2nd International Engineering Conference (IEC2017)
Federal University of Technology, Minna, Nigeria

Evaluation of Speech Quality Based on Qos Key Performance Index (KPI): A Survey

*Gbadamosi, A. S[1], & Maliki, D[2]
[1]Electrical and Electronics Engineering Department, Federal University of Technology,PMB 65 Minna Niger State, Nigeria
[2]Department of Computer Engineering, Federal University of Technology,PMB 65 Minna Niger State, Nigeria
*Corresponding author email: g_safiu.its@futminna.edu.ng, +2348060511079

ABSTRACT

The utmost contest between cellular operators have left the customers highly selective with the quality of speech based on the quality of service (QoS) experienced. In order to quantitatively and qualify the network performance based on speech quality, different key performance indicators (KPI) used, are classified to aid data collection independently of the cellular providers. In this paper, a survey of different techniques of KPI data collection methods that are used to reduce impairment as its affect speech are discussed and classified according to their operational mode. It provides a detailed report about the perceptuality of speech quality and its statistical expectation of the network. Each of these techniques has its advantages and disadvantages as applied to underscore the network performance from a user's point of view. These techniques evaluate degradation caused by various impairments in cellular networks. It is our hope that, if user get hold of these applications, it will go a long way to keep cellular operators on their toes.

Keywords: *Key performance Indicator (KPI); Quality of Service (QoS); Speech quality; Techniques*

1 INTRODUCTION

The experience of user's with the quality of speech over GSM services varies with network operators. These experiences range from cross talk, loss of speech signals, mixture of unwanted signals when receiving calls, echoing, background noise and very low output speech signal from the users' terminals among others. These experiences have created opportunities for users to make decisions on either to continue to use a particular service provider or not. The speech quality is affected by various impairments that depends on the transmission and switching technology used, which affect its performance in various ways (Celinus Kiyea, 2014).

The main goal of any service provider is to achieve QoS that is satisfactory for the users (Adegoke, A. S, & Babalola, I. T, 2011). But this has become a serious issue since the view and requirement of a user on what is termed satisfactory differs in terms of experience and what was billed for the service.

This discrepancy has raised many questions that have led to various researches by different groups. Among the question raised were [ECC report 51];

• What degree of level of quality can actually translate into a substantive understanding of QoS to the users?

• What encrypting methods are available to scale down the effect of network deterioration in speech quality?

• What voice processing applications can enhance the quality of speech for real time, considering various impairments that distort the speech signal.

• How efficient are these techniques (applications) when the challenges lies on the user's terminals or access system [ECC report 51].

There are various applications, algorithms and models based research paper in literature that have been discussed. The Motivation of this paper, is to classify all these algorithms and applications based on their predictive and evaluation capacity of speech in relation to QoS approach. Also, the performance, advantages and disadvantages of these applications were mentioned as its affect the enhancement of speech quality.

The rest of the paper is organized as follows. Section 2 describes the classification and discussion of Speech quality based on QoS KPI data collection approach. Section 3 describes the previous algorithms used in reducing impairment in speech such as noise. Section 4 concludes the work of the paper.

2 CLASSIFICATION OF SPEECH QUALITY BASED ON QoS KPI

Several parameters that influence the quality of service (QoS) in relation to the experience of users with the quality of voice (Speech) over GSM services can be classified into three (3) namely;

Network Management System (NMS), Speech Quality Index (SQI) and Field test and Survey.

Figure1 below illustrates the classification of speech quality based on QoS KPI.

represent the overall network performance, but do not include the speech quality key performance indicator (KPI).

In a GSM stable network, Dropped Call Rate (DCR) is always near zero percent (0%) and hence other parameters most especially BER, RxQual, etc.; are used to determine the user's perceived speech quality (Barile *et al*., 2006; Kumar & Saini, 2010). In this paper, brief explanation of these parameters is discussed.

