บทความวิจัย

วารสารวิชาการเทคโนโลยีอุตสาหกรรม (The Journal of Industrial Technology)
ISSN (Print): 1686-9869, ISSN (online): 2697-5548
DOI: 10.14416/j.ind.tech.2020.03.003

Audiovisual Quality Assessment: A Study of Video Calls Provided by Social Media Applications

Phisit Pornpongtechavanich[1] and Therdpong Daengsi[2*]

[1] Faculty of Industry and Technology, Rajamangala University of Technology Rattanakosin (Wang Klai Kangwon Campus), Prachuap Khiri Khan, Thailand

[2] Faculty of Engineering, Rajamangala University of Technology Phra Nakhon (North Bangkok Center), Bangkok, Thailand

[*] Corresponding author, E-mail: therdpong.d@rmutp.ac.th

Received: 7 October 2019; Revised 25 November 2019; Accepted: 12 December 2019
Online Published: 20 March 2020

Abstract: This study presents a comparative study about audiovisual quality from Facebook Messenger, WhatsApp, and LINE when using them specifically to make video calls from smartphones. The first part of this study is based on subjective tests. After gathering the results, audiovisual quality from each social media application has been calculated before conducting statistical analysis. From the results and analysis, it has been found that Messenger can provide slightly better audiovisual quality with Mean Opinion Score (MOS) of 4.02-4.10 than WhatsApp and LINE with MOS of 3.75-3.85 and 3.63-3.90 respectively, although there is no significant difference after testing the hypotheses by using Analysis of Variance (ANOVA) technique. Besides, this study has shown that the Hands' additive linear model and the Belmudez & Moller's full model are reliable to be utilized for short conversation - audiovisual quality assessment, which is one of the contributions of this study.

Keywords: Audiovisual; Video Call; Social Media Application; QoE; MOS; Subjective Method

วารสารวิชาการเทคโนโลยีอุตสาหกรรม (The Journal of Industrial Technology)
ISSN (Print): 1686-9869, ISSN (online): 2697-5548
DOI: 10.14416/j.ind.tech.2020.03.003

1. Introduction

Nowadays, smartphones are being used extensively worldwide, particularly in developed countries. Statistical estimates were that there are more than 2.5 billion smartphone users in the world in 2018 and it was predicted that there will be almost 2.9 billion smartphone users worldwide in 2020 [1]. Smartphones are not only used for voice communication but also for multimedia communication (e.g., video calls). Because of the proliferation of smartphones, the market for applications for smartphones would grow as well, including social media applications. These applications support various services such as instant messaging, file sending, Voice over Internet Protocol (VoIP) and video calls. Dreamgrow also presented the top 10 social media applications globally in 2018: WhatsApp and Messenger are at the first and second place respectively with 1.5 billion monthly active users (MAUs) and 1.3 billion MAUs, while LINE is in the eighth place with about 220 million MAUs [2].

Important functions of these social media applications include VoIP and video calls. The obvious benefit of these functions is providing free calls using the internet access of smartphones. Particularly video calls require high quality of service (QoS) from the overall performance of the network as seen by the users of the communication network. In order to meet user's needs and expectations delay of end-to-end and drop rates have to be reduced [3]. In times of poor internet connection, poor network parameters usually occur, resulting in poor quality issues for video calls. Video call feature provides not only voice but also the video streaming or images in almost real-time. However, internet issues such as inconsistent IP networks would result in poor voice quality as well as poor video quality. Several steps in processes from the start to the end may degrade the quality perceived by end-users while network impairments would reduce the quality of experience for end-users. Therefore, monitoring and evaluating the quality perceived by users, known as quality of experience (QoE) is important and necessary in order to gain or maintain high satisfaction of users [4]. For video calls, these services are classified as parts of audiovisual services. Therefore, audiovisual quality assessment can be applied to evaluate the quality of the video call service provided by social media applications. The assessments of video call quality from those applications can be done on voice and visual qualities separately by using the same concept as in [5].

In this study, the subjective assessment methods for voice quality and video quality from three social media applications; consisting of WhatsApp, Messenger, and LINE, have been

วารสารวิชาการเทคโนโลยีอุตสาหกรรม (The Journal of Industrial Technology)
ISSN (Print): 1686-9869, ISSN (online): 2697-5548
DOI: 10.14416/j.ind.tech.2020.03.003

conducted before utilizing them together to find overall audiovisual quality using assessment methods proposed by researchers previously. However, some data from Messenger and LINE have been applied from [6]. Finally, the audiovisual quality of video calls provided by the three applications has been compared and analyzed using the statistical technique, in order to see whether the audiovisual quality from three applications is the same or not. Furthermore, this study also tries to find appropriate mathematical models for audiovisual quality assessment referring to short conversation tests.

