For example specifying maxframes=N and interleaving=L, the maximum size of an interleave group would be N*(L+1).

B.2.1 MIME Registration

MIME-name for the AMR-WB codec is allocated from IETF tree since AMR-WB is expected to be widely used speech codec in VoIP applications.

Media Type name: audio

Media subtype name: AMR-WB

Required parameters: none

Optional parameters:

mode-set: Requested AMR-WB mode set. Restricts the active codec mode set to a subset of all modes. Possible values are comma separated list of modes: 0,...,8 (see [18]). If not present, all speech modes are available.

mode-change-period: Defines a number N which restricts the mode changes in such a way that mode changes are only allowed on multiples of N, initial state of the phase is arbitrary. If this parameter is not present, mode change can happen at any time.

mode-change-neighbor: If present, mode changes SHALL only be made to neighboring modes in the active codec mode set. If not present, change between any two modes in the active codec mode set is allowed.

maxframes: Maximum number of AMR speech frames in one RTP packet. The receiver may set this parameter in order to limit the buffering requirements or delay.

crc: If present, transmission of CRCs in the payload is supported, otherwise not supported.

robust-sorting: If present, robust payload sorting is supported, otherwise not supported and simple payload sorting SHALL be used.

Interleaving: Indicates that the frame interleaving is supported and defines a maximum value for interleaving length field ILL. If this parameter is not present, the interleaving is not supported.

B.2.2 Mapping to SDP Parameters

Parameters are mapped to SDP as usual.

Example usage in SDP:

```
m=audio 49120 RTP/AVP 97
a=rtpmap:97 AMR-WB/16000
a=fmtp:97 mode-set=2,3,4,5,6; maxframes=1
```

Annex C (normative): ITU-T H.263 MIME media type registration

Note: The intention is to replace this normative annex with the IETF RFC defining the H.263 video codec MIME media type registration when the RFC is available.

H.263 video codec MIME media type is specified as follows:
MIME media type name: video
MIME subtype name: H263-2000
Required parameters: None
Optional parameters:
profile: H.263 profile number, in the range 0 through 8, specifying the supported H.263 annexes/subparts.
level: Level of bitstream operation, in the range 0 through 99, specifying the level of computational complexity of the decoding process. When profile and level parameters are not specified, Baseline Profile (Profile 0) Level 10 are the default values.
The profile and level specifications can be found in [19]. Note that the RTP payload format for H263-2000 is the same as for H263-1998 published in RFC 2429, but additional annexes/subparts are specified along with the profiles and levels.

Annex D (normative): AMR-NB RTP payload and MIME type registration

This section specifies the AMR-NB speech codec RTP payload and MIME type registration.

Note: The intention is to replace this normative annex with the IETF RFC defining the AMR Narrowband RTP payload and MIME media type registration when the RFC is available.

D.1 AMR-NB RTP payload

The AMR-NB payload format supports transmission of multiple frames per payload, the use of fast codec mode adaptation, and robustness against packet losses and bit errors.

The AMR payload format is designed to be flexible, ranging from very low overhead to an extended format with the possibility to increase bit error robustness and pack several speech frames in one packet.

The payload format consists of one payload header, a table of content, optionally one CRC per payload frame and zero or more payload frames. The payload format is bandwidth efficient. This is achieved by not using octet alignment for the payload header, table of content or the payload frames, but the full payload is octet aligned. If the option to transmit a robust sorted payload is enabled and employed, the full payload SHALL finally be ordered in descending bit error sensitivity order to be prepared for unequal error protection or unequal error detection schemes. The AMR encoded bit streams are defined in sensitivity order in Annex B of [2], the original order as delivered from the speech encoder is defined in [1]. The last octet of an AMR payload packet MUST be padded with zeroes at the end if not all bits are used.