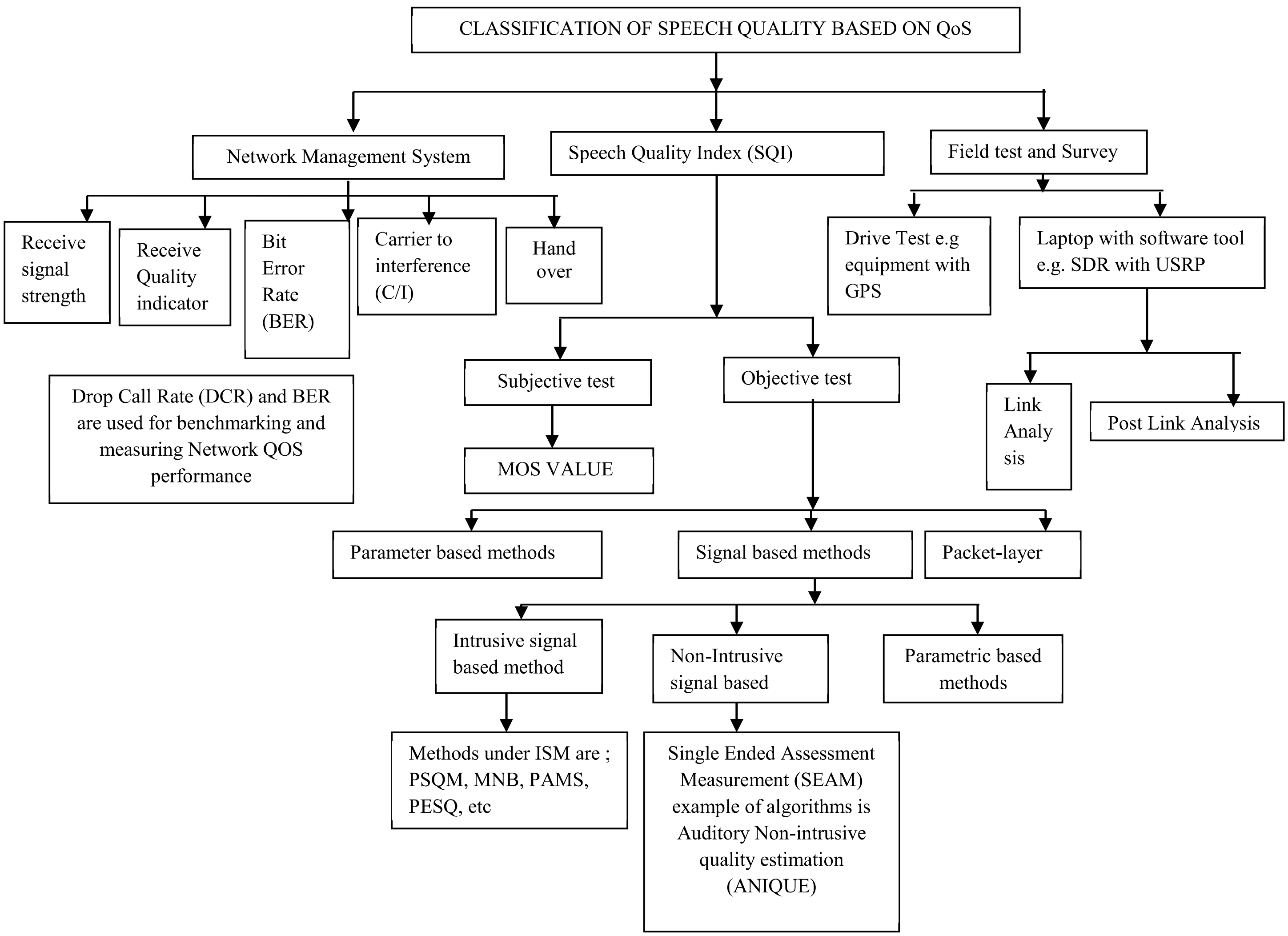

Figure 1 CLASSIFICATION OF SPEECH QUALITY BASED ON QoS APPROACH
Source: Cote, (2011); Počta *et al.*(2010); Mahdi and Picovici, D. (2007); Kumar and Saini (2010)

2.1 NETWORK MANAGEMENT SYSTEM (NMS)

Cellular network provider employs effective tools that are moderate to eliminate problems and enhance the network. Optimization consists of authentication, observation and improving the efficiency of the radio network, which are collected either from Network Management system (NMS) statistics or from Drive tests. NMS statistics

2.1.1 BIT ERROR RATE (BER)

By GSM network, the amount of bits of a data stream over a channel of communication that have been distorted with noise, bit synchronization errors and interference is defined as BER. It is defined as the quotient of bit errors over the whole number of transmitted bits during the period of study.

$$BER = \frac{bit \quad error}{number \quad of\ bits} \qquad (1)$$

The efficiency of BER is expressed as a percentage unit less value (Gbadamosi et al., 2015; Shannon, 1948). BER becomes a problem to the service provider when the value of BER is higher than the set threshold value, which in this case will affect speech quality. Unless the BER is maintained, the problem of call clarity cannot be attained.