2. Background and Related Studies

2.1 Social Media Applications: WhatsApp, Messenger and LINE

Nowadays, social networks have grown enormously and provided people a new world where they can share everything that happens around them [7]. These networks are defined assets of people or groups of people with some pattern of contacts or interactions between them [8]. People or users from everywhere can easily access social networks via social networking sites (SNS) or social media applications. Social media applications, such as Mobile Instant Messaging (MIM) applications offer not only sending and receiving messages but also provides several other services (e.g., video calls). Many popular social media applications are free to use. However, there are only three applications associated with the tests in this study, consisting of WhatsApp, Messenger, and LINE.

WhatsApp is the most popular with monthly active users in more than 180 countries [2, 9]. It is a free social media application, offering simple, secure and reliable messaging, VoIP calling and video calling. It is available not only in smartphones but also on personal computers (PCs) [9]. It supports Android, iOS, PCs and even Windows Phone. It supports sending and receiving a variety of media, including text, photo, video, document, location, voice and video call.

Messenger is under the umbrella of Facebook. It is an application used for instant messaging, photo/video/sticker sending, audio and video calling and SMS [10-12]. It supports Android, iOS, PCs and Windows Phone. It has been claimed to be the leading application in social media for text and video chats that have been downloaded from the App Store [6, 13]. In addition to its functions and features, it can collaborate with Facebook, the biggest social networking site (SNS), with over 2,000 million active users [1].

LINE is another popular application created after the great earthquake and tsunami in March 2011. As stated in [6], it was developed and officially launched a few months later in order to

วารสารวิชาการเทคโนโลยีอุตสาหกรรม (The Journal of Industrial Technology)
ISSN (Print): 1686-9869, ISSN (online): 2697-5548
DOI: 10.14416/j.ind.tech.2020.03.003

help people chat with each other using short messages over the limited infrastructure of telecom networks [14-15]. It is available for smartphone devices (e.g., iPhone, Android, Windows Phone, BlackBerry, and Nokia) and PCs. It has been cited in [5] that it became the major social media platform daily in Thailand market with 41 million users, and was being used by 92% of connected consumers in Thailand [16]. Its popularity could be attributed to its most highly used feature - the stickers.

2.2 Audiovisual quality and its metrics

Audiovisual quality assessments are generally classified into two main methods, subjective and objective assessment [17]. Subjective quality assessment methods may involve making video calls between pairs of participants. Then, their judgments of the video call quality are gathered and utilized as a quality metric. Objective quality assessment methods, however, do not rely on human judgment but usually involve automated procedures to determine quality metrics [17]. Both subjective and objective methods are referred to as Quality of Experience (QoE) assessment methods. It can be the key determining factor in the success of multimedia applications or services [18]. Unlike QoS that evaluate networks and/or systems parameters (e.g., loss and delay), QoE contains more factors and covers a wider range of aspects which can directly reflect users' acceptance of the service, therefore it is more suitable for the evaluation of service quality in networks, systems or even applications [19].

For QoE, Mean Opinion Score (MOS) is the key quality indicator, the metric that directly shows the QoE level [20]. It is the 5-point scale metric used to quantify the level of multimedia quality perception (5=excellent, 4=good, …, and 1=bad) that is used widely [21]. It directly presents the QoE level provided by real-time applications (e.g. VoIP and Video Streaming), thus it is useful to evaluate applications and/or networks from users' point of view [22]. Nowadays, several kinds of services/applications are also utilized by MOS for quality assessment. For example, in the case of video streaming, even the Peak Signal-to-Noise Ratio (PSNR), the most common and widely used metric for video quality assessment in the past, is also converted to MOS [23].

2.3 Hypothesis Tests

A hypothesis test is a type of statistical inference [15]. It uses the data from a sample to decide between a null hypothesis (H_0), and an alternative hypothesis (H_1), which makes a specific claim about the parameters and describes that the null hypothesis is false. It is necessary to determine the p-value from the

วารสารวิชาการเทคโนโลยีอุตสาหกรรม (The Journal of Industrial Technology)
ISSN (Print): 1686-9869, ISSN (online): 2697-5548
DOI: 10.14416/j.ind.tech.2020.03.003

gathered data to accept or reject H0. If the p-value is less than the significant level (e.g., 0.05), H0 is rejected and H1 is accepted. In contrast, if the p-value is more than the significant level, H0 is accepted. Thus, the significant level is the breakpoint to accept or reject the null hypothesis. In this paper, the t-test as in [24] cannot be used because multiple comparisons are required, therefore, ANOVA, as mentioned in [25], has been applied.

2.4 Previous related studies

Focusing on audiovisual quality assessment, it has been found that there are several interesting models as shown in Table 1 [26-29]. Those mathematical models may be utilized for video call quality calculation provided by social media applications. However, before applying those models, each model has been applied to five scenarios called S1-S5 as shown in Table 2, with different values of MOS_A and MOS_V to calculate for MOS_{AV} using (1)-(8), in order to verify if there is any issue about reliability.

