Mesures de qualité vocale sur les derniers réseaux mobiles basés sur IP: la meilleure façon de mettre en œuvre les standards de test

Par Lee Roberts, Business Development Manager, Anritsu (UK)

Les utilisateurs de téléphones mobiles attendent de leur services de phonie une grande qualité vocale lors des appels vocaux. Il est donc crucial pour les opérateurs de réseaux sans fil de savoir si leurs réseaux sont capables de fournir une excellente qualité vocale et pour les fabricants de vérifier la qualité audio des équipements, et plus particulièrement celle des téléphones portables..

Mais comment vérifier la qualité vocale? Quelles mesures de performances les opérateurs et les fabricants peuvent-ils utiliser pour mesurer la reproduction d'une conversation d'un utilisateur ?

Les ramifications de ces questions se sont récemment étendues avec l'évolution vers une architecture IMS (IP Multimédia Service) au travers de laquelle la voix est transmise en tant que paquets de données IP (Internet Protocol) dédiés entre les terminaux portables. La transmission de la voix sur IP dans les technologies telles que VoLTE (Voice over LTE) introduit de nouvelles formes de distorsions qui doivent être mesurées correctement pour donner une image précise de l'expérience réelle de l'utilisateur.

Cet article récapitule les indicateurs les plus utiles permettant de tester la qualité de la voix et décrit les méthodes de mise en œuvre des tests qui assurent à la fois des résultats précis et répétables et qui rentabilisent au mieux et le plus efficacement le temps de travail des ingénieurs.

Les normes de tests de qualité vocale

Le but d'un test de qualité audio est de convertir les caractéristiques de mesure d'un signal (avec des indicateurs objectifs tels que le temps de latence, le bruit, la gigue et les variations de fréquences) en un nombre représentant le jugement subjectif de l'utilisateur moyen à propos de la qualité de reproduction audio qu'il ou elle entend.

Bien entendu, les appels vocaux sont sujets à des dégradations nombreuses et variées (voir figures 1, 2 et 3) qui comprennent :

- Distorsions des codecs
- Erreurs de transmission
- Retards
- Bruit de fond
- Désadaptation entre les volumes sonores de deux interlocuteurs
- Echo acoustique
- Modifications anormales des variations habituelles de la hauteur de la voix, ou 'pertes spectrales'

D'une façon générale, ces distorsions peuvent être caractérisées soit comme « additives » (telles que le bruit, des clicks ou des pops) soit comme « soustractives » (telles que les pertes spectrales ou les trous dans le discours).

Un test de qualité vocale doit également fournir un moyen d'évaluer l'amplitude et la fréquence de nombreux éléments de distorsions et de pondérer leurs effets en accord avec la perception humaine, car certaines formes de distorsions sont plus gênantes que d'autres pour les utilisateurs. En utilisant des modèles basés sur des recherches menées avec l'aide de vrais utilisateurs, une valeur moyenne

de qualité peut en être dérivée, reflétant la perception des effets de distorsion par l'utilisateur ainsi que la qualité vocale absolue.

