Fraunhofer IIS | Am Wolfsmantel 33 | 91058 Erlangen, Germany

Telecom Regulatory Authority of India (TRAI)
Shri Asit Kadayan
Advisor (QoS)
New Delhi

Fraunhofer Institute for Integrated Circuits IIS

Management of the Institute
Prof. Dr. Albert Heuberger (executive)
Dr. Bernhard Grill

Am Wolfsmantel 33
91058 Erlangen, Germany

Dr. Alfonso Carrera
Key Account Manager Telecom Market
Audio & Multimedia
Phone + 34 655 717235
alfonso.carrera@iis.fraunhofer.de
www.iis.fraunhofer.de

Erlangen, April 6, 2018

Fraunhofer's Response to TRAI's Consultation Paper on Voice Services to LTE users

1. Fraunhofer's Background

The Fraunhofer-Gesellschaft (FhG) is Europe's largest application-oriented research organization. Its research activities are conducted by 72 institutes and research units at locations throughout Germany. Since decades, Fraunhofer's Institute for Integrated Circuits (IIS) is a leading contributor to audio and speech codecs such as mp3, AAC, xHE-AAC, MPEG-H 3D-Audio. Fraunhofer IIS was also a main developer of 3GPP's state-of-the-art speech codec for Enhanced Voice Services (EVS, standardized in 2014) [1]. EVS has been specifically designed for VoLTE and offers significant improvements with respect to quality of service compared to its predecessors AMR and AMR-WB. In GSMA's VoLTE specification IR.92 [2], EVS is specified as mandatory codec for super-wideband VoLTE services. However, EVS also provides major advantages in speech quality and error robustness for wideband and narrowband VoLTE services and is already used by various mobile operators worldwide.

2. Response to QoS and Measurement Tool Issues

Given its background as co-developer of 3GPP's most recent speech codec, Fraunhofer IIS would like to respond to TRAI's *Consultation Paper on Voice Services to LTE users (including VoLTE and CS Fallback)* as follows:

Issues 1-4:
1. Whether prescribed QoS parameters, as per existing QoS Regulations, are sufficient to effectively monitor QoS of VoLTE/CSFB calls? Please provide suggestions with justifi- cations.
2. If existing QoS parameters are not sufficient to monitor QoS of VoLTE/CSFB calls, then what new parameters can be introduced? Please provide details with justifications.

Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e. V., München
Executive Board
Prof. Dr.-Ing. habil. Prof. E. h. Dr.-Ing. E. h. mult. Dr. h. c. Dr. h. c. Reimund Neugebauer, President
Prof. Dr. rer. publ. ass. iur. Alexander Kurz
Dipl.-Kfm. Andreas Meuer
Prof. Dr. rer. nat. Georg Rosenfeld

Cheques and transfers payable to:
Deutsche Bank, München
Account 752193300 BLZ 700 700 10
IBAN DE86 7007 0010 0752 1933 00
BIC (SWIFT-Code) DEUTDEMM
V.A.T. Ident No. DE129515865
Tax Number 143/215/20392

3. How to define instance of silence/voice mute? How many such instances may be accepted during voice call? Whether existing parameters like packet loss, jitter, latency, end-to-end delay are sufficient to identify or measure silence/voice mute or some other parameters are also need to be factored to measure it? Please provide details with justifications.
4. How to measure report and evaluate network or service from perspective of silence/voice mute problem? Which ITU measurement tools can be used to prepare framework for measurement of silence/voice mute problem? Please provide details with justifications.