The AMR frame types, or modes, are defined in [2]. Frame type 15, no transmission, is needed to indicate not transmitted frames or lost frames. Not transmitted could mean both no data produced by the speech encoder for this frame or no data transmitted in this payload, i.e. valid data for this frame could be sent in another payload. For example, when multiple frames are sent in each payload and comfort noise starts. A frame type sequence in a payload with 8 frames, speech frames with AMR mode 7 are interrupted by CN in the fifth frame, could look like: {7,7,7,7,8,15,15,8}. The AMR SCR/DTX is described in [4].

The AMR payload format supports robust transmission, multiple frames in one payload packet, and the use of fast codec mode adaptation.

Robustness against packet loss can be accomplished by using the possibility to retransmit previously transmitted frames together with the current frame or frames.

The AMR performance over error tolerant links can be be improved by delivering also speech frames with bit errors. Unequal error detection is needed since bit errors SHOULD only be allowed in the least error sensitive bits.

D.1.1 The payload header

The length of the payload header is 6 bits. The bits in the header are specified as follows:

S (1bit): Indicates if set that the payload is robust sorted, otherwise simple payload sorting is employed. Note that this bit can be set only if the receiver has signaled support for the OPTIONAL robust payload sorting.

C (1 bit): Indicates the existence of optional CRC fields in the payload table of content. Note that this bit can be set only if the receiver has signaled support for the OPTIONAL CRC.

R (1 bit): Indicates, if set, that the Codec Mode Request (CMR) is valid.

CMR (3 bits): this field is only valid if the R bit is set(R=1). Codec Mode Requested (CMR) for the other communication direction. It is only allowed to request the one of the speech modes, frame type index 0-7 see Table 1a in [2]. If R=0 the CMR bits SHALL be set to zero, other values are for future use.

```
0
0 1 2 3 4 5
```

```
+-+-+-+-+-+-+

|S|C|R| CMR |

+-+-+-+-+-+-+
```

Figure D.1: AMR payload header

D.1.2 The payload table of content and CRCs

The table of content (ToC) consists of one table of content entry for each speech frame in the payload. A table of content entry includes several specified fields as follows:

F (1 bit): Indicates if this frame is followed by further frames. F=1 further frames follow, F=0 last frame.

Q (1 bit): The payload quality bit indicates, if not set, that the payload is severely damaged and the receiver should set the RX_TYPE, see [4], to SPEECH_BAD or SID_BAD depending on the frame type (FT).

FT (4 bits): Frame type indicator, indicating the AMR speech coding mode or comfort noise (CN) mode. The mapping of existing AMR modes to FT is given in Table 1a in [2]. If FT=15 (No transmission) no CRC or payload frame is present.

```
 0

 0 1 2 3 4 5

+-+-+-+-+-+-+

|F|Q|  FT   |

+-+-+-+-+-+-+
```

Figure D.2: Table of content entry field

CRC (8 bits): OPTIONAL field, exists if the payload header bit C is set (C=1). The 8 bit CRC is used for error detection. These 8 parity bits are generated according to section 4.1.4 in [2].

```
 0

 0 1 2 3 4 5 6 7

+-+-+-+-+-+-+-+-+

|      CRC      |

+-+-+-+-+-+-+-+-+
```

Figure D.3: CRC field

The ToC and CRCs are arranged with all table of content entries fields first followed by all CRC fields. The ToC starts with the frame data belonging to the oldest speech frame.

```
 0                   1                   2                   3

 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

|F|Q|  FT   |F|Q|  FT   |F|Q|  FT   |      CRC      |      CRC  |

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

