

Recommendations for Big Data in Online Video Quality of Experience Assessment

Ethan Court, Kapilan Radhakrishnan, Kemi Ademoye, Stephen Hole

School of Applied Computing (SOAC), University of Wales Trinity Saint David, Swansea, UK
Email: ethan.court@uwtsd.ac.uk, kapilan.radhakrishnan@uwtsd.ac.uk, kemi.ademoye@uwtsd.ac.uk, stephen.hole@uwtsd.ac.uk

Received 24 March 2016; accepted 19 May 2016; published 26 May 2016

Abstract

Real-time video application usage is increasing rapidly. Hence, accurate and efficient assessment of video Quality of Experience (QoE) is a crucial concern for end-users and communication service providers. After considering the relevant literature on QoS, QoE and characteristics of video transmissions, this paper investigates the role of big data in video QoE assessment. The impact of QoS parameters on video QoE are established based on test-bed experiments. Essentially big data is employed as a method to establish a sensible mapping between network QoS parameters and the resulting video QoE. Ultimately, based on the outcome of experiments, recommendations/requirements are made for a Big Data-driven QoE model.

Keywords

Quality of Experience, QoE, Big Data, Online, Video, Traffic

1. Introduction

This paper presents a brief outline of Quality of Experience (QoE) in video traffic and describes how big data can provide a possible solution to the challenges in video QoE assessment. We have carried out an experiment that is used to gain an understanding of how we can apply the enormous amounts of data available to us when video is delivered through the Internet. Using this we place recommendations for the creation of future QoE models, particularly with big data in mind.

Video traffic is forecast to make up 80% - 90% of global consumer video traffic by 2018 [1]. Using this prediction, authors of [2] present various case studies and summarise the following benefits of QoE analysis in video.

- **Identifying, isolating and fixing problems**—Using effective QoE measurement can help end-users to derive if the problems exist in their home network, providers or third party application services. From an operators perspective a complimentary understanding of end-user experience can help identify and fix network issues quicker and also leads to better, more concise notification of affected end-users.
- **Design and planning**—With monitoring end-user experience providers can design and plan their networks in accordance to levels of user expectations and service level agreements. The information gained from QoE

assessment can also be used in proactive activities within design and planning. Expanding on this point, [3] states that quality assessment methods are extremely useful for in-service quality monitoring and management, codec optimization and quality design of networks and terminals.

- **Understanding the quality experienced by customers**—Network operators can gain a better insight into the end-to-end performance experienced by its customers. This allows operators to provide better services to their customers and creates a better understanding for senior managers who make investment decisions.
- **Understanding the impact and operation of new devices and technology**—As new products or technologies are deployed into network infrastructures it is essential that its operational impact can be measured and evaluated. Also quantifying these new implementations can lead to more informed decision making for larger, wider spread rollouts.

2. Quality of Service (QoS)

2.1. Background

Quality of Service (QoS) has been defined by the International Telecommunications Union (ITU) as the, “*Totality of characteristics of a telecommunications service that bear on its ability to satisfy stated and implied needs of the user of the service*” [4]. From this definition it can be stated that QoS is the ability of an Internet service such as email, web browsing or a video conference call to provide the minimum level of quality so that the service can be completed and meet the needs of the end-user.

2.2. How Do We Measure Quality of Service

As mentioned previously QoS is mostly thought to be inferred from network performance indicators. In [5], four parameters identifying the treatment of packets through the IP network are given:

- **Bit Rate**—Also known as through-put or more commonly bandwidth, this defines the total speed capable of data transfer based on an end-to-end link.
- **Delay**—This is the time it takes for a packet to traverse the network or a segment of a network. It is often expressed as latency, which is the time it takes for a data packet to get from one designated point of a network to another [6].
- **Jitter**—This is the full range of packet delay from the maximum amount to the minimum amount [6]. In [7], the Internet Engineering Task Force (IETF) states that “jitter” is used differently by various groups so the term Delay Variance should be used for a more clear and concise understanding.
- **Packet Loss**—Most commonly shown as a percentage. Packet loss refers to the number of packets lost over a period of transmission. Packet loss also tends to be broken down into two forms, burst and random [8].