2.1. Suggested QoS parameter Mean Opinion Score (MOS)

In Fraunhofer IIS's experience, apart from basic parameters like call failure ratio, call setup failure and unintended call drops, it is important to measure the quality of an ongoing call by means of a Mean Opinion Score (MOS). Mean Opinion Scores are ideally derived by listening tests (usually conducted according to ITU-T Recommendation P.800 [3], but this is only practical during codec development/standardization. In order to measure and monitor quality in a running network, ITU-T Recommendation P.863 specifies the measurement tool POLQA (Perceptual Objective Listening Quality Assessment) [4]. By comparing a transmitted speech sample with the original, POLQA is able to provide a reasonable estimation of the MOS of a transmission. Comparing that with the MOS under ideal conditions (i.e. a transparent transmission channel) allows detection of problems in the connection. Network operators can obtain POLQA-based measurement systems from various vendors [5] to do monitoring in house or contract a service provider [6] to do the measurements for them.

Measurement of MOS is essential, since the quality of an ongoing call heavily depends on the following factors:

a) choice of codec/bandwidth & bitrate
b) transmission impairments (e.g. packet loss in case of VoLTE)

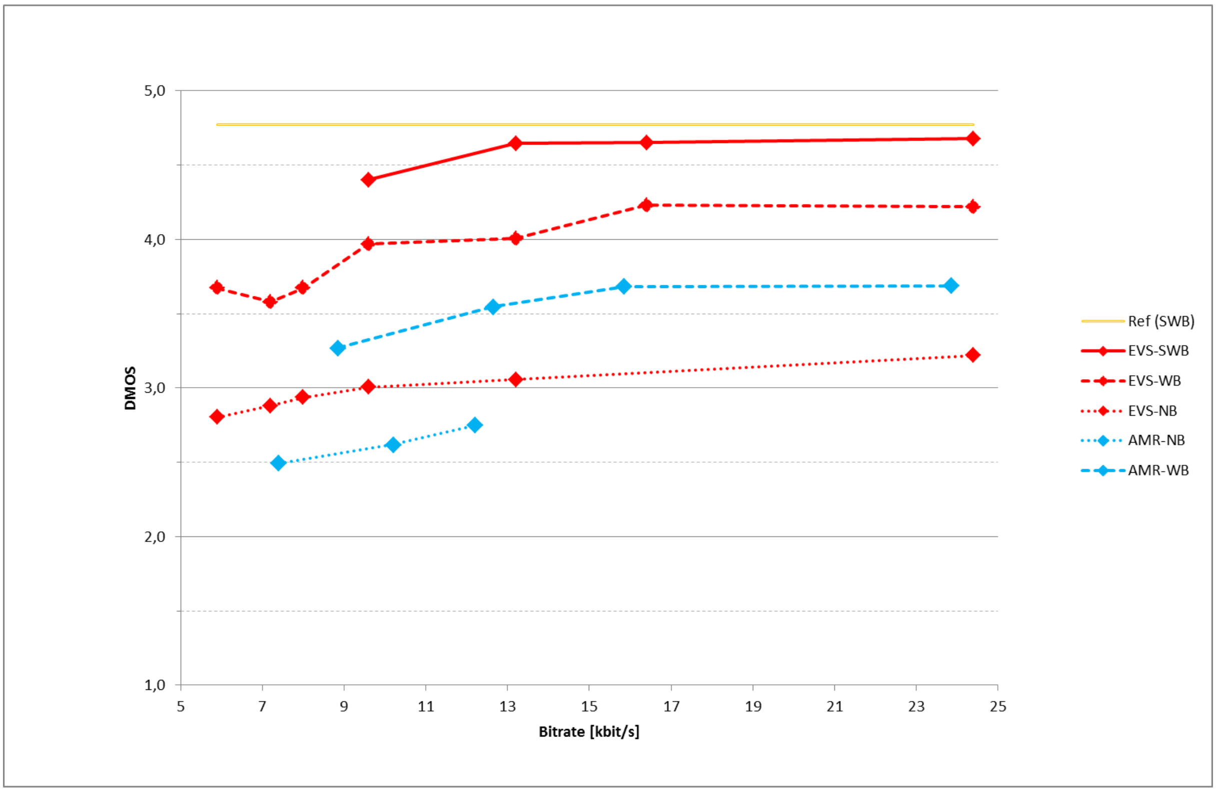
If a call takes place between different networks/technologies (e.g. 4G to 2G, 4G to fixed line, 4G to 4G without IP interconnect), transcoding losses may affect quality in addition (this is not further elaborated in this document for the sake of simplicity).

