Second International Balkan Conference on Communications and Networking Podgorica, Montenegro, June 6-8,2018

VoLTE Performance Analysis and Evaluation in Real Networks

Bujar Krasniqi
Faculty of Electrical and Computer Engineering
University of Prishtina
Prishtina, Kosovo
bujar.krasniqi@uni-pr.edu

Gentian Bytyqi
Faculty of Electrical and Computer Engineering
University of Prishtina
Prishtina, Kosovo
gentian.bytyqi@uni-pr.edu

Driton Statovci
Institute of telecommunications
TU WIEN
Vienna, Austria
driton.statovci@nt.tuwien.ac.at

***Abstract*—Moving voice traffic from circuit-switched domain to packet-switched domain seems to be the only rational way to improve quality of experience. Voice over Long-Term Evolution (VoLTE) is a technology that enables transmission of voice calls over LTE network by using IP Multimedia Subsystem (IMS). VoLTE is bandwidth efficient and offer better speech quality than legacy technologies such as Universal Mobile Telecommunications System (UMTS) and Global System for Mobile communication (GSM). In this paper, we analyze the impact of radio propagation conditions, codec type implementation, and mouth-to-ear delays on VoLTE service performance. In particular, we consider drive test, stationary test, and train test measurements in real VoLTE network deployments. Testing results indicate that when both end users are using VoLTE technology, the obtained speech quality is impressive compared to the case when one user is using legacy technologies (UMTS or GSM) or performing Circuit Switched Fall Back (CSFB) to legacy. Furthermore, results show that VoLTE call setup time and mouth-to-ear delays are shorter while call reliability is comparable with legacy technologies. However, the percentage of VoLTE dropped calls is high and this indicates a potential for further optimization in the current implementation of VoLTE.**

Keywords—VoLTE, LTE, CSFB, SRVCC, UMTS, GSM, handover

I. Introduction

The cellular communications industry has witnessed extensive growth since the mid of 1990's. The demand for higher speeds and better Quality of Experience (QoE) in mobile communications network is continually increasing. Long-Term Evolution (LTE) supports only packet-switched network across an all-IP system, whereas previous cellular networks GSM or UMTS support both packet- and circuit-switched network. In the beginning of LTE deployment, all voice traffic is handled by legacy Circuit-Switched (CS) networks, while data traffic is handled by LTE packet-switched networks. Some solutions have been proposed in order to deliver voice services in LTE. In practice mostly it is used the so-called Circuit Switched Fall Back (CSFB) approach, which enables voice transmission in LTE via legacy networks.

Even LTE was originally designed as packet-switched network, its Quality of Service (QoS) and capacity provide mobile subscribers significant improvement in QoE for voice services, such as High Definition (HD) voice. To reliably deliver voice calls over all IP network, mobile operators have adopted, deployed, and recently started with service called Voice-over-LTE (VoLTE). Until the end of 2016, around 156 mobile operators in 73 countries have invested in VoLTE including 102 operators that have launched HD voice service using VoLTE [1]. The VoLTE solution introduces the voice functionality in the LTE network by using new IP Multimedia-Subsystem (IMS) framework [2].

A performance analysis and optimization of CSFB from LTE to UMTS is performed by authors in [3]. Differently, the authors in [4] have analyzed the VoLTE performance of user equipment (UE) using carrier aggregation. A deeper analysis and evaluation of CSFB and VoLTE in terms of end-to-end assessment of call setup delay under different radio conditions has been done by authors in [5]. Our main contributions in this paper are the analysis and evaluation of CSFB performance, VoLTE call setup, VoLTE speech quality, end-to-end delay and call reliability for all possible scenarios such us drive test, stationary test, train test and overall.

The rest of the paper is organized as follows. In Section II we shortly introduce VoLTE technology and the key performance indicators to analyze the performance of VoLTE. In Section III we perform a detailed performance analysis of VoLTE under different test scenarios. Conclusions of our work are drawn in Section IV.