2.3. Effects of Final Video Output

The previously mentioned parameters can greatly impact end-user video output. End users would often describe four of these undesired effects as the following:

- **Blocking**—Video coding is block based. This means loss of data or perhaps coding errors due to network performance issues will result in this.
- **Blurring**—This can be seen through the loss of spatial information/features, edges around a scene or object tend to become indistinguishable.
- **Edginess**—Specifically referring to the edges in comparison to the original video. Objects within the content has irregular edges.
- **Motion Stutter**—Usually evaluated with real-time against video time via sequence numbering. Often content will freeze or skip segments, relates to FPS.

Figure 1 and **Figure 2** look at the impact of network parameters on final video output. When we compare, a noticeable visual disparity is present. The previously described effects on final video output can be seen *i.e.*, blocking, edginess and blurring around the child’s structure. **Figure 1** has been influenced with 100ms delay and 5% packet loss. In comparison, **Figure 2** which was streamed in optimal conditions with 10ms delay and 0% packet loss and shows an obvious visual superiority in its output. From the previously shown figures it is clear to see the potential visual impacts network conditions can have on final video outputs.



Figure 1. Video: poor output.



Figure 2. Video: optimal conditions.

3. Quality of Experience (QoE)

3.1. Background

The term Quality of Experience has seen increased usage in research, consumerism and industry. The phrase in itself indicates an impact on end-users, meaning how an Internet service is experienced. The ITU define QoE as, “*The overall acceptability of an application or service, as perceived subjectively by the end-user*” [9]. At an initial glance QoE is seemingly seen to heavily overlap with QoS. However, there are various other factors that this term encompasses that QoS does not. As previously identified QoS is concerned with the delivery of services to end-users. QoE on the other hand seeks to evaluate the perceived quality experienced of a service. Take an IPTV stream for example, QoS would seek to ensure the service provides the necessary hardware and/or software capabilities/functions so that the service can be provided to an end-user in the highest possible quality. Quality of experience in contrast, seeks to evaluate the actual level of subjective quality experienced by end-users.

3.2. Subjective QoE

With end-users playing such a key role in the assessment of QoE, subjective testing is a natural progression. Perceived video quality by nature is a subjective area. In order to grasp an actual end-user’s perceived quality the most obvious and simple way to do this is to ask them. As described in [10], subjective testing consists of

firstly, building a panel with real human subjects. These subjects will evaluate a series of videos, usually small sequences that reflect a larger video. Subjects will then give a score based on their notion of quality.

The ITU has standardised subjective testing methods for multimedia application in P.910 [11]. The most commonly used subjective scoring method is Absolute Category Rating (ACR). ACR is a method of judgement where test sequences are presented one at a time and rated independently. It is a scale of 1 - 5 of a viewers judgement of video, 1 being poor, 5 being excellent.

3.3. Discussion

Although subjective testing provides the most accurate indication of user-perceived quality it also has various issues that should be considered. The most commonly associated issues with subjective testing are related to time and man-power [10]. Extensive preparation is required in that they firstly require a video database to be established that relates to the aims and objectives of a test. A panel must then be gathered, educated and carry out the test, all resulting in time lost that could have been allocated elsewhere. Expanding on this point, subjective testing is also usually done in a laboratory environment and is thus restricted in terms of test conditions, video types and viewer demographics [12]. The previously mentioned issues mean that subjective testing is not being viable in real-life scenarios due to time constraints and scalability. Further to this, precedence has been placed on the limited real-time application of subjective video quality assessment. This means that it cannot be applied in monitoring.

3.4. Objective QoE

In order to create reliable QoE prediction, as well as eliminating the negative aspects of subjective testing, objective QoE models are used. Objective models are computational based but still retain the primary goal of predicting perceived end-user video quality. Authors in [12] and [13] have previously categorised methods.

The most commonly used reference classification approach is seen in layer **Figure 3**. Classifications here are based on if the model requires an original source video. They are as follows:

- Full Reference (FR): Full access to original unaltered source video sequence.
- Reduced Reference (RR): Partial video information is required. Usually destination output.
- No Reference (NR): No access to original source video is required.

As both FR and RR models require access to video output they are usually based around a comparison approach where the original sequence is compare against a processed video sequence. Due to this they are often considered intrusive methods [14], meaning that they have an impact on end-user services. Conversely, NR models are associated with non-intrusive testing meaning they have very limited impact on end-user outputs. NR models are classified when no access to original video source is present/required.

Figure 3 is a small extract of the model of classification we have created. The model extends to layer 2 and 3 where we identify the approach and data used in models. From this we gathered that essentially all data can be used in objective QoE estimation models. This is where the initial interest in applying big-data to QoE prediction began to take prevalence.