POLQA measurements are useful to detect voice quality and network deficiencies regardless of whether the voice service is based on circuit switched (2G, 3G) or VoLTE (4G). Therefore, the measurements can well cover circuit switched fall back.

In the below, we will show how codec, bandwidth, bitrate and channel impairments affect the MOS. With respect to channel impairments, this document focuses on packet loss typical for packet switched VoLTE networks.

2.2. Impact of codec, bandwidth and bitrate on quality

The following diagram illustrates the influence of choice of codec and bitrate on quality:



The diagram shows the result of an independent P.800 listening test performed during the standardization of EVS (Experiment M1 of the EVS Performance Characterization [7], North American English Clean Speech, DTX on, no channel impairments). As explained above, POLQA measurements will show similar results (usually, however, somewhat compressed with respect to the quality scale).

The test shows that, even under perfect transmission conditions, call quality varies from 2.5 MOS to 4.7 MOS depending on the choice of codec and bitrate. With respect to newer codecs, the EVS Codec in wideband mode (WB) is about 0.5 MOS better, the EVS codec in super wideband mode (SWB) is about 1.0 MOS better than AMR-WB. Thus, the quality of a call can be largely improved by using the newer codec.

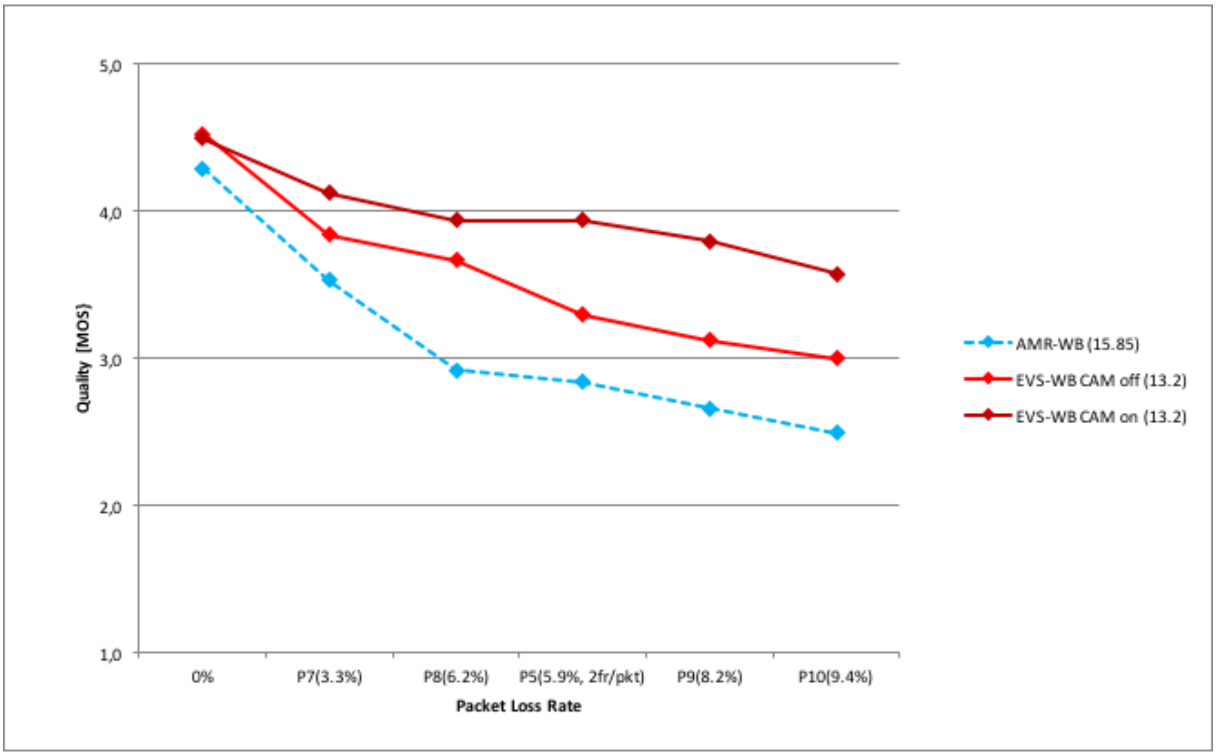
Alternatively, the EVS codec can provide the same quality as AMR-WB at a much lower bitrate, leading to more stable connections and/or higher network capacity.

