Video stalling detection method

Shushi Yu* and Na Wang
College of Information Engineering, Communication Engineering, Shenzhen University, China

*Email: 739536243@qq.com

Abstract. ISP providing network access services has a huge user group; therefore the competition between the ISP providers is extremely fierce. Under the situation that video occupies the most of network flow, providing high satisfaction network services for video users can effectively improve the competitiveness of ISP, in which video Stalling plays an important index. For protecting the privacy of users, most video websites use HTTPS as the application layer protocol for video transmission, so ISP could not get the playback of video users by parsing HTTP packets, and could not directly detect the video users. In this assay, a new method, protocol transport layer Video Stalling Detection Method based on HTTPS is proposed. This method can effectively match the YouTube, Youku, Letv, Tencent video and other major video sources of the stalling event.

1. Introduction

The main medium of internet communication in 3G era was text, and 4G era was pictures and small videos. With the growing of 5G technology which proposed faster speed of the internet and proposed the release of the internet flow, video will become the mainstream media in the future. It can be predicted that video flow will account for a larger share in near future, with data reporting that the proportion of video flow will rise from 60% in 2016 to 78% in 2021 [1].

As a provider of Internet services, ISP has always been highly competitive [2]. The arrival of 5G will be an opportunity for them to re-divide the market, the one who provide a better user experience who will be able to have more users, so it is particularly important to improve the competitiveness of ISP. As video is the main body of network traffic, the user group behind it is huge. For these users, faster network speed means less tolerance of video stalling, therefore improving the watching experience of video users is an effective means to improve the competitiveness of the ISP industry. Video stalling detection becomes extremely important at the meanwhile. In reference [3], the author estimate the time and frequency of the stalling time (QOS metric) by analysis the packets captured in the transport layer, and modelling the download time, the throughput of network packets and the video buffer. The modelling accuracy of video buffer is the highest; its main method is to compare the playback times of video frames with the time stamps of received packets by parsing data packets. On this basis, reference [4] adds some additional parameters, including video playback time, playback delay and so on, to improve the estimation accuracy and improve the model. In reference [5], the author makes statistical analysis on average packet loss, loss period, loss distance, RTT and other user satisfaction (QOS) metrics by using machine learning algorithms such as support vector machine (SVM), neural network (KNN) and Logistic regression algorithm, and eventually realized the classification of user satisfaction, but this method needs to add the observation flow actively and
consume extra bandwidth.

All of the above methods need to parse the HTTP package. With the user privacy has been concerned, more and more video providers abandon the former HTTP as the application layer protocol, and use HTTPS as the application layer protocol. This makes it infeasible that obtain the user buffer state by parsing the HTTP to evaluate the video Stalling probability of the user. In Reference [6-8], scholars use the relevant models and algorithms of machine learning to apply to video stalling detection. However, machine learning methods often require a large number of training samples in order to achieve a higher prediction accuracy of the model, which means that a large number of tagged video is required, that is undoubtedly a huge labor and material resources of the project.

In this paper, we propose a new method for detecting and evaluating video stalling based on HTTPS video of ISP. By capturing the TCP/IP packet information of user's video in real time and using the Stalling matching algorithm, the outlier points of video stalling can be matched. The decision analysis is added to the matching segment to realize the matching of the stalling fragment. By collecting and experimenting on various video sources of various video websites, the final results show that this method has a good effect on the estimation of stalling. The main features of this method are:

• Able to adapt to HTTPS protocol.
• Use simple and efficient models.

2. Video stalling detection method
Video Stalling refers to the slow buffering of video in the process of watching video, such as network fluctuations, which makes the downloaded video segment consume faster than the cache speed, and leads to the end of the client buffer depletion and then video pause. The video data message is downloaded from the video source server to the client buffer, and it will go through a dynamic interactive process from transmission to play. Next, we will analyze the phenomenon of video stalling in the bottom layer of the client and the transport layer.

