Proceedings of APCC2015 copyright © 2015 IEICE 14 SB 0087

Simple MOS Estimation Model for Skype Referring to Packet Loss Effects: Development Using Conversation-like Tests

Montri Rungruangthum and Pongpisit Wuttidittachotti
Faculty of Information Technology
King Mongkut's University of Technology North Bangkok
Bangkok, Thailand

Therdpong Daengsi
Enterprise Services
JADS Comm Ltd.
Bangkok, Thailand

***Abstract*—This paper proposed a simple mathematical model that can be used to estimate VoIP quality, called mean opinion score (MOS), provided by Skype. This model had been developed using the data from the informal interview tests, called conversation-like tests, referring to packet loss of 0%, 5%, 10%, 15%, 20% and 25%. The data were gathered from 144 subjects, who participated in the tests. After using curve fitting technique and analysis, the 2nd polynomial equation was selected as the representative of the simple model that is one contribution of this work. After evaluating with the test set of 32 subjects obtaining from packet loss of 0%, 15% and 25%, this selected model shows Mean Absolute Percent Error (MAPE) of 11.97% which means it is a good and simple MOS model for Skype.**

***Keywords*—VoIP; MOS; SILK codec; Packet loss.**

I. Introduction

Voice over IP (VoIP) is a communication technology that works over Internet Protocol (IP) Network. It converts voice signals into voice packets, to be transmitted over IP networks. One advantage of VoIP is cost saving. Many VoIP applications, such as Skype, a popular and the most mature VoIP application at present, can be used without cost in several cases if an IP network is available [1-2]. However, VoIP quality issue may occur, if there are loss, delay and jitter in IP network [3]. The metric that is used to indicate VoIP quality level is called Mean Opinion Score (MOS) [3-6]. To ensure that VoIP quality of each VoIP application is good enough, it is necessary to know by measuring or estimating MOS values.

Focusing on Skype which is very popular and the most mature VoIP application [1-2], it has been mainly researched particularly with SILK which is used inside Skype, based-on objective measurement [7-9]. Thus, this study using subjective tests referring to packet loss effects has been conducted. Then, the simple mathematical model, which is one contribution of this paper, was proposed before evaluation and analysis.

II. Background

A. Importance of Subjective Assessment and Thai Users

Although there are objective evaluation methods for VoIP (e.g., E-model), it has been mentioned that subjective tests with a large group of subjects is the ground truth measurement for voice quality because subjective methods give more truthful results than objective methods [6]. For example, it was stated that E-model does not provide values of MOS that correlate very well with the values of subjective MOS [10], which is consistent with the statements in previous works [11-12]. However, subjective MOS may be impacted by language and cultural variation and/or balance of conditions [13].

B. Native Thai Users

It was stated in [14-15] that Thai users in Thailand use Thai, which is a tonal language, like Vietnamese and Chinese but exactly different in terms of the sound system. There are five distinct tones in Thai, which refers to change of fundamental frequency (F_0). Different tones change the meaning of Thai words. Besides, according to national cultural dimensions of Hofstede, Thai people have their own culture, which is different to other cultures [14]. Moreover, based-on the way Thai users respond to situations, they usually do not respond in good or bad conditions, whereas, westerners generally emphasis on black and white or winning and losing.

C. VoIP Quality Evaluation and MOS

In fundamental, VoIP quality evaluation is based on subjective tests [4]. Conversation test is one of VoIP quality evaluation methods recommended by ITU-T because it can reach the same standard of realism, whereas Interview test is optional [15-17]. However, the disadvantages of the conversation test are requiring two low background noise rooms, high cost, high effort, and good collaboration and management skills, while this test wastes time [15]. The result from subjective tests is usually called subjective MOS that has been obtained from the averaged result, voted by a group of subjects using 5-point scale as shown in Table I [4][15].

D. Skype and SILK codec

Skype relies on a peer-to-peer (P2P) infrastructure and uses "supernodes" for message relaying and handling metadata such as user profile and presence information [18-19]. Also, Skype nodes include clients, and servers for updates and authentication. Skype can multiplex different service flows on an established connection: voice calls to another Skype node, video conferencing, chat, file upload/download.