2.1.2 RECEIVED SIGNAL STRENGTH INDICATOR (RSSI)

The most essential indicator of network is a strength of signal. It is defined as the amount of power in a received radio signal. It is hidden to the person with the device, but visible to the network administrators (ITU-Recommendation (G.1000), 2001). In GSM the Rxlev is often used to replace RSSI. The difference between the two terms is that, RSSI was employed in analog signal while Rexlev is used in digitalized networks.

2.1.3 RECEIVE SIGNAL QUALITY (RXQUAL)

The simplest form of measurement indicator on GSM network is the RXQUAL value. It expresses the quality of received signal. These values depend on all frames of the slow access channel (SACCH) multi-frames, in which the bits might have been altered along the radio frequency (RF) path or damage due to interference or fading. This demonstrates that, if the use of discontinuous transmission downlink (DTX DL) was employed, false values would be generated for that period. Since it involves bit error measurements at periods when no data was sent. This result to very high BER and only manifests the expected value of BER over a time period of 0.5s. The amount of bit errors was accrued to the BER sum for each SACCH multi-frame. The results were categorized from 0 to 7 in accordance with BER- RxQual conversion table. RxQual was used in GSM as a part of Network Measurement Reports (NMR) (Kumar and Saini, 2010).

2.1.4 CARRIER TO INTERFERENCE (C/I) RATIO

The ratio between the expected value of received modulated carrier power C and the expected value of received co-channel interference power I, i.e., the presence of an unwanted signal, from other transmitters than the useful signal. (Kumar and Saini, 2010).

2.1.5 HANDOVER

The portion of fortunate outgoing handover attempts is called Handover success rate. An utmost value of the handover failure would have higher consequences on the dropped calls Rate (DCR). DCR is the quotient between the drop calls in traffic channels during the conversation to the amount of successful "seizers" on the cells or areas (Kumar and Saini, 2010).

2.1.6 LAYER 3

The essential roles for circuit switched call control are managed by L3 messages. The roles executed at this level include: in-call modification; control procedure for DTMF transmission; call establishment procedures for mobile-originated calls; mobile-terminated calls; dual tone multi frequency (DTMF) and call re-establishment. The contents of this include: supplementary service (SS), call control (CC), and short message service (SMS) (Kumar and Saini, 2010). Typical illustration of L3 message is paging procedure drawn in Figure 2.

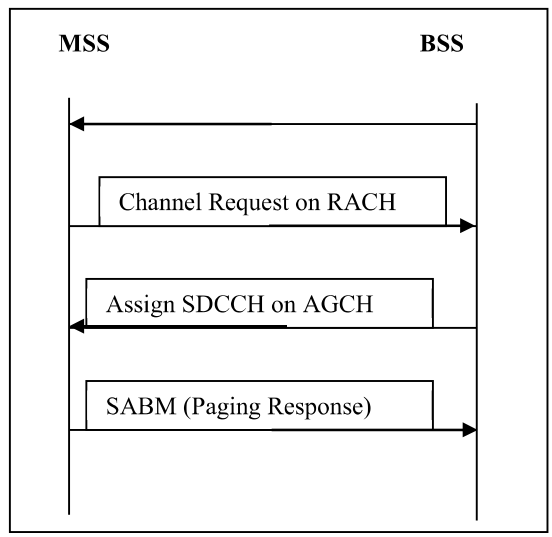

Fig 2 Paging Procedure

Source: Kumar and Saini, (2010

2.2 SPEECH QUALITY INDEX

The Speech Quality Index is a technique use in measuring and monitoring the quality of speech in the network, using the quality of radio information (Werner et.al, 2003). It is classified into two methods, namely: the Subjective Test and Objective Test respectively.

2.2.1 SUBJECTIVE TEST

ITU-T recommendation P.800 defined a universal standard for subjective speech quality as the expected opinion value (also known as Mean Opinion Score (MOS) scale). In this MOS test, an individual that listens to the short speech sentences is asked to grade the sentence from one (1) to five (5) (Stankiewicz et al., 2011; Möller et al., 2011). One (1) is ranked as the lowest quality and 5, the highest perceived quality. The procedure is simple, but it usually requires a great amount of time and cost. There are two subjective speech quality measures used frequently to estimate performance of telecommunication systems. These are; Mean Opinion Score (MOS, also known as an absolute category rating (ACR) (Ribeiro et al., 2011), and Degradation Mean Opinion Score (DMOS, also known as a degradation category rating) (Alavi and Nikmehr, 2010; Thorpe and Shelton, 1993; Dimolitsas, Corcoran and Ravishankar, 1995). The difference between MOS and DMOS is that, MOS predicts speech quality without making use of any reference while DMOS rate annoyance or degradation level by comparing the speech utterance being tested to the original (reference). The inherent problem in subjective MOS measurement is that, it is expensive and time-consuming, lacks repeatability and cannot be used for long-term or large scale voice quality monitoring in an operational network infrastructure (Barile et al., 2006; Mahdi and Picovici, 2009; Rix and Hollier, 2006).