II. Voice over LTE

The absence of CS network in LTE has led the industry and standardization bodies to propose various solutions to support voice over LTE network. The CS fallback solution is defined in [6] and supports voice services in LTE systems by reusing the existing GSM/UMTS network. Mobile users connected to a LTE network for data services are obligated to fall back on a legacy network when they make or receive a voice call. A CS fallback enabled mobile device connected to LTE, may use UMTS/GSM network to connect to the CS domain. Thus, users have to perform both registrations in LTE and GSM/UMTS networks in order to proceed with the call setup. The registration is done by mobility management entity.

A. *VoLTE Key Performance Indicators*

The most important performance indicators in VoLTE service are: speech quality, call setup time, call setup success ratio, call reliability, handover success ratio and dropped calls ratio [7]. Speech quality depends deeply on the voice codec sampling rate and the resulting audio bandwidth. Adaptive Multi-Rate Narrow Band (AMR-NB) provides audio bandwidth spectrum from 80 Hz to 3700 Hz while Adaptive Multi-Rate Wide Band (AMR-WB) extends the audio bandwidth spectrum from 50 Hz to 7000 Hz as illustrated for particular examples in Fig. 1 and Fig. 2. Higher voice bandwidth of AMR-WB codec results in better quality and more natural sound.and therefore the VoLTE users using this type of codec will experience better QoE compared to voice in CS networks.

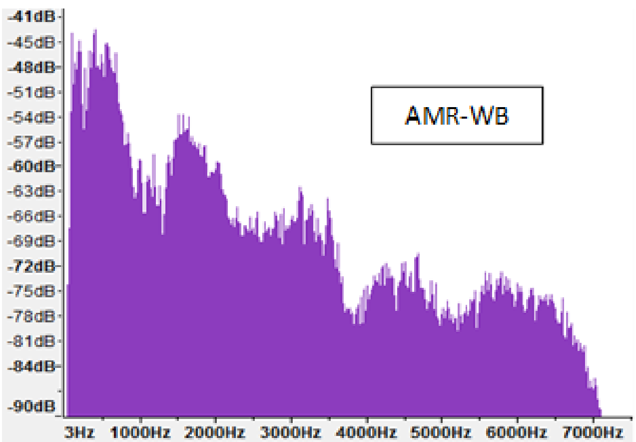

Fig. 1. An example of audio bandwidth of Adaptive Multi-Rate Wide Band (AMRWB) codec.

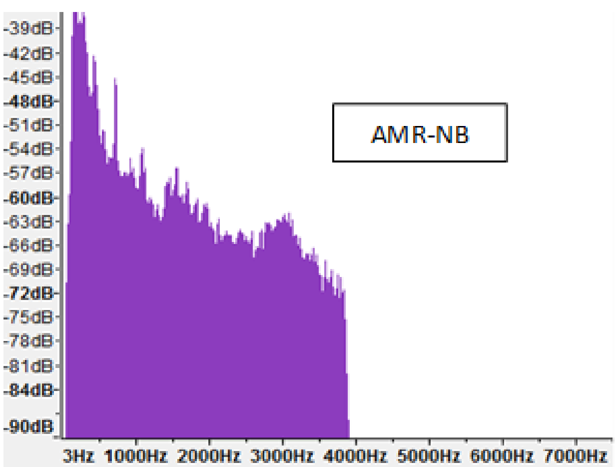

Fig. 2. An example of audio bandwidth of Adaptive Multi-Rate Narrow Band (AMR-NB) codec.

Additionally, another type of codec which offers better quality than AMR-WB is the evolved HD voice, which is also called Enhanced Voice Service (EVS). EVS extends the HD voice service experience with natural voice and music by supporting the full human voice frequency range. In order to benefit from this codec, the user equipment has to support supper wideband (up to 14 kHz) or full band (up to 20 kHz) audio frequency range [8]. Speech quality is evaluated using Perceptual Objective Listening Quality Analysis (POLQA) wideband algorithm [9].

