Paper Summary
Estimating WebRTC Video QoE Metrics Without Using Application Headers

Sarthak | 2020CS10379

June 2024

Problem Statement

The paper addresses the challenge of estimating Quality of Experience (QoE) metrics for video conferencing applications (VCAs) without relying on application headers, which are often inaccessible due to privacy and security concerns

Motivation

- VCAs must keep a short jitter buffer and thus are susceptible to a wide range of incidents that can disrupt or degrade network quality
- Existing VCA QoE estimation methods use passive measurements of application-level Real-time Transport Protocol (RTP) headers
- Cannot always parse application level headers, especially when encrypted, routed over VPN, uses complex encapsulation

Key Idea

To use Heuristics or Machine Learning with a combination of flow statistics (e.g., throughput) and features derived based on the mechanisms used by the VCAs to identify frame boundaries. Then, obtain a sequence of frames along with their sizes and use this information to estimate the key QoE metrics (Video bitrate, Frame rate, Framejitter, resolution etc.) over a window W.

Framework & Methodology

Media Classification

As voice samples can be encoded in fewer bits than images, hence audio packets are smaller than video ones. Thus, Any packet with size greater than or equal to

V_{min} is tagged as a video packet, while the remaining packets are not considered

Inference

Two approaches were developed to infer QoE metrics :

Heuristic Based

- Patterns in inter-arrival times do not represent patterns in inter-departure times
- Packet sizes tend to resemble those within same frame and differ from those in consecutive frames
- Use a packet size difference threshold Δ^{max}_{size} and declare frame boundary if the size difference between N^{max} consecutive packets is greater than Δ^{max}_{size}. N^{max} and Δ^{max}_{size} being VCA specific parameters

IP/UDP ML

- VCA semantics-based features include number of unique packet sizes observed and number of microbursts of packets in the prediction window W
- Flow-level statistics include number of bytes and packets per second as well as five statistics (mean, var, min, max, median) on packet sizes and inter-arrival times
- The paper experimented with three classical supervised ML models :
  - Support Vector Machines (SVMs)
  - Decision Trees
  - Random Forests: Selected for the final implementation as they yielded highest accuracy

Contributions

- Developed a machine learning-based method to fragment a frame into packets and infer VCA QoE metrics at finer time granularity using only the IP/UDP headers
- Developed an automated browser-based, VCA data collection framework and used it to collect data under controlled in-lab network conditions as well as from 15 households
- Characterized the network conditions under which the model has high errors and the potential reasons for errors

Strengths

- Uses only IP/UDP headers, not requiring hard to obtain/encrypted application level header
- The predictive model trained on data from controlled lab settings transfered easily to real-world networks
- The proposed method shows high accuracy, with frame rate estimates within 1.50 frames of the ground truth QoE

Weaknesses & Limitations

- Heuristic based approach breaks under certain conditions e.g. under high latency jitter or packet loss
- Methodology is primarily evaluated on WebRTC-based VCAs, need to generalize to other VCAs
- Need to develop methodology for various application modes, such as multiparty conferencing and screen sharing