2.2.2 OBJECTIVE TEST

An Objective test is usually used to predict MOS values that have close proximity to the values (ratings) obtained from subjective tests for various adverse speeches distortion conditions. These objective tests substitute human panel by computational algorithms that compute a MOS value by observing a sample of the speech in question (Mahdi and Picovici, 2009). Therefore, the performance evaluation, accuracy and effectiveness of these objective tests can be determined by the correlation of objective scores with subjective MOS scores. Automatically, objective methods enable standards to be efficiently maintained together with effective assessment of systems and network during design, commissioning and operation (Mahdi and Picovici, 2009). Objective test can be classified based on an assessment paradigm into three different classes: Parameter based method, Signal-based method and Packet-layer method as indicated in Figure 1.

A. Parameter Based Method

The qualities of elements of the transmission path are characterized by parameters, to plan the future transmission networks. The entire quality of the whole transmission path can be accessed from the characteristics of every network element. A correlation between the physical features of every element and the equivalent perceived quality is established via the use of parameter sets, identifying every element of the transmission system from the speaker to the listener's ear. Therefore, the quality of speech can be predicted for future network with the adoption of parameter-based models for the system under investigation using a set of parameters. The corresponding relationships between the physical parameters of either the transmission network or user device and the quality of expected speech are exploited using parameters like Loudness rating, Opinion Models and E-Model (Côté, 2011).

B. Packet-Layer Method

A quality model can be designed to assess certain processing conditions or networks, e.g. for network monitoring purposes. A High number of assessment requests, needed to monitor the whole transmission system implies simplifications in the algorithm complexity. Packet-layer models enable one to consider both these constraints and the need of a reliable instrumental model. These methods measure, in gateways, or at the listener's side, several network-related and IP packet pattern-based parameters such as transmission delay, packet-loss percentage and burst ratio. Compared to intrusive methods, packet-layer models are less complex and require less memory. In addition, packet-layer quality measurement methods associate the advantages of parametric models to those of signal-based models. They use several parameters provided by the packet-switched network and estimations from simple non-intrusive models. The current ITU-T standard is the ITU–T Rec. P.564 (2007) (Côté, 2011).

C. Based Method

Signal-based models employ transmitted speech signals or transformed speech by processing systems to estimate quality. Most of the signal based models approximate quality as related to the Absolute Category Rating (ACR) listening quality scale defined in (Quackenbush *et al.,*2001). However, recently, other models have been designed to predict individual quality features (Rec, 2006; Raake, 2007; Scholz, 2008). Three types of signal-based models can be observed in Fig 1 and Fig 2: intrusive or double-ended (also known as "full-reference") models, depend on a reference (system input) speech signal and a corresponding degraded (system output) speech signal. Reference-free ("Non-intrusive" or "single-ended") models, depend only on the degraded signal while the Parametric based models has no sound signal to process (and so make limited use of perceptual techniques), but instead estimate MOS from measured properties of the underlying transport and/or terminal, such as echo, delay, speech levels and noise (Wältermann *et al.*, 2008).

1 INTRUSIVE OBJECTIVE MEASURE

The Intrusive objective measure is a method that uses a known, controlled test signals which are processed through the condition under test. Both the original and processed signals are available to the model, and it is also typically assumed that the original signal is of near-perfect quality itself. The problem in this case is to estimate the MOS of the degraded signal from the differences between the original and degraded signals. The development of intrusive models has been closely related to that of low bit-rate speech and audio codec's (Mahdi and Picovici, 2009). A lot of intrusive models share a similar measurement structure that involves two main processes as shown in Figure 3.

Figure 3 depicts the processes that are involved in the estimation of an intrusive objective measure. The first process indicates the preprocessing of the speech signal and extraction of relevant speech parameters. Here, the original (input) speech signal and the signal degraded by the system under test, i.e., the output signal, are transformed into a relevant domain, such as temporal, spectral or perceptual domain.

