Article

Statistical Error Propagation Affecting the Quality of Experience Evaluation in Video on Demand Applications

Abdul Wahab [1]*, John Schormans [2] and Nafi Ahmad [3]

[1,2,3] School of Electronic Engineering and Computer Science, Queen Mary University of London, Mile End Rd, London, United Kingdom

* Correspondence: A.wahab@qmul.ac.uk

Abstract: In addition to the traditional QoS metrics of delay, delay jitter, and packet loss probability (PLP), Quality of Experience (QoE) is now widely accepted as a numerical proxy for actual user experience. The literature has reported many mathematical mappings between QoE and QoS. These QoS parameters are measured by the network providers using sampling. There are some papers studying sampling errors in QoS measurements; however there is no account of propagation of these sampling errors to QoE evaluation. In this paper, we used industrially acquired measurements of PLP and jitter to evaluate the sampling errors and correlation in measurements. Focussing on Video-on-demand (VoD) applications, we use subjective testing and regression to map QoE metrics onto PLP and jitter. The resulting mathematical functions of QoE and theory of error propagation was used to evaluate the propagated error in QoE, and this error was represented as confidence i nterval. Using the guidelines of UK government for sampling, our results indicate that confidence intervals around estimated QoE in a busy hour can be between MOS=1 to MOS=5 at targeted operating point of QoS parameters. These results are a new perspective on QoE evaluation, and are of great significance to all organisations that need to estimate the QoE VoD applications precisely.

Keywords: Quality of Experience; Quality of Service; QoE evaluation video on demand; Quality of Service; QoS correlation; subjective testing

---

1. Introduction

Industry and academia are focusing on QoE for quality evaluation of network applications, rather than just the more traditional QoS metrics of delay, delay jitter, and loss. Huawei and Telefonica are collaborating on projects such as end-to-end QoE evaluation, in which user perception will be key in shaping and improving their future services [1]. Other projects like Video QoE using U-vMOS, are aiming at comprehensive real-time QoE evaluation for different network service providers [2]. Recently Witbe launched "Witbe on the go" that enables the user to determine the QoE at any location in London, for any service provider and video on demand VoD [3].

Considerable research has already shown how QoS parameters such as packet loss, jitter and delay affect the perception of the user (QoE) for VoD applications. QoE can be evaluated by subjective testing for VoD applications and the acquired QoE results can be used to map QoE to PLP and jitter [4-6]. The resulting mathematical relationship between QoE and QoS can be used to assess QoE at different magnitudes of QoS parameters.

The high dependence of QoE on QoS parameters motivate the network service providers to measure the network parameters to ensure the provision of the magnitudes above the satisfaction threshold. These measurements are made both passively and actively. The latter employs sampling, and will be subject to sampling errors [7-10]. In this paper, we use the data acquired by our industrial collaborator Teragence. The data constitute the measurements of packet loss and delay jitter for cellular networks based in London. This data was used to evaluate the sampling error in QoS measurements, and to determine the Pearson correlation coefficient between PLP and jitter measurements.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 December 2019 doi:10.20944/preprints201912.0148.v1

We propose a novel technique to quantify the sampling error in the evaluated QoE, which we later refer to as uncertainty. The technique we use is in line with earlier work on how a dependent variable is affected by errors in the measurement of the independent variable(s) - the propagation of uncertainty. The sampling error (uncertainty) in QoE is expressed as 95 % confidence intervals. We also consider the correlation between the independent variable (loss and jitter); very little has been published on the correlation between these QoS metrics.

The remaining paper is organised as follows: In Section 2, we presented the related works elaborating the dependence of QoE of VoD on QoS parameters and highlights the possible causes of sampling errors in QoS measurements. Section 3 presents the related literature supporting the need for uncertainty analysis. Section 4 describes the methodology followed in this research. Results and discussion of the outcomes of this research was presented in Section 6. Finally, Section 7 concludes this paper and discusses the scope of future works.

2. Related Works

As reported by Cisco [11], video data transmission is variable-bit-rate and hence is regarded as bursty. Conventionally, User Datagram Protocol (UDP) is more suitable for the task of video transmission. However as the type of video traffic became more dynamic with the introduction of new types of codecs and an increase in the user bandwidth and larger receiver buffers for accommodating extra TCP traffic, UDP has been replaced by TCP for new applications like YouTube [12]. VoD applications working over HTTP/TCP show great dependence on the network QoS parameters. As reported in [4-6] and [13-14], the increase in PLP and jitter results in degradation of video quality.

Researchers in [4] and [6] have demonstrated the effect of PLP on the perceived QoE for VoD applications and concluded that QoE for VoD degrades for increasing PLP. The authors of [14] show how PLP affects different VoD frames transmitted on different codecs. The author of [4], determined that QoE for Video application is an exponential function of PLP and presented their results for users of different age groups. Similarly the authors of [6] determine that QoE is a power-law function of PLP for MPEG-4 video codec over UDP, with packets being transmitted through a traffic shaper. The consequence of degrading parameters like PLP and jitter on recently developed codecs like VP-9 and H.265 was studied in [15] and they determined that QoE is an exponential of these QoS parameters. Similarly, the authors in [16] captured the data for different types of data traffic from mobile applications and presented QoE of different application for different QoS parameters like PLP, delay and throughput. They also address the limitations of presenting QoE as MOS and presented their QoE results as slope of QoE for a certain QoS parameter for different multimedia services.