Inside Skype, SILK is the important part. It is the flexible audio codec for real-time communications [20]. Therefore, Skype can adjust bit rate and sampling rate, as shown in Table II [9]. Recently, SILK_V3 codec was developed with the advantage about capacity to check packet loss rates from call technical information for real-time communication [21]. Furthermore, Forward error correction (FEC) mechanism has been used in Skype to recover the lost packets that some packets will piggyback the previous packets based on the redundancy ratio. Nevertheless, it was mentioned in [9] that Skype can provide brilliant voice calls service under the situations of packet loss rate ranging from 0% to 10%. The maximum MOS from SILK of Skype was about 4.5 from Dynastat MOS test with 32 subjects, see Fig. 1 [20], which is consistent with MOS of 4.4-4.5 as shown in [22].

E. Packet loss rates

In general, VoIP packets transmitted are sometimes lost IP networks. Packet loss mainly occurs when packets are sent but some of them are not received at the destination endpoint due to some events occur in the network (e.g., router failures and fiber link down) [23]. Packet loss can be both random and burst [24-25]. It was mentioned in [14] that the maximum loss of VoIP packets between two endpoints should be 1% or less for very good voice quality, whereas 3% or less is acceptable for business quality. However, packet loss of 5% can be acceptable for some manufacturers.

F. Related Research

One study was conducted for VoIP quality evaluation from a few codecs with 32 subjects, referring to packet loss effects. It reported that SILK provided the best VoIP quality when compared to AMR-WB codec and Speex codec. SILK showed MOS values of 4.5, 3.2, 2.5 and > 2.0 approximately, from packet loss of 0%, 2%, 5% and 10% [20].

Xiaomin studied the performance evaluation of VoIP quality from Speex and SILK referring to several conditions, including different complexity, different bit rates, and different buffer size [26]. It was found that SILK tended to provide better VoIP quality than Speex.

Goudarzi et al. presented a regression-based model to estimate MOS values of the wideband (WB) and narrowband (NB) SILK [27]. The developed model uses the network parameter and the application parameter to estimate MOS values. Subjective tests were also conducted to validate the model. They found that the model showed very good performance (97% for WB and 91% for NB).

Assem et al. proposed an algorithm that performs in-call selection of the most appropriate codec given prevailing conditions on the network path between the endpoints of VoIP calls [22]. They tested the algorithm on different packages that contain a selection of several codecs, including SILK. The results showed that their algorithm significantly produced improvement in VoIP quality as compared to the use of a codec selected at the start of a call and maintained for the call duration, whereas the combination of the PCMU and SILK provided better performance than other commonly used codecs.

TABLE I. RANGES OF BIT RATE AND SAMPLE RATE FOR SILK OPERATING MODES

Score	Meaning	User Opinion
5	Excellent	User Satisfied
4	Good	Satisfied
3	Fair	Some Users Dissatisfied
2	Poor	Most Users Dissatisfied
1	Bad	Nearly all Users Dissatisfied

TABLE II. RANGES OF BIT RATE AND SAMPLE RATE FOR SILK OPERATING MODES

Mode	Bite Rate (Kbps)	Sample Rate (KHz)
Narrowband	6-20	8
Mediumband	7-25	8, 12
Wideband	3-30	8, 12, 16
Super Wideband	12-40	8, 12, 16, 24

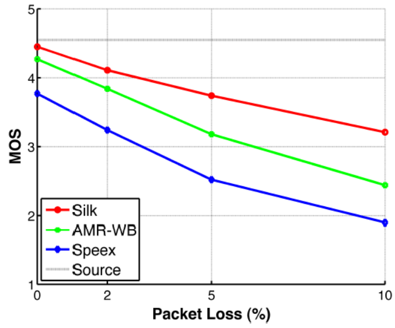

Fig. 1. Subjective MOS provided by SILK, comparing to other codecs.