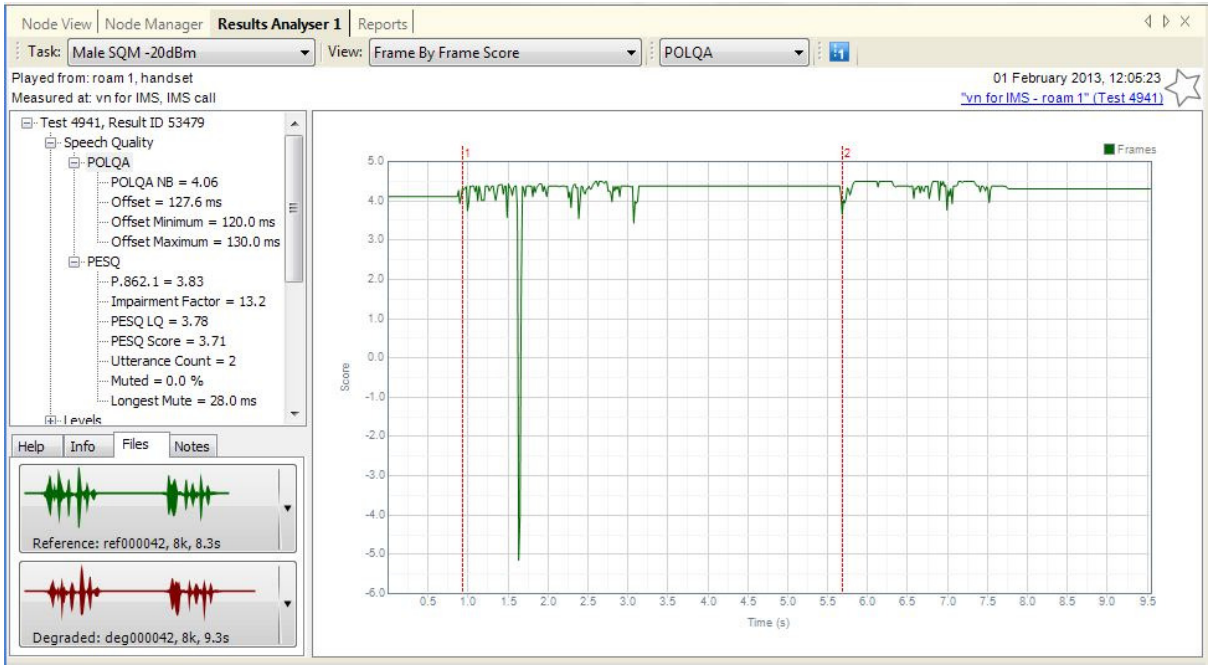


Fig. 1: Le diagramme de la valeur POLQA trame par trame dans le logiciel de test de qualité de la voix de Malden montre une perte de qualité de la voix aux alentours de 1,6s

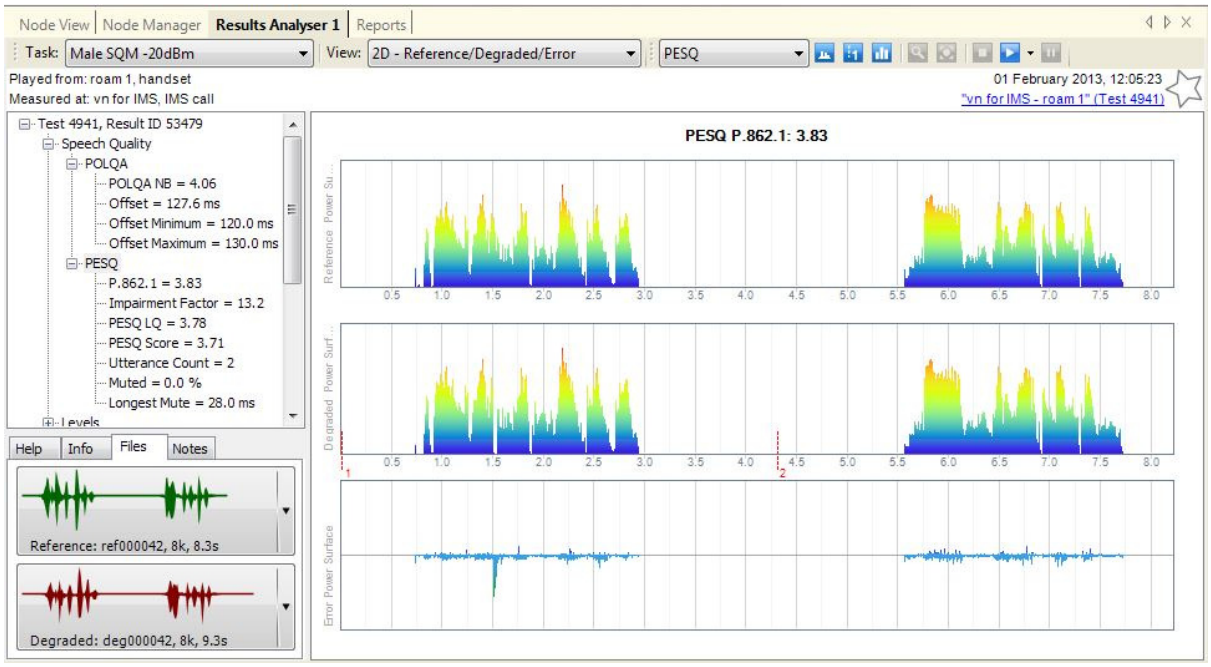


Fig. 2: La vue en « surface 2D » PESQ de Malden montre un trou dans le signal affiché en surface dégradée et la perte affichée dans la « surface d'erreur » à environ 1,5s.