As reported in [15], the QoE for VoD using the VP-9 codec degrades rapidly for small increases in jitter, and they determined the relationship as an exponential function. In contrast to this, QoE showed less sensitivity to increasing jitter magnitudes and decays slowly when jitter was increased, as shown in [5], where the authors determined that the QoE is a power-law function of PLP. This difference shows that different codecs, testbeds and regression algorithms yield different QoE and jitter models. Moreover, attempts were made to study the combined effect PLP and jitter on QoE of VoIP in [17]. The authors fit their subjectively acquired MOS to PLP and jitter using polynomial regression and presented their results as 1st and 2nd order polynomials. We tried to reproduce the results of the polynomials by using different operating points of PLP and jitter but the published polynomial results were not reproducible and hence the model is not used in this project for further analysis. No similar attempt was seen for mapping QoE to a multi-variable function of both PLP and jitter for VoD in literature.

The quantification of QoE as MOS has many limitations as highlighted by the authors in [18], and they represented QoE as Quantiles to elaborate the meaning of the perception to QoE evaluators. Evaluating perception just based on MOS is not fair as MOS represents a mean of the score and hides significant information about the underlying data set and hence presenting the QoE as Quantiles or distributions as done in [19] is more beneficial than using MOS alone. We propose a related idea here: to present the QoE as confidence intervals at certain QoS operating points for better understanding of perception of VoD applications and to overcome the limitations imposed by use of MOS alone.

Network measurement techniques are technically mature and discussed in detail in [9-10] and [20] but their accuracy is still the subject of ongoing research. Network measurement techniques were examined for statistical

 doi:10.20944/preprints201912.0148.v1

accuracy in [7-8]. The measurements of packet delay result in low accuracy when measuring a congested network using packet probes [7] and [20]. The author of [21] attempted to measure the packet loss using probes and reported that increasing the probe frequency compromises the accuracy of the measurement. A similar challenge was presented in [22] which reported that increasing the sampling rate causes probes to interfere with application traffic thus making the measurements prone to error. Roughan [23] and [24] has also reported bounds on the accuracy and precision of using probes for measuring packet loss, and has tried to quantify the error in packet loss using statistical techniques. The size of the probes, sampling frequency of the probes and the load on the network can also cause inaccurate measurements as presented in [7-10], [20], [22] and [25].

Since QoE can be presented as the mathematical function of one or more QoS parameters, it is possible to work out the sampling error propagated to QoE, if the sampling error in the measurements of QoS parameters is known. The detailed study of analytical propagation of error in a function is reported in [26]. It is concluded that if the independent variables have errors in their measurements, then these measurement errors propagate to the dependent variable(s) and therefore it is vital to quantify this uncertainty when quantifying the magnitude of the dependent variables.

3. Uncertainty analysis

Uncertainty analysis is vital for all measurements and as reported by the National Physical Laboratory, measurements are only complete with a complete reference to the uncertainty in the measurand [27]. In common with [7-10], we consider the uncertainty that derives from measuring QoS parameters using sampling, e.g. [20]. Since QoE can be represented as a mathematical expression in terms of QoS parameters using the uncertainty propagation techniques presented in [25], the uncertainty in QoE is investigated. The evaluation of the propagated uncertainty is commonly used in different fields of biochemistry [26], physics [27] and psychology [28].

In this paper, the propagation of the uncertainty depends upon the mathematical model that relates QoE to QoS metrics, and whether the QoS parameters under observation are considered to be independent of each other or correlated. Our analysis of real network measured data has shown that they are in fact correlated. A viable uncertainty analysis for correlated variables requires the distribution of the uncertainty in each variable.

4. Methodology

We present a methodology that was adapted to quantify the propagated uncertainty in QoE due to statistical errors in QoS parameter measurements. The QoS parameters considered in this paper are PLP and jitter.

4.1. Capturing QoS measurements

The cellular measurement data used in this paper was captured by active measurements using the state-of-art platform of our industrial collaborator Teragence. The captured data consists of measurements of PLP, delay, jitter, the location of the mobile device and the mobile operator. All the measurements used in this paper were treated as anonymous in accordance to the privacy policy of UK.

The measurements are made actively by probing the packets from the android mobile application installed in the user device. The application sends set of 20 packets every 10ms referred as a ‘packet volley’ to the Teragence data servers installed in different locations in UK. The probing packets travels through the network and are captured at the servers, where the packets are processed to retrieve different measurements and are stored in the database. The administered access to the database was provided to us by Teragence to use their measurement data for analysis in this paper.