As shown in Table 3, if the error from each model is accepted within the error range of ± 0.5 only, one can see that there is an unreasonable MOS_{AV} provided by several audiovisual models. The details can be presented as follows:

- (5), (6) and (8) are unreliable for S1
- (1), (2) and (6) are unreliable for S2, S3, and S4.
- (1) and (4) are unreliable for S5

After analyzing those models based on the error assumption using a simple approach, it has been found that there are only 2 models (the additive linear model as in (3) and the full model as in (7) which provides results without errors. Therefore, these two models have been applied in Section 3.

3. Methodology

This study is based on subjective tests, see Figure 1. It focuses on video telephony utilizing audiovisual short conversation test (AVSCT) [29]. For the free conversation topics, for example; studies in their classes, TV series, TV stars, and lottery. For video telephony applications, WhatsApp and Messenger were chosen due to the ranking in the Top 10 most popular social media applications, while Line was also selected in this study because of its high usage in Thailand [16].

Each pair of subjects is required to sit in separate rooms. The smartphones used in this study were a pair of iPhone 6 (model) with an iOS 11.4 operating system. Their displays were 4.7 inches with aspect ratio 16:9 and a resolution of 1334x750. Also, special headsets with foam ear cushions were used to get rid of noises. Both

วารสารวิชาการเทคโนโลยีอุตสาหกรรม (The Journal of Industrial Technology)
ISSN (Print): 1686-9869, ISSN (online): 2697-5548
DOI: 10.14416/j.ind.tech.2020.03.003

smartphones connected to the IP network via WiFi or wireless network (IEEE802.11n standard) inside the Faculty of Industry and Technology, Rajamangala University of Technology Rattanakosin (Wang Klai Kangwon Campus).

The IP network inside the campus was specifically configured for this study. The Internet link with a bandwidth of 1 Gbps was provided by one major network operator using a fiber link to provide the Internet network. For testing, there were 30 subjects per social media application [30-31], see Table 4 for more background information. Those subjects are represented as non-expert users. The same selection criteria were adopted to remove bias (e.g., they studied in the same faculty and had the same background knowledge about science and technology while most of them were about 21 years old).

Table 1 List of previous research works

Author(s)	Proposed concept	
ITU-T [26]	MOS_{AV}= -0.3255MOS_A+0.3309MOS_V+0.1494$MOS_A MOS_V$+0.5457	(1)
	MOS_{AV}= -0.1638MOS_A+0.3626MOS_V+0.1291$MOS_A MOS_V$+0.5456	(2)
	Where (1) is for the video display size of 4.2 inches, whereas (2) is for the display size of 2.1 inches	
Hands [27]	MOS_{AV}=0.17$MOS_A MOS_V$+1.15	(3)
	MOS_{AV}=0.25MOS_V+0.15$MOS_A MOS_V$+0.95	(4)
Winkler and Faller [28]	MOS_{AV}=1.98+0.103$MOS_A MOS_V$	(5)
	MOS_{AV}=0.456MOS_A+0.770MOS_V-1.51	(6)
Belmudez and Möller [29]	MOS_{AV}=0.584MOS_A+0.357MOS_V+0.0013$MOS_A MOS_V$+0.127	(7)
	MOS_{AV}=0.030MOS_A+0.079MOS_V+0.123$MOS_A MOS_V$+1.374	(8)

Table 2 The scenarios under tests for the model evaluation

Scenario	Quality		
	Audio	Video	Average
S1	1	1	1
S2	2	2	2
S3	3	3	3
S4	4	4	4
S5	5	5	5

วารสารวิชาการเทคโนโลยีอุตสาหกรรม (The Journal of Industrial Technology)
ISSN (Print): 1686-9869, ISSN (online): 2697-5548
DOI: 10.14416/j.ind.tech.2020.03.003

Table 3 The simple scenarios under tests for the model verification

Scenario	MOS			Acceptable Error	Calculated MOS-Audiovisual							
	Audio	Video	Average		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
S1	1	1	1.00	±0.50	0.70	0.87	1.32	1.35	2.08*	-0.28*	1.07	1.61*
S2	2	2	2.00	±0.50	1.15*	1.46*	1.83	2.05	2.39	0.94*	2.01	2.08
S3	3	3	3.00	±0.50	1.91*	2.30*	2.68	3.05	2.91	2.17*	2.96	2.81
S4	4	4	4.00	±0.50	2.96*	3.41*	3.87	4.35	3.63	3.39*	3.91	3.78
S5	5	5	5.00	±0.50	4.31*	4.77	5.40	5.95*	4.56	4.62	4.86	4.99

Remark: * = unacceptable calculated MOS

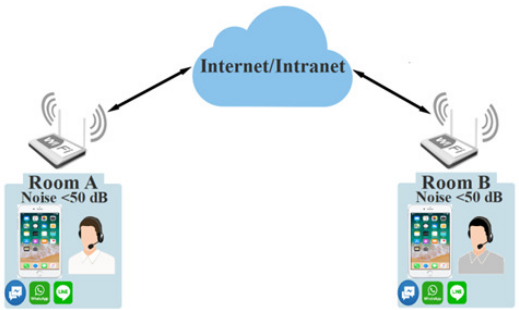

Fig. 1 Overview of the audiovisual subjective tests

Table 4 Background information of participants

Information of Participants	Social Media Platform		
	WhatsApp	Messenger	LINE
Male Subjects	14	17	15
Female Subjects	16	13	15
Average Age (Years)	21.33±0.82	21.10±0.99	21.00±1.11

