

International Telecommunication Union

ITU-T

P.565.1 Implementer's Guide

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

(17 June 2022)

SERIES P: TERMINALS AND SUBJECTIVE AND
OBJECTIVE ASSESSMENT METHODS

Methods for objective and subjective assessment of
speech quality

**Implementer's guide for Recommendation
ITU-T P.565.1**

Summary

This Implementer's Guide provides information about the handling of the wide-band modes of voice codecs EVS and Opus at coding rates above 20 kbits/s by Recommendation ITU-T P.565.1.

This document contains all updates submitted up to and including those at Study Group 12 meeting in June 2022.

This document was agreed by ITU-T Study Group 12 on 17 June 2022 and is the initial version of this implementer's guide for Recommendation ITU-T P.565.1.

Note

This is an informative ITU-T publication. Mandatory provisions, such as those found in ITU-T Recommendations, are outside the scope of this publication. This publication should only be referenced bibliographically in ITU-T Recommendations.

Change log

2022-06 Initial version

Contact information

ITU-T Study Group 12 / Question 15 Co-Rapporteur	Sebastian Möller Technische Universität Berlin Germany	Tel.: +49 30 8353 58465 E-mail: sebastian.moeller@tu-berlin.de
ITU-T Study Group 12 / Question 15 Co-Rapporteur	Joachim Pomy Opticom GmbH Germany	Tel.: +49 177 78 7195 E-mail: consultant@joachimpomy.de
ITU-T Study Group 12 / Question 15 Co-Rapporteur	Vincent Barriac Orange Innovation France	Tel.: +33 296 071 810 E-mail: vincent.barriac@orange.com

Table of Contents

	Page
1 Scope	1
2 Introduction	1
3 Defect resolution procedure.....	1
4 References.....	1
5 Nomenclature.....	1
6 Technical corrections to the application of Recommendation ITU-T P.565.1	2
Annex – P.565.1 (2021-11) Defect Report Form.....	4

Implementer's guide for Recommendation ITU-T P.565.1

1 Scope

This guide resolves defects in the following categories:

- ambiguities

In addition, the implementer's guide may include explanatory text found necessary as a result of interpretation difficulties apparent from the defect reports.

This guide will not address proposed additions, deletions or modifications to the Recommendations that are not strictly related to implementation difficulties in the above categories. Proposals for new features should be made through contributions to ITU-T.

2 Introduction

This implementer's guide is a compilation of reported defects for all versions of Recommendation ITU-T P.565.1. In this edition of the guide, reported defects identified as of June 2022 are given for:

- Recommendation ITU-T P.565.1 (2021).

The guide must be read in conjunction with Recommendation ITU-T P.565.1 (2021) to serve as an additional source of information for implementers. The changes, clarifications and corrections defined herein are expected to be included in future versions of the affected Recommendations.

3 Defect resolution procedure

Upon discovering technical defects with any components of the texts covered by this implementer's guide, please provide a written description directly to the editors of the affected Recommendation(s) with a copy to the respective Rapporteur (See contacts above on page ii). The template for a defect report is located at the end of this guide. Return contact information should also be supplied so a dialogue can be established to resolve the matter and an appropriate reply to the defect report can be conveyed. This defect resolution process is open to any interested party. Formal membership in the ITU is not required to participate in this process.

4 References

This implementer's guide refers to the following ITU-T Recommendations:

- Recommendation ITU-T P.565 (2021), *Framework for creation and performance testing of machine learning based models for the assessment of transmission network impact on speech quality for mobile packet-switched voice services.*
- Recommendation ITU-T P.565.1 (2021), *Machine learning model for the assessment of transmission network impact on speech quality for mobile packet-switched voice services.*
- Recommendation ITU-T P.863 (2018), *Perceptual objective listening quality prediction.*

5 Nomenclature

In addition to traditional revision marks, the following marks and symbols may be used to indicate to the reader how changes to the text of a Recommendation should be applied:

Symbol	Description
<u>[Begin Correction]</u>	Identifies the start of revision marked text based on extractions from the published Recommendations affected by the correction being described.
<u>[End Correction]</u>	Identifies the end of revision marked text based on extractions from the published Recommendations affected by the correction being described.
...	Indicates that the portion of the Recommendation between the text appearing before and after this symbol has remained unaffected by the correction being described and has been omitted for brevity.
--- SPECIAL INSTRUCTIONS --- {instructions}	Indicates a set of special editing instructions to be followed.