For English clean speech, as in the test shown above, there is little to gain by using bitrates above 13 kbps. Similar tests exist for noisy speech and mixed content and music (see Annex). The basic codec ranking is the same, however there is more to gain by providing bitrates above 13 kbps for these content types. With respect music and other non-speech content, EVS is the first 3GPP conversational coder to provide reasonable quality.

Fraunhofer IIS

2.3. Impact of channel impairments on quality

The second important factor is the impact of transmission problems on call quality. The test below shows the effect of increasing packet loss rates:



This result is again taken from the EVS Performance Characterization [7], Characterization Experiment W1, Clean Speech North American English. All codecs run in WB mode, i.e. the additional gain by going to SWB is not taken into account in this test. AMR-WB operates at 15.85 kbps, EVS-WB at 13.2 kbps. "CAM on" stands for the special "Channel Aware Mode" of EVS, specifically designed to improve quality in packet loss conditions. "CAM off" is EVS without using those tools. P5-P10 refers to simulated VoLTE loss profiles which result into the packet loss rate given in brackets.

Key findings when comparing the error free channel with P10 (9.4% packet loss rate) are:

	MOS at 0% PL	MOS Degradation	MOS at 9.4% PL
AMR-WB	4.3	-1.8	2.5
EVS-WB	4.5	-1.5	3.0
EVS-WB Channel Aware	4.5	-0.9	3.6

Specifically, the Channel Aware Mode of EVS improves the quality in packet loss scenarios by a large margin compared to AMR-WB. As a rough rule of thumb, EVS-CAM at 6% packet loss rate performs similar to AMR-WB at 1% packet loss rate. Investigations by Qualcomm [8] show that this is equivalent to a 2.5 dB gain in transmission SNR and can be used to increase general network robustness and to expand coverage (including indoor coverage).

Many more tests with other languages, background noises, music a.s.o. are provided in the EVS Performance Characterization [7]

3. Response to Issue 7

Issue 7: Any other issue which is relevant to this subject?

The listening tests cited herein show that the quality of a VoLTE network can take a huge lift by using EVS, 3GPPs latest codec, instead of AMR-WB.

EVS provides
- higher efficiency
- higher robustness against packet loss, and
- higher quality

all at the same time. This is even true if EVS is running in wideband mode on a mobile phone not equipped for super wideband audio.