III. VoLTE Performance Analysis

In this section, we study the effects of radio propagation conditions, codec type implementation, and end-to-end delays on VoLTE service. Furthermore, we analyze the differences in speech quality between end-to-end VoLTE call and a voice call which combines VoLTE with UMTS/GSM or is employing Single Radio Voice Call Continuity (SRVCC) and CSFB. The VoLTE speech quality, mouth-to-ear delay, call setup time, and call reliability are analyzed and compared with legacy technologies.

A. *Measurement Methodology*

To perceive the end user experience while they are using VoLTE service, both mobility and stationary test measurements are conducted gathering real-time data from the network. Numerical values of relevant parameters and their statistical distributions have been extracted from the measurement campaigns carried out in different mobile operators in Europe where the VoLTE technology have been already widely deployed. Measurements have been conducted across the big metropolitan areas, the connecting roads between those cities and smaller towns close to the connecting roads. Technical performances of the networks have been tested from the point of view of the end consumer on VoLTE capable smartphones. The measurements were conducted between two independent mobile-to-mobile systems. Both mobiles, A-party and B-party, were calling each other using Samsung Galaxy S5 LTE Cat.4. Voice test calls were mobile-to-mobile while the measurement cars were moving and the duration of each call was 100 seconds.

To perform measurements, different Rohde&Schwarz test-solutions have been used. Rohde&Schwarz ROMES software platform in combination with other test and measurement equipment have been involved in recording and visualization of test parameters. In addition to signal strength, the Received Signal Strength Indicator (RSSI) and the Reference Signal Received Quality (RSRQ) are displayed for every cell as well as the reference Signal to Interference-plus-Noise Ratio (SINR).

B. *Speech Quality Test Results for Different Technologies*

This section presents the differences in the call quality between end-to-end VoLTE call and a voice call which combines VoLTE with UMTS and GSM or employing CSFB and SRVCC. The speech quality is evaluated in real VoLTE and GSM/UMTS network deployments. Drive test measurements were performed using two measurement cars. In both cars are located mobile phones, A-party and B-party, where both mobiles are calling each other. Speech quality was measured by playing reference voice samples on talking side and recording the transmitted samples on listening side. POLQA wideband algorithm was applied to derive the average

Mean Opinion Score (MOS) values. For each single call eight voice samples were recorded and total number of samples is showed in Table 1.

TABLE I. NUMBER OF RECORDED SAMPLES FOR DIFFERENT SCENARIOS

Different call scenario	**Technologies**		
	VoLTE-VoLTE	***VoLTE-UMTS***	***VoLTE-GSM***
Nr. Of recorded samples	1394	1998	330

The cumulative distribution of speech quality between end-to-end VoLTE and VoLTE to UMTS/GSM calls is shown in Fig. 3. The results indicate that when both end users are using VoLTE technology, the obtained speech quality is impressive comparing with the case when one mobile is using legacy or performing CSFB to legacy. The small number of samples for VoLTE to GSM call scenario may not accurately quantify the real-word impact and can be further improved with more recorded samples. In VoLTE to VoLTE call scenario the average of speech quality is 4.11 MOS; while in VoLTE to UMTS/GSM is 3.35 MOS and 2.77 MOS respectively. Voice calls involving multiple technologies may result in using codecs in tandem. For example, in the case of VoLTE to GSM call, the output of AMR-NB codec voice data for purpose of interworking must beconverted into another format e.g., Pulse Code Modulation (PCM), which further degrades voice quality.

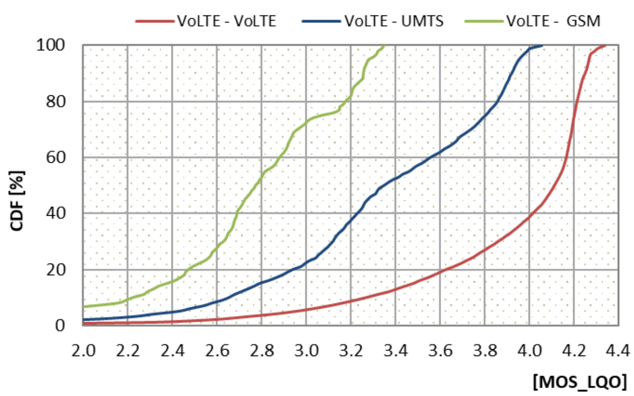

Fig. 3. Cumulative Distribution Function (CDF) of speech quality for different call scenarios.