The performance of video client-side is that the client-side buffer segment is depleted and the client-side buffer fragment is less than the maximum playback threshold, which results in the video not being able to play normally [3]. As shown in figure 1, the horizontal axis represents the timestamp and the vertical axis indicates the client video buffer size.

• Video starts playing from Click to download when the playback threshold $\theta$ is reached. Before that, the video was in a stall state with time $\Delta t_1$.
• Video stalling during the period from video buffer exhaustion to download to playback threshold $\theta$ with time $\Delta t_2$.

In these two stages, the video is in a stalling state. As shown by the red dotted line in the figure 1.

![Figure 1. Video buffer changes.](image)

Combining buffer changes, download speeds, and message accumulations. In figure 2, the horizontal axis represents the timestamp, and the three zigzag line segments represent the TCP_Segment_Sum, Buffer_Size, and Download_Speed. We can see that after 140 seconds of speed
limits, the client buffer changes within a minimum range and tends to 0 at some point, according to the
description of figure 1, when the video appears to be a stalling phenomenon. The description of the
video stalling is:
1) The buffer is depleted;
2) The download speed is too low for a certain period of time;
3) The slope of the curve in the Cumulative Variation of the Video Packet diagram begins to
decrease.

Figure 2. Buffer size, video download speed, the cumulative variation of the video packet.

The description in 3) means that the download of TCP values $Q_d$, in cumulative graph will
decrease in a certain period of time T. It's different from the front area, we call this region an abnormal
segment. Next we prove that video Stalling is related to unit time T and the download values $Q_d$.

In the case of speed limit, we assume that the video buffer from the beginning to the time
occurrence of stalling is $T$ and the average download speed of video is $V_d$, the average video playback
speed is $V_p$, the client buffer threshold is $Q$. Then, $T$ shows below:

$$T = \frac{Q}{V_p - V_d}$$  \hspace{1cm} (1)

During video playback, for a given $Q$ and $V$ values, when the network fluctuates, causing $V_d$
Smaller, then:

$$T = \begin{cases} 
\infty, \text{while } V_p - V_d < \exists, \text{where } \exists \text{ is Infinite small, and } \exists > 0 \\
\partial, \text{while } V_p > V_d, \text{ where } \partial \text{ is Constant}
\end{cases}$$  \hspace{1cm} (2)

$T$ shows the time of stalling happening. When the network fluctuates, the average download speed
is less than the average playback speed, and the video will occur after the time $T$, $T$ is a constant.

Besides, the formula can be written in another form. We assume that in duration time $T$, the video
download values is $Q_d$, the amount of video playback is $Q_p$, the buffer threshold value is $Q$, and
the Stalling event is $S$ which is a Boolean value. When $S$ equalling 1 means a Stalling occurs, then:

$$S = \begin{cases} 
1, \text{while } Q_v - Q_d > Q \\
0, \text{while } Q_v - Q_d \leq Q
\end{cases}$$  \hspace{1cm} (3)

From the estimation of $T$ in the above formula and the Boolean judgement of the Stalling event, it
can be seen that whether the video stalling occurs or not, relates to the download quantity $Q_d$ within
$T$ in this time period. So we can do this by $T$ and $Q_d$ to carry out parameter estimation to fit
the stalling event and realize the stalling matching. Of which $L_{vc}$ it represents the cumulative length
of TCP fragments in a certain period of time, and lambda denotes the threshold value. When $T$ value
is greater than the threshold $T_0$, it is considered as an abnormal fragment, that is, the fragment of Stalling
may occur. Here, why do we achieve the best match, we construct a matching function $F$. $F$ to achieve
the best segment matching for the input fragment.
a. TCP Message Accumulation Graph.  
b. Matching TCP Message Accumulation Graph.