```
|   |      CRC      |
+-+-+-+-+-+-+-+-+-+-+-+
```

Figure D.4: The ToC and CRCs for a payload with three speech frames

D.1.3 AMR speech frame

An AMR speech frame represent one encoded speech frame encode with the mode according to the ToC field FT. The length of this field is implicitly defined by the AMR mode in the FT field. The bits SHALL be sorted according to Appendix B of [2].

D.1.4. Compound AMR payload

The compound AMR payload consists of one AMR payload header, the table of content and one or more AMR payload frames. These can be put together with robust or simple payload sorting. The payload header bit S indicates the method used.

Definitions for describing the compound AMR payload:

b(m) - bit m of the compound AMR payload
t(n,m) - bit m in the table of content entry for speech frame n
p(n,m) - bit m in the CRC for speech frame n
f(n,m) - bit m in speech frame n
F(n) - number of bits in speech frame n, defined by FT
h(m) - bit m of payload header
C - number of CRC bits , 0 or 8 bits
N - number of payload frames in the payload
S - number of unused bits

Payload frames f(n,m) are ordered in consecutive order, where frame n=1 is preceding frame n=2. Within one payload all frames between the oldest and most recent must be present. If speech data is missing for one frame, due to e.g. DTX, send the NO_TRANSMISSION frame type.

D.1.4.1. Robust payload sorting

A bit error in a more sensitive bit is subjectively more annoying than in a less sensitive bit. Therefore, to be able to protect only the most sensitive bits in a payload packet with a forward error detection code, e.g. a CRC outside RTP, the bits inside a frame are ordered into sensitivity order. The protection SHOULD cover an appropriate number of octets from the beginning of the payload, covering at least the AMR payload header, ToC and class A bits (see [2]). Exactly how many octets that needs protection depends on the network and application. To maintain sensitivity ordering inside the AMR payload, when more than one speech frame is transmitted in one payload, reordering of the data is needed.

The reordering to maintain the sensitivity ordered AMR payload SHALL be performed on bit level. The AMR payload header, ToC and CRCs SHALL still be placed unchanged in the beginning of the payload. Thereafter, the payload frames are sorted with one bit alternating from each payload frame.

The robust payload sorting algorithm is defined in C-style as:

```
/* payload header */
k=0;
for (i = 0; i < 6; i++){
  b(k++) = h(i);
```

```
   }
   /* table of content */
   for (j = 0; j < N; j++){
     for (i = 0; i < 6; i++){
       b(k++) = t(j,i);
     }
   }

/* CRCs */
   for (j = 0; j < N; j++){
     for (i = 0; i < C; i++){
       b(k++) = p(j,i);
     }
   }
   /* payload frames */
   max = max(F(0),..,F(N-1));
   for (i = 0; i < max; i++){
     for (j = 0; j < N; j++){
       if (i < F(j)){
         b(k++) = f(j,i);
       }
     }
   }
   /* padding */
   S = 8 - k%8;
   if (S < 8){
     for (i = 0; i < S; i++){
       b(k++) = 0;
     }
   }
```

D.1.4.2. Simple payload sorting

If multiple new frames are encapsulated into the payload and robust payload sorting is not used. The payload is formed by concatenating the payload header, the ToC, optional CRC fields and the speech frames in the payload. However, the bits inside a frame are ordered into sensitivity order as defined in [2].

The simple payload sorting algorithm is defined in C-style as:

```
   /* payload header */
   k=0;
   for (i = 0; i < 6; i++){
     b(k++) = h(i);
   }
   /* table of content */
   for (j = 0; j < N; j++){
     for (i = 0; i < 6; i++){
       b(k++) = t(j,i);
     }
   }
   /* CRCs */
   for (j = 0; j < N; j++){
     for (i = 0; i < C; i++){
       b(k++) = p(j,i);
     }
   }