Each pair of subjects tested with only one application randomly in order to avoid biases. In addition, the balance between male and female subjects was tried by monitoring and control. The subjective tests were conducted between the second week of January up to the second week of February 2018. After finishing the tests, the data was gathered and then analyzed. The results are presented in Section IV. Further, an additional task to investigate data consumption was also conducted. For this task, a third phone was used to take a 160-second video clip of the caller's phone after the call connection between the caller's mobile phone and the receiving mobile phone was established. Each application was tested 10 times. The average data consumption is also presented in the next section.

4. Results and Analysis

After gathering the total results of the subjective tests with the 90 non-expert subjects, (as described in Section III), MOS for Audio or MOS-Audio, and MOS for Video or MOS-Video, from the three social media applications, then the audiovisual quality values were calculated by

วารสารวิชาการเทคโนโลยีอุตสาหกรรม (The Journal of Industrial Technology)
ISSN (Print): 1686-9869, ISSN (online): 2697-5548
DOI: 10.14416/j.ind.tech.2020.03.003

using three approaches. They consisted of average calculation, the Hands' model (H1) as shown in (3) and Belmudez and Möller's model (BM1) as in (7). The results can be presented as in Figure 2.

One can see that the subjective results from all three applications show that MOS-Audio is higher than MOS-Video. Overall, Messenger is the best for audio quality and video quality when compared to WhatsApp and LINE. When focusing on audio quality, the result shows that Messenger with MOS-Audio of 4.27 is higher than LINE and WhatsApp with MOS-Audio of 4.17 and 3.90 respectively. For video quality, Messenger with MOS-Audio of 3.93 is slightly higher than WhatsApp and LINE with MOS-Audio of 3.80 and 3.63 respectively. For the average MOS from MOS-Audio and MOS-Video, it can be seen that MOS-Average provided by Messenger with MOS-Average of 4.10 is slightly greater than LINE and WhatsApp with MOS-Average of 3.90 and 3.85 respectively.

For the MOS-Audiovisual values, or MOS-AV for short, it was found that the MOS score from Hands' model that has been calculated using (3), called MOS-AV(H1), provided by Messenger with the score of 4.02 is greater than WhatsApp and LINE with the scores of 3.75 and 3.63 respectively. The MOS score from Belmudez and Möller's model that has been calculated using (7), called MOS-AV(BM1), provided by Messenger with a score of 4.04 is higher than LINE and WhatsApp with the score of 3.85 and 3.78 respectively. Nonetheless, in order to investigate if there is a significantly different quality provided by those three applications, the hypotheses H1 – H8 must be tested using statistical technique [15].

$H1_0$: The audio quality perception of the participants from WhatsApp, Messenger, and LINE is the same

$H1_1$: The audio quality perception of the participants from WhatsApp, Messenger, and Line is not the same

$H2_0$: The video quality perception of the participants from WhatsApp, Messenger, and LINE is the same

$H2_1$: The video quality perception of the participants from WhatsApp, Messenger, and LINE is not the same

$H3_0$: The average quality (between audio and video quality) from WhatsApp, Messenger, and LINE is the same

$H3_1$: The average quality (between audio and video quality) from WhatsApp, Messenger, and LINE is not the same

$H4_0$: The audiovisual quality using Hands' model from WhatsApp, Messenger and LINE is the same

วารสารวิชาการเทคโนโลยีอุตสาหกรรม (The Journal of Industrial Technology)
ISSN (Print): 1686-9869, ISSN (online): 2697-5548
DOI: 10.14416/j.ind.tech.2020.03.003

$H4_1$: The audiovisual quality using Hands' model from WhatsApp, Messenger and LINE is not the same

$H5_0$: The audiovisual quality using Belmudez and Möller's model from WhatsApp, Messenger, and LINE is the same

$H5_1$: The audiovisual quality using Belmudez and Möller's model from WhatsApp, Messenger, and LINE is not the same

$H6_0$: The audiovisual quality perceived from WhatsApp calculated by different equations is the same

$H6_1$: The audiovisual quality perceived from WhatsApp calculated by different equations is not the same

$H7_0$: The audiovisual quality perceived from Messenger calculated by different equations is the same

$H7_1$: The audiovisual quality perceived from Messenger calculated by different equations is not the same

$H8_0$: The audiovisual quality perceived from LINE calculated by different equations is the same

$H8_1$: The audiovisual quality perceived from LINE calculated by different equations is not the same

Then, the test results of the hypotheses H1-H8 are presented as in Table 5. After the hypothesis testing with the confidence interval of 95 %, one can see that there is no significant difference in every tested case. In addition, according to the task about playing video clip (3 minutes per session, including 10 seconds before playing and 10 seconds after playing) in front of the caller's mobile phone that is monitored by using the GlassWire application [32], the results can be seen as in Figure 3 that is based on the same quality of 480p playing on the third mobile phone, Messenger consumes the highest data (about 5.78 MB per round), whereas LINE and WhatsApp consume only 3.21 MB per round and 2.79 MB per round approximately. It means Messenger consumes about 2 times of the data required by LINE and WhatsApp.