Liu and Sun studied Skype referring to packet loss of 0%, 8%, 12% and 20% [9]. They reported that the objective MOS values were about 3.83, 3.00, 2.80 and 2.33 respectively, whereas the subjective MOS values were 5, 4.67, 2.67 and 1.17 respectively. However, there were only 3 listeners in their work and did not described the details about their methodology (e.g., monaural or binaural). They also stated that the voice quality of Skype was unacceptable with the packet loss rate of 20%.

Takahashi et al. proposed a simple mathematical model of MOS-CQE for Japanese, based-on E-model result [28]. The proposed model involved 40 native Japanese speakers. Of course, the proposed model showed good correlation with the subjective MOS from conversation tests.

Sun and Ifeachor proposed a new methodology for developing perceptually accurate models for nonintrusive prediction of VoIP quality [29]. Also, they presented efficient regression models for predicting conversational voice quality non-intrusively for G.729, G.723.1, AMR and iLBC. The results showed the models provided accuracy close to the combined ITU PESQ/E-model method using real Internet traces with correlation coefficient of over 0.98.

To improve E-model, Ding and Goubran [30] proposed the extended E-model using modified the equipment impairment factor (Ie) and the proposed jitter impairment factor (Ij). While Ren et al. proposed the different Ij in [31]. Moreover, Ren et al, proposed the language impairment factor (Il) in their previous work [32].

Wuttidittachotti et al. conducted listening tests using IP phones to obtain MOS values from G.729 referring to packet loss rate of about 0% - 15%, with over 100 native Thai subjects [33]. Then they proposed a mathematical model with good performance with MAPE of 16%, and better than PESQ.

Daengsi et al. proposed the enhanced E-model using bias factor with the combination of E-model and the mathematical model obtained from conversation test with 400 native Thai subjects [34]. It showed better performance than the standard E-model about > 20%. Also, Daengsi et al. proposed mathematical models referring to packet loss and delay effects, called ThaiVQE models, for two standard codecs, G.711 and G.729 [14]. The data for modeling were from conversation tests with a total of > 700 subjects. ThaiVQE model showed error reduction of over 13% and 28% for G.711 and G.729 respectively, when compared to E-model.

Hines et al. proposed an objective speech quality model, called the Virtual Speech Quality Objective Listener (ViSQOL), which is a signal-based, full-reference, intrusive metric using a spectro-temporal measure of similarity between a reference and a test speech signal [35]. That algorithm was compared to the ITU-T standard metrics PESQ and POLQA. The results and analysis showed that both ViSQOL and POLQA had some performance weaknesses and under-predict perceived quality in certain VoIP conditions.

Triyason and Kanthamanon [36] proposed E-model modification for 8 languages (including Thai) using the language impairment factor (Il) obtained from PESQ results, called EL-model. After evaluation, it was found that the EL-model showed about 80% improvement when compared to G.107 model.

In summary for this section, related works about Skype referring to network parameter and other effects and mathematical models for VoIP quality measurement were covered. However, there was no model that proposed a MOS estimation model for Skype using subjective results intensively. Thus, it became the reason to conduct this study.

III. Preparation and Subjective Tests

A. Subjective test design

1) The environment for conducting this study was one zone at the 7th floor of the Central Library of KMUTNB, because of low background noise requirements [4][17].

2) For the subjective test method in this study, it was an informal subjective method called conversation-like test method, which had been applied from the interview test method and the conversation test method [16-17]. Instead of using Richard's task, the 'guess my birthday task' had been used. Mainly, each participant or subjects must ask 3 questions for clues before guessing the birthday date of the interviewer.

3) All participants or subjects were students in KMUTNB. Mainly, they were both undergraduate students and graduate students. To avoid gender bias, numbers of male and female subjects were balanced as many as possible.

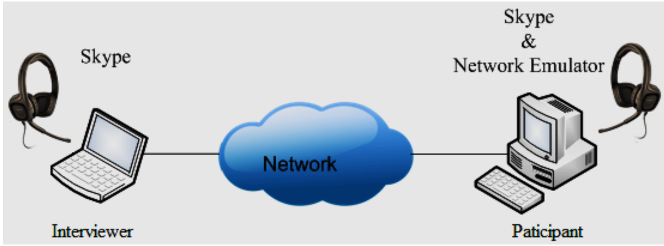

Fig. 2. Overview on the testbed system

4) Two computers and two good quality of headsets, see Fig. 2, were used instead of two IP telephones. Skype had been installed in two computers Skype (version 7.2.59.103).