4.2. Sampling error in packet loss probability

In this paper, we assume that PLP is a random process that follows a Bernoulli distribution. The probability for the packet to be dropped is p and the probability for the packet being successfully being received is $q= (1-p)$. Since the whole population of the packets are not available, and the measurements are taken by sampling, the asymptotic standard deviation of X∼ B (1, p) is used, see equation (1).

 doi:10.20944/preprints201912.0148.v1

$$\sigma_{PLP} = \sqrt{\frac{p(1-p)}{n}} \tag{1}$$

where, p= packet loss probability and n is the number of samples.

4.3. Sampling error in jitter

In this paper, as in [5] and [17] we focus on jitter as being the standard deviation of the delay distribution. We used measurement data from Teragence to determine the delay distribution of the samples. It was found that the delay distribution follows a lognormal distribution.

In [29] the authors derived an expression for the standard deviation of the estimated standard deviation of a sample. Since the standard deviation is a measure of variation and in such case, the type of distribution does not affect the analysis [28]. From this, we used equation 2 to determine the uncertainty in jitter (standard deviation of the standard deviation)

$$\sigma_{jitter} = \frac{jitter}{\sqrt{2(n-1)}} \tag{2}$$

In equation (2), n represents the sample size and σ_{jitter} shows the standard deviation of jitter. The standard deviation of the jitter will be used in the appropriate formula as, the measure of uncertainty for jitter in this paper.

4.4. Correlation of packet loss and jitter

As described earlier QoS parameters can exhibit correlation and this correlation can be quantified using Pearson correlation coefficient. We calculated Pearson's Correlation coefficient from the QoS measurement data acquired from the measurements for different UK network service providers. Pearson correlation coefficient can be any value between -1 and 1, where -1 shows that two quantities are negatively correlated and 1 shows that they are positively correlated while 0 means that there is no correlation between two quantities.

We used a million data points for each of the 4 leading UK network providers over a period of almost one month and then calculated the Pearson correlation coefficient between the PLP and jitter measurements. The data sets were anonymous in alignment to data privacy rules and the measurements of jitter and PLP were used to calculate results presented in Table 1.

Table 1. Pearson Correlation Coefficient between PLP and jitter for leading UK network Providers

Service Providers	Correlation Coefficient between PLP and jitter
A	0.22
B	0.19
C	0.28
D	0.35

Table 1 shows that for all leading UK network providers the jitter and PLP exhibit non-trivial levels of correlation .This makes it critical to take this correlation in account while investigating the propagation of the uncertainty in QoE.

4.5. QoE models for video on demand applications

VoD applications like video streaming employing UDP were used to map the effect of changing QoS parameters to QoE. We used the Linux-based network emulator Netem to change the magnitudes of PLP and jitter. Our testbed includes a host Dell computer streaming the video to the subject Dell computer via Netem through Ethernet cables. VLC was used as a media player to stream the video between the host and subject machines.

A 2 minute long video of a car advertisement was used in this study fulfilling the spatial and temporal information requirements mentioned in ITU recommendation P.910 [30]. The video was encoded using H.264 and a frame rate of 30fps was configured. The resolution was set to 720p and fixed bitrate of 1200kbps was used in all experiments. Moreover, the lighting of the room of was set according to ITU recommendation BT.500 [31] and the subjective testing was carried out using ITU recommendation P.910 for subjective multimedia testing.

We used subjective testing to evaluate the QoE of video streaming at different PLP and jitter magnitudes. 11 subjects were used in this investigation with ages between 18 and 31 and has no visual disabilities. Each subject was shown the video at five different operating points of PLP and jitter without telling them network configuration to avoid bias. At the end of the video, the subjects were asked to rate the video according to Absolute Category Rating scale and the rating were used to calculate the MOS at that certain network configuration. Only one network parameter was changed at one time to see the effect of each parameter on QoE individually. No prior training was required and the total time taken by 1 subject was almost 25 minutes.

The collected data of MOS and regression analysis was used to derive the QoE expressions in terms of PLP and jitter. Since this is not the main focus of this research and we will not go in details of the performance of these models and see [4] for details. Equation 4 and 5 presents QoE in terms of PLP and jitter respectively.

$$QoE = 3.2 * exp^{-11.4*PLP} + 2 * exp^{-1065*PLP} \tag{4}$$

In equation(4) PLP is measured as a probability not a proportion. The fitting parameters for regression were Sum of Squares Due to Error (SSE) =0.002, R-square=0.99 and adjusted R-square as 0.99.

$$QoE = -0.9 * Jitter^{0.3} + 6 \tag{5}$$

In equation (5) jitter is measured in milliseconds and the fitting parameters for the regression were SSE=0.001, R-square=.098 and adjusted R-Square=0.99.

The SSE shows the error between the original QoE data and the fitted curve. We understand these errors will also result in variations of QoE at certain operating points of PLP and jitter but we will not focus on these fitting errors in this research and will only emphasise on the sampling errors in QoS measurements that can propagate to QoE.

