Informational F.Wang
Internet Draft Beijing Institute of Technology
Intended status: Informational Z.Fei
Expires: Nov 26, 2015 Beijing Institute of Technology
May 26, 2015

# QoE Evaluation for HTTP Adaptive Streaming
# draft-wang-tsvwg-qoe-evaluation-has-01.txt

Abstract

This document describes a method to evaluate the Quality of Experience (QoE) of real-time video delivered over HTTP Adaptive Streaming (HAS) technology. Not only the end points but also the content providers and network operators can acquire the QoE of HAS by implementing this method.

Table of Contents



## 1. Introduction

HTTP Adaptive Streaming (HAS) has been used widely to deliver real-time video streams, including Apple's HTTP Live Streaming (HLS), Microsoft Smooth Streaming (MSS), Adobe's HTTP Dynamic Streaming (HDS), and 3GPP's standardized solution 3GP-DASH [DASH].

Most HAS protocols have defined Manifest File, i.e. Media Presentation Description (MPD) of 3GP-DASH and Playlist file of HLS, which lists the location of various chunks, as well as some other informational tags set such as the way the content has been chunked. Each chunk is specified by its address and associated informational tags set. There will be no change in quality for a single chunk after it has been delivered due to the reliability of HTTP, thus the quality of each chunk can be measured at the source point before multimedia delivering.

Based on the above analysis, the quality of each chunk can be as a tag and embedded into the associated informational tags set of Manifest Files. Then the overall QoE for the whole video can be obtained by a pooling model which takes the quality of each received chunk into consideration.

### 1.1. Method of evaluate QoE for HAS

Since, for HAS all Manifest files must be download before the multimedia data playing, by capturing and parsing these files, quality of each chunk can be get easily. For each playing chunk, its corresponding information can be acquired from the HTTP request information, because a URL that included in the HTTP request information, corresponds to a specific presentation.

By the above analysis, the quality of a playing chunk can be obtained in real time at any point, and the quality for a certain time can be predicted through pooling methods which take all the segments within the prediction period into consideration. The simplest pooling method is averaging over all played chunks. In this draft, we have proposed

a liner model which has been demonstrated has a higher accuracy in [Pooling method].

## 2. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

In this document, these words will appear with that interpretation only when in ALL CAPS. Lower case uses of these words are not to be interpreted as carrying RFC-2119 significance.

A list of acronyms and abbreviations used in this document are presented below.

- HAS: HTTP Adaptive Streaming
- HLS: HTTP Live Streaming
- DASH: Dynamic Adaptive Streaming over HTTP
- QoE: Quality of Experience
- MOS: Mean of Score
- PSNR: Peak Signal to Noise Ratio
- dPSNR: differential PSNR

## 3. Chunk-quality metrics

### 3.1. The analysis of Quality metrics

The quality of each chunk can be measured by subjective or objective methods. It is recognized that subjective methods are time and

manpower consuming, which makes these methods hard to utilize in practice. The limitation of subjective methods drives the development of objective methods. Usually the objective methods can be divided into three categories according to the dependency of original video: 1) full-reference methods (FR, in which need complete original content), 2) reduced-reference methods (RR, in which need partial information about original content) 3) no-reference methods (NR, in which need nothing about the original content).

FR methods are more simple and accurate than RR and NR methods but not practical for most delivery scenarios due to their dependence of the complete original content. But for HAS delivery scenario, FR methods can be used, since the quality of each segment can be

measured at the source point that with original content available.

### 3.2. Metric chosen for QoE evaluation

Peak Signal to Noise Ratio (PSNR), Structural Similarity (SSIM) and Video Quality Metric (VQM) are three commonly used FR objective metrics for quality analysis. But these methods are sensitive to video content and often used as a measurement metric for different compression versions with a same given original content. The sensitivity of these methods are caused by the different spatial-temporal characteristics in different video content, which will lead to different PSNR range for different content although they have the same QoE range.

In order to solve the above problem, this draft present the concept of differential PSNR (dPSNR) based on PSNR. For each segment, a suitable presentation is chosen as the benchmark, then the dPSNR of all presentations of this segment are obtained by subtracting the PSNR value of the benchmark.