5) There were six packet loss conditions in this study. Those loss conditions are 0%, 5%, 10%, 15%, 20% and 25%. For the network emulator that had been installed in the receiver computer of the testbed system for generating packet loss, it is the Network Emulator for Windows Toolkit (version 2.1.0003.0), following [37]. For the broadband Internet that was used to register Skype outside, it had been checked and found that its uplink speed was about 70±15 Mbps, whereas its downlink speed was 30±20 Mbps approximately.

B. Data Gathering

Each subject was asked to participate the conversation-like test. After asking about basic information by the interviewer, the 'guess my birthday task' was started. Before finishing the conversation, a participant must give the voice quality score. Of course, all data was recorded by the interviewer using a paper based method. The subjects were grouped as groups of six subjects. Firstly, the 1st group was asked to join in one test condition (e.g., loss of 0%), one by one subjects. Next, the 2nd group was invited to join in the next test condition (e.g., loss of 5%). Of course, the tests conducted until the last condition had been completed, before restart again with new groups until obtaining at least 24 subjects per condition, totally about 150 subjects both male and female subjects. Moreover, at least 30 subjects from testing with the conditions of 3%, 15% and 25% were also conducted, for model evaluation purpose.

IV. Modeling, Evaluation and Model Selection

After discarding the outliers, 144 subjects totally (as shown in Table III) with the average age of 21.46±3.10 years old were ready for creating the MOS Estimation model, using curve fitting technique in Excel (following [33]).

From this procedure, four equations had been obtained, as shown in (1)-(4), whereas MOS is the estimated MOS, and x is the packet loss rate (%).

$$MOS = 4.4826e^{-1.273x} \quad (1)$$

$$MOS = -4.5476x + 4.4643 \quad (2)$$

$$MOS = -16.071x^2 - 0.5298x + 4.3304 \quad (3)$$

$$MOS = -49.383x^3 + 2.4471x^2 - 2.2211x + 4.3489 \quad (4)$$

$$MOS = -1111.1x^4 + 506.17x^3 - 82.87x^2 + 1.7471x + 4.337 \quad (5)$$

Besides, R^2 value can be the good fit indicator for each equation also provided from Excel, as shown in Table IV. However, it is not every case that the highest R^2 means the best equation and should be the most appropriate mathematical model because it might be the impact from the over-fitting behavior [36]. However, instead of following the technique as in [36], all equations with R^2 of over 0.31 were analyzed using Mean Absolute Percent Error (MAPE) and statistical tools, ANOVA [14][15].

MAPE results were calculated with the test set obtained from 3 test conditions of packet loss effects with totally 32 subjects (as shown in Table V). This group of subjects had the average age of 21.13±1.86 years old. The MAPE results from (3)-(5) are presented in Table VI. As shown in the table, (3)-(5) can provide almost the same MAPE result, although (4) becomes the best model with the lowest MAPE of 11.75%. However, if there is no significant different among those equations, (3)-(5), the simplest equation in this comparison, (3), should be the representative of the MOS estimation model in this study, its MAPE is 11.97%. With this assumption, (3)-(5) had been used for calculation for MOS values referring to packet loss of 0%, 2%, 4%, ..., 24% and then plotted as shown in Fig. 3. Next, ANOVA technique was conducted, with the following hypotheses:

$$\left.\begin{array}{l} H_0\text{: } \mu_{(3)} = \mu_{(4)} = \mu_{(5)} \\ H_1\text{: Not all of the means are equal} \end{array}\right\} \quad (6)$$

Where H_0 and H_1 are the null hypothesis and the alternative hypothesis respectively, and $\mu_{(3)}$, $\mu_{(4)}$, and $\mu_{(3)}$ are the means of the results from (3)-(4). If the p-values is more than the significance level α of 0.05, with 95% confidence interval, the null hypothesis is accepted and then reject the alternative hypothesis instead.