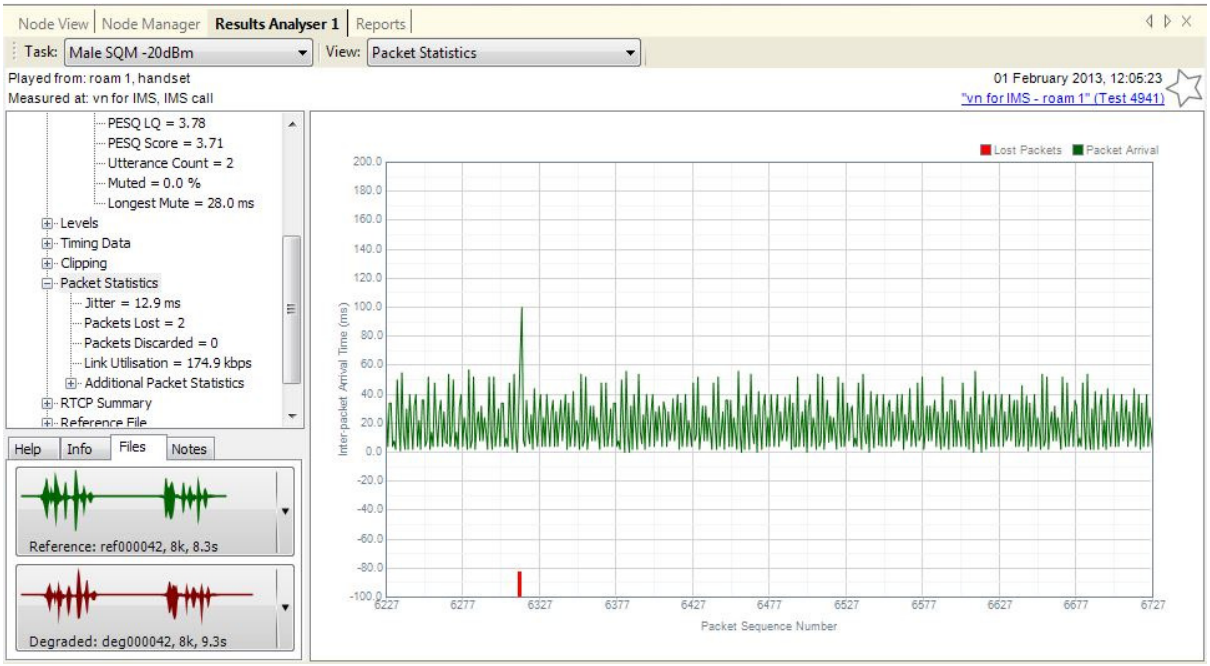


Fig. 3: La vue des statistiques des paquets de Malden montre une gigue significative et une perte de paquets autour du paquet numéro 6300

Dans les normes industrielles telles que PESQ (Perceptual Evaluation of Speech Quality) et POLQA (Perceptual Objective Listening Quality Assesment), des algorithmes comparent un signal de référence propre (appel vocal émis) avec sa version distordue après transmission (c'est à dire du côté de celui qui écoute). A partir de cette comparaison, ils génèrent une valeur unique MOS (Mean Opinion Score), classant la qualité de la voix généralement de 1 (mauvais) à 5 (excellent).

En plus de fournir une valeur de MOS claire, les tests POLQA et PESQ fournissent aussi des moyens d'analyse des caractéristiques de la distorsion. Ceci est utile car les dégradations de la qualité vocale peuvent être complexes par nature et difficiles à isoler et à corriger. Une information précise à propos des causes de distorsion permet aux opérateurs de prendre les mesures nécessaires telles que l'amélioration de l'utilisation des capacités du réseau.

Réussir à obtenir une grande qualité vocale sur les réseaux basés sur IP
Afin de supporter des services voix tels que la VoIP (Voix sur IP) et la VoLTE (Voix sur LTE), le consortium de l'industrie mobile 3GPP a introduit l'IMS (IP Multimedia Subsystem). Ce système fournit une architecture combinée sur une plateforme commune aux services vocaux et multimédia, qui facilite de plus la convergence des systèmes de communications filaires ou sans fil indépendants..

La VoLTE (aussi appelée Voice over IMS Profile et IMS VoIP) est la norme supportant les communications vocales sur IMS ainsi que le contrôle des communications utilisant les messages SIP (Session Initiation Protocol).