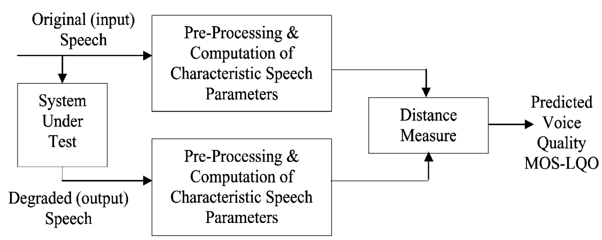

Figure 3 Basic structure of an intrusive (input-to-output) objective voice quality measure.
Source: Mahdi and Picovici, D. (2007).

The second process involves a distance measure, whereby the distortion between the input and output speech signals is computed using an appropriate quantitative measure (Mahdi, 2007) depending on the domain transformation used. These models include Perceptual Speech Quality Measure (PSQM) (Singh and Dubey, 2012), Measuring Normalizing System (MNB) (Beerends and Stemerdink, 1994), Perceptual Analysis Measurement System (PAMS) (Rix and Hollier, 2006), and Perceptual Evaluation of Speech Quality (PESQ) (Rix *et al.,* 2006; Rix *et al.*, 2001). Among them, PSQM, and PESQ were standardized by ITU-T as P.861 (Möller *et al.,* 2008) and P.862 (Möller and Raake, 2002) respectively. The major drawback of intrusive measurements is that, only pre-defined test calls can be evaluated because both the original and degraded speech samples are needed. Additional network load is generated by these measurements and it is sometimes difficult to see which part of the transmission chain contributes to the resulting end-to-end quality. More importantly, PESQ has limitation when applied in the area of noise reduction, echo cancellations, listening level, loudness and so on (Möller *et al*., 2011).

2 NON-INTRUSIVE OBJECTIVE MEASURE

The introduction of non-intrusive objective measure was proposed to solve two problems which determine the performance and accuracy of intrusive objective measurement. The first reason was the lack of perfect synchronization of time alignment between input and output speech vectors which are difficult to achieve, because of fading or error burst that occur in wireless systems. Secondly, it is difficult to obtain reference free-signal without being distorted by either echo, delays or background noise and measuring the distortion between the input and output speech will not provide a true indication of speech quality in communication systems. It is also not always possible to have access to both ends of the network connection to perform speech quality measurement using the intrusive objective measure as there might be too many connections to be monitored and again, far end location may be unknown. Due to the problems above, non-intrusive objective measure proffers an alternative to intrusive objective measure as it only deals with output (degraded) speech signals as shown in Figure 4. Table 1.0 depicts the difference between Intrusive based and Non-intrusive based methods. It can be used to measure in-service network or real time traffic in live networks. Two approaches have been identified in carrying out non-intrusive measurement processes. Priori-based and Source based non-intrusive objective measure as illustrated in Figure 1.

Priori-based Non-intrusive objective measure uses a set of well-characterized distortions. It learns a statistical relationship between the finite set and subjective opinions. While source-based non-intrusive measure uses an artificial reference signal selected from parameters characterizing the degraded speech signal. The selected artificial reference is compared to the degraded signal (Mahdi and Picovici, 2009; ITU-T RECOMMENDATION, 2004; Falk and Chan, 2006; Kim, 2005).

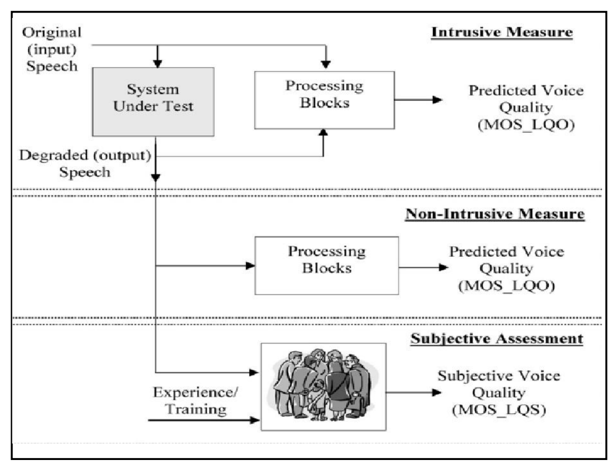

Figure 4 Intrusive and nonintrusive voice quality measures source: Mahdi, A. E. and Picovici, D. (2009).