As shown in Table VII, the p-value is 0.994. It means there is no significant difference because H_0 has been strongly accepted. Therefore, instead of selecting (4), (3) becomes the most appropriate mathematical model with the simplest equation when compared to (4) and (5).

V. Discussion

Mainly, there are two issues to be discussed, as follows:

1) Based on subjective tests, called conversation-like tests, it was a surprise that the result from this study with Skype that used SILK codec inside, is not consistent with the results from two previous study with SILK, both subjective study in [20] and objective study in [22]. It can be seen obviously in Table III that Thai subjects rated rather good VoIP Quality with MOS of about 3.6 although there was packet loss of 20%. The result might be impacted from language and cultural effects because Thai speech sounds are tonal and Thai users usually do not respond to either good or bad network conditions extremely. Besides, it might be from the system preparation that provide good quality headset for testing in diotic mode (two ears) that can be considered as the balance of conditions [13]. This issue should be investigate in the future.

2) For the mathematical modeling, it can be seen in Fig. 3, the 2nd, 3rd and 4th polynomial equations provide almost the same result. It has been confirmed by the result from a hypothesis test using t-test that there is no significant difference among thos equations. Therefore, it is the reason that the 2nd polynomial becomes the representative of the MOS estimation model for Skype in this study, which provides good performance with MAPE of < 12% (0% < MAPE < 10% is excellent, 10% < MAPE < 20% is good) [14][33].

TABLE III. Results from the conversation-like tests referring to loss effects

Condition		No. of Subjects		MOS-CQS	Standard Deviation
		Male	Female		
Packet Loss (%)	0%	12	12	4.33	0.48
	5%	12	12	4.29	0.55
	10%	10	14	4.04	0.55
	15%	11	13	3.92	0.72
	20%	12	12	3.63	0.58
	25%	12	12	3.17	0.70

TABLE IV. R^2 for each equation obtaining from the curve fitting

Equation	R^2	Remark
(1); Exponential	0.2918	Don't care.
(2); Linear	0.2957	Don't care.
(3); Polynomial order 2nd	0.3154	
(4); Polynomial order 3rd	0.3162	
(5); Polynomial order 4th	0.3175	

TABLE V. Test set information

Condition		No. of Subjects		MOS-CQS	Standard Deviation
		Male	Female		
Packet Loss (%)	3%	7	5	4.25	0.45
	15%	4	6	3.92	0.67
	25%	5	5	3.25	0.62

TABLE VI. MAPE results of the focused equations

Equation	MAPE	Remark
(3); Polynomial order 2nd	11.97%	
(4); Polynomial order 3rd	11.75%	The lowest is the best performance.
(5); Polynomial order 4th	12.03%	

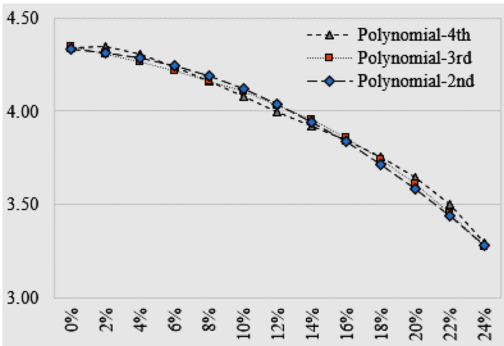

Fig. 3. Curves of the focused mathematical models.

TABLE VII. Hypothesis test result using ANOVA analysis

Equation	p-value
H_0: $\mu_{(3)} = \mu_{(4)} = \mu_{(5)}$	0.994
H_1: Not all of the means are equal	

VI. Conclusion and Future Work

After creating MOS estimation model for Skype using the subjective results that were obtained from the conversation-like tests with 69 male and 75 female native Thai subjects, the 2^{nd} polynomial equation, which is the most simplest mathematical model with MAPE of 11.97%, has been analyzed by comparing to other equations and then considered as the representative of the simple subjective MOS estimation model for Skype. The MAPE of about 12% from the model evaluation means this is a good model, particularly to Thai users. The model has shown that some Thai users may be satisfied VoIP quality with MOS of about 3.6 from Skype referring to packet loss of about 20%. For future work, the subjective MOS values from this study seem to be high when packet loss rates are high (e.g., 20% - 25%). Therefore, this should be investigated or re-conducted subjective tests with the similar or the different experimental design.