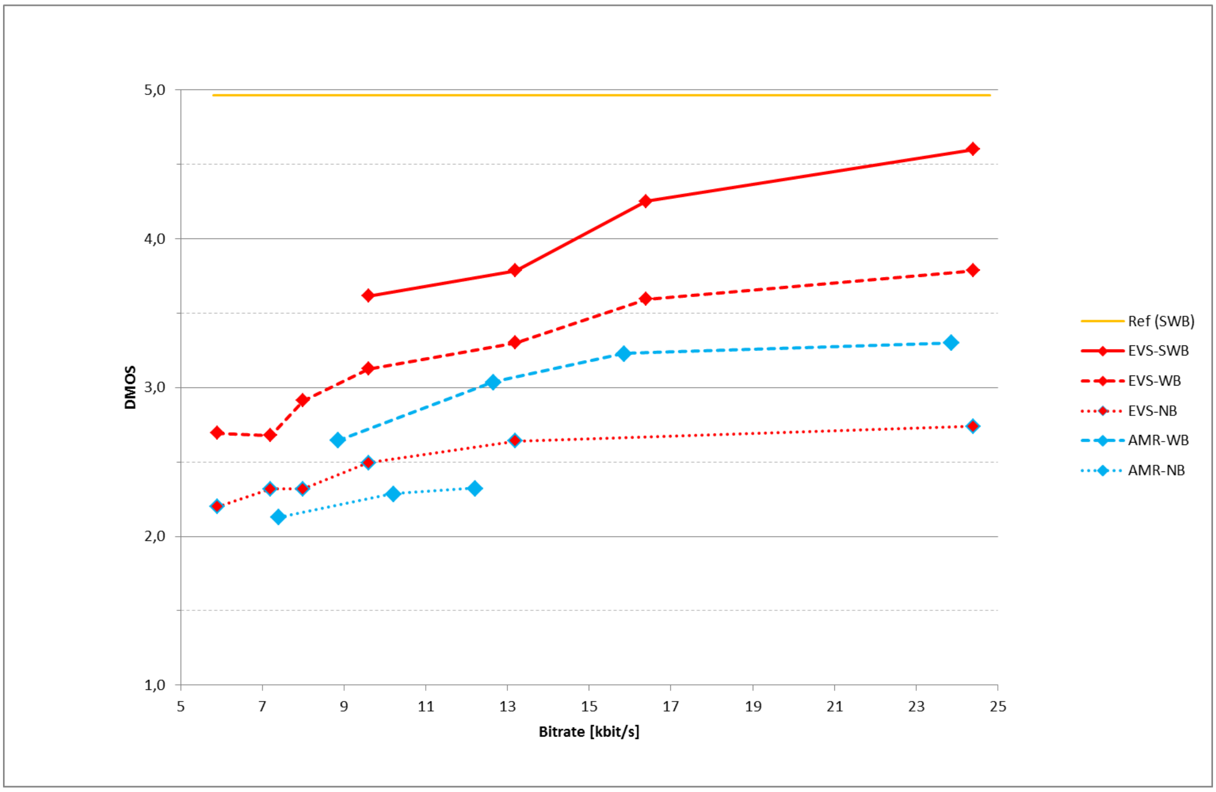
In addition to other features necessary for 3GPP networks, such as improved Voice Activity Detection, Discontinuous Transmission and Comfort Noise Generation (VAD/DTX/CNG), EVS also comes with a built-in AMR-WB-compatible mode, the main purpose of which is to provide seamless switching to AMR-WB in case of a SRVCC scenario.

Designed by 3GPP specifically for VoLTE, EVS is easy to enable in VoLTE networks. EVS-capable network equipment (e.g. gateways) can be obtained from all major vendors. On the handset side, EVS is broadly implemented on mobile chipsets and thus available in many recent mobile phones. 3GPP is also providing specifications allowing the use of EVS in 3G networks [9]. However, to Fraunhofer's knowledge, as of today this is not in use.