2.3 FIELD TEST AND SURVEY

The most optimal way to test the suitability for individual application is to perform the test in a real environment. In this case the quality of the whole system, not just the speech quality, is usually tested. The test could perform with the use of a drive test or software products utilizing a laptop which has the power to pick up and communicate signals to user mobile devices.

TABLE 1:DIFFERENCE BETWEEN INTRUSIVE BASED METHODS AND NON-INTRUSIVE

Intrusive	Non-intrusive
Compared two signals- reference signal and degraded signal using proper algorithms, e.g ITU-T P862 (PESQ)	Estimates the quality just from the degraded samples
Identify the audible distortions based on two signal incorporating human auditory models	Generate an artificial reference ie (undistorted signal) from degraded speech signal and use this reference in a signal-comparison approach
Give precise outcomes that can be compared with the quality obtained from the listener expected value that is acquired from listening tests.	The Results obtained can further be modified by a parametric degradation analysis and integrated into the assessment of overall quality. E.g algorithms include auditory Non-intrusive quality estimation (ANTIQUE) and Standardized P.563.
It is difficult to monitor the entire network	The entire network can be monitored at a glance
It involves high cost as it involves two signals.	It is cost efficient and it is conducted in real network data and states
Due to the high cost and time consumption, it may not be feasible.	It is feasible, since entire network can be monitored
It is sometimes difficult to differentiate between the end to end quality contributions of the various transmission chain elements.	Measurement can be an unlimited number of speech samples, before the quality of the total sample can be determined.

3 OTHER PREVIOUS ALGORITHMS EMPLOYED

Previous techniques used in reducing noise are; statistical models, transform domain analysis or Advanced filtering techniques. These techniques have tremendously improved the quality of speech and speech enhancement. The achievement of statistical model on image processing technique in reducing artifact on clinical images without loss in resolution was highly appreciated (Zhou, 2011; Hu and Loizou, 2007). The limitation of filters and orthogonal transform techniques was that, they blur the edges of the signals or amplifies the noise as well as in the case of high pass filters (HPF), used in many applications to remove noise by averaging over many identical signal frames (Ersoy, O. 1997).

Review from Hu and Loizou, 2007, listed best techniques for reducing noise. The techniques are: subspace approach (Hu and Loizou, 2003), the perceptually-based subspace approach (Jabloun and Champagne, 2003), the log minimum mean square error (logMMSE) algorithm (Ephraim and Malah, 1985), the logMMSE algorithm with speech-presence uncertainty (Cohen and Berdugo, 2002), the spectral subtraction algorithm based on reduced-delay convolution (Gustafsson *et al.*, 2001), the multiband spectral-subtraction algorithm (Kamath and Loizou, 2002), the Wiener filtering algorithm based on wavelet-threshold multitaper spectra (Hu and Loizou, 2004), and the traditional Wiener algorithm (Scalart and Filho, 1996). A Common feature of the above techniques is that, the performance technique depends on Mean, Variance and maximum amplitude of the error. The processing time of the signal and complexity of practical implementation of circuits was also considered as a measuring tool for the performance of the above techniques (Hu and Loizou, 2003). Therefore, to allow real time implementation of these algorithms and to achieve a trade-off between high quality noise reduction and low computational load, an algorithm that reduce impairment in speech, while retaining the subjective properties of the speech signal is of great important.

4 CONCLUSION

Brief survey on evaluation of the quality of speech with emphasis on the QoS KPI was provided using a chart of classification as illustrated in Figure 1. Users can employ some of these techniques to evaluate the network performance of any service provider in order to quantify and qualify the network based on these KPIs. The results of any selected technique can be used to underscore the particular network. Therefore, the user is at liberty to either use or select another network depending on the result. Furthermore, research is ongoing to hybridize some of these techniques. Although, research that will demonstrate 100% successful evaluation of the quality of speech comparison between objective and subjective methods as well as the validation of different implementations has not been achieved yet (Becvar, Z., *et al.,*2006).

REFERENCES

Adegoke, A. S., & Babalola, I. T. (2011). Quality of service analysis of GSM telephone system in Nigeria. American journal of scientific and industrial research © 2011, science huβ, http://www.scihub.org/ajsir Issn: 2153-649x doi:10.5251/ajsir.2011.2.5.707.712

Alavi, M., & Nikmehr, H. (2010). A New Computational Model to Evaluate the Quality of Perceptual Voice Using E-Model in VOIP Communications Networked Digital Technologies, *Springer 4(7)*, 594-603.