วารสารวิชาการเทคโนโลยีอุตสาหกรรม (The Journal of Industrial Technology)
ISSN (Print): 1686-9869, ISSN (online): 2697-5548
DOI: 10.14416/j.ind.tech.2020.03.003

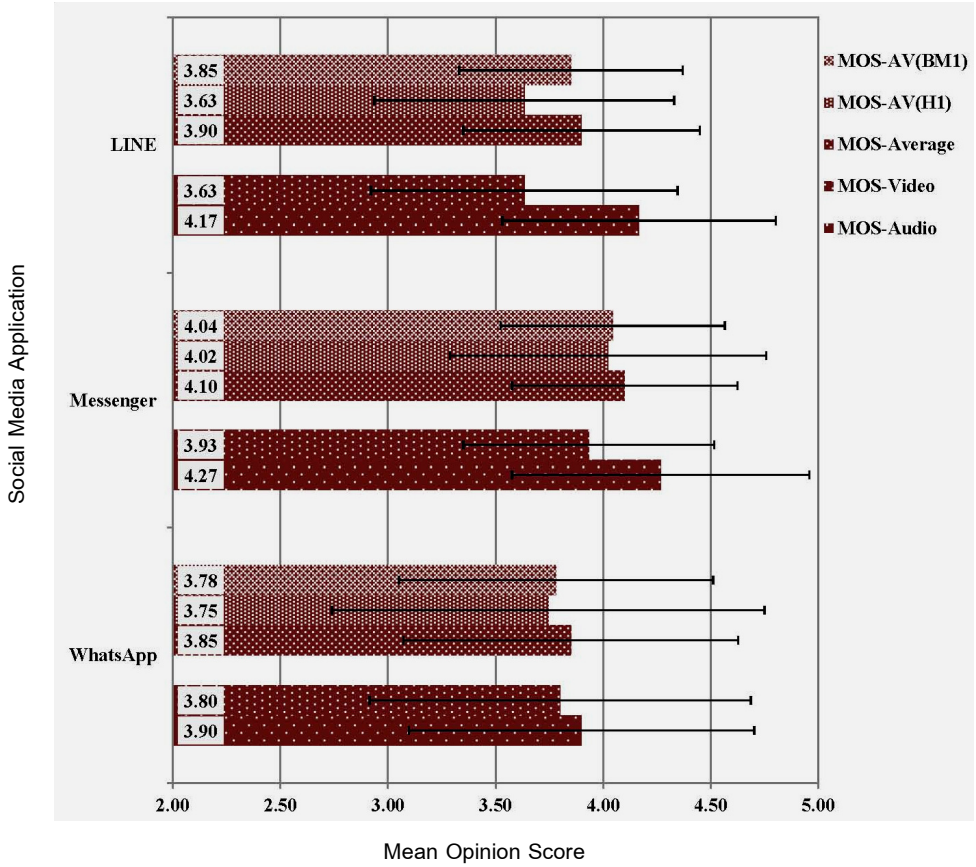

Fig. 2 Results from the social media applications under tests

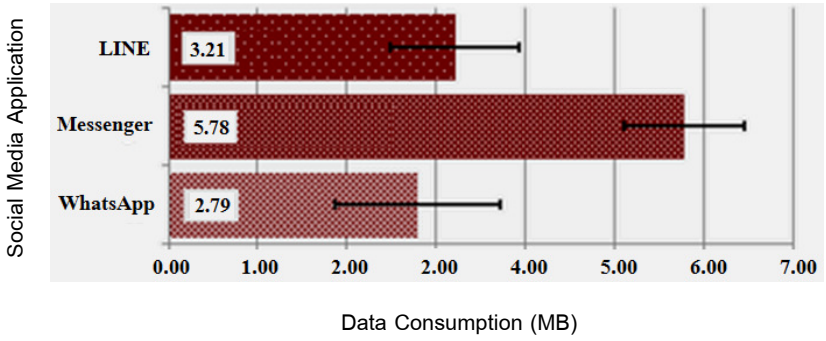

Fig. 3 The monitoring results from 160-second video clip playing (the data consumption tested)

วารสารวิชาการเทคโนโลยีอุตสาหกรรม (The Journal of Industrial Technology)
ISSN (Print): 1686-9869, ISSN (online): 2697-5548
DOI: 10.14416/j.ind.tech.2020.03.003