For the VoLTE to VoLTE call scenario only AMR-WB codec is used and PCM based transport cannot be used with the wideband codec because PCM only applies to a narrowband voice. Therefore, AMR-WB codec use Transcoder Free Operation (TrFO) [10]. TrFO is a solution where encoded speech is transmitted through the network as a packet, without the need to decode it into PCM format, in this way the quality of mobile-to-mobile calls is improved.

It is important to mention that in VoLTE to UMTS call scenario, AMR-WB and AMR-NB codecs were used in both sides, but in 85% of the calls the codec rate was less than 12.65 kbps. Whereas, for VoLTE to GSM only AMR-NB was used with highest codec rate of 12.2 kbps. The codec operation used in VoLTE should follow the limitations on codec mode changes in UMTS and GSM networks.

The difference in quality between VoLTE drive and stationary tests measurements is shown in Fig. 4. The total number of recorded voice samples for drive test was 13594 while for stationary test was 6717 voice samples.

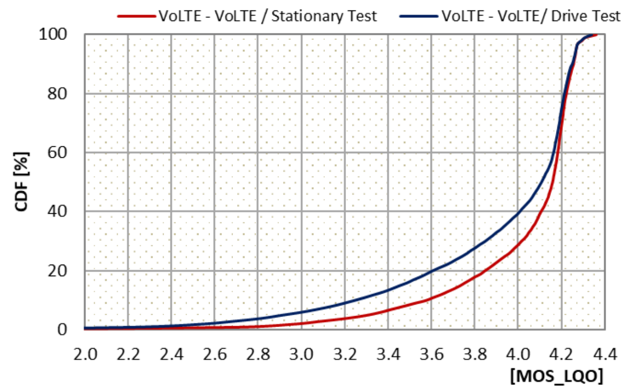

Fig. 4. CDF of VoLTE speech quality for drive tests and stationary tests.

Measurement results indicate that the speech quality is lower when both VoLTE UEs are in mobility. For drive test measurements the probability that speech quality will take values less than 4.0 MOS is 39 %, while for stationary test it is 29 %. This is due to many factors; one of them is the higher probability for making handover in mobility scenario. VoLTE handover is SRVCC, which seamlessly handovers the VoLTE call session to GSM/UMTS networks. If the path handover interruption is high then the speech frames will be lost and this will impact the speech quality. Finally, the degradation of speech quality during and after SRVCC handover is represented in Fig. 5.

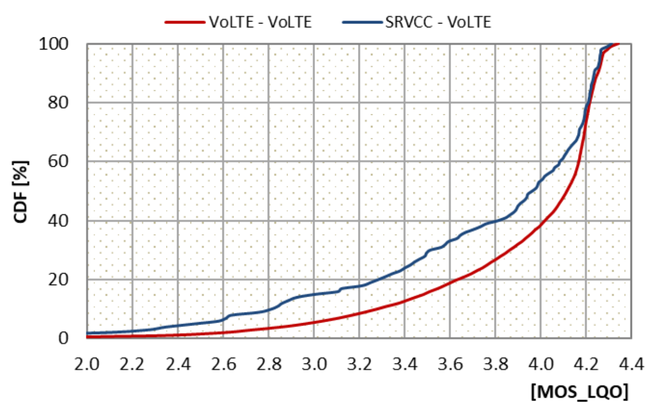

Fig. 5. CDF of speech quality: VoLTE-VoLTE and SRVCC-VoLTE

Only drive test measurements were conducted for two different call scenarios. In the first scenario, both mobiles are using VoLTE without SRVCC handover. In the second scenario during the VoLTE call mobiles in car 1 were performing SRVCC handover to UMTS/GSM technology, while mobiles in car 2 were using VoLTE and were not performing SRVCC handover.

