

# NVSU VoIP: A Mean Opinion Score (MOS) Assessment

**Dennis Hope R. Directo**

Department of Computer Science, Nueva Vizcaya State University, Bambang, Nueva Vizcaya, Philippines 3702

---

## **Keywords:**

MOS test, VoIP, subjective listening test

## **ABSTRACT**

Subjective listening test is one of the most reliable and definitive way of assessing the effectiveness of Voice over Internet Protocol Technology. One of the various way to assess Voice over Internet Protocol is the Mean Opinion Score test. This study aimed to measure the Mean Opinion Score of the implemented Voice over Internet Protocol system in Nueva Vizcaya State University – Bambang Campus. The study utilized the descriptive method of research to verify theories that refers to the present situation. The proponent used several techniques such as the use of questionnaires, interview, observation and a rubric to gathering facts and information.

The study shows that the call quality produced by the implemented Voice over Internet Protocol system is comparable to the quality produced by traditional telephone and cellular phone system having an average Mean Opinion Score of 4.49 which is slightly higher than the industry standard. It is therefore recommended to fully implement such technology to various offices of the university to minimize the subscription of telephone lines and maximize existing telephone systems.

---

## **INTRODUCTION**

Consumers and enterprises are rapidly adopting the latest trend in voice technologies to date, the Voice over Internet Protocol (VoIP). A technology that revolutionizes on how people communicate with each other. VoIP possesses all the features and benefits of a conventional telephone systems offering more flexibility and more advance feature compared to its predecessors (Venkatesh *et al.*, 2012).

VoIP is a technology that enables analog voice signal to be transmitted over computer networks rather than the conventional transmission of voice signals over analog or telephone lines. Since VoIP utilizes computer networks, the implementation of such technology does not require a new infrastructure thus making

implementation cost to at least the minimum. (Zhang *et. al.*, 2004).

Several studies were conducted with regards to VoIP implementation but only few researches being conducted focus on the assessment on the quality of the sound or voice being transmitted over the computer network. It is therefore the intention of this research to conduct such, to measure the Mean Opinion Score (MOS) of the implemented VoIP system in the Nueva Vizcaya State University, Bambang Campus.

The MOS test is a subjective listening test which is used to measure and determine the speech quality being transmitted over the network. Subjective listening tests are generally regarded as the most reliable and definitive way of assessing audio quality (Riberio *et al.*, 2011). In a MOS test, the test persons listen to short speech samples, where

every speech sample consists of two to five sentences. The total MOS is then the mean of all individual results. Since laboratory quality studies are time consuming and expensive, researchers often run small studies with less statistical significance or use objective measures which only approximate human perception.

In general, subjective quality measures require that (1) there are enough listening subjects of sufficient diversity to deliver statistically significant results; (2) experiments are conducted in a controlled environment with specific acoustic characteristics and equipment; (3) every subject receives the same instructions and stimuli.

The main objective of this study is to measure the MOS of the implemented VoIP system at the Nueva Vizcaya State University – Bambang Campus using the International Telecommunications Union (ITU) scale. Specifically, the study aims to: (1) conduct MOS test to various end users of the VoIP System; (2) measure and compute the MOS score of various groups as to gender and technical knowledge on VoIP technology; (3) compare and interpret the computed average MOS score to existing industry standards; and (4) check the configured parameter of the VoIP system if it is acceptable to end users.

## RESEARCH METHODOLOGY

The study utilized the descriptive method of research to verify theories that refer to the present situation. Data were gathered through interviews, research and survey questionnaire, and a rubric to test the call quality. The researcher made use of the following statistical treatments: weighted mean, Likert scale, percentage distribution to interpret the data gathered.

Due to the limited number of offices installed with a VoIP system, the proponent conducted the pilot test to the

whole university. All offices installed with the VoIP system were selected, 40 offices of the campus composing of faculty members with administrative positions/ designated officials and administrative non-teaching staff to as respondent of the study grouped as Non-Technical Group and ten Information Technology, Computer Science, Computer Engineering and Electronic and Communications Engineering faculty members as respondents of the Technical Group.

In the conduct of MOS test, listeners were placed in a “quiet room”, where no disruptions or interference as the conduct of test is being performed and scored the call quality as they perceived it as per ITU -T P.800 scale presented in Table 1.

The MOS was computed by transmitting several audio file of a male and female voice to each respondent. For each audio the respondent rated the quality and the average MOS was then calculated from the group of individual scores.

Each of the respondents received a call from the researcher using the IP Phone system and the MOS test was conducted. The respondents received instructions coming from the researcher explaining the reason why the test needed to be conducted, the objective of the MOS test and how they are going to rate the MOS test.