Table 5 Statistical analysis results

Hypotheses	p-value	Remark
$H1_0$: $MOS\text{-}Audio_{WhatsApp} = MOS\text{-}Audio_{Messenger} = MOS\text{-}Audio_{LINE}$ $H1_1$: $MOS\text{-}Audio_{WhatsApp} \neq MOS\text{-}Audio_{Messenger} \neq MOS\text{-}Audio_{LINE}$	0.152	Insignificant
$H2_0$: $MOS\text{-}Video_{WhatsApp} = MOS\text{-}Video_{Messenger} = MOS\text{-}Video_{LINE}$ $H2_1$: $MOS\text{-}Video_{WhatsApp} \neq MOS\text{-}Video_{Messenger} \neq MOS\text{-}Video_{LINE}$	0.310	Insignificant
$H3_0$: $MOS\text{-}Average_{WhatsApp} = MOS\text{-}Average_{Messenger} = MOS\text{-}Average_{LINE}$ $H3_1$: $MOS\text{-}Average_{WhatsApp} \neq MOS\text{-}Average_{Messenger} \neq MOS\text{-}Average_{LINE}$	0.296	Insignificant
$H4_0$: $MOS\text{-}AV(H1)_{WhatsApp} = MOS\text{-}AV(H1)_{Messenger} = MOS\text{-}AV(H1)_{LINE}$ $H4_1$: $MOS\text{-}AV(H1)_{WhatsApp} \neq MOS\text{-}AV(H1)_{Messenger} \neq MOS\text{-}AV(H1)_{LINE}$	0.367	Insignificant
$H5_0$: $MOS\text{-}AV(BM1)_{WhatsApp} = MOS\text{-}AV(BM1)_{Messenger} = MOS\text{-}AV(BM1)_{LINE}$ $H5_1$: $MOS\text{-}AV(BM1)_{WhatsApp} \neq MOS\text{-}AV(BM1)_{Messenger} \neq MOS\text{-}AV(BM1)_{LINE}$	0.259	Insignificant
$H6_0$: $MOS\text{-}Average_{WhatsApp} = MOS\text{-}AV(H1)_{WhatsApp} = MOS\text{-}AV(BM1)_{WhatsApp}$ $H6_1$: $MOS\text{-}Average_{WhatsApp} \neq MOS\text{-}AV(H1)_{WhatsApp} \neq MOS\text{-}AV(BM1)_{WhatsApp}$	0.886	Insignificant
$H7_0$: $MOS\text{-}Average_{Messenger} = MOS\text{-}AV(H1)_{Messenger} = MOS\text{-}AV(BM1)_{Messenger}$ $H7_1$: $MOS\text{-}Average_{Messenger} \neq MOS\text{-}AV(H1)_{Messenger} \neq MOS\text{-}AV(BM1)_{Messenger}$	0.874	Insignificant
$H8_0$: $MOS\text{-}Average_{LINE} = MOS\text{-}AV(H1)_{LINE} = MOS\text{-}AV(BM1)_{LINE}$ $H8_1$: $MOS\text{-}Average_{LINE} \neq MOS\text{-}AV(H1)_{LINE} \neq MOS\text{-}AV(BM1)_{LINE}$	0.728	Insignificant

Note: it is insignificant if p-value > 0.05 for 95% confidence interval

5. Discussion

Several issues can be discussed as follows:

1) It was found that some of the assessment models provide low reliability in some cases, for example, the model as in (1), (2), (5) and (6). (1), (2) and (6) seem inappropriate to be used for audiovisual quality calculation in S2, S3, and S4 (see the '*' in Table 3). One obvious issue might be about different sizes of the mobile phone display because (1) and (2) have been developed based on the display sizes of 4.2 inches and 2.1 inches respectively, while, (5) and (6) shows errors when they have been used to compute the audio quality in S1. However, for the assumption in this study, the reliability of the model is low if the result of MOS is higher or lower than the average score of more than 0.5. It means, this assumption should be investigated further in the future.

วารสารวิชาการเทคโนโลยีอุตสาหกรรม (The Journal of Industrial Technology)
ISSN (Print): 1686-9869, ISSN (online): 2697-5548
DOI: 10.14416/j.ind.tech.2020.03.003

2) For the audio, video and audiovisual qualities, although Messenger provides slightly better quality than WhatsApp and LINE, there is no significant difference among three social media applications (see the p-values from H1-H5 tests that are higher than 0.05).

3) For model performance, it has been investigated by utilizing the results from the following models: Hands' model, Belmudez and Möller's model, and the average calculation. The selected technique is ANOVA. From the results, it has been found that there is no significant difference among the three models or equations for audiovisual quality calculation because the p-values from H6-H8 tests are higher than 0.05. Thus, audiovisual quality from video calls can be roughly estimated by using the average calculation from the audio quality and video quality.