The drive test results indicate that SRVCC handover to UMTS or GSM has a large impact on speech quality. The path interruption time (voice interruption) due to handover was on average 150 ms. For voice services such as VoLTE, the size of packets is small, and the inter-arrival time is 20 ms. The path interruption of 150 ms due to SRVCC handover will cause packet loss, which degrades the speech quality. From the results in Fig. 5 can be concluded that the average of speech quality is 4.11 MOS without handover, while 3.9 MOS when the SRVCC is employed. Both inter- Radio Access Technology (RAT) handover and session transfer in the core network contribute in the voice interruption time. To minimize the voice interruption, the SRVCC initiates the inter-RAT handover and session transfer simultaneously so they can run in parallel.

C. VoLTE End-to-End Delay Test Results

VoLTE end-to-end delay (mouth-to-ear delay) is one of the most important metric when we measure VoLTE performance. Mouth-to-ear delay is latency between the speaker utters a word and the listener actually hears it [11], containing both one-way latency in the network and the time spent on encoding and decoding audio packets. The maximum mouth-to-ear delay in good quality communication should be lower than 250 ms [12]. According to ITU-R and 3GPP standard, the 153 ms mean mouth-to-ear delay of VoLTE suggests excellent user satisfactory. To calculate the delay budget the assumption from [12] is used. The speech coder takes 20 ms of speech samples and encodes into a speech frame. The voice encoding delay is assumed to be 30 ms, including 20 ms frame size. The uplink transmission takes 50 ms, while Evolved Node B (eNB) processing take 5 ms and gateway processing 1 ms. The transport delay is assumed to be 10 ms, while downlink transmission takes 50 ms. Decoding and processing delay in the receiver UE is assumed to be 5 ms. Under these assumptions the mouth-to-ear delay is approximately 150 ms and it is illustrated in Fig. 6.

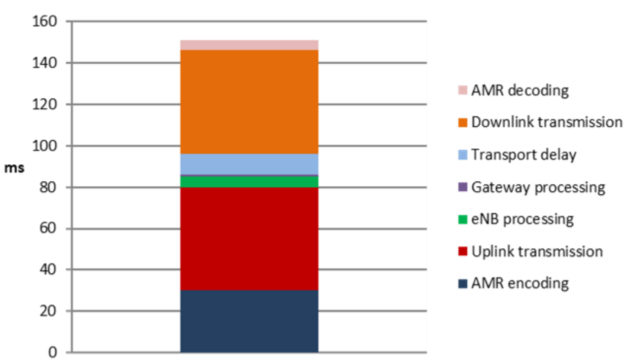

Fig. 6. Average speech quality: VoLTE-VoLTE and SRVCC-VoLTE [12]

VoLTE uses dedicated bearer offering Guaranteed Bit Rate (GBR) for transporting voice media to minimize delays. In another hand, eNB may decrease the GBR rate for UEs in undesirable network conditions. To proof the QoS guarantee provided by dedicated bearer, VoLTE calls are performed under various channel conditions. To calculate the mouth-to-ear delay, a connection box (GPS-time synchronized) is used. Fig. 7 shows mouth-to-ear delay when both mobiles are using VoLTE. Fig. 8 presents mouth-to-ear delay when one mobile is using VoLTE the other one is using legacy or when both mobiles are using legacy.

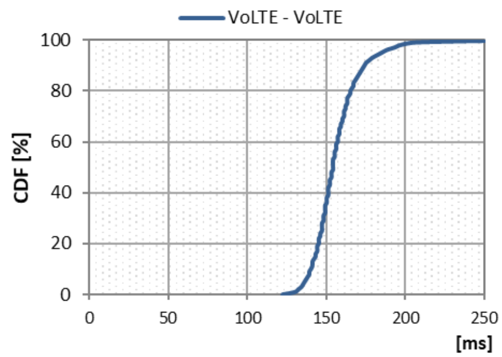

Fig. 7. Mouth-to-ear delay in VoLTE-VoLTE.