La VoLTE diffère des services d'appels et des applications établies (telles que Skype) basées sur IP qui utilisent des technologies VoIP simples: leurs paquets sont envoyés sur la base du 'best effort' d'une manière similaire à celle d'un trafic non vocal tel que des pages web. Par ailleurs, VoLTE utilise des indicateurs de Qualité de Service (QoS) intégrés dans le réseau radio en parallèle avec des protocoles et des procédures IMS spécifiques pour s'assurer que les ressources adéquates du réseau soient allouées aux appels vocaux. Ceci aide à conserver une latence des paquets minimale

et fournit des méthodes de correction d'erreurs améliorées. Ainsi, l'expérience utilisateur de la qualité vocale en VoLTE devrait être excellente.

En pratique, la cause la plus classique de dégradation de la voix dans les services de commutation de paquets tels que la VoIP et la VoLTE est la gigue: c'est une latence variable due, par exemple, aux temps de propagations et aux problèmes de files d'attente des données. Elle peut être la cause de paquets retardés ou perdus.

La gigue est généralement due à des problèmes de congestion du réseau, des décalages temporels et de files d'attente des données. Le symptôme au niveau de l'équipement de l'utilisateur est un discours erratique ou même embrouillé.

La latence est généralement le résultat des délais de transmission. Le décalage temporel entre l'émission à partir de la source du signal et sa réception, normalement mesuré en millisecondes, devient perceptible par l'utilisateur s'il s'allonge.

Ceci veut dire que les différentes façons dont la capacité du réseau IP est utilisé, les ressources réseau allouées, la mise en œuvre des protocoles VoLTE sur les mobiles IP ainsi que les codecs vocaux utilisés auront tous une effet marqué sur la perception de la qualité vocale par l'utilisateur. Par conséquent, comment les mesures PESQ et POLQA reflètent-elles la qualité vocale perçue ?

Comment fonctionnent les algorithmes de mesure vocale
PESQ compare un signal original en entrée avec un signal de sortie aligné temporellement qui a été dégradé après être passé au travers d'un système de communications ou DUT (Device Under Test). POLQA est son successeur qui est utilisé pour l'analyse comparative avancée et l'évaluation de la qualité vocale sur les réseaux fixes et sans-fil basés sur IP. POLQA a été standardisé par l'ITU (International Telecommunication Union) sous la recommandation P.863 et peut s'appliquer à l'analyse de qualité vocale des réseaux mobiles 3G et LTE (4G).

L'algorithme POLQA atténue les limitations de PESQ fournissant de nouvelles caractéristiques significatives et améliorant les capacités de traitement des signaux audio avec de plus grandes bandes passantes de données et une meilleure gestion des variations des retards.

De plus, il possède un double mode de fonctionnement qui permet aussi d'évaluer des signaux vocaux Haute Définition qui utilisent de plus grandes bandes passantes que les appels vocaux habituels. Dans ce cas, au lieu d'utiliser une fréquence d'échantillonnage de 8 kHz comme avec le PESQ, POLQA utilise comme référence une fréquence d'échantillonnage de la parole de 48 kHz.

Il supporte également la multitude de codecs existants et nouveaux utilisés dans les systèmes GSM, W-CDMA, LTE, VoIP, IMS et VoLTE. POLQA possède aussi la capacité de fonctionner efficacement sous l'influence de nouveaux types de distorsions telles que les interruptions spécifiques liées aux réseaux, ainsi que les distorsions dues aux variations de conception des interfaces acoustiques comme les hauts-parleurs et les micros.

Mise en œuvre des tests vocaux dans un environnement simulé
Comme il est décrit ci-dessus, l'évaluation des équipements du réseau et des terminaux des utilisateurs est essentielle à la fois pour les fabricants de mobiles et les opérateurs de réseau. Quand ils sont effectués en laboratoire, les appels de test de qualité vocale implémentés dans les équipements de test mettent en œuvre les tests PESQ et POLQA d'une manière simple et intuitive pour l'ingénieur de test permettant une reconfiguration rapide des paramètres de test.