6 Technical corrections to the application of Recommendation ITU-T P.565.1

Technical background

The speech quality predicting model described in Recommendation ITU-T P.565.1, following the framework defined in Recommendation ITU-T P.565, was validated in against reference mean opinion scores (MOSs) produced by Recommendation ITU-T P.863.

When the current version of ITU-T P.863 (edition 3, 2018) was approved, the scope of the Fullband (FB) mode of the model in the Recommendation was broadened to address a new generation of codecs that can work in either narrowband (NB), wideband (WB) or super-wideband (SWB) / fullband (FB) audio modes (the exact definitions for these working modes of voice codecs can be found in Recommendation ITU-T P.10/G.100). This generation comprises mostly the enhanced voice service (EVS) and Opus codecs.

At the time of this approval, the assumptions of the ecosystem concerning the audio bitrates of these new generation codecs under WB and SWB modes were that WB would require around 10 kbits/s to outperform older WB codecs (such as AMR WB) and that higher bitrates above 10 kbit/s would always enable SWB or even FB transmission mode. For instance, most network equipment and devices implementing EVS are usually working in SWB mode at 24.4 kbit/s or in 13.2 kbit/s or very rarely in WB mode at 9.6 kbit/s.

In logical accordance with this, the subjective databases used to train and validate the FB mode of edition 3 of Recommendation ITU-T P.863 contained speech samples representative of these assumed working conditions.

In a contribution received by SG12 in June 2022 (see SG12-C29¹ of the 2022-2024 study period), Orange reported an unexpected usage of both EVS and Opus in their WB modes at bitrates above 20 kbit/s. Concerning EVS, the reported conditions are observed on VoLTE communications. Concerning Opus, they are observed on over-the-top (OTT) VoIP applications.

These conditions were not present in the subjective databases used to train and validate edition 3 of Recommendation ITU-T P.863.

¹ <https://www.itu.int/md/T22-SG12-C-0029/en>

This contribution reported also that the prediction of speech quality made under these high bitrate wideband conditions for new generation codecs may be too optimistic. Indeed, reported MOS-LQOF values are slightly lower, if not equal, to those corresponding to the SWB modes of these codecs at similar bitrates. This observation is independent on the used reference speech sample.

This result does not seem to be in line with the informal listening of corresponding speech samples by expert and naïve listeners. No formal listening speech quality subjective test result is available for the time being to quantify the exact perceived difference between these conditions.

Conclusion

The FB mode of Recommendation ITU-T P.863 has to be considered as non-validated for conditions with new generation voice codecs (such as EVS and Opus) in WB mode with a bit rate above 20 kbit/s.

As a logical consequence, the model of Recommendation ITU-T P.565.1 has to be considered too as non-validated for these conditions.

Implementers of ITU-T P.565.1 for VoLTE (EVS mode) and OTT VoIP services are requested to check both audio mode (WB or SWB) and bitrate of the tested communications before implementing any post-processing or reporting action on measurement results. In particular, comparison between MOS-LQOF scores obtained under WB and SWB modes respectively of these codecs must be avoided if bitrates of the conditions in WB mode are above 20 kbit/s.

Annex: P.565.1 (2021-11) Defect Report Form
--

DATE:	
CONTACT INFORMATION NAME: COMPANY: ADDRESS: TEL.: FAX: E-MAIL:	
AFFECTED RECOMMENDATIONS:	
DESCRIPTION OF PROBLEM:	
SUGGESTIONS FOR RESOLUTION:	

NOTE – Attach additional pages if more space is required than is provided above.