In order to solve the above problem, this draft present the concept of differential PSNR (dPSNR) based on PSNR. For each segment, a suitable presentation is chosen as the benchmark, then the dPSNR of all presentations of this segment are obtained by subtracting the PSNR value of the benchmark.

E.g. if different contents at a content provider are encoded with the same bitrate levels set, for each segment, we can choose the presentation with the highest bitrate level as the benchmark.

### 3.3. The chunk-quality tag

Manifest files contain the URLs and other informational tags for various chunks. In this draft, the following attributes are defined and added to the basic Manifest files:

QUALITY TYPE

The value is the type of quality metric has been used.

VALUE

It denotes the value of quality under the given quality metric which can be get from the QUALITY TYPE attribute.

DESCRIPTION

It denotes some other information for presentation-quality, e.g. if we have chosen dPSNR as the presentation-quality metric, it can describe whether a presentation is chosen as a benchmark, if QUALITY DESCRIPTION=1 the presentation is a benchmark, otherwise not.

In this draft, dPSNR has been used as the presentation-quality metric. Thus the QUALITY TYPE: dPSNR. The dPSNR value of each

multimedia presentation can be calculated at the content provider point.

A simple example of media play list file for HLS after we define and embed the chunk-quality tag #EXT-X-QoE:

```
#EXTM3U
#EXT-X-VERSION: 3
#EXT-X-TARGETDURATION: 8
#EXT-X-MEDIA-SEQUENCE: 2680
#EXT-X-QoE: QUALITY TYPE=dPSNR, VALUE=20, DESCRIPTION=0
#EXTINF: 7.975,
https://priv.example.com/fileSequence2680.ts
#EXT-X-QoE: QUALITY TYPE=dPSNR, VALUE=18, DESCRIPTION=0
#EXTINF: 7.941,
https://priv.example.com/fileSequence2681.ts
#EXT-X-QoE: QUALITY TYPE=dPSNR, VALUE=15, DESCRIPTION=0
#EXTINF: 7.975,
https://priv.example.com/fileSequence2682.ts
```

## 4. Pooling method

Since many tests have proved that for a certain time, the quality perceived by the end users not only depends on mean video quality, but also depends much on the segments with the best or worst quality that may leave a deep impression to end users, and the occurrence of quality switching that will distract viewers' attention. The liner model is proposed based on the above analysis, which take more influenced factors into consideration.

A simple example to evaluate quality for a certain period is shown as follows,

PMOS=a*mean+b*max+c*min+d*std,

where a,b,c,d are parameters associated with content type, encoded type and prediction period, mean, max, min, std are calculated influenced factors, which measure the mean quality, maximum quality, minimum quality and standard deviation quality respectively.

## 5. Security Considerations

Since the protocol relies on HTTP Live Streaming, most of the same security considerations apply. See section 11 of draft-pantos-httplive-streaming-13.

## 6. IANA Considerations

Same IANA considerations of HTTP Live Streaming apply. See section 10 of draft-pantos-http-live-streaming-13.

# 7. References

## 7.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2616] Fielding, R., Gettys, J., Mogul, J., Frystyk, H.,Masinter, L., Leach, P. and T. Berners-Lee, "Hypertext Transfer Protocol -- HTTP/1.1", RFC 2616, June 1999.

## 7.2. Informative References

[DASH] 3GPP TS 26.247 v12.2.0, "Progressive Download and Dynamic Adaptive Streaming over HTTP (3GP-DASH)", Release 12, mar 2014

[Pooling method] Xiaolin Deng, Liang Chen, Fei Wang "A Novel Strategy to Evaluate QoE for Video Service Delivered over HTTP Adaptive Streaming", unpublished

Authors' Addresses

Fei Wang
Beijing Institute of Technology
5 South Zhongguancun Street, Haidian District, Beijing, China

Email: fei_wang@bit.edu.cn

Zesong Fei
Beijing Institute of Technology
5 South Zhongguancun Street, Haidian District, Beijing, China

Email: feizesong@bit.edu.cn