4) This study was conducted based on AVSCT only which is a common scenario for video calls. It means the average calculation, Hands' model and Belmudez and Möller's model as applied in this study may provide different results if they are used for other scenarios (e.g., IPTV)

5) This study was conducted with native Thai subjects who have different cultures and languages than westerners. Thus it may be inferred as per human influential factors [33], that the results might be different because different subjects may have variable behavior in a conversation due to culture and personality. For example, this could create greater variability in subjects' responses in the evaluation of speech quality [31].

6) All three social media applications may use the different proprietary codecs for voice and video transmission over IP networks. This could be the reason why they provide different audiovisual quality. However, it is difficult to investigate the codecs used by these applications because they are normally encrypted., Skype, on the other hand, has revealed that it uses SILK codec [30].

8) Focusing on the relation of quality and the quantity of data used for each session of video calls, it has been found from this study that Messenger consumes about 2 times of data than LINE and WhatsApp. This probably is the reason why it provides the best audio quality, video quality, and audiovisual quality when compared to WhatsApp and LINE.

6. Conclusion and Future Work

From this study, it was found that Messenger provides better audiovisual quality than LINE and WhatsApp, where MOS-AV values from the average calculation are 4.10, 3.90 and 3.85 for Messenger, LINE and WhatsApp respectively.

วารสารวิชาการเทคโนโลยีอุตสาหกรรม (The Journal of Industrial Technology)
ISSN (Print): 1686-9869, ISSN (online): 2697-5548
DOI: 10.14416/j.ind.tech.2020.03.003

From the calculation using the Hands' additive linear model, it was found that Messenger provides better audiovisual quality than LINE and WhatsApp, where MOS-AV values from the average calculation are 4.02, 3.75 and 3.63 for Messenger, WhatsApp and LINE respectively. Lastly, from the calculation using the Belmudez & Moller's full model, it was found that Messenger provides better audiovisual quality than LINE and WhatsApp, where MOS-AV values from the average calculation are 4.04, 3.85 and 3.78 for Messenger, LINE and WhatsApp respectively. However, the better quality with insignificant difference provided by Messenger compared with the other applications is a trade-off given it has 2 times the data consumption for making video calls with the same condition under tests. Therefore, users may use the conclusion from this study to apply selection criteria for the right application. Moreover, this study has shown that the Hands' additive linear model and the Belmudez & Moller's full model are appropriate for audiovisual quality evaluation based on a short conversation. Thus, these two models can be used to assess the audiovisual quality provided by other social media applications.

This study presented the utilization of the existing models only. A new QoE model with high reliability and accuracy for audiovisual quality measurement by focusing on audiovisual short conversation test (AVSCT), might be developed as future work. The intensive subjective study with the overall quality from audiovisual communication provided by social media applications should be conducted and compared to the existing models. Besides, the study with different sizes of display and other conversation scenarios might be conducted in the future as well.

7. Acknowledgment

Thank you to the Faculty of Engineering, Rajamangala University of Technology Phra Nakhon and the Faculty of Industry and Technology, Rajamangala University of Technology Rattanakosin (Wang Klai Kangwon Campus) for supporting. Especially, thank you to all students who were participants or involved. Lastly, special thanks to Mrs.Anong Lertrakskun and Mr.Bhaskar Laha for English Editing.

8. References

[1] www.statista.com/statistics/203734/globalsmartphone-penetration-per-capita-since-2005 (Accessed on 20 January 2020)

[2] www.dreamgrow.com/top-15-most-popular-social-networking-sites/ (Accessed on 20 January 2020)

วารสารวิชาการเทคโนโลยีอุตสาหกรรม (The Journal of Industrial Technology)
ISSN (Print): 1686-9869, ISSN (online): 2697-5548
DOI: 10.14416/j.ind.tech.2020.03.003

[3] S. Jana, A. Pande, A. Chan and P. Mohapatra, Mobile video chat: Issues and Challenges, IEEE Communications Magazine, 2013, 51, 144 – 151.

[4] I. Saidi, L. Zhang, V. Barriac and O. Deforges, Audiovisual Quality Study for Videoconferencing on IP Networks, IEEE Workshop on Multimedia Signal Processing (MMSP 2016), 2016. DOI: 10.1109/MMSP.2016.781337.

[5] K. Yamagishi, T. Tominaga, T. Hayashi and A. Takahashi, Objective Quality Evaluation Model for Videophone Services, NTT Technical Review, 2007, 5(6), 1 – 5.

[6] T. Daengsi and P. Pornpongtechavanich, Comparison of Video Call Quality from Social Media Applications: Messenger vs LINE, The International Conference on Science, Technology and Innovation for Sustainable Well-Being (STISWB 2019), Proceeding, 2019, 578 – 583.

[7] P. Zou, H. Li, W. Wang, C. Xin, and B. Zhu, Finding Disjoint Dense Clubs in a Social Network, Theoretical Computer Science, 2018, 734, 15 – 23.

[8] Md.O. Faruq, A. Al-Reza, Md. M. Rahman, and M.R. Alam, Impact of Social Networking Sites in Bangladesh: Few Possible Solutions, International Journal of Intelligent Systems and Applications, 2017, 9(4), 53 – 60.