C’est pourquoi Anritsu a amélioré son testeur d’application populaire MD8475A pour qu’il soit compatible avec les logiciels de test PESQ et POLQA de Malden, spécialiste des tests de qualité vocale. Le système MultiDSLA de Malden est un outil de mesure et d’analyse de la qualité vocale bidirectionnel qui intègre les capacités complètes PESQ et POLQA

Le MD8475A dispose d’un serveur SIP intégré ainsi que des services IMS, simplifiant la configuration d’un réseau simulant n’importe quelle configuration choisie par l’utilisateur. Puisque le MD8475A supporte le système MultiDSLA de Malden incluant l’option VoxPort Packet (VPP) dans le domaine IP et l’interface à deux canaux du DSLA dans le domaine analogique, il est également facile de préparer et d’établir un appel VoLTE entre un ordinateur hôte utilisant le logiciel Malden et un Dispositif Sous Test (DST/DUT) tel qu’un smartphone LTE via le réseau sans fil simulé par le MD8475A (voir fig. 4 et 5)

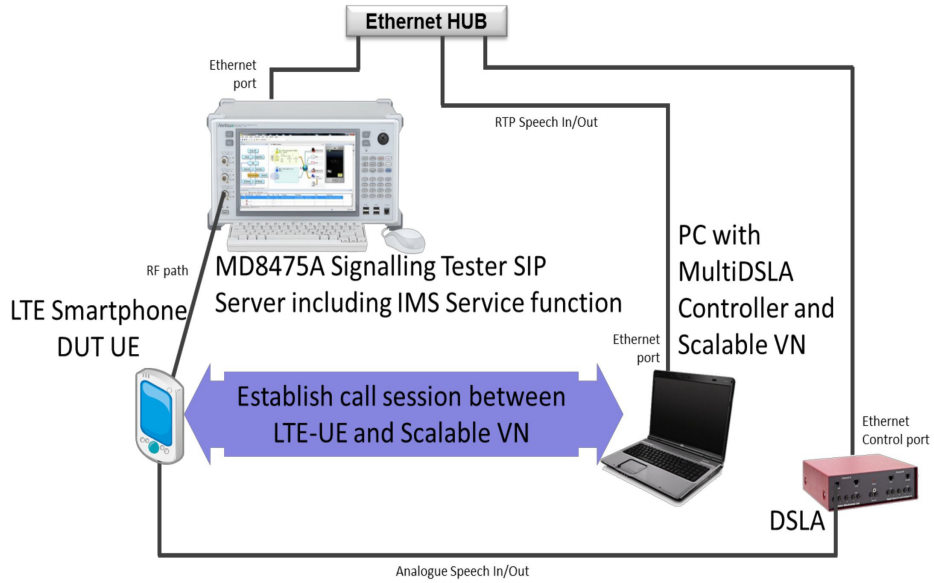


Fig. 4: Exemple de configuration de test pour un test POLQA de qualité vocale d’un smartphone LTE



Fig. 5: Configuration de test de la figure 4

Si l'on prend l'exemple du test d'un smartphone LTE , le simulateur de signalisation d'Anritsu contrôle la configuration du test, l'enregistrement et l'établissement d'un canal de connexion LTE avec le terminal sous test. L'application de services IMS intégrée dans le MD8475A établit une connexion avec le serveur SIP en utilisant les messages du protocole de signalisation habituels échangés entre le DST et l'application VPP. Par ailleurs, VPP est configurée pour permettre l'enregistrement sur le serveur IMS/SIP.

Un appel VoLTE peut alors être configuré et initié depuis l'ordinateur hébergeant VPP vers le DST via le simulateur de signalisation. L'outil MultiDSLA de Malden affiche ensuite les valeurs MOS pour les algorithmes PESQ et POLQA basées sur les résultats capturés depuis le DST.

Comme on peut le voir, la mise en place simple de ce test, requiérant uniquement l'utilisation du testeur de signalisation ainsi que le matériel et le logiciel Malden pour le test de qualité vocale, constitue l'ensemble de ce qui est nécessaire pour capturer des résultats de test PESQ et POLQA précis et répétables. Cette configuration de test offre des possibilités de variations infinies de la configuration du réseau pour examiner le comportement du DST ou bien un appel complet de bout en bout selon les conditions de test choisies.

En mettant en œuvre les tests de qualité vocale de cette façon, les opérateurs et les fabricants de mobiles s'assurent que les utilisateurs puissent apprécier une qualité de reproduction vocale sur leurs téléphones mobiles au moins aussi bonne que leur expérience sur les équipements filaires.