4.6. Our approach to evaluating Uncertainty in QoE due to statistical errors in QoS measurements

The main equation for evaluating uncertainty in QoE due to sampling error in QoS metrics is given here as equation 6.

$$\sigma_{QoE} = \sqrt{(\frac{\partial QoE}{\partial PLP})^2 * \sigma^2_{PLP} + (\frac{\partial QoE}{\partial Jitter})^2 * \sigma^2_{jitter} + 2 * corr(PLP, jitter) * \frac{\partial QoE}{\partial PLP} * \sigma_{PLP} * \frac{\partial QoE}{\partial jitter} * \sigma_{jitter}} \tag{6}$$

In Equation 6, corr(PLP,jitter) is the Pearson Correlation coefficient between PLP and jitter calculated from the data set. It can be seen from this equation that the partial differential of QoE with respect to PLP and jitter is a key factor in evaluating the uncertainty in QoE, here evaluated as asymptotic standard deviation of QoE.

It can be seen from Figure 1 that, if the shape of the functional relationship between QoE and QoS is known, than for any small change in QoS, the resultant change in QoE can be found. To find the propagated uncertainty we calculate partial derivative of QoE w.r.t QoS parameters and the asymptotic standard deviations of PLP and jitter.

The approach to calculate the propagated uncertainty also depends upon whether the dependent variables are completely independent of each other or have some correlation between them. If the measured values of PLP and jitter are considered to be independent of each other than the model proposed in Equation 7 can be used to work out the propagated statistical error in QoE metric.

$$\sigma_{QoE} = \sqrt{(\frac{\partial QoE}{\partial PLP})^2 * \sigma^2_{PLP} + (\frac{\partial QoE}{\partial Jitter})^2 * \sigma^2_{jitter}} \tag{7}$$

Since, we have already established that PLP and jitter has considerable (20-30%) correlation between them, we work out the asymptotic standard deviation of QoE using Equation 6.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 December 2019 doi:10.20944/preprints201912.0148.v1

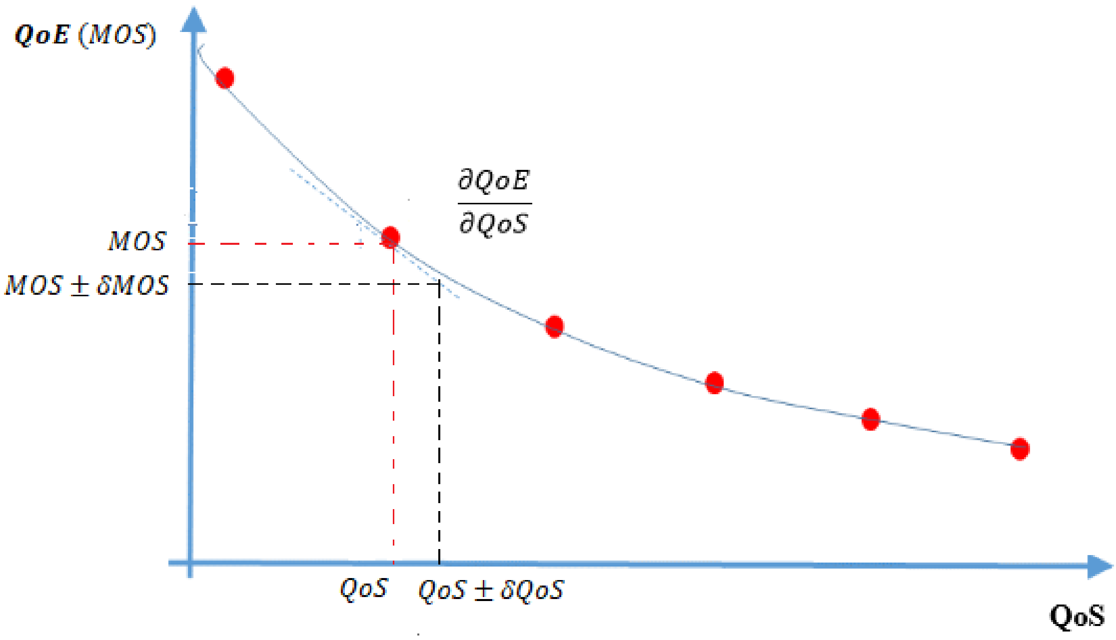

Figure 1. Propagation of uncertainty in QoE due to fluctuation in QoS.

4.7. Using Confidence Intervals to model uncertainty in QoE

The Confidence interval is used to show the spread around mean value of the evaluated QoE. Working out the mean value at the operating point is not the focus of this research but can be found using QoE relationships in Equation 4 and 5. In this paper, we used 95% confidence interval, see Equation 8.

$$ConfidenceInterval = Q\bar{o}E \pm Z * \sigma_{QoE} \tag{8}$$

Where: $Q\bar{o}E$ is the mean QoE as evaluated at the operating point, Z is the critical value for standard normal distribution and is 1.96 for 95% confidence interval and σ_{QoE} is the asymptotic standard deviation of QoE.