[9] www.whatsapp.com/about/?lang=en (Accessed on 20 January 2020)

[10] https://crambler.com/truth-about-facebook-messenger-app-privacy (Accessed on 20 January 2020)

[11] www.webwise.ie/parents/explained-what-is-messenger (Accessed on 20 January 2020)

[12] https://play.google.com/store/apps/details?id=com.facebook.orca&hl=en (Accessed on 20 January 2020)

[13] https://itunes.apple.com/us/app/messenger/id454638411 (Accessed on 20 January 2020)

[14] www. businessinsider.com/history-of-line-japan-app-2014-1 (Accessed on 20 January 2020)

[15] P. Wuttidittachotti and T. Daengsi, Quality Evaluation of Mobile Networks Using VoIP Applications: A Case Study with Skype and LINE based-on Stationary Tests in Bangkok, International Journal of Computer Network and Information Security, 2015, 7(12), 28 – 41.

[16] www.pressreader.com/thailand/Bangkok-post/20170321/281925952832752 (Accessed on 20 January 2020)

วารสารวิชาการเทคโนโลยีอุตสาหกรรม (The Journal of Industrial Technology)
ISSN (Print): 1686-9869, ISSN (online): 2697-5548
DOI: 10.14416/j.ind.tech.2020.03.003

[17] P. Calyam and C. Lee, Characterizing Voice and Video Traffic Behaviour Over the Internet, The 20th International Symposium on Computer and Information Sciences (ISCIS'05), 2005, 91 – 95.

[18] A. Khan, L. Sun, E. Jammeh and E. Ifeachor, Quality of Experience-Driven Adaptation Scheme for Video Applications Over Wireless Networks, IET Communications, 2010, 4(11), 1337 – 1347.

[19] B. Jia, L. Hao, C. Zhang, Zhang, H. Zhang and M. Khan, An IoT Service Aggregation Method Based on Dynamic Planning for QoE Restraints, Mobile Networks and Applications, 2019, 24, 25 – 33.

[20] H. Batteram, G. Damm, A. Mukhopadhyay, L. Philippart, R. Odysseos, and C. Urrutia-Valdés, Delivering quality of experience in multimedia networks, Bell Labs Technical Journal, 2010, 15(1), 175 – 193.

[21] ITU-T Recommendation P.800, Methods for Subjective Determination of Transmission Quality, 1996.

[22] P. Wuttidittachotti, P. Khaoduang and T. Daengsi, MOS Eestimation Model Development using ACR Listening-Opinion Tests with Thai Users Referring to Loss Effects: A Case of G.726 and G.729", Multimedia Systems, 2018, 24(3), 285 – 295.

[23] T. Zinner, O. Abboud, O. Hohlfeld, T. Hossfeld and P. Tran-Gia, Towards QoE Management for Scalable Video Streaming, The 21st ITC Specialist Seminar on Multimedia Applications - Traffic, Performance and QoE (ITC-SS21), 2010, 64 – 69.

[24] M. Piriyakul, Analysis of Moderation Effects Through PROCESS, The Journal of Industrial Technology, 2019, 15(2), 144 – 161.

[25] K. Songsorn, K. Sriprateep and S. Rittidechhttp, Optimization of Metal Inert Gas Pulse Brazing Process on Galvanized Steel Sheets Based on Taguchi Method, The Journal of Industrial Technology, 2017, 13(2), 78 – 95.

[26] ITU-T Recommendation G.1070, Opinion Model for Video-Telephony Applications, 2018.

[27] D. S. Hands, A Basic Multimedia Quality Model, IEEE Transactions on Multimedia, 2004, 6(6), 806 – 816.

[28] S. Winkler and C. Faller, Perceived Audiovisual Quality of Low-Bitrate Multimedia Content, IEEE Transactions on Multimedia, 2006, 8(5), 973 – 980.

วารสารวิชาการเทคโนโลยีอุตสาหกรรม (The Journal of Industrial Technology)
ISSN (Print): 1686-9869, ISSN (online): 2697-5548
DOI: 10.14416/j.ind.tech.2020.03.003

[29] B. Belmudez and S. Möller, Audiovisual Quality Integration for Interactive Communications, EURASIP Journal on Audio Speech and Music Processing, 2013, Article No. 2013: 24, 1 – 23.

[30] P. Wuttidittachotti and T. Daengsi, Subjective MOS Model and Simplified E-Model Enhancement for Skype Associated with Packet Loss Effects: A Case Using Conversation-like Tests with Thai Users, Multimedia Tools and Applications, 2017, 76(15), 16163 – 16187.

[31] ITU-T Recommendation P.805, Subjective Evaluation of Conversational Quality, 2007.

[32] https://www.glasswire.com/?os=win (Accessed on 23 August 2018)

[33] M. Matulin and Š. Mrvelj, Modelling User Quality of Experience from Objective and Subjective Data Sets Using Fuzzy Logic, Multimedia Systems, 2018, 24, 645 – 667.