Barile, M, C. P., Aquila, R. D., & Vitti, N. (2006). Parametric Models for Speech Quality Estimation in GSM Networks. In Software in Telecommunications and Computer Networks, 2006. SoftCOM 2006. *International Conference on*, *204-208.IEEE.*

Beerends, J. G., Busz, B., Oudshoorn, P., Van V., Jeroen, A. K., & Niamut, O. (2007). Degradation decomposition of the perceived quality of speech signals on the basis of a perceptual modeling approach. *Journal of the Audio Engineering Society, 55(12)*, 1059-1076.

Beerends, J. G., Hekstra, A. P., Rix, A. W., & Hollier, M. P. (2002). Perceptual evaluation of speech quality (pesq) the new itu standard for end-to-end speech quality assessment part ii: psychoacoustic model. *Journal of the Audio Engineering Society, 50(10)*, 765-778.

Beerends, J. G., & Stemerdink, J. A. (1994). A perceptual speech-quality measure based on a psychoacoustic sound representation. *Journal of the Audio Engineering Society, 42(3),* 115-123.

Becvar, Z., Zelenka, J., Brada, M., & Valenta, T. (2006). Comparison of Subjective and Objective Speech Quality Testing Methods in the VOIP Networks. Zdenekbecvar.org.

Côté, N. (2011). Conclusions and Outlooks. In Integral and Diagnostic Intrusive Prediction of Speech Quality. 213-217. Springer Berlin Heidelberg.

Celinus, K. (2014). Performance Analysis of Quality of Service of GSM/CDMA Mobile Networks in Zaria International Journal of Science and Research (IJSR), Volume 3 Issue 10, October 2014. www.ijsr.net. ISSN (Online): 2319-7064

Dimolitsas, S., Corcoran, F. L., & Ravishankar, C. (1995). Dependence of opinion scores on listening sets used in degradation category rating assessments. *Speech and Audio Processing, IEEE Transactions on, 3(5),* 421-424.

Ersoy, O. (1997). Fourier-related transforms, fast algorithms and applications. Prentice-Hall, Inc..

Falk, T. H., & Chan, W. Y. (2006). Single-ended speech quality measurement using machine learning methods. *Audio, Speech, and Language Processing, IEEE Transactions on, 14(6),* 1935-1947.

Recommendation, I. T. U. T. G. (2001). 1000, Communications quality of service: A framework and definitions. International Telecommunication Union.

Gandhimathi, G., & Jayakumar, S. (2012). Efficient method of pitch estimation for speech signal using MATLAB. *Special Issue IICCT, 61(2)*, 395-403.

Hu, Y., & Loizou, P. C. (2007). A comparative intelligibility study of single-microphone noise reduction algorithms. *The Journal of the Acoustical Society of America, 122(3),* 1777-1786.

Hu, Y., & Loizou, P. C. (2004). Speech enhancement based on wavelet thresholding the multitaper spectrum. *Speech and Audio Processing, IEEE Transactions on, 12(1),* 59-67.

Union, I. T. (2004). Objective Quality Measurement of Telephone Band (300-3400 Hz) Speech Codecs. *ITU-T Recommendation,* 861.

. Kim, D. S. (2005). ANIQUE: An auditory model for single-ended speech quality estimation. *Speech and Audio Processing, IEEE Transactions on, 13(5)*, 821-831.

Kumar, R., & Saini, S. (2010). Measuring Parameters for speech quality in cellular networks, Computer Science and Application Department, Kurukshetra University, Kurukshetra, India, 77-82.

Mahdi, A. E., & Picovici, D. (2009). Advances in voice quality measurement in modern telecommunications. *Digital Signal Processing, 19(1),* 79-103.

Mahdi, A. E. (2007). Voice quality measurement in modern telecommunication networks. In Systems, Signals and Image Processing, 2007 and 6th EURASIP Conference focused on Speech and Image Processing, Multimedia Communications and Services. *14th International Workshop on,* 25-32. IEEE.

Möller, S., Chan, W. Y., Côté, N., Falk, T. H., Raake, A., & Wältermann, M. (2011). Speech quality estimation: *Models and trends. Signal Processing Magazine, IEEE, 28(6),* 18-28.

Möller, S., Kim, D. S., & Malfait, L. (2008). Estimating the quality of synthesized and natural speech transmitted through telephone networks using single-ended prediction models. *Acta Acustica united with Acustica, 94(1),* 21-31.