3.5. Discussion

The previous section provided a brief overview of QoE models. Authors of [12] present some challenges directly relating to the categorised assessment methods are as follows:

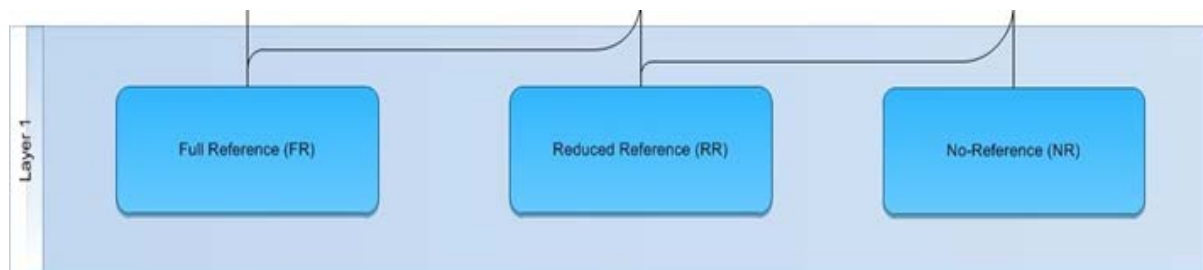


Figure 3. Extract of our overview of objective model classifications.

- 1) FR Models cannot to be implemented in real-time due to complexity and a full-reference sequence being needed.
- 2) RR models, although not needing an original sequence still requires resources such as side channels to extract information of the video sequence.
- 3) Models based on subjective testing and the Human Visual System (HVS), although accurate require extensive preparation and validation and are often very complex.
- 4) In contrast, NR/engineering approaches have a lower complexity, but have reduced accuracy and are only accurate to specific data sets.
- 5) Evaluation of models are usually based on the data sets they have been created on. Adding to this the subjective tests they are based on often are only specific to certain criteria, ie viewer demographic, viewing time, etc.

3.6. Big Data-Driven QoE

Increased delivery of video over the Internet has created a surplus of data available for analysis. Improving end-user QoE has also become a crucial aspect of service agreements. So, with the utilization of this surplus of data it will provided an added monetization incentive that can give increased benefit for both deliverer and receiver.

Data available involves aspects of the viewer. Metrics have been defined as:

Viewer-Session Metrics (VSMs):

- **Viewing time per view**—time user watches video, expressed in a ratio to full video time.
- **Abandoned view ratio**—percentage of views that are abandoned compared to those initiated, expressed as a percentage.

Viewer-Level Metric (VLMs):

- **Number of views**—number of views a certain video has, at a current time if real-time.
- **Viewing time per visit**—ratio viewing time compared to initiated time.
- **Return/refresh rate**—as an indication of viewer frustration of reduced QoE.
- **Video rating**—User rating at end of transmission.

Aside from the viewer metrics available, data-driven QoE has focused more on the use of extensive QoS metric available, these include:

- **Startup delay**—the time between a user request and initiation of a video.
- **Re-buffering**—how long a video stream is paused to ensure content is delivered, otherwise known as stuttering. How often and for how long are considered.
- **Average bitrate**—How fast the video content is displayed on screen. Dependent on video encoding/decoding, network and possible hardware statistics.
- **Previous QoS metrics**—As discussed in Section 2.2.

Video QoE analysis has seemingly come full circle with the use of network QoS to QoE mapping again taking priority so that a real-time and real-world analysis can take place. The priority now is accuracy and efficiency.

4. Experiment/Results

An experiment was carried out in order to confirm the influence of QoS on end-user QoE and also gain an idea of the challenges and process of applying big data to QoE assessment with the data that is available to us. We utilize the easy to influence and monitor QoS metrics as stated in Section 2.2. Specifically we influence video output with delay and packet loss. In total 6 videos ranging from 10 - 12 seconds were streamed in various conditions. The videos ranged from 176×144 to 640×480 resolution. QoE score was then evaluated using a FR objective method described in Section 3.4. The scale used for video QoE is 0 - 50, translating to 0 - 5 from an end-user subjective standpoint as described in Section 3.2.

We mapped results against increasingly degraded network conditions as seen in **Figures 4-7**. Overall the outcome is as expected whereas when the network QoS conditions deteriorate, we see a decrease in end-user QoE scores. Notable, delay and packet loss as a single affecting network QoS parameters see very similar outputs in end-user QoE. Combining them sees an increased impact on video QoE where the worst condition we tested averages an overall QoE rating of 5.2. Something to consider is the initial streaming implication where coding, compression and decompression have an important impact on video QoE scores.

