A Comprehensive Analysis of Video Service Quality on IQIYI from Large-Scale Data Sets

Yao Guo[1](✉), Qiujian Lv[2], Fang Liu[1], Jie Yang[1], and Zhe Gao[3]

[1] Beijing Laboratory of Advanced Information Networks,
Beijing Key Laboratory of Network System Architecture and Convergence,
Center for Data Science, BUPT, Beijing 100876, China
{2013211938,lindaliu,janeyang}@bupt.edu.cn

[2] Institute of Information Engineering, Chinese Academy of Sciences,
Beijing 100093, China
lvqiujian@iie.ac.cn

[3] Technology Research Institute, Aisino Corporation, Beijing 100086, China
gaozhe@aisino.com

Abstract. With the proliferation of online video, measuring the quality of the video service has become a vital aspect for improving user's experience. Recent work shows that measurable quality metrics such as buffering, bitrate, and video resolutions impact user's experience, but none of them reveal the real relationships between these metrics and user's actual experience. This paper attempts to solve the problem above. We use IQIYI as the sample, and our large-scale dataset consists of 7 days real Internet traffic in a northern city of China. We quantify user's experience at per-video level (or view). Using Apache Spark, we extract some video events and calculate several quality metrics. In order to investigate the relationship between the metrics and user's experience, we use the FP-Growth algorithm to mining the implicit association rules and get some interesting results.

Keywords: IQIYI · Apache Spark · Video quality
User's experience · FP-Growth algorithm

1 Introduction

In the past few years, video streaming has become one of most popular Internet services. In China, the number of online video service users has reached 565 million, accounting for 75.2% of the total number of Internet users in 2017 [1]. IQIYI is one of the most popular online video service providers in China. According to a report [2] issued in 2016, IQIYI has accounted for 51% of the market share in the online video industry, having the largest amount of monthly active users comparing other companies such as Tencent and Youku.

L. Wang et al. (Eds.): QShine 2017, LNICST 234, pp. 83–93, 2018.
https://doi.org/10.1007/978-3-319-78078-8_9

As video distribution over the Internet becomes mainstream, user's demands on video qualities have dramatically increased. Different from other services on the Internet, online video streaming tends to occupy more traffic and bandwidth, and is easier to be disturbed by external factors such as network congestion, causing terrible experience. In this context, analysis of video service quality is paramount for content providers.

Quality of Experience (QoE) in HTTP video streaming is a well-known and largely investigated topic. Recent work [3,4] shows that measurable quality metrics, such as buffering, joining time, bitrate, and frequency of bitrate switching, impact user experience. Unfortunately, converting these observations into a quantitative QoE metric turns out to be challenging since these metrics are interrelated in complex, and can be unpredictable [5]. Some researchers [5] address these challenges through casting QoE inference as a machine learning problem. However, they have not revealed the implicit interrelation between the video metrics and user's experience yet, and the understanding of video QoE is just limited to a simple qualitative understanding of how individual metrics impact engagement (e.g., playing time) in some other works [6].

To solve the problems above, this paper quantifies user's experiences at a per-video level and introduce the FP-Growth algorithm to mine the implicit association rules between the quality metrics and user's experience. To improve the quality of the analysis, this paper uses large-scale traffic records from a leading Chinese Online video service IQIYI, which is similar to Youtube. Considering the huge size of data, we apply the Apache Spark to analyze the quality of IQIYI video service. Overall, the contributions of this paper can be summarized as follows:

(1) We analyze the interactive processes of watching an IQIYI video in detail. We distinguish the video events using Spark from massive data, which are extracted from real network traffic records collected from a metropolitan city that represent the activities of more than 223,800 people in seven days.
(2) We extract some video parameters and calculate video quality metrics about IQIYI;
(3) Using FP-Growth algorithm, we get some association rules between quality metrics and user's experience, and we verify the findings.

The rest of paper is structured as follows. We present some related work in Sect. 2. After this, we analyze the process of playing an IQIYI video and give some features of its request URI, then we filtered the HTTP packets of IQIYI. Based on the work before, we calculate some video quality metrics in Sect. 3. In Sect. 4, we try to analyze some relationships between different video quality metrics and give some conclusions.

2 Related Work

At present, with regard to video quality, there are two main research directions, which focus on the network transmission side and the user side respectively.

QoS-based video quality prediction: The measurements of QoS focus on the network transmission side. QoS metrics uses available network diagnostic parameters, such as traffic jitter, arrival time, and packet loss, which are easy to measure at any point inside the delivery network [7], and ITU-T has given some standards [8]. The advantage of QoS-based metrics is that these models are designed to be easily deployed at any point in the distribution network. The disadvantage is that QoS models do not have access to two pieces of important information: (1) how the video originally looked and sounded, and (2) what the end-user sees and hears. This limits the ability of QoS models to predict quality as perceived by the end user.

Measuring end-user's quality perception: The measurements of QoE focus on the user side. Subjective video quality tests and objective video quality metrics provide established techniques for end-user point-of-view. Subjective methods for video quality assessment are considered the most accurate and reliable way [9], and ITU-T has already given some standards [10]. However, these methods are non-real-time, terrible in portability and applicability between different service providers. Many researchers turn to the objective assessment approach and get some results. Some found that the initial buffering duration, buffering length, buffer location and bit rate were the key factors affecting video quality [3,4]. Authors in [6,11,12] showed that the percentage of time spent in buffering had the largest impact on user engagement in some special video contents, while authors in [13] showed that initial buffering latency was preferred to stalling by around 90% of users. Unfortunately, the state of the art in their understanding of video QoE was limited to a simple qualitative understanding of how individual metrics impacted engagement. In this context, the researchers continued to introduce different types of regression models [14] and classification models [15], and further explore the user's experience. Besides, some researchers have attempted to map the objective quality of service (QoS) metrics acquired in the network transmission side directly to the subjective QoE evaluation [9]. However, they all didn't point out the implicit association rules between the quality metrics and user's experience.

Apart from the previous works, this paper is to investigate the implicit relationships between quality metrics and how these metrics influence user's experience in user side.

3 Analysis of IQIYI Video Metrics