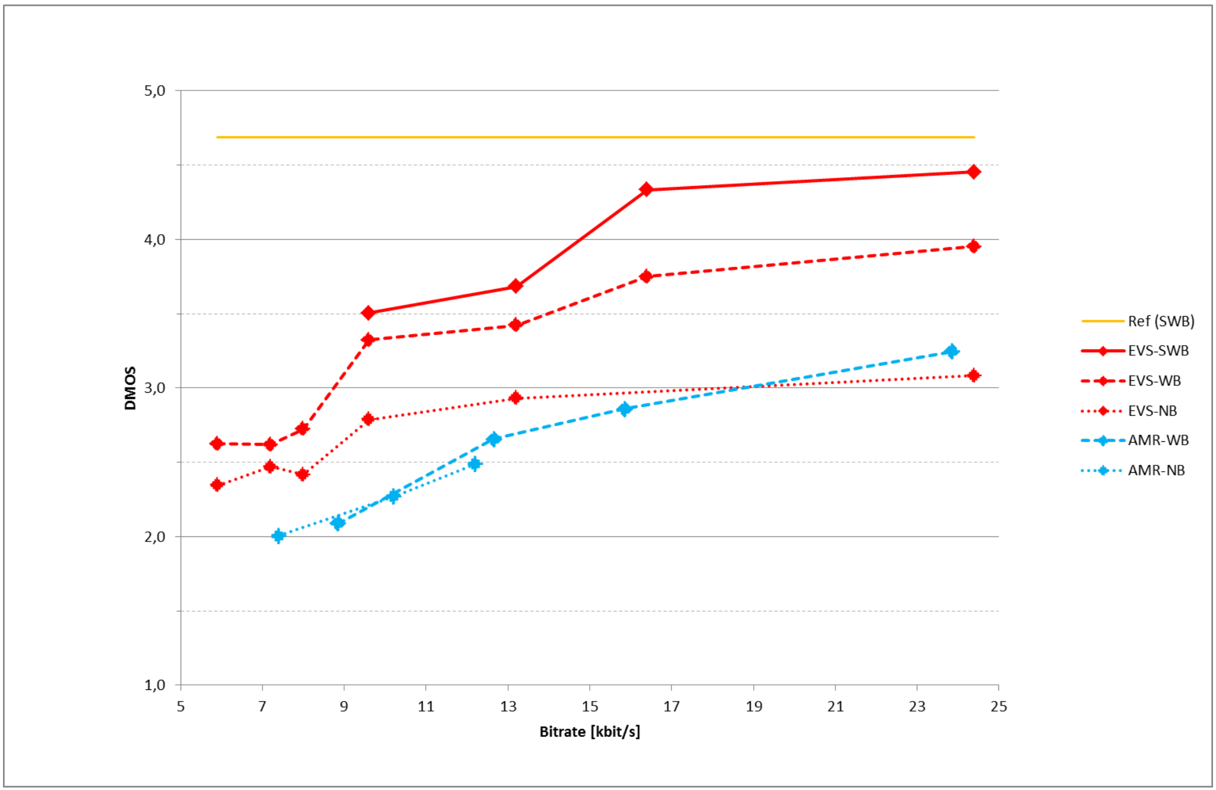
Due to all its features, EVS has the capability to advance the quality of service of any VoLTE network. Fraunhofer therefore recommends to mandate the use of EVS for VoLTE networks, even if super wideband speech is not used or only recommended.

Annex – Noisy Speech and Mixed and Music Listening Tests

The following two test have also been taken from the EVS Characterization Report [7]



Experiment M2, Noisy Speech (Car Noise 20dB), DTX on, Finnish



Experiment M3b, Mixed and Music, DTX on, North American English

References

[1] 3GPP TS 26.441 – 26.451, “Codec for Enhanced Voice Services (EVS)”
[2] GSMA PRD IR.92, “IMS Profile for Voice and SMS”
[3] Recommendation ITU-T P.800, "Methods for subjective determination of transmission quality"
[4] ITU-T P.863, “Perceptual Objective Listening Quality Assessment (POLQA)”
[5] POLQA Equipment vendors: E.g. Rohde&Schwarz UPV Audio Analyzer, Spirent Umetrix Voice
[6] Quality Measurement Service Providers: e.g. Sigos
[7] 3GPP TR 26.952, “Codec for Enhanced Voice Services (EVS); Performance Characterization”
[8] 3GPP R2-165194, ”On eNB awareness of EVS codec for coverage and mobility enhancements”
[9] 3GPP TS 26.453, “Codec for Enhanced Voice Services (EVS) – Speech codec frame structure” and 3GPP TS 26.454, “Codec for Enhanced Voice Services (EVS) – Interface to Iu, Uu, Nb and

Yours sincerely,

A. Carrera

Dr. Alfonso Carrera