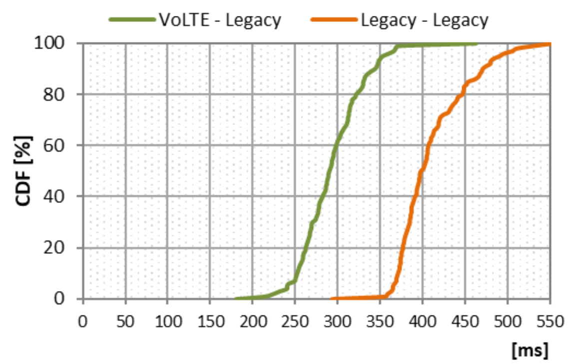

Fig. 8. Mouth-to-ear delay in VoLTE-Legacy and Legacy-Legacy.

From the test results shown in Fig. 7, one can see that VoLTE calls have low mouth-to-ear delay. The average of mouth-to-ear delay for VoLTE is found to be 154 ms. This value is 4 ms greater than the ITU-R requirements for one-way VoLTE end-to-end delay to experience high quality voice call. In another hand when users are performing CSFB to legacy calls or both users are using legacy (cf. Fig. 8) the average of mouth-to-ear delay is 280 ms and 410 ms, respectively. This will have large impact on the user experience. Radio frequency conditions will also impact the mouth-to-ear delay. Fig. 9 represents the RSSI impact in mouth-to-ear delays.

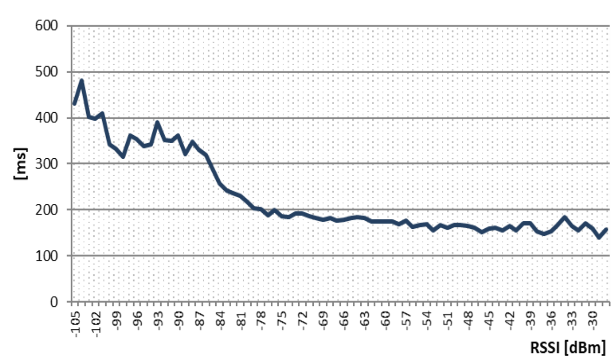

Fig. 9. Mouth-to-ear delay versus RSSI values.

Results indicates that when the RSSI is greater than –78 dBm the mouth-to-ear delay is constantly less than 200 ms, while for RSSI lower than –78 dBm the mouth-to-ear delay might reach up to 480 ms.

D. VoLTE Call Setup Time and Reliability

The equal number of call attempts and the same mobility route was taken for VoLTE and legacy (GSM/UMTS or CSFB). Calls were mobile-to-mobile and both mobiles were in mobility or stationary, except for railway test, one mobile was in train and the other one in highway. Fig. 10 shows the call setup time for VoLTE and legacy for mobility and stationary tests.

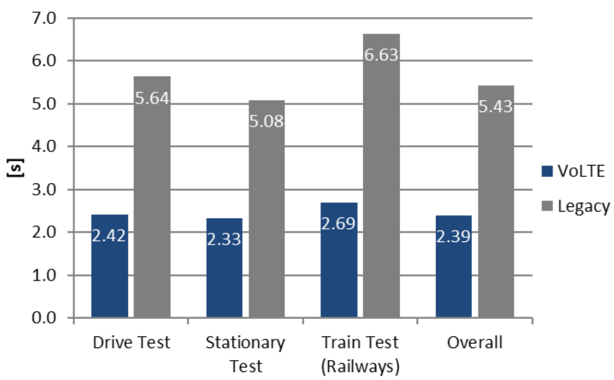

Fig. 10. Call setup time for mobility and stationary tests.

VoLTE call setup time shows an impressive improvement comparing with call setup time in legacy. The mean VoLTE call setup time is 2.39 s, which is much better than legacy calls. Employing CSFB while making a call lead a higher call setup time and indicates a potential for optimization. A call is defined as a successful call if the call is successfully established and maintained for at least 100-seconds during the test, so the call reliability can be measured as the probability of making a successful call. The comparison results of call reliability for VoLTE and legacy in both stationary and mobility cases are shown in Fig. 11.