Acknowledgment

Thanks all participants and the Central Library, KMUTNB.

References

[1] T. De Pessemier, I. Stevens, L. De Marez, L. Martens, and W. Joseph, "Quality assessment and usage behavior of a mobile voice-over-IP service," Telecommunication System, 2015. http://dx.doi.org/10.1007/s11235-014-9961-9.

[2] VoIP Supply, "Benefits of Skype," http://www.voipsupply.com/benefits-of-skype.

[3] F. De Rango, M. Tropea, P. Fazio, and S. Marano, "Overview on VoIP: Subjective and Objective Measurement Methods", IJCSNS, Vol. 6(1B), 2006.

[4] ITU-T Recommendation P.800, "Methods for subjective determination of transmission quality", Aug. 1996.

[5] ITU-T Recommendation P.800.1, "Mean Opinion Score (MOS) terminology," Jul. 1996.

[6] A. Hines, J. Skoglund, A. C. Kokaram, and N. Harte, "ViSQOL: an objective speech quality model," EURASIP Journal on Audio, Speech, and Music Processing, Vol. 13, 2015. http://dx.doi.org/10.1186/s13636-015-0054-9.

[7] M. Goudarzi, L. Sun and E.C. Ifeachor, "Modelling Speech Quality for NB and WB SILK Codec for VoIP Applications," NGMAST 2011, Cardiff, UK, pp. 42-47, Sep. 2011.

[8] C. Chen, C. Chu, S. Yah, H. Chu, and P. Huang, "Modeling the QoE of Rate Changes in SKYPE/SILK VoIP Calls," IEEE/ACM Transactions on Networking, Vol. 22(6), pp. 1781-1793, Nov. 2013.

[9] L. Liu, and L. Sun, "Performance Analysis of Video Calls Using Skype," 2013. https://www.cscan.org/download/?id=924.

[10] M. Y. Arafat, M. M. Alam, and F. Ahmed, "SIP-Based QoS in IP Telephony," Journal of Networks, Vol. 9(12), pp. 3415-3426, Dec. 2014.

[11] M. Narbutt, and M. Davis, "Assessing the quality of VoIP transmission affected by playout buffer scheme," 4th MESAQIN 2005, Prague, Czech Republic. Jun. 2005.

[12] M. Goudarzi, "Evaluation of Voice Quality in 3G Mobile Networks," Thesis, University of Plymouth, Plymouth, UK, Jun. 2008.

[13] A. W. Rix, "Comparison between subjective listening quality and P.862 PESQ score," Psytechnics, Sep 2003.

[14] T. Daengsi, N. Khitmoh, and P. Wattidittachotti, "VoIP quality measurement: subjective VoIP quality estimation model for G.711 and G.729 based on native Thai users," Multimedia Systems. http://dx.doi.org/10.1007/s00530-015-0468-3.

[15] T. Daengsi, C. Wutiwiwatchai, A. Preechayasomboon, and S. Sukparungsee, "A Study of VoIP Quality Evaluation: User Perception of Voive Quality from G.729, G.711 and G.722," 9th IEEE-CCNC 2012 (SS-QOE), Lasvegas, NV, pp. 342-345, Jan. 2012.

[16] T. Daengsi, C. Wutiwiwatchai, A. Preechayasomboon, and S. Sukparungsee, "Speech Quality Assessment of VoIP: G.711 VS G.722 Based on Interview Tests with Thai Users," IJITCS, Vol. 4(2), pp. 19-25, Mar. 2012.

[17] ITU-T Recommendation P.805, "Subjective evaluation of conversational quality," Apr. 2007.

[18] B. Trammell et al., "Identifying Skype traffic in a large-scale flow data repository. Springer Berlin Heidelberg," LNCS, pp. 72-85, 2011.

[19] M. Korczynski, and A. Duda, "Classifying Service Flows in the Encrypted Skype Traffic," IEEE ICC 2012, Ottawa, Canada, pp. 1064-1068. Jun. 2012.