5. Results and Discussion

Initially, we calculated the first derivative of the QoE (PLP) function and QoE (jitter) function presented in section 3.4 as Equation 4 and 5 respectively. At targeted operating points for PLP and jitter, we calculated the first derivative of the functions. Our preliminary analysis reveals that uncertainty in QoE is highly dependent upon PLP and is almost completely unaffected by the shape of jitter function, as it decays slowly. Consequently, we used 1000 operating point ranging from 0 to 0.1 for PLP and used only 4 operating points for jitter, We presented the first derivative w.r.t PLP at five operating points of PLP in Table 2 and the first derivative w.r.t jitter at operating points of jitter in Table 3.

Table 2. First derivative of QoE w.r.t jitter at operating points

PLP Operating Points	$\partial QoE/\partial PLP$
0.0005	-1286.6
0.001	-770.3
0.01	-32.6
0.05	-20.6
0.1	-11.7

Table 3. First derivative of QoE w.r.t jitter at operating points

Jitter Operating Points	$\partial QoE / \partial jitter$
10	-0.050
100	-0.009
400	-0.003
800	-0.002

The results presented in Table 2 and 3 supports the observation made in the preliminary analysis that QoE of VoD applications is highly dependent upon the QoE's response to change in PLP in comparison to the QoE's response to change in jitter. Another interesting finding from Table 2 and 3 is that, for both PLP and jitter, when the magnitude of the operating points increases the first derivative decreases that back up the general shape of the QoE vs QoS function such as one displayed in Figure 1 in section 3.5.

Another important factor that affects the propagated uncertainty in QoE due to variations in PLP and jitter measurements is the sample size. So, we investigate the effect of different sample sizes on the magnitude of the resulting uncertainty in QoE.

5.1. Statistical error propagation in QoE considering correlation between PLP and jitter

We use 400, 800 and 1200 samples because the UK Government [32] recommends 800 samples per busy hour, and we consider it prudent to also consider a range above and below that figure.

Analysis was carried out that also includes the correlation values presented in Table 1, which reveals that on average the UK network service providers have around 20-40% Pearson correlation between their measured PLP and jitter. This is significant, and is included in our analysis that evaluates the propagated asymptotic standard deviation of QoE for VoD applications. Sample sizes of 400, 800 and 1200 were taken from the population of the measured data provided by the collaborating commercial organisation. The measurements were anonymised by following the privacy policy and they were distributed over a period of almost 50 hours. The calculated asymptotic standard deviation in QoE was used to calculate the confidence interval width of the evaluated QoE at different operating points of PLP and jitter. The resulting confidence interval width of QoE at a certain operating point of jitter was plotted against PLP and presented in Figures 2 and 3. Based on further analysis we found that the confidence interval width is not affected by the jitter magnitude for the power-law model presented in Equation 5. Hence, we only used the upper and lower limiting two values of jitter (10ms and 400ms) in our analysis for this model.

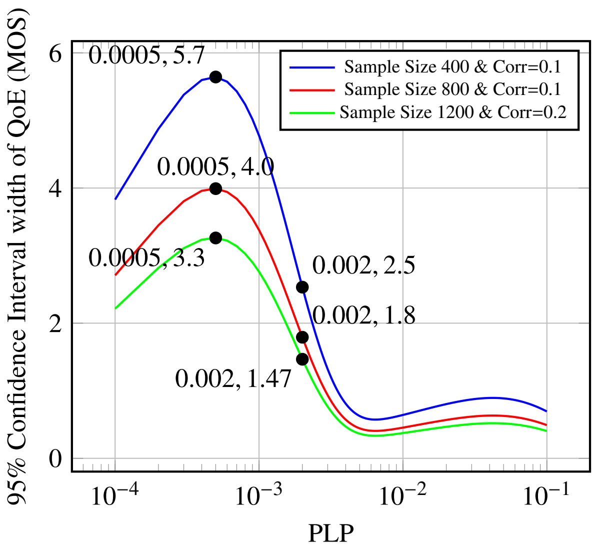

Figure 2. Propagated uncertainty in QoE vs PLP at jitter=10ms with Correlation for power-law jitter model.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 December 2019 doi:10.20944/preprints201912.0148.v1

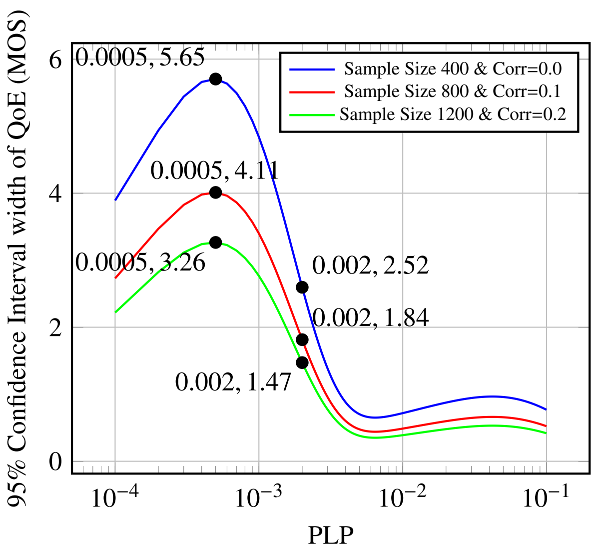

Figure 3. Propagated uncertainty in QoE vs PLP at jitter=400ms with Correlation for power-law jitter model.