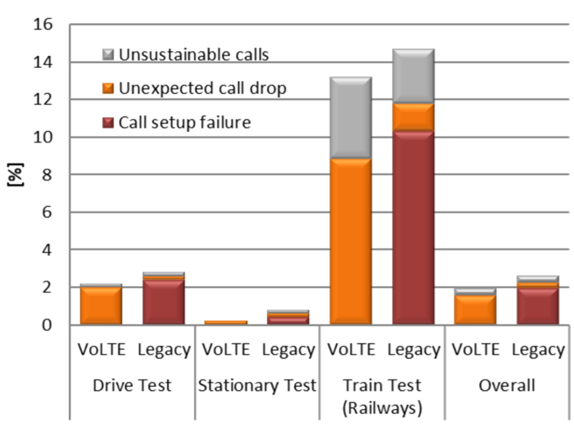

Fig. 11. Call reliability for mobility and stationary tests.

VoLTE shows an improvement in call setup time, while high number of dropped calls leaves a room for further optimization. In the case of mobility tests, VoLTE shows higher number of dropped calls, especially for train tests.

IV. CONCLUSION

In this paper, we presented a study about Voice over LTE (VoLTE) with focus on performance analysis in real network deployments. We considered both mobility and stationary test measurements to gather real-time data from the network. From the test results, we found that that for good radio channel conditions the best possible speech quality is delivered using Adaptive Multi-Rate Wide Band (AMR-WB) with coding rate of 23.85 kbps, while for bad radio conditions (Received Signal Strength Indicator (RSSI) < –85 dBm) the best speech quality is achieved using AMR-WB with coding rate 12.65 kbps. The average mouth-to-ear delay was 154 ms. In another hand, when RSSI < –78 dBm the mouth-to-ear delay has been increased drastically. Finally, the call reliability and call setup time for VoLTE and legacy calls was analyzed. Test results indicate clearly that VoLTE shows an improvement in call setup time while the number of dropped calls is higher than in legacy technologies.

V. REFERENCES

[1] Juniper Research “VoLTE - Voice over LTE: Key Market Movements 2017,” tech. rep., Juniper, 2017.

[2] Ayman ElNashar, Mohamed El-saidny, Mahmoud Sherif. Design, Deployment and Performance of 4G-LTE Networks: A Practical Approach, May 2014

[3] J. Bautista, S. Sawhney, M. Shukair, I. Singh, V. Govindaraju and S. Sarkar, “Performance of CS Fallback from LTE to UMTS,” IEEE Comm. Magn. September 2013

[4] J. Sonkusare, and S. Dhage, “LTE UE Performance Evaluation for Carrier Aggregation and Voice-over-LTE,” IEEE Applied and Theoretical Computing and Communication Technology (iCATccT), 2015 International Conference on. October 2015

[5] A. Elnashar, M. El-Saidny and M. Mahmoud, “Practical Performance Analyses of Circuit-Switched Fallback and Voice Over LTE,” IEEE Transactions on Vehicular Technology, Vol. 66, No. 2, February 2017

[6] 3GPP 3rd Generation Partnership Project; Technical Specification 23.272, "Circuit Switched (CS) fallback in Evolved Packet System (EPS)", Stage 2 (Release 13).

[7] J. rankin, A Costaiche, J. Zeto and K. O’Neil “Validating VoLTE A Definitive Guide to Successful Deployments”, IXIA 2014.

[8] Ericsson white paper, “Evolved HD Voice for LTE”, October 2014.

[9] International Telecommunication Union, ITU-T P.863, Perceptual Objective Listening Quality Assessment, September 2014.

[10] C. Birkehammar, S. Bruhn, P. Eneroth, K. Hellwig and S. Johansson, “New high-quality voice service for mobile networks”2006.

[11] C.Agastya, D.Mechanic, and N.Kothari.”Mouth-to-ear latency in popular voip client”.2009.

[12] 3GPP 3rd Generation Partnership Project; Technical Specification 23.216, “Single Radio Voice Call Continuity (SRVCC)”, Stage 2 (Release 13)

[13] Shyam Chakraborty, Tomas Frankkila, Janne Peisa and Per Synnergen “IMS Multimedia Telephony over Cellular Systems”, 2007.