Möller, S., & Raake, A. (2002). Telephone speech quality prediction: towards network planning and monitoring models for modern network scenarios. *Speech Communication, 38(1),* 47-75.

Počta, P., Bilšák, M., & Rouseková, J. (2010). Impact of fragmentation threshold tuning on performance of voice service and background traffic in IEEE 802.11b WLANs. *In Radioelektronika (RADIOELEKTRONIKA), 2010 20th International Conference,* 1-4. IEEE.

Počta, P., Holub, J., Vlčková, H., & Polkova, Z. (2010). Impact of Different Active-Speech-Ratios on PESQ's Predictions in Case of Independent and Dependent Losses (in Presence of Receiver-Side Comfort-Noise). *Radioengineering, 19(1),* 79.

Quackenbush, S. R., Barnwell, T. P., & Clements, M. A. (2001). Objective Measures of Speech Quality, Prentice-Hall.

Raake, A. (2007). Speech quality of VoIP: assessment and prediction: John Wiley & Sons. NY.

Rec, I. T. U. T. (2006). P. 800: Methods for subjective determination of transmission quality. *International Telecommunication Union*, Geneva.

Ribeiro, F., Florêncio, D., Zhang, C., & Seltzer, M. (2011). Crowdmos: An approach for crowdsourcing mean opinion score studies. *In Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, 2416-2419. IEEE.

Rix, A. W., Beerends, J. G., Hollier, M. P., & Hekstra, A. P. (2001). Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs. In Acoustics, Speech, and Signal Processing, 2001 Proceedings.*(ICASSP'01), 2001 IEEE International Conference on, vol.2,* 749-752. IEEE.

Rix, A. W., Beerends, J. G., Kim, D. S., Kroon, P., & Ghitza, O. (2006). Objective Assessment of Speech and Audio Quality; Technology and Applications. Audio, Speech, and Language Processing, *IEEE Transactions on, 14(6)*, 1890-1901.

Rix, A. W., & Hollier, M. P. (2000). The perceptual analysis measurement system for robust end-to-end speech quality assessment. In Acoustics, Speech, and Signal Processing, 2000. ICASSP'00, Proceedings. *2000 IEEE International Conference on, vol. 3,* 1515-1518. IEEE.

Gothenburg. (July 2004). ECC report on " Voice Quality Over IP based Networks".

Gbadamosi S. A., Ugweje, O. C., Onumanyi A. J., & Onwuka, E. N. (2015). Non-Intrusive Noise Reduction In Gsm Voice Signal Using Non-Parametric Modeling Technique, *IEC 2015,* 57-63.

Scholz, K. (2008). Instrumental quality assessment of telephone-band speech based on quality attributes. Ph.D. thesis, (in German), *Christian-Albrechts-University of Kiel, D-Kiel,* 45-67.

Shannon, CE. (1948). BA mathematical theory of communication, Bell System Tech: NJ.

Singh, M. M., & Dubey, M. R. K. (2012). Non-Intrusive Speech Quality with Different Time Scale. *IJIS,* 233-236.

Stankiewicz, R., Cholda, P., & Jajszczyk, A. (2011). QoX: what is it really? Communications Magazine, *IEEE, 49(4)*, 148-158.

Thorpe, L. A., & Shelton, B. R. (1993). Subjective test methodology: MOS vs. DMOS in evaluation of speech coding algorithms. In Speech Coding for Telecommunications, 1993. Proceedings. *IEEE Workshop on*, 73-74. IEEE.

Thorpe, L., & Shelton, B. R., (1993). Subjective test methodology: MOS vs. DMOS in evaluation of speech coding algorithms. In Speech Coding for Telecommunications, 1993 Proceedings. *IEEE Workshop on*, 73-74. IEEE.

Wältermann, M., Scholz, K., Möller, S., Huo, L., Raake, A., & Heute, U. (2008). An instrumental measure for end-to-end speech transmission quality based on perceptual dimensions: Framework and realization. *In 9th Annual Conference of the International Speech Communication Association (ISCA)*, 44-52.

Werner, M., Kamps, K., Tuisel, U., Beerends, J. G., & Vary, P. (2003). Parameter-based speech quality measures for GSM. In Personal, Indoor and Mobile Radio Communications, 2003. PIMRC 2003. *14th IEEE Proceedings on, Vol. 3,* 2611-2615. IEEE..

Zhou, W. (2011). Information Content Weighting for Perceptual Image Quality Assessment. Image Processing, *IEEE Transactions on, 20(5),* 1185-1198.