/* payload frames */
   for (j = 0; j < N; j++){
     for (i = 0; i < F(j); i++){
         b(k++) = f(j,i);
       }
     }
   }
   /* padding */
   S = 8 - k%8;
   if (S < 8){
     for (i = 0; i < S; i++){
       b(k++) = 0;
     }
   }
```

D.2 RTP header usage

The RTP header marker bit (M) is used to mark (M=1) the packages containing the first speech frame after CN. For all other packages the marker bit is set to 0 (M=0).

The timestamp corresponds to the sampling instant of the first sample encoded for the first frame in the packet. A frame can be either encoded speech, comfort noise parameters, or NO_TRANSMISSION. The timestamp unit is in samples. The duration of one AMR speech frame is 20 ms and the sampling frequency is 8 kHz, corresponding to 160 encoded speech samples per frame. Thus, the timestamp is increased by 160 for each consecutive frame. All frames in a packet MUST be successive 20 ms frames.

D.3 Examples

D.3.1 Simple example

In the simple example we just send one frame in each RTP packet, no valid Codec Mode Request CMR is sent (R=0), the payload was not damaged at IP origin (Q=1) and no CRC is used. The AMR mode is the 5.9 kbps mode (FT=2). The speech encoded bits are put into f(0) to f(117) in descending sensitivity order according to [2]. Simple payload sorting is used, S=0.

```
   |                           Bit no.                           |
Oct|   0       1       2       3       4       5       6       7   |
---+-------+-------+-------+-------+-------+-------+-------+-------+
 0 |  S=0  |  C=0  |  R=0  |   0   |   0   |   0   |  F=0  |  Q=1  |
---+-------+-------+-------+-------+-------+-------+-------+-------+
 1 |   0   |   0   |   1   |   0   | f(0)  | f(1)  | f(2)  |  ...  |
---+-------+-------+-------+-------+-------+-------+-------+-------+
16 | f(116)| f(117)|   0   |   0   |   0   |   0   |   0   |   0   |
---+-------+-------+-------+-------+-------+-------+-------+-------+
```

Figure D.5: One frame per packet example.

D.3.2 Example with CRCs

In this example the two frames with 6.7 kbps mode (FT=3) are sent in the payload. A mode request is sent(R=1), requesting the 10.2 kbps mode for the other link(CMR=6). CRC is used (C=1). Frame one (134 bits) is f1(0..133) and frame 2 f2(0..133). For each payload frame a CRC is calculated p1(0..7) for frame 1 and p2(0..7) for frame 2. Simple payload sorting is used, S=0.

```
   |                           Bit no.                           |
Oct|   0       1       2       3       4       5       6       7   |
---+-------+-------+-------+-------+-------+-------+-------+-------+
 0 |  S=0  |  C=1  |  R=1  |   1   |   1   |   0   |  F=1  |  Q=1  |
---+-------+-------+-------+-------+-------+-------+-------+-------+
```

```
 1 |   0   |   0   |   1   |   1   |  F=0  |  Q=1  |   0   |   0   |
---+-------+-------+-------+-------+-------+-------+-------+-------+
 2 |   1   |   1   | p1(0) | p1(1) | p1(2) | p1(3) | p1(4) | p1(5) |
---+-------+-------+-------+-------+-------+-------+-------+-------+
 3 | p1(6) | p1(7) | p2(0) | p2(1) | p2(2) | p2(3) | p2(4) | p2(5) |
---+-------+-------+-------+-------+-------+-------+-------+-------+
 4 | p2(6) | p2(7) | f1(0) | f1(1) |  ...  |  ...  |  ...  |  ...  |
---+-------+-------+-------+-------+-------+-------+-------+-------+
20 |  ...  |  ...  |  ...  |  ...  |  ...  |  ...  |f1(132)|f1(133)|
---+-------+-------+-------+-------+-------+-------+-------+-------+
21 | f2(0) | f2(1) |  ...  |  ...  |  ...  |  ...  |  ...  |  ...  |
---+-------+-------+-------+-------+-------+-------+-------+-------+
37 |  ...  |  ...  |  ...  |f2(131)|f2(132)|f2(133)|   0   |   0   |
---+-------+-------+-------+-------+-------+-------+-------+-------+
```

Figure D.6: Example with CRCs.

D.3.3 Example with multiple frames per payload and robust sorting

In this example two 5.9 kbps mode (FT=2) frames are sent in one payload. No CRC is used (C=0). A mode request is sent(R=1), requesting the 7.95 kbps mode for the other link(CMR=5). The first frame is represented by the 118 bits f(0) to f(117) and the subsequent frame by g(0) to g(117). Robust sorting is used.