**Figure 3.** Matching analysis.

Through the matching algorithm, the optimal matching is achieved, as shown in figure 3. The horizontal axis represents the timestamp, the vertical axis represents the cumulative length of the TCP/IP package. Figure 3a means when the client plays, the accumulated information of video message length captured in the experiment. Figure 3b is the result of the Stalling matching algorithm for the accumulated message shown in figure 3a. After estimating the parameters of the time length T in the stalling matching algorithm and the download quantity Q, the best match is realized, and leads to the matching of stalling by getting the length of the download interval in the playback process. The experimental results show that the accuracy of this method is over 97% for stalling event matching.

3. Experiment and the result

3.1. Video source and collection distribution

In order to verify the effectiveness of the method, we sampled videos of different duration of major mainstream websites with equal probability, as shown in figure 4. In view of the characteristics of the method, we only need to detect the short-time Stalling. From figure 5, we can see that 50% of the Stalling time distribution is about 1 second; 90% of the Stalling length distribution is less than 15 seconds; less than 2% of video stalling are longer than 50 seconds, mainly because some of them are too long due to some uncontrollable factors during the experiment:

**Figure 4.** Stalling times distribution.  
**Figure 5.** Stalling times distribution.

3.2. Stalling detection accuracy

Next, we will show the results of the whole experiment. The main purpose of this experiment is to detect different Stalling time, and the experimental objects include all the main video sources. For ISP, the key point of this paper is whether or not we can effectively identify the Stalling events of video users. We illustrate this by analyzing the detection accuracy of different video sources (domestic and foreign) with different effective playback time distribution.

The following is a histogram of the accuracy rate of different video websites and different video Stallings in the course of the experiment:
Figure 6. The horizontal axis shows the video duration, the vertical axis represents the accuracy of the Stalling decision, and the left and right histograms represent the YouTube and other video sources, respectively.

We draw columnar diagrams for different video transmission technologies at home and abroad, and for different playback time long video sources. As can be seen from figure 6, this method has a very high accuracy for different video sources and different video types.

4. Conclusion
The experiment is to have the stalling detection of the video transmitted by HTTPS protocol based on ISP. In the process of the experiment, through the Wireshark capture video message to simulate the ISP side, in the case of non-parsing video messages we only use the message basic information to realize video stalling detection. After that, through the video capture and experiment of 4 mainstream video sources, a total of 66 videos were collected including YouTube, and various user behaviours were added several times. The result of abnormal analysis shows that the total matching rate of stalling is 98.42%, and the matching rate of domestic video is 97.99% and that of foreign video is 99.04%.

References
[1] L. Ioannis, et al, Artificial Neural Networks for Traffic Prediction in 4G Networks (International Wireless Internet Conference Springer, Cham, 2014).
[2] D. Jamie, COMPETITIVE EFFECTS ON QUALITY COMPETITION: A Study of the U.S. Internet Service Provider Market (Economics, 2018).
[3] S. Raimund, T. Hobfeld, and P. Casas, Passive YouTube QoE Monitoring for ISPs. (IEEE, Sixth International Conference on Innovative Mobile & Internet Services in Ubiquitous Computing, 2012).
[4] C. Pedro, R. Schatz, and T. Hoßfeld, Monitoring YouTube QoE: Is Your Mobile Network Delivering the Right Experience to your Customers? (Wireless Communications & Networking Conference IEEE, 2013).
[5] M. Umar Nasir, Monitoring Network Congestion in Wi-Fi, based on QoE in HTTP Video Steaming Services (Lunds universitet, elektro- och informationsteknik, 2016).
[6] D. Giorgos, et al, Measuring Video QoE from Encrypted Traffic (The 2016 ACM ACM, 2016).
[7] D. Tsilimantos, T. Karakgioules, S. Valentin, Classifying flows and buffer state for YouTube's HTTP adaptive streaming service in mobile networks, arXiv:1803.00303(2018).
[8] A. Vaneet, et al, Prometheus: toward quality-of-experience estimation for mobile apps from passive network measurements (Workshop on Mobile Computing Systems, 2014).