QoE Metrics Estimation for Video Conferencing using Passive Network Measurements

Tarun Mangla, Junchen Jiang, Nick Feamster, Arpit Gupta

Abstract

As users continue to depend on video conferencing applications (VCAs) for remote participation in work, education, healthcare, and recreation, ensuring a high quality of experience (QoE) when using VCAs is critical. Although QoE depends to some degree on the specific circumstances of end users, network operators can often play important role in mitigating QoE degradation resulting from poor local network conditions. A network operator who can observe a VCA's QoE metrics may be able to diagnose and react to QoE degradation, potentially preventing even transient congestion events from affecting user experience. Unfortunately, network operators lack direct access to application QoE, and must infer QoE from the encrypted application traffic as it traverses the network.

Therefore, we consider the question *if and how we can enable network operators to infer QoE*[1] *for VCAs at per-second time granularity, from passive measurements of encrypted network traffic.* Recent work has proposed data-driven techniques, often based on machine learning, to estimate VCA QoE metrics from network-layer metrics [2, 6, 8]. Yet, most existing techniques produce coarse-grained inference, such as average frame rate or mean opinion score of a video session (VCA call). Therefore, we study whether it is possible to estimate these metrics *at a one second timescale*, and ultimately to detect transient QoE degradation events that would be otherwise undetected in coarser-grained QoE metrics. This fine-grained visibility enables network operators to perform better post-hoc analysis and network intervention.

Past work also assumes operators can access and parse application-level headers which may not be practical in many cases. Certain VCAs such as Zoom use custom application protocols, thus making it challenging to extract any information using standard network monitors [5]. Moreover, given trends in traffic encryption, all application headers may eventually be encrypted [9]. Thus, *we also attempt to estimate video QoE using only IP and UDP headers*, and compare the accuracy of the models that use only these headers to those that also rely on application headers. An additional advantage of this is that existing network monitoring systems can extract such information at scale [7].

With these design goals, we develop a QoE inference algorithm that leverages the semantics of video delivery in VCA network protocols. A key technical insight in this research is that due to VCA's real-time nature, each video frame is encoded and its packets are sent as a group—sometimes even in a microburst [3]. As a result, it is possible to group packets by frame and thus estimate key QoE metrics at a fine time granularity. We demonstrate this capability using the application-layers headers of unencrypted traffic as well as using only IP/UDP headers for encrypted traffic.

We conduct an in-lab evaluation of our inference technique using data collected from two popular VCAs, Meet and Teams. We assume a 2-person call setting and collect network traces and ground truth data for such setting under diverse emulated network conditions. Our evaluation demonstrates that we can estimate key QoE metrics with high accuracy.

Future Work: Our work provides a promising direction to infer fine-grained VCA QoE using passive network measurements, that too only IP and UDP headers. There are key challenges/questions that still remain, including:

- Our preliminary evaluation is under controlled settings and over only two VCAs. The inference techniques need to be evaluated under real user setting. This includes consideration of different usage modality (e.g., screen sharing, multi-party call) and multiple VCAs.
- Inference techniques need to adapt to changes over time in networks and VCA implementations. Thus, a framework is needed for continual validation and calibration of inference techniques.
- It is important to develop scalable QoE monitoring solutions that can provide QoE estimates in real time. Such solutions require innovation in developing efficient network monitoring systems as well as developing lightweight inference techniques.

[1] Although QoE is subjective by definition [4], our focus in this paper is on inferring objective application metrics such as frame rate and frame jitter. Existing work provides methods to model user experience using these application metrics [1]. In addition, although VCA performance is determined by both audio and video, in this paper we study only video. This is in accordance with previous research in this area [2, 6].

References

[1] Amin Banitalebi-Dehkordi, Mahsa T Pourazad, and Panos Nasiopoulos. “The effect of frame rate on 3D video quality and bitrate”. In: *3D Research* (2015).

[2] Giovanna Carofiglio, Giulio Grassi, Enrico Loparco, Luca Muscariello, Michele Papalini, and Jacques Samain. “Characterizing the Relationship Between Application QoE and Network QoS for Real-Time Services”. In: *Proceedings of the ACM SIGCOMM 2021 Workshop on Network-Application Integration*. 2021.

[3] Sadjad Fouladi, John Emmons, Emre Orbay, Catherine Wu, Riad S. Wahby, and Keith Winstein. “Salsify: Low-Latency Network Video through Tighter Integration between a Video Codec and a Transport Protocol”. In: *15th USENIX Symposium on Networked Systems Design and Implementation (NSDI 18)*. USENIX Association, 2018.

[4] Tobias Hossfeld, Poul E Heegaard, Martin Varela, and Sebastian Möller. “Formal definition of QoE metrics”. In: *arXiv preprint arXiv:1607.00321* (2016).

[5] Citizen Lab. *Zoom uses a bespoke extension of RTP*. `https://citizenlab.ca/2020/04/move-fast-roll-your-own-crypto-a-quick-look-at-the-confidentiality-of-zoom-meetings/`. 2020.

[6] Ashkan Nikravesh, David Ke Hong, Qi Alfred Chen, Harsha V Madhyastha, and Z Morley Mao. “QoE inference without application control”. In: *Proceedings of the 2016 workshop on QoE-based Analysis and Management of Data Communication Networks*. 2016.

[7] John Sonchack, Oliver Michel, Adam J Aviv, Eric Keller, and Jonathan M Smith. “Scaling Hardware Accelerated Network Monitoring to Concurrent and Dynamic Queries With {* Flow}”. In: *2018 USENIX Annual Technical Conference (USENIX ATC 18)*. 2018, pp. 823–835.

[8] Suying Yan, Yuchun Guo, Yishuai Chen, Feng Xie, Chenguang Yu, and Yong Liu. “Enabling QoE learning and prediction of WebRTC video communication in WiFi networks”. In: *Proceedings of the ICC*. 2017.

[9] Ming-hour Yang. “Security analysis of application layer protocols on wireless local area networks”. In: *Journal of Shanghai Jiaotong University (Science)* (2011).