[20] K. Vos, S. S. Jensen, and K. V. Sørensen, "SILK Speech/Audio Codec," http://www.ietf.org/proceedings/75/slides/codec-3.pdf.

[21] P. Wuttidittachotti, W. Akapan, and T. Daengsi, "Comparison of VoIP-QoE from Skype, LINE, Tango and Viber over 3G Networks in Thailand," presented in ICUFN 2015, Supporo, Japan, Jul 2015.

[22] H. Assem, M. Adel, B. Jennings, D. Malone, J. Dunne, and P. O'Sullivan, "A generic algorithm for mid-call audio codec switching," IFIP/IEEE International Symposium onIntegrated Network Management (IM 2013), Ghent, Belgium, pp. 1276-1281. May 2013.

[23] A. Markopoulou, G. Iannaccone, S. Bhattacharyya S, and C.-N. Chuah, C. Diot, "Characterization of Failures in an IP Backbone," IEEE INFOCOM 2004, Hong Kong, China, pp. 2307-2317, 2004.

[24] ITU-T Recommendation G.107 "The E-model: a computational model for use in transmission planning." 2011.

[25] H. Zhang, L. Xie, J. Byun, P. Flynn, and C. Shim, "Packet Loss Burstiness and Enhancement to the E-Model," SNPD/SAWN 2005, Towson, MD, pp. 214-219, 2005.

[26] L. Xiaomin, "Performance Evaluation of Speech Quality for VoIP on the Internet," MS Thesis, National University of Singapore, Singapore, 2011.

[27] M. Goudarzi, L. Sun and E.C. Ifeachor, "Modelling Speech Quality for NB and WB SILK Codec for VoIP Applications," 5th NGMAST 2011, Cardiff, UK, pp. 42-47, Sep. 2011.

[28] A. Takahashi, A. Kurashima, and H. Yoshino, "Objective Assessment Methodology for Estimating Conversational Quality in VoIP," IEEE Transactions on Speech Audio and Language Processing, Vol. 14(6), pp. 1983-1993, Nov. 2006.

[29] L. Sun, and E. C. Ifeachor, "Voice quality prediction models and their application in VoIP networks," IEEE Transactions on Multimedia, Vol. 8(4), pp. 809-820, Aug. 2006.

[30] L. Ding and R. A. Goubran, "Speech quality prediction in VoIP using the extended E-model," IEEE GLOBECOM 2003, San Francisco, CA, Vol. 7, pp. 3974-3978, Dec. 2003.

[31] J. Ren, C. Zhang, W. Huang and D. Mao, "Enhancement to E-model on standard deviation of packet delay," ICIS 2010, Chengdu, China, pp. 256-259. Jun. 2010.

[32] J. Ren, H. Zhang, Y. Zhu and C. Gao, "Assessment of effects of different language in VOIP," ICALIP2008, Shanghai, Chinan, pp. 1624-1628, Jul 2008.

[33] P. Wuttidittachotti, P. Khaoduang, and T. Daengsi, "Development of a MOS estimation model for G.729 using listening-opinion tests with Thai speech referring to packet loss effects," ISCAIE 2014, Penang, Malaysia, pp. 29-32, Apr. 2014.

[34] T. Daengsi, and P. Wuttidittachotti, "VoIP quality measurement: Enhanced E-model using bias factor," GLOBECOM 2013, Atlanta, GA, pp. 1329-1334, Dec. 2013.

[35] A. Hines, J. Skoglund, A. C. Kokaram, and N. Harte, "ViSQOL: an objective speech quality model," EURASIP Journal on Audio, Speech, and Music Processing, Vol. 13, 2015. http://dx.doi.org/10.1186/s13636-015-0054-9.

[36] T. Triyason, and P. Kanthamanon, "E-Model Modification for Multi-Languages over IP," Elektronika Ir Elektrotechnika, Vol. 21(1), pp. 82-87, 2015.

[37] X. Zhang, Y. Xu, H. Hu, Y. Liu, Z. Guo, and Y. Wang, "Profiling skype video calls: Rate control and video quality," INFOCOM 2012, Orlando, FL, pp. 621-629, Mar. 2012.