The procedure of the test was explained to each of the respondents. After explaining the procedures of the MOS

**Table 1: MOS / Call quality level scale**

MOS (ITU Scaled)	User Satisfaction
4.34 – 5.00	Very Satisfied
4.03 – 4.33	Satisfied
3.60 – 4.02	Some Users Satisfied
3.10 – 3.59	Many Users Dissatisfied
2.58 – 3.09	Nearly All Users Dissatisfied
1.00 – 2.59	Not Recommended

**Table 2. Summary of MOS**

<b>Gender / Group</b>	<b>Mean of the Technical Group</b>	<b>Mean of the Non – Technical Group</b>	<b>Ave rage Mean</b>	<b>Interpretation</b>
Male	4.48	4.61	4.55	Very Satisfied
Female	4.28	4.60	4.44	Very Satisfied
<b>Overall</b>	<b>4.38</b>	<b>4.60</b>	<b>4.49</b>	<b>Very Satisfied</b>

test, the researcher asked the respondents if they have understood the rationale of the upcoming test and if they are willing and are ready to proceed to the next step. After positive affirmation and confirmation by the respondents, the respondents heard a male voice informing them to rate the quality of the call after hearing the male correspondent reading the first sentence.

Then, the respondents were asked if they have understood the lines which was preciously read and if they already finished rating the first sentence. After receiving confirmation from the respondent, the second sentence was read and the same process as the latter were performed until the end of the fifth sentence. After the male voice was heard, the researcher explained to the respondent that the next voice they are going to hear is a female voice and the same process shall take place. After all processes were performed, the researcher informed the respondents that the MOS test was concluded.

The following lines of words were read or transmitted over the network in a male voice followed by a female voice.

1. You will have to be very quiet.
2. There was nothing to be seen.
3. They worshipped wooden idols.
4. I want a minute with the inspector.
5. Did he need any money?

## **RESULTS AND DISCUSSIONS**

After the conduct of the MOS test, the results were coded, tallied, tabulated,

analyzed and subjected to thorough statistical treatment using spreadsheet software with data analysis tool and the following findings as presented in Table 2.

Table 2 shows the evaluation of the MOS or on how the respondents rated the quality of the voice being transmitted over the communication lines that both groups were very satisfied with the quality of voice transmitted on both male and female voices. The average MOS for the male voice was 4.55 and 4.44 for the female voice.

The table also presents that the Technical Group garnered an overall mean of 4.38 with a descriptive interpretation that the users were very satisfied with the quality of the voice being transmitted over the network while the Non – Technical Group attained a mean rating of 4.60 with a descriptive rating of very satisfied.

According to [www. voip-info.org](http://www.voip-info.org), VoIP calls in the industry often have a MOS range of around 3.5 to 4.2. The computed average MOS of 4.49 signifies that the implemented VoIP system in the university is comparable with the industry standard.

## **CONCLUSION AND RECOMMENDATION**

### **Conclusion**

In this paper, the proponent described the MOS test which is a subjective measure to determine the quality of sound or voice being transmitted over the network and determined the MOS of the implemented VoIP system at

the Nueva Vizcaya State University Bambang Campus.

1. A total of 40 respondents were subjected to MOS test and results were carefully computed.
2. A computed average MOS of 4.49 was derived using the implemented VoIP system which is slightly higher as compared to the average MOS of various telecommunications company.
3. The quality of audio/voice transmitted over the network is slightly clearer thus such systems is a viable substitute as a cost effective inter-office communication system.
4. The respondents were very satisfied with the quality of the voice being transmitted over the network as per ITU scale. This result proves that the call quality produced by the VoIP system is comparable to the quality produced by traditional telephone and cellular phone system.
5. It was also concluded that configured parameters at the VoIP server are acceptable and is comparable to Industry standards as manifested by the computed average MOS of the respondents.

### **Recommendation**

1. A follow up study be conducted to measure the quality of audio being transmitted using signal analyzer or the likes.
2. It is highly recommended to fully implement such technology in the university to minimize the subscription of telephone lines and maximize existing telephone systems.
3. Similar study be conducted using the Voice over Wireless Local Area Network.
4. Compare results when implemented using Analog Trunk Lines.
5. The system must use backup power in case of power failure.

### **LITERATURE CITED**

- Venkatesh, V., Thong, J. Y., & Xu, X. (2012). Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS quarterly*, 36(1), 157-178.
- 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), IEEE International Conference, 2416-2419. IEEE.
- Zhang, J., Chan, S., & Fang, X. (2004). Enterprise user adoption of VoIP. School of Computer Science, Telecommunications, and Information Systems. Chicago, DePaul University, 10.