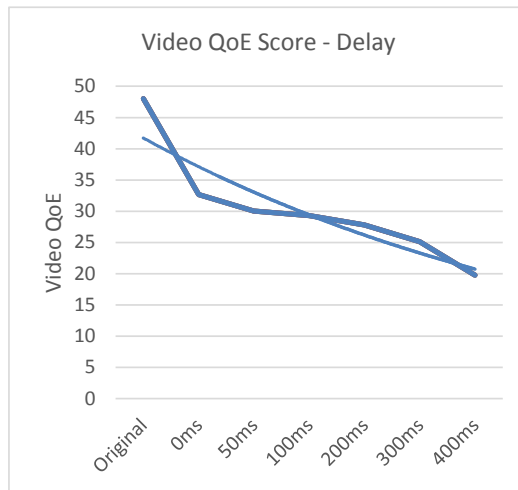


Figure 4. Delay QoE influence.

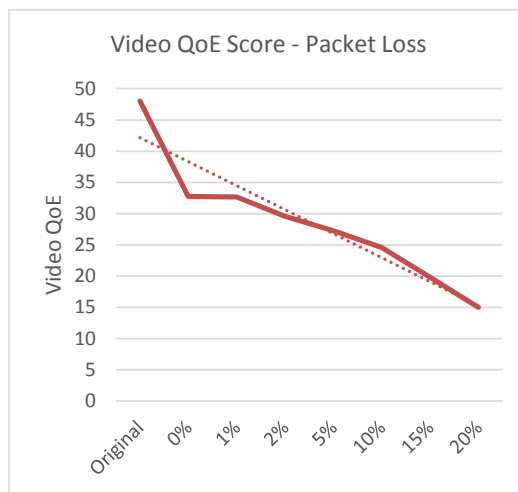


Figure 5. Packet loss QoE influence.

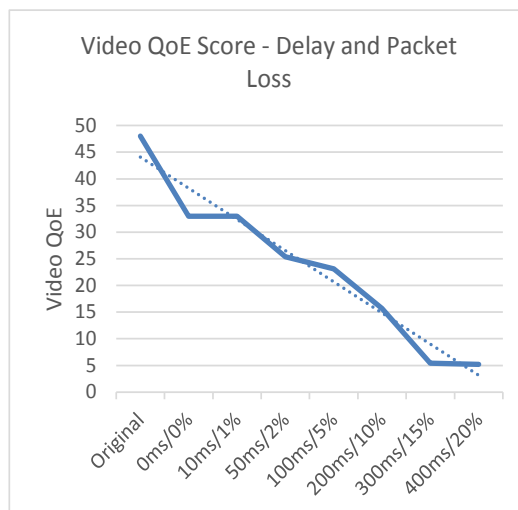


Figure 6. Delay and packet loss QoE influence.

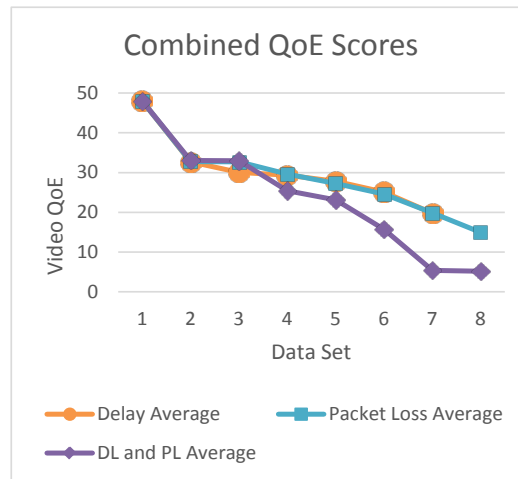


Figure 7. Combined QoE influence.

5. Recommendations for Big Data-Driven QoE Model

With an insight to QoE models gained we follow the ideas set out by works [12] and add the experience gained through the experiment described previously to establish core recommendations for a big data-driven QoE assessment model.

- Requirements for a QoE Metric:
 - **Quantifiable**—Easily viewed and quantified in real-terms.
 - **Accurate**—The output should be an accurate representation of end-user QoE.
 - **Informative**—It has some real-world use to industry and is indicative of what it is representing.
 - **Fit for Purpose**—Question the purpose of the output, what/who does it serve and is it meeting the specific needs.
- Requirements for a QoS to QoE Mapping Model based on Big Data:
 - **Consistent**—Is the output consistent with the input into the model and expected results.
 - **Expressive**—Is the relationship between QoS and QoE shown appropriately and accurately.
 - **Real-time**—Video is inherently real-time, the solution should translate to this.
 - **Scalable**—The Internet is ever growing, the model should be adaptable and translate to this fact.
 - **Correct Flagging**—When is a QoE result considered an issue? This should be accounted for.
 - **Simplicity**—The relationship of QoS to QoE can become very complex, it should be kept as simple as possible, whilst retaining accuracy.

6. Conclusion

The main goal of the paper was to determine how big data can be used in achieving the goal of accurately assessing end-user perceived quality, without the usual negative drawbacks. The experiment uses accessible QoS parameters to gain understanding of applying what data is available if a big-data QoE model is created. With insight from literature and the experience gained in the experiment, we achieve the goal of the paper by providing core recommendations for a big-data driven QoE model. The scope of experiment findings were limited as we only include two parameters in testing, but the progress gained still provides a very effective foundation as the recommendations placed can be followed when advancing a new big data-driven QoE model. Future work will entail increasing the parameters used, extending testing to higher resolution videos and adapting the QoE output to predict end-user quality with input of various real-time obtainable parameters.

References

- [1] Cisco (2014) White Paper: Cisco Visual Networking Index: Forecast and Methodology, 2013-2018. Technical Report, June 2014.
- [2] Linsner, M., Eardley, P., Burbridge, T. and Sorensen, F. (2015) Large-Scale Broadband Measurement Use Cases

draft-ietf-lmap-use-cases-06. ONLINE, February.

- [3] Takashi, A., Hands, D. and Barriac, V. (2008) Standardization Activities in the ITU for a QoE Assessment of IPTV. *IEEE Communications Magazine*, **46**.
- [4] ITU-T (2009) Recommendation E.800: Definitions of Terms Related to Quality of Service. ONLINE, April.
- [5] Gozdecki, J., Jajszczyk, A. and Stankiewicz, R. (2003) Quality of Service Terminology in IP Networks. *IEEE Communications Magazine*, **41**, 153-159. <http://dx.doi.org/10.1109/MCOM.2003.1186560>
- [6] ITU-T (2011) Recommendation G.1050: Network Model for Evaluating Multimedia Transmission Performance over Internet Protocol. ONLINE, March 2011.
- [7] IETF (2002) IP Packet Delay Variation Metric for IP Performance Metrics (RFC 3393). ONLINE, November 2002.
- [8] You, F.H., Zhang, W. and Xiao, J. (2009) Packet Loss Pattern and Parametric Video Quality Model for IPTV. *Eighth IEEE/ACIS International Conference on Computer and Information Science*, June 2009, 824-828. <http://dx.doi.org/10.1109/icis.2009.24>
- [9] ITU-T (2007) Recommendation P.10/G.100: Vocabulary for Performance and Quality of Service. ONLINE, July 2007.
- [10] Cancela, H., Rodriguez-Bocca, P. and Rubino, G. (2007) Perceptual Quality in P2P Multi-Source Video. *IEEE Global Telecommunications Conference*, Washington DC, 26-30 November 2007, 2780-2785.
- [11] ITU-T (2008) Recommendation P.910: Subjective Video Quality Assessment Methods for Multimedia Applications. ONLINE, April 2008.
- [12] Chen, Y.J., Wu, K.S. and Zhang, Q. (2015) From QoS to QoE: A Tutorial on Video Quality Assessment. *IEEE Communications Surveys Tutorials*, **17**, 1126-1165. <http://dx.doi.org/10.1109/COMST.2014.2363139>
- [13] Chikkerur, S., Sundaram, V., Reisslein, M. and Karam, L.J. (2011) Objective Video Quality Assessment Methods: A Classification, Review, and Performance Comparison. *IEEE Transactions on Broadcasting*, **57**, 165-182. <http://dx.doi.org/10.1109/TBC.2011.2104671>
- [14] Cole, R.G. and Rosenbluth, J.H. (2001) Voice over IP Performance Monitoring. *SIGCOMM Comput. Commun. Rev.*, **31**, 9-24. <http://dx.doi.org/10.1145/505666.505669>