```
   |                            Bit no.                            |
Oct|   0       1       2       3       4       5       6       7   |
---+-------+-------+-------+-------+-------+-------+-------+-------+
 0 |  S=1  |  C=0  |  R=1  |   1   |   0   |   1   |  F=1  |  Q=1  |
---+-------+-------+-------+-------+-------+-------+-------+-------+
 1 |   0   |   0   |   1   |   0   |  F=0  |  Q=1  |   0   |   0   |
---+-------+-------+-------+-------+-------+-------+-------+-------+
 2 |   1   |   0   | f(0)  | g(0)  | f(1)  | g(1)  |  ...  |  ...  |
---+-------+-------+-------+-------+-------+-------+-------+-------+
31 |  ...  |  ...  | f(116)| g(116)| f(117)| g(117)|   0   |   0   |
---+-------+-------+-------+-------+-------+-------+-------+-------+
```

Figure D.7: Example two frames per payload and robust sorting.

D.4 The AMR MIME type registration

This chapter defines the MIME type for the Adaptive Multi-Rate (AMR) speech codec [1]. The data format and parameters are specified for both real-time transport and for storage type applications (e.g. e-mail attachment, multimedia messaging). The former is referred as RTP mode and the latter as storage mode.

AMR implementations according to [1] MUST support all eight coding modes. The mode change can occur at any time during operation and therefore the mode information is transmitted in-band together with speech bits to allow mode change without any additional signalling.

In addition to the speech codec, AMR specifications also include Discontinuous Transmission / comfort noise (DTX/CN) functionality [11]. The DTX/CN switches the transmission off during silent parts of the speech and only CN parameter updates are sent at regular intervals.

D.4.1 RTP mode

It is possible that the decoder may want to receive a certain AMR mode or a subset of AMR modes, due to link limitations in some cellular systems, e.g. the GSM radio link can only use a subset of maximum four modes. Therefore, it is possible to request a specific set of AMR modes in capability description and the encoder MUST abide this request. If the request for mode set is not given any mode may be used or requested.

The AMR codec can in principle perform a mode change at any time between any two modes. To support interoperability with GSM through a gate-way it is possible to set limitations for mode changes. The decoder has possibility to define the minimum number of frames between mode changes and to limit the mode change to happen into neighboring modes only.

It is also possible to limit the number of AMR frames encapsulated into one RTP packet. This is an optional feature and if no parameter is given in capability description, the transmitter can encapsulate any number of AMR speech frames into one RTP packet.

The payload CRC UED MUST only be used if the receiver has signalled support for this functionality in the capability description.

To support unequal error protection and/or detection the payload format supports robust payload sorting. The robust payload sorting is an OPTIONAL feature and MUST only be used if the receiver has signalled support for this functionality in the capability description.

D.4.2 Storage mode

The AMR storage mode is used for storing AMR frames, e.g. as a file or e-mail attachment. Frames are stored in consecutive order in octet aligned manner. This implies that the first octet after the last octet of frame n must be the first octet of frame n+1. Each stored AMR frame consists of a Q bit and the 4-bit FT field, followed by the AMR encoded speech bits. The last octet of each frame is padded with zeroes, if needed, to achieve octet alignment. An example is given in figure D.8.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Q|  FT   |                                                     |
+-+-+-+-+-+                                                     +
|                                                               |
+                AMR speech bits for frame n                    +
|                                                               |
```

```
+                                                           +-+-+-+-+-+
|                                                           | Padding |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Q|  FT   |                                                           |
+-+-+-+-+-+                                                           +
|                                                                     |
+                AMR speech bits for frame n+1                        +
|                                                                     |
+                                                           +-+-+-+-+-+
|                                                           | Padding |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure D.8: An example of storage format with two AMR 5.9 kbit/s frames (118 speech bits). Note that bits marked as 'padding' must be set to zero.

Frames lost in transmission and non-received frames between SID updates during non-speech period must be stored as NO_TRANSMISSION frames (frame type 15, see definition in [2]) to keep synchronization with the original media.

The receiving entity (AMR decoder) MUST be able to decode all eight coding modes as well as the AMR DTX/CN [6]. Since no exchange of particular coding considerations can be signalled before downloading or receiving stored AMR data, the optional features (robust sorting, CRC) specified for RTP mode MUST NOT be used with storage mode.

D.4.3 MIME Registration

MIME-name for the AMR codec is allocated from IETF tree since AMR is expected to be widely used speech codec in VoIP applications. Some parts of this chapter will distinguish between RTP and storage modes.

Media Type name: audio

Media subtype name: AMR

Required parameters: none

Optional parameters for RTP mode:

mode-set: Requested AMR mode set. Restricts the active codec mode set to a subset of all modes. Possible values are comma separated list of modes: 0,...,7 (see Table 1a [2] an example is given in section 8.4). If not present, all speech modes are available.

mode-change-period: Defines a number N which restricts the mode changes in such a way that mode changes are only allowed on multiples of N, initial state of the phase is arbitrary. If this parameter is not present, mode change can happen at any time.

mode-change-neighbor: If present, mode changes SHALL only be made to neighboring modes in the active codec mode set. If not present, change between any two modes in the active codec mode set is allowed.

maxframes: Maximum number of AMR speech frames in one RTP packet. The receiver may set this parameter in order to limit the buffering requirements or delay.

crc: If present, transmission of CRCs in the payload is supported, otherwise not supported.

robust-sorting: If present, robust payload sorting is supported, otherwise not supported and simple payload sorting SHALL be used.

Optional parameters for storage mode: none

Additional information for storage mode:
Magic number: none
File extensions: amr, AMR
Macintosh file type code: none
Object identifier or OID: none

D.4.4 Mapping to SDP Parameters

Please note that this chapter applies to the RTP mode only.

Parameters are mapped to SDP as usual.
Example usage in SDP:
m=audio 49120 RTP/AVP 97
a=rtpmap:97 AMR/8000
a=fmtp:97 mode-set=0,2,5,7; maxframes=1

D.5 References

[1] 3G TS 26.090, "Adaptive Multi-Rate (AMR) speech transcoding".

[2] 3G TS 26.101, "AMR Speech Codec Frame Structure".[3] IETF RFC 2119, "Key words for use in RFCs to Indicate Requirement Levels".

[4] 3G TS 26.093, "AMR Speech Codec; Source Controlled Rate operation".

[5] GSM 06.60, "Enhanced Full Rate (EFR) speech transcoding".

[6] TIA/EIA -136-Rev.A, part 410 - "TDMA Cellular/PCS – Radio Interface, Enhanced Full Rate Voice Codec (ACELP). Formerly IS-641. TIA published standard, 1998".

[7] ARIB, RCR STD-27H, "Personal Digital Cellular Telecommunication System RCR Standard".

[8] IETF RFC1889, "RTP: A Transport Protocol for Real-Time Applications".

[9] IETF draft-westberg-realtime-cellular-01.txt, "Realtime Traffic over Cellular Access Networks".

[10] IETF draft-larzon-udplite-03.txt, "The UDP Lite Protocol".

[11] GSM 06.92, "Comfort noise aspects for Adaptive Multi-Rate (AMR) speech traffic channels".

[12] M. Handley and V. Jacobson, "SDP: Session Description Protocol", RFC 2327, April 1998

Annex E (informative): Change history

Change history							
Date	TSG #	TSG Doc.	CR	Rev	Subject/Comment	Old	New