The trend in confidence interval width of QoE vs PLP at all operating points of jitter is the same. We observe that:

- Uncertainty (measured as CI width) in QoE rises to a peak between PLP=0.0001 and PLP=0.001, and then rapidly diminishes as PLP either increases or decreases
- This shape is constant regardless of the jitter value, from very small (10ms) to very large (400ms) jitter values.
- uncertainty (measured as CI width) in QoE has a peak somewhere between PLP=0.0001 and PLP=0.001: this has great significance for network and service operators, as the mean PLP written into most SLAs is around the value of PLP=0.001 [33]
- The absolute predicted values of 95% CI width in QoE is large: for 800 samples CI width peaks at around 4 units of MOS.

Figures 2 and 3 supports our preliminary analysis, and it can be seen that with such big increase in jitter values from 10ms to 400ms there is no significant change in the confidence interval. This doesn't mean that jitter does not affect the QoE of VoD applications, as increase in jitter magnitude results in degradation of QoE for VoD as demonstrated in Figure 1. Rather it means that the statistical errors in QoE depends upon the variations in QoE with PLP (which is large) while the variation of QoE with jitter is relatively small.

We extend this analysis further and considered a more sensitive QoE (jitter) model for VoD. This published model presented in Equation 9 shows a steeper degradation for increasing jitter values [26]. Another reason for choosing this model is that it is also an exponential function, similar to the QoE (PLP) function used in this investigation.

$$QoE = 11.62 * exp^{(-3.386 * jitter)} - 4.408 * exp^{(-0.3477 * jitter)} \tag{9}$$

We probed this model using the same methodology and used Equation 6 to work out the asymptotic standard deviation and then confidence interval width of QoE for different operating points of PLP and jitter. The results were plotted against PLP and presented in Figures 4 and 5.

The results in Figures 4 and 5 show insignificant variation from results in Figure 2 and 3 and hence which shows that confidence interval width of QoE for VoD is not significantly affected by jitter, despite VoD actually being sensitive to jitter. We examined this by using two different models: one in which QoE degrades gradually with increasing jitter and second in which QoE degrades sharply with increasing jitter. In both cases the statistical errors in QoE remains the same.

Although, there is around 10-40% correlation observed in all cases, in none of the cases the propagated confidence interval width of QoE changes significantly as a function of correlation. This can be explained using

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 December 2019 doi:10.20944/preprints201912.0148.v1

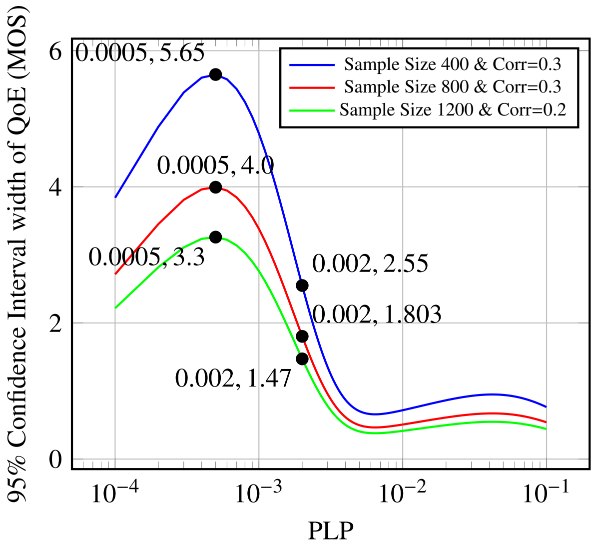

Figure 4. Propagated uncertainty in QoE vs PLP at jitter=1ms with Correlation for exponential jitter model.

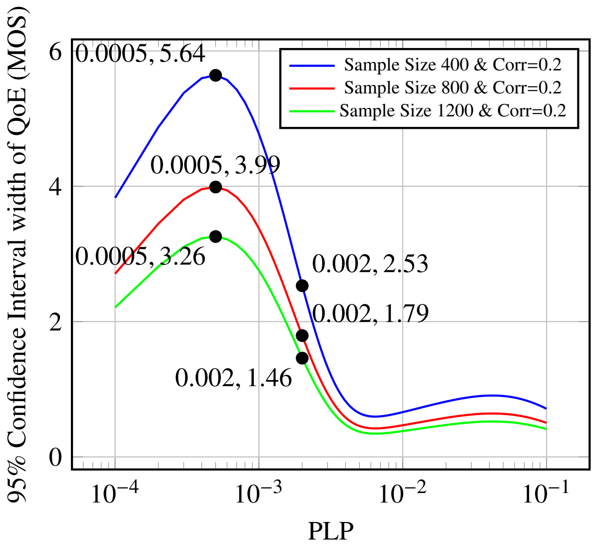

Figure 5. Propagated uncertainty in QoE vs PLP at jitter=5ms with Correlation for exponential jitter model.

Equation 7 for both power-law and exponential jitter models, where we can see that $\partial QoE/\partial PLP$ is squared and as described before the magnitude of $\partial QoE/\partial PLP$ is significantly higher than $\partial QoE/\partial jitter$ and hence the propagation of uncertainty for VoD QoE is mainly dependent upon the rapid change in QoE with PLP (about the operating point) rather than much slower change in QoE with jitter or correlation about the chosen operating points.

6. Conclusion

QoE is now widely accepted as providing a numerical proxy for actual user experience. While the literature has reported many objective mathematical mappings between QoS and QoE, there has been limited evaluation of the statistical errors that arise when estimating QoS by sampling, and previously none on how this propagates statistical error into the estimated QoE.

Using state-of-the-art commercially measured loss and jitter values, including loss-jitter correlation, we evaluated the error that arises during QoS measurements. We then used the theory of error propagation to evaluate the error that arises in the estimated QoE, and represented this error as confidence interval width. To ensure legitimate results, we employed UK Government guidelines for sampling.

 doi:10.20944/preprints201912.0148.v1

Our analysis shows that the statistical errors expressed as CI width in QoE is almost entirely the result of sampling error in PLP and not the sampling error in jitter. Importantly the magnitude of the sampling error (expressed as QoE CI) changes significantly at different operating points.

For instance, at very high or very low PLP and jitter operating points the uncertainty propagated is very small; however at the key operating point around PLP=0.001 the uncertainty propagated to QoE from the sampling errors in the QoS metrics of jitter and PLP is very much larger. Most commercial Service Level Agreements (SLAs) specify a target PLP of 0.001 (on average) [31], so our discovery that this is a critical point around which QoE uncertainty is at its largest is of great significance. In general, our results indicate that the confidence intervals around estimated QoE in a busy hour can be huge, which are results of great significance to all organisations that needs to precisely estimate an objective QoE metric.

Author Contributions: The idea of this project was proposed by J.S and A.W designed the experiments and carried out the analysis.N.A contributed towards data analysis and correlation.All authors contributed towards writing and reviewing of the final manuscript.

Funding: No External funding was given for this research

Acknowledgments: We are grateful to the School of Electronic Engineering and computer at the Queen Mary University of London for their help and support in this project. We extend our thanks to Dr Raul Mondragon and Dr Karen Shoop for their contributions in the project. We will also like to thank Teragence for their collaboration and providing us with the active measurement data.

Conflicts of Interest: There is no conflict of interest declared by this research or authors.

Abbreviations

The following abbreviations are used in this manuscript:

QoE	Quality of Experience
QoS	Quality of Service
VoD	Video on Demand
PLP	Packet loss Probability
CI	Confidence Interval
MOS	Mean Opinion Score
UDP	User Datagram protocol
TCP	Transmission control protocol

References

1. Telefonica: One step ahead in E2E QoE assurance. https://www.huawei.com/uk/industry-insights/outlook/customer-talk/hw_279053. accessed: 24 August 2019
2. Wireline Video QoE using U-vMOS. https://ovum.informa.com/~/media/informa-shop-window/tmt/files/whitepapers/huawei_wireline_video_online.pdf, accessed: 25 August 2019
3. Witbe QoE Snapshot #5 - What Is The BEst neighborhood in London to watch videos on mobile devices?', http://www.witbe.net/2019/05/14/snapshot5/ accessed 20 August 2019
4. Roshan,M. 'Experimental Approach to Evaluating QoE. PhD thesis. Queen Mary University of London, 2018
5. Ickin,S.; Vogeleer,K.D.; Fiedler,M. The effects of Packet Delay Variation on the perceptual quality of video. IEEE Local Computer Network Conference, Denver, CO, 2010, pp. 663-668.
6. Khorsandroo,S.; Noor,R.; Khorsandroo,S. A generic quantitative relationship between quality of experience and packet loss in video streaming services. 2012 Fourth International Conference on Ubiquitous and Future Networks (ICUFN), Phuket, 2012, pp. 352-356.
7. Timotijevic,T.; Leung,C.M.; Schormans,J. Accuracy of measurement techniques supporting QoS in packet-based intranet and extranet VPNs. *IEE Proceedings - Communications*.**2004**, 151,1, pp. 89-94.
8. Hasib,H.; Schormans,J.; Timotijevic,T. Accuracy of packet loss monitoring over networked CPE. *IET Communications*.**2007**, 1, 3, pp. 507-513.
9. Parker,B.; Gilmour,S.G; Schormans,J. Measurement of packet loss probability by optimal design of packet probing experiments. *IET Communications*.**2009**, 3, 6, 979-991.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 December 2019 doi:10.20944/preprints201912.0148.v1

10. Parker,B.; Gilmour,S.G.; Schormans,J. Optimal design of measurements on queueing systems, *Queueing Systems.* **2015** ,79, 3-4, 365-390.
11. Video Quality of Service (QOS) Tutorial. https://www.cisco.com/c/en/us/support/docs/quality-of-service-qos/qos-video/212134-Video-Quality-of-Service-QOS-Tutorial.pdf. accessed 02 August 2019
12. Verma,S., Flow control for Video Applications, in *Internet Congestion Control* (Morgan Kaufmann Publishers, 2015.); pp. 173–180
13. Hossfeld,T; Tran-Gia,T.; Fiedler,M. Quantification of Quality of Experience for Edge-Based Applications, in *Managing Traffic Performance in Converged Networks: 20th International Teletraffic Congress* (Springer International Publishing, 2007.); pp. 361–373
14. Saputra,Y.M; Hendrawan, The effect of packet loss and delay jitter on the video streaming performance using H.264/MPEG-4 Scalable Video Coding, 2016 10th International Conference on Telecommunication Systems Services and Applications. (TSSA), Denpasar, 2016, 1-6.
15. Pai,D.; Vanijja,V. Effect of network QoS on user QoE for a mobile video streaming service using H.265/VP9 codec. *Procedia Computer Science*.**2017**,111, 214-222.
16. Xu,Z.;Zhang,A. Network Traffic Type-Based Quality of Experience (QoE) Assessment for Universal Services. *MDPI applied sciences.* **2019**,9(19), 1-17.
17. Orosz,P.; Skopko,T.; Nagy,Z. A No-reference Voice Quality Estimation Method for Opus-based VoIP Services. *International Journal on Advances in Telecommunications*.**2014**, 7, 1-2, pp.12-21.
18. Hossfeld,T.; Heegard,P.E; Varela,M; QoE beyond the MOS: Added Value Using Quantiles and Distributions.2015 Seventh International Workshop on Quality of Multimedia Experience (QoMEX), Pylos-Nestoras, 2015, pp. 1-6
19. Seufart,M. Fundamental Advantages of Considering Quality of Experience Distributions over Mean Opinion Scores. 2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX), Berlin, Germany, 2019, pp. 1-6.
20. Mondragon,R.; Moore,A.; Pitts.J; Schormans,J. Analysis, simulation and measurement in large-scale packet networks *IET Communications.***2009**, 3, 6, 887-905.
21. Sommers,J.; Barford,P.; Duffield,N. Improving accuracy in end-to-end packet loss measurement. Proceedings of the 2005 conference on applications, technologies, architectures, and protocols for computer communications, Philadelphia, Pennsylvania, USA. August 2005, pp.157-168,
22. Schormans,J.; Pitts,J.M; "So you think you can measure IP QoS?," 2004 IEE Telecommunications Quality of Services: The Business of Success QoS 2004, London, UK, 2004, pp. 151-155.
23. Nguyen,H.X.; Roughan,M. Rigorous Statistical Analysis of Internet Loss Measurements, *IEEE/ACM Transactions on Networking*. **2013**, 21, 3, 734-745.
24. Roughan,R. Fundamental bounds on the accuracy of network performance measurements. Proceeding SIGMETRICS '05 Proceedings of the 2005 ACM SIGMETRICS international conference on Measurement and modeling of computer systems, Banff, Alberta, Canada, June 2005, pp.253-264,
25. Wahid,A.A; Schormans,J. Bounding the maximum sampling rate when measuring PLP in a packet buffer. 2013 5th Computer Science and Electronic Engineering Conference (CEEC), Colchester, 2013, pp. 115-118.
26. Ferrance,I.;Frenkel,R. Uncertainty of Measurement: A Review of the Rules for Calculating Uncertainty Components through Functional Relationships. *The Clinical biochemist. Reviews / Australian Association of Clinical Biochemists.***2012**, 33, 2, 49-75.
27. Bell,S. Beginner's Guide to Uncertainty of Measurements. (National Physics laboratory 2013) pp. 1-10.
28. Harding,B.; Tremblay,C.; Cousineau,C. Standard errors: A review and evaluation of standard error estimators using Monte Carlo simulations. *The Quantitaive Methods for Psychology.***2014**, 10, 2, 107-123.
29. Ahn,S.; Fessler,J.A. Standard Errors of Mean, Variance, and Standard Deviation Estimators. 2003 Available from: https://web.eecs.umich.edu/ fessler/papers/files/tr/stderr.pdf, accessed: 15th July 2019
30. ITU-R Rec.P.910 Subjective video quality assessment method for multimedia applications, ITU 2018
31. ITU-R Rec.BT.500 Methodologies for the subjective assessment of the quality of television images, ITU 2019
32. PSN Quality of Service (QoS) testing and reporting. https://www.gov.uk/government/publications/psn-quality-of-service- qos-testing-and -reporting/psn-quality-of-service-qos-testing-and-reportingtest-requirements. Accessed 20 August 2019
33. ITU telecommunication Development Sector. Quality of Service Regulation Manual. Geneva, Switzerland, ITU 2017