Real Time Video QoE Analysis of RTMP Streams

Holly French, Jie Lin, Tung Phan, Amy Csizmar Dalal
Department of Computer Science, Carleton College, Northfield, MN, USA
Email: {frenchh, linji, phant, adalal}@carleton.edu

***Abstract*—We aim to develop self-healing networks that can detect degradation of streaming video quality of experience (QoE), react, and correct the pathology on the network. We present an architecture to assess real time video QoE of RTMP streams. Results from a small set of preliminary experiments demonstrate that we can predict video QoE with 70-80% accuracy based on stream state measurements and previous users' ratings, using at little as 20 seconds of stream state information.** [1]

***Index Terms*—Video streaming, quality of experience (QoE), quality of service (QoS), multimedia applications, measurement, subjective quality**

I. Introduction

Self-healing networks detect existing or potential pathologies and fix or mitigate them with minimal to no human intervention. Because such pathologies are visibly evident in the end-users' quality of experience (QoE), self-healing networks are by definition responsive to the needs of networked applications, such as streaming video. By understanding how video applications are affected by network conditions, we can design network protocols and structure new networks to better support these and future (high quality, high-definition, immersive) video applications. Similarly, understanding how network conditions impact application performance leads to the design of more robust application protocols that can better utilize existing network resources.

Several studies have attempted to make explicit connections between video stream state information, such as bandwidth and frame rate, and end user QoE, most notably [1]–[3]. In previous work, we have demonstrated an explicit tie between stream state data (retransmitted packets, for UDP streams, and a combination of bandwidth and frame rate, for TCP streams) and user QoE, using data mining techniques to infer QoE ratings solely from the stream state data [4], [5]. We now study the tie between stream state data for web-based, RTMP video streams and end user QoE, as web-based video streams make up the majority of video streams.

We describe the major modifications we made to our existing stream quality assessment system [5] to enable us to infer end user QoE of RTMP video streams. We present our modified Flash player and data collection infrastructure, and report the results of preliminary experiments run with a small set of participants (N = 10). Our findings indicate bitrate in combination with either frame rate or bandwidth serves as an accurate indicator of QoE for RTMP videos.

[1]This work is sponsored by grants from the Clare Boothe Luce Foundation and the Howard Hughes Medical Foundation.

II. System architecture

Figure 1 illustrates the stream quality assessment system architecture. The front end consists of a web page with an embedded Flash player and a ratings slider bar that allows a user to rate video quality on a continuous scale from 0 (lowest quality) to 100 (highest quality). The embedded Flash player is an ActionScript plugin that interfaces with the Flash video player Flowplayer. The plugin polls Flowplayer every second for current, player-reported statistics on video playback: bandwidth, frame rate, and bitrate. The video playback data, and the current position of the ratings slider bar, are logged to a server-side database every second. Currently, this data is analyzed off-line, but the system supports real-time analysis of this data as well.

The backend consists of a Flash server, which serves up the media content; a web server, which hosts the web page and the plugin; and an analysis server, which hosts the database and the stream quality analyzer. The Flash server resides on a controlled network, behind a router running netem [6], allowing us to introduce specific loss and delays onto the network between the Flash server and the viewing clients. The stream quality analyzer and the web server reside outside of the controlled network, so that they are not affected by the introduced losses and delays. The stream quality analyzer, described in [5], uses data mining techniques to infer video QoE solely from stream state data. The algorithm uses k-nearest neighbors with dynamic time warping as a distance metric to compare the stream state data of the current stream to be analyzed with all of the streams previously analyzed. To assign a rating to the current stream, we calculate the median of the ratings of the k nearest streams.

III. Experiments

Our initial data set consists of stream state data and video quality ratings collected from ten participants during July 2011. Table I describes the four source videos for this experiment. The "trailer" video was used to norm our participants' expectations of video quality. Each participant watched on average ten video streams[2], giving us data for approximately 100 video streams. As participants watched the streams, they moved the ratings slider bar whenever they perceived a change in video quality.

We introduced a random amount of packet loss and delay at the netem router, affecting packets traveling from the Flash

[2]A stream is one viewing by one participant from start to finish of one video

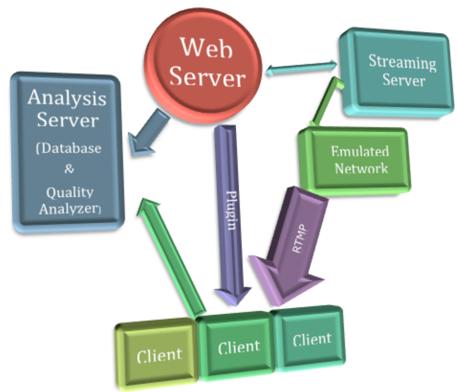

Fig. 1. Video stream quality assessment architecture

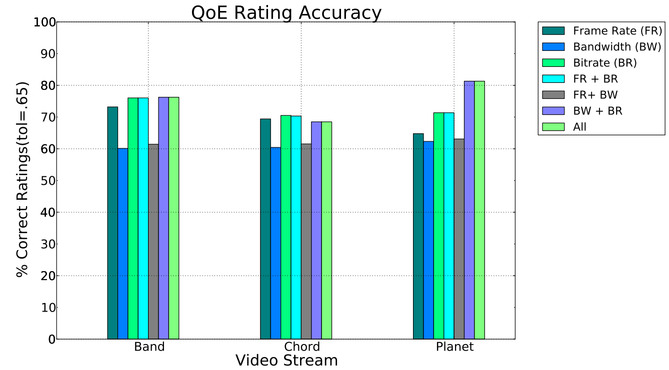

Fig. 2. The percentage of correctly-assigned ratings, per 20-second stream subset, for different combinations of stream state data.

TABLE I
CHARACTERISTICS OF THE SOURCE VIDEOS

Name	Duration (mm:ss)	Description	High/Standard Definition	Bit rate (kbps)
planet	4:42	low action nature scenes	high	2584
chord	5:43	medium action music performance	standard	664
band	4:33	high action animated short	standard	720
trailer	1:57	high action movie trailer	standard	570

server to the clients. Packet loss varied from 0 to 15%, and packet delay varied from 0ms to 100ms. The participants were unaware which streams were affected by packet loss and/or delay.

Because user ratings reflect individual preferences and experiences, we normalize the slider bar ratings using the z-score, $z_s = \frac{r_s - \bar{r}}{\sigma_r}$, where r_s is the user's quality rating for stream subset s, $\bar{r}$ is the average of the user's quality ratings on all stream subsets, and σ_r is the standard deviation of the user's quality ratings. The stream quality analyzer assigned ratings based on this normalized scale. We compared the ratings assigned by the analyzer to the actual ratings assigned by our participants. If the assigned rating fell within 0.65 of the actual rating (approximately equivalent to one perceptible quality degradation level, determined experimentally), we counted the assigned rating as correct.

To mimic real-time QoE assessment, we divided video streams into smaller stream subsets ranging from 5 to 50 seconds, and had the analyzer assign a rating for each subset (corresponding to the average rating over that subset). We found that changing the subset size produced generally similar results; we present the results for the 20-second subsets here.

IV. RESULTS AND FUTURE WORK

Figure 2 highlights the accuracy with which the stream quality analyzer assigns ratings to the video stream subsets using all videos as the training set. We achieve correct ratings over 80% of the time with the high definition video when using either bandwidth and bitrate together or in combination with frame rate. We achieve over 70% accuracy in assigning QoE ratings to the standard definition videos using bitrate alone or in combination with frame rate. There are several combinations of stream state measurements that accurately indicate video QoE on short time scales, giving us a lot of flexibility in choosing which stream state data to use in designing a stream quality assessment system. Further, the fact that we can assign accurate QoE ratings over 70% of the time with just 20 seconds of stream state information demonstrates that the system we present here can report on stream quality degradations, as they are reflected in QoE, quickly, and possibly use this information to mitigate them before they worsen.

Our goal is to develop self-healing networks that can detect degradation of streaming video quality, react, and correct the pathology on the network. We can predict video QoE with 70-80% accuracy based on stream state measurements and previous users' ratings, on short subsets of the source videos. Future work includes conducting a larger set of experiments, developing heuristics to map detected stream quality degradations from the predicted QoE to the network events that cause them, and developing strategies to mitigate future stream quality degradation before such degradation is apparent to the end user.

REFERENCES

[1] Y. Wang and M. Claypool, "RealTracer - tools for measuring the performance of RealVideo on the internet," *Kluwer Multimedia Tools and Applications*, vol. 27, no. 3, December 2005.

[2] M. Li, F. Li, M. Claypool, and R. Kinicki, "Weather forecasting - predicting performance for streaming video over wireless LANs," in *Proceeedings of NOSSDAV*, Stevenson, Washington, June 2005.

[3] P. Calyam, M. Sridharan, W. Mandrawa, and P. Schopis, "Performance measurement and analysis of H.323 traffic," in *Proc. PAM 2004*, Antibes Juan-les-Pins, France, April 2004.

[4] A. Csizmar Dalal, A. Bouchard, S. Cantor, Y. Guo, and A. Johnson, "Predicting video stream quality in real time using objective stream state measurements," Carleton College, Tech. Rep., October 2010.

[5] A. Csizmar Dalal, E. Kawaler, and S. Tucker, "Towards real-time stream quality prediction: Predicting video stream quality from partial stream information," in *Proc QShine*, Las Palmas de Gran Canaria, Spain, November 2009.

[6] Linux Foundation, "netem," http://www.linuxfoundation.org/collaborate/workgroups/networking/netem.