ETSI TS 126 193 V17.0.0 (2022-05)

Digital cellular telecommunications system (Phase 2+) (GSM);
Universal Mobile Telecommunications System (UMTS);
LTE;
5G;
Speech codec speech processing functions;
Adaptive Multi-Rate - Wideband (AMR-WB) speech codec;
Source controlled rate operation
(3GPP TS 26.193 version 17.0.0 Release 17)

Reference
RTS/TSGS-0426193vh00

Keywords
5G,GSM,LTE,UMTS

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - APE 7112B
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° w061004871

Important notice

The present document can be downloaded from:
http://www.etsi.org/standards-search

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the prevailing version of an ETSI deliverable is the one made publicly available in PDF format at www.etsi.org/deliver.

Users of the present document should be aware that the document may be subject to revision or change of status.
Information on the current status of this and other ETSI documents is available at
https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

If you find errors in the present document, please send your comment to one of the following services:
https://portal.etsi.org/People/CommiteeSupportStaff.aspx

If you find a security vulnerability in the present document, please report it through our
Coordinated Vulnerability Disclosure Program:
https://www.etsi.org/standards/coordinated-vulnerability-disclosure

Notice of disclaimer & limitation of liability

The information provided in the present deliverable is directed solely to professionals who have the appropriate degree of experience to understand and interpret its content in accordance with generally accepted engineering or other professional standard and applicable regulations.
No recommendation as to products and services or vendors is made or should be implied.
No representation or warranty is made that this deliverable is technically accurate or sufficient or conforms to any law and/or governmental rule and/or regulation and further, no representation or warranty is made of merchantability or fitness for any particular purpose or against infringement of intellectual property rights.
In no event shall ETSI be held liable for loss of profits or any other incidental or consequential damages.

Any software contained in this deliverable is provided "AS IS" with no warranties, express or implied, including but not limited to, the warranties of merchantability, fitness for a particular purpose and non-infringement of intellectual property rights and ETSI shall not be held liable in any event for any damages whatsoever (including, without limitation, damages for loss of profits, business interruption, loss of information, or any other pecuniary loss) arising out of or related to the use of or inability to use the software.

Copyright Notification

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (https://ipr.etsi.org/).

Pursuant to the ETSI Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP™** and **LTE™** are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM**® and the GSM logo are trademarks registered and owned by the GSM Association.

Legal Notice

This Technical Specification (TS) has been produced by ETSI 3rd Generation Partnership Project (3GPP).

The present document may refer to technical specifications or reports using their 3GPP identities. These shall be interpreted as being references to the corresponding ETSI deliverables.

The cross reference between 3GPP and ETSI identities can be found under http://webapp.etsi.org/key/queryform.asp.

Modal verbs terminology

In the present document "**shall**", "**shall not**", "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Contents

Foreword

This Technical Specification has been produced by the 3GPP.

The present document describes the operation of the Adaptive Multi Rate Wideband speech codec during Source Controlled Rate (SCR) operation 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 x.y.z

where:

x the first digit:

1 presented to TSG for information;

2 presented to TSG for approval;

3 or greater 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;

1 Scope

This document describes the Source Controlled Rate (SCR) operation of the Adaptive Multi-Rate Wideband speech Codec. The implementation of this SCR operation is mandatory in all UMTS equipment implementing the Adaptive Multi-Rate Wideband speech Codec.

The description is structured according to the block diagram in Figure 1. This structure of distributing the various functions between system entities is not mandatory for implementation, as long as the operation on the speech decoder output remains the same.

Annex A describes the Discontinuous Transmission (DTX) operation of the Adaptive Multi-Rate Wideband speech Codec in Codec Type FR_AMR-WB for the GSM system.

2 Normative references

This document incorporates by dated and undated reference, provisions from other publications. These normative references are cited at the appropriate places in the text and the publications are listed hereafter. For dated references, subsequent amendments to or revisions of any of these publications apply to this document only when incorporated in it by amendment or revision. For undated references, the latest edition of the publication referred to applies.

[1] 3GPP TS 26.171 : "AMR Wideband Speech Codec; General description".

[2] 3GPP TS 26.173 : "AMR Wideband Speech Codec; ANSI-C code".

[3] 3GPP TS 26.174 : "AMR Wideband Speech Codec; Test sequences".

[4] 3GPP TS 26.190 : "AMR Wideband Speech Codec; Transcoding functions".

[5] 3GPP TS 26.191 : "AMR Wideband Speech Codec; Error concealment of lost frames".

[6] 3GPP TS 26.192 : "AMR Wideband Speech Codec; Comfort noise aspects".

[7] 3GPP TS 26.194 : "AMR Wideband Speech Codec; Voice Activity Detector (VAD)".

[8] 3GPP TS 26.201 : "AMR Wideband Speech Codec; Frame structure".

3 Definitions, symbols and abbreviations

3.1 Definitions

For the purpose of this document, the following definitions apply.

frame: Time interval of 20 ms, corresponding to the time segmentation of the Adaptive Multi-Rate Wideband speech Codec, also used as a short term for a traffic frame.

traffic frame: Block of 132..477 information bits transmitted on the speech traffic channels.

SID frame: Frame that conveys information about the acoustic background noise.

speech frame: Traffic frame that has been classified as SPEECH_GOOD, SPEECH_BAD or SPEECH_LOST frame.

VAD flag: Boolean flag, generated by the VAD algorithm indicating the presence ("1") or the absence ("0") of a speech frame.

RX_TYPE: classifies the received frame.

TX_TYPE: classifies the frame to be transmitted.

hangover period: A period of frames added at the end of a speech burst in which VAD flag ="0" and TX_TYPE is ="SPEECH_GOOD", this period provides the encoder with an extra window to derive the Comfort Noise parameters .

3.2 Symbols

For the purpose of this document, the following symbols apply.

$N_{elapsed}$	Number of elapsed frames since the last updated SID frame.

3.3 Abbreviations

For the purpose of this document , the following abbreviations apply.

AN	Access Network
SCR	Source Controlled Rate operation
TS	Telecommunication Standard, Technical Specification
GSM	Global System for Mobile Telecommunication
UE	User Equipment
RAN	Radio Access Network
RX	Receive
SID	Silence Descriptor
TX	Transmit
VAD	Voice Activity Detector

4 General

Source Controlled Rate operation (SCR) is a mechanism for the AMR Wideband Speech Codec, which allows to encode the input signal at a lower average rate by taking speech inactivity into account. The SCR scheme may be used for the following purposes:

-to save power in the User Equipment;

-to reduce the overall interference and load in the networks.

SCR in the transmitting path (uplink) shall be in operation in UEs, if commanded so by the network. The UE shall handle SCR in the receiving path (downlink) at any time, regardless, whether SCR in the transmitting path is commanded or not.

4.1 General organisation

The default SCR mechanism described in this document requires the following functions:

-a Voice Activity Detector (VAD) on the transmit (TX) side;

-evaluation of the background acoustic noise on the transmit (TX) side, in order to transmit characteristic parameters to the receive (RX) side;

-generation on the receive (RX) side of a similar noise, called comfort noise, during periods where the transmission is switched off.

The Voice Activity Detector (VAD) is defined in [7] and the comfort noise functions in [6]. Both are based partly on the speech Codec and its internal variables, defined in [4].

In addition to these functions, if the parameters arriving at the RX side are detected to be seriously corrupted by errors, the speech or comfort noise must be generated from substituted data in order to avoid seriously annoying effects for the listener. These functions are defined in [5].

An overall description of the speech processing parts can be found in [1]. An overview of one link SCR operation is shown in Figure 1.

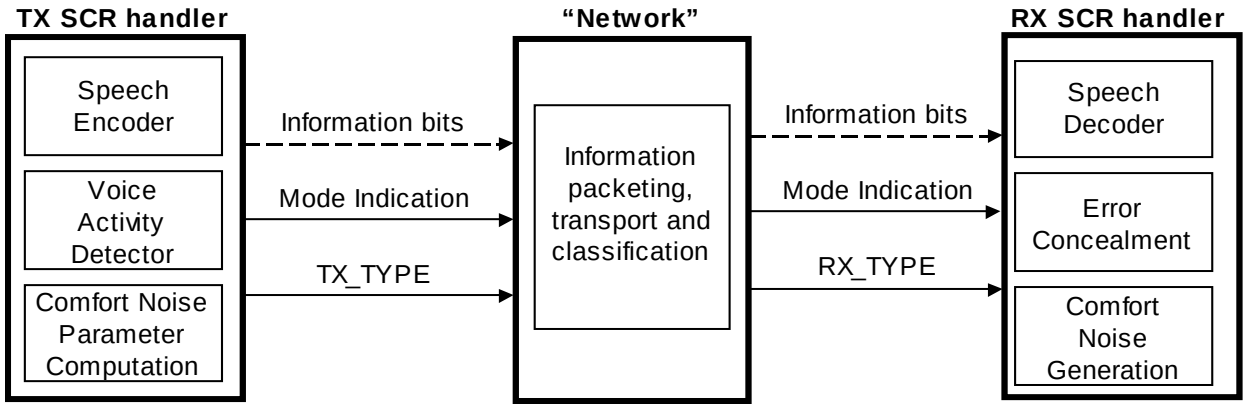

Figure 1: Block diagram of one link SCR operation

5 AMR-WB SCR operation

5.1 Transmit (TX) side

A block diagram of the transmit side SCR functions is shown in Figure 2.

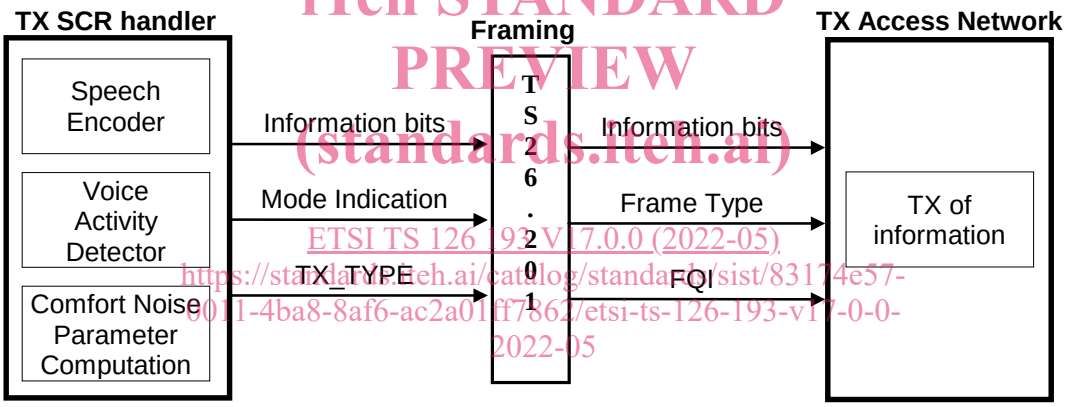

Figure 2: Block diagram of SCR functions at the TX side

5.1.1 General operation

The TX SCR handler passes traffic frames, individually marked by TX_TYPE, to the Framing unit. Each frame consists of bit fields containing the information bits, the codec mode indication, and the TX_TYPE. TX_TYPE shall be used to specify the contents of the frame. The table below provides an overview of the different TX_TYPEs used and explains the required contents in the information bit and the mode indication bit fields.

Table 1: SCR TX_TYPE identifiers for UMTS_AMR-WB and FR_AMR-WB

TX_TYPE	Information Bits	Mode Indication
SPEECH_GOOD	Speech frame, size 132..477 bits, depending on codec mode	Current codec mode
SPEECH_BAD	Corrupt speech frame (bad CRC), size 132..477 bits, depending on codec mode	Current codec mode
SPEECH_LOST	No useful information. (Note: If implementation does not support the SPEECH_LOST, SPEECH_BAD shall be used instead)	No useful information
SID_FIRST	Marker for the end of talkspurt, no further information, all 35 comfort noise bits set to "0"	The codec mode that would have been used if TX_TYPE had been "SPEECH_GOOD"
SID_UPDATE	35 comfort noise bits	The codec mode that would have been used if TX_TYPE had been "SPEECH_GOOD"
SID_BAD	Corrupt SID update frame (bad CRC)	The codec mode that would have been used if TX_TYPE had been "SPEECH_GOOD"
NO_DATA	No useful information, nothing to be transmitted	No useful information

TX_TYPE = "SPEECH_LOST" indicates that the Information Bit and Codec Mode fields do not contain any useful data (but still should be transmitted over AN). The purpose of this TX_TYPE is indicate that the frame was transmitted but lost on some previous phase. This TX_TYPEs may occur only in TFO and TrFO situations. Note, that it is possible to replace SPEECH_LOST with SPEECH_BAD but this may degrade the quality of the error concealment in the receiving end because concealment may try to use part of the received parameters from the frame which do not contain any useful information.

TX_TYPE = "NO_DATA" indicates that the Information Bit and Codec Mode fields do not contain any useful data (and should not be transmitted over AN). The purpose of this TX_TYPE is to provide the option to save network transmission between the transcoder and AN.

Note, the TX_TYPEs "SPEECH_BAD", "SPEECH_LOST" and "SID_BAD" may occur in TFO and TrFO situations.

The scheduling of the frames for transmission on the Access Network is controlled by the TX SCR handler by the use of the TX_TYPE field.

5.1.2 Functions of the TX SCR handler

If TX SCR operation is disabled, the TX SCR handler continuously generates speech frames, i.e. frames marked with TX_TYPE="SPEECH_GOOD".

If the TX SCR operation is enabled, the VAD flag controls the TX SCR handler operation as described in the following paragraphs.

5.1.2.1 AMR-WB SCR Timing procedures

To allow an exact verification of the TX SCR handler functions, all frames before the reset of the system are treated as if there were speech frames of an infinitely long time. Therefore, and in order to ensure the correct estimation of

comfort noise parameters at RX SCR side, the first 7 frames after the reset or after enabling the SCR operation shall always be marked with TX_TYPE= "SPEECH_GOOD", even if VAD flag ="0" (hangover period, see figure 3).

The Voice Activity Detector (VAD) shall operate all the time in order to assess whether the input signal contains speech or not. The output is a binary flag (VAD flag ="1" or VAD flag ="0", respectively) on a frame by frame basis (see [7]).

The VAD flag controls indirectly, via the TX SCR handler operations described below, the overall SCR operation on the transmit side.

Whenever VAD flag ="1", the speech encoder output frame along with mode information shall be passed directly to the AN, marked with TX_TYPE =" SPEECH_GOOD "

At the end of a speech burst (transition VAD flag ="1" to VAD flag ="0"), it takes eight consecutive frames to make a new updated SID analysis available (see [6]). Normally, the first seven speech encoder output frames after the end of the speech burst shall therefore be passed directly to the AN, marked with TX_TYPE =" SPEECH_GOOD " ("hangover period").

The end of the speech is then indicated by passing frame eight after the end of the speech burst to the AN, marked with TX_TYPE = "SID_FIRST" (see figure 3). SID_FIRST frames do not contain data.

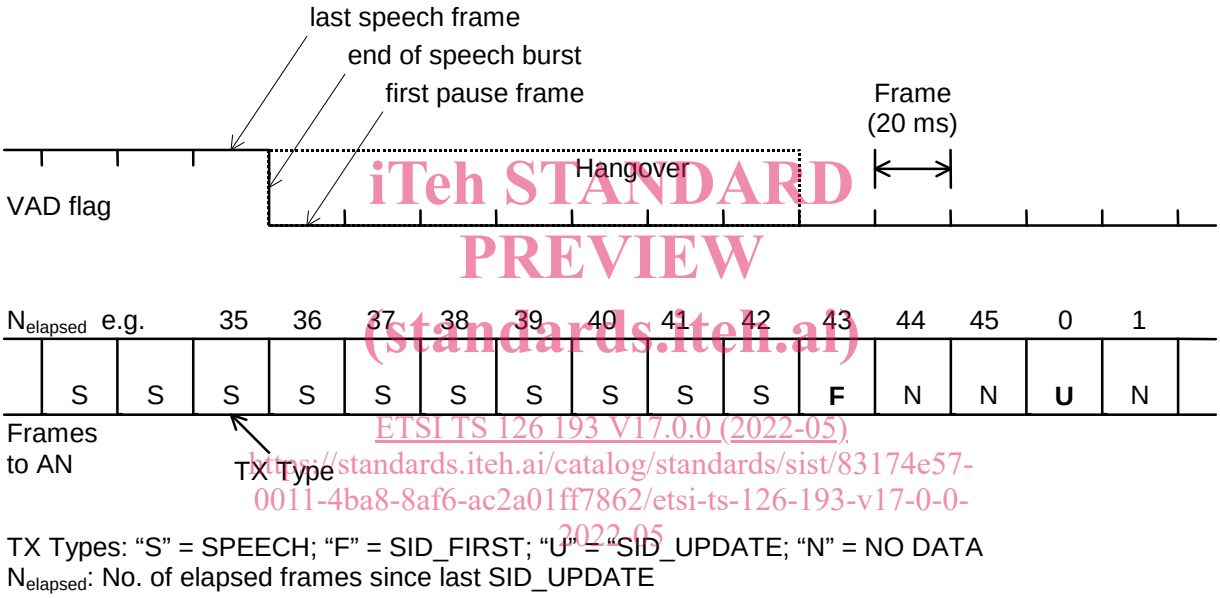

Figure 3: Normal hangover procedure for AMR-WB ($N_{elapsed}$ > 23)

If, however, at the end of the speech burst, less than 24 frames have elapsed since the last SID_UPDATE frame was computed, then this last analysed SID_UPDATE frame should be passed to the AN whenever a SID_UPDATE frame is to be produced, until a new updated SID analysis is available (8 consecutive frames marked with VAD flag ="0").This reduces the load on the network in cases where short background noise spikes are taken for speech, by avoiding the "hangover" waiting for the SID frame computation.

Once the SID_FIRST frame has been passed to the AN, the TX SCR handler shall at regular intervals compute and pass updated SID_UPDATE (Comfort Noise) frames to the AN as long as VAD flag = "0". SID_UPDATE frames shall be generated every 8th frame. The first SID_UPDATE shall be sent as the third frame after the SID_FIRST frame.

The speech encoder is operated in full speech modality if TX_TYPE = " SPEECH_GOOD " and otherwise in a simplified mode, because not all encoder functions are required for the evaluation of comfort noise parameters and because comfort noise parameters are only to be generated at certain times.

5.1.3 The TX part of the AN

The TX part of the AN has the following overall functionality. The transmission is cut after the transmission of a SID_FIRST frame when the speaker stops talking. During speech pauses the transmission is resumed at regular intervals for transmission of one SID_UPDATE frame, in order to update the generated comfort noise on the RX side. The operation of

the TX part of the AN is controlled by the TX SCR handler via the TX_TYPE.

All frames, marked with SPEECH_GOOD, SID_FIRST or SID_UPDATE shall be transmitted by the TX part of the AN.

5.2 Receive (RX) side

A block diagram of the receive side SCR functions is shown in Figure 3 below.

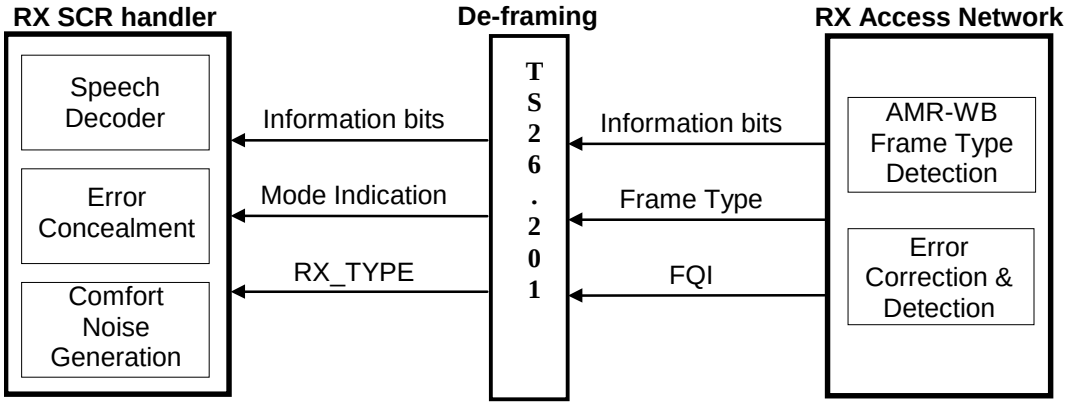

Figure 4: Block diagram of the receive side SCR functions

5.2.1 General operation

The AN passes all the received traffic frames to the RX SCR handler, classified with RX_TYPE, as described in Table 2 (see TS 26.201). The RX SCR handles the frame accordingly.

Table 2: RX_TYPE identifiers for AMR-WB

RX_TYPE	Information Bits
SPEECH_GOOD	Speech frame without detected errors.
SPEECH_BAD	(likely) speech frame with bad CRC (or estimated to be very bad by the RX part of the AN)
SPEECH_LOST	No frame received. Indicates that this frame was transmitted, but never received.
SID_FIRST	This SID-frame marks the beginning of a comfort noise period.
SID_UPDATE	Correct SID update frame
SID_BAD	Corrupt SID update frame (bad CRC; applicable only for SID_UPDATE frames)
NO_DATA	Nothing useable was received. The synthesis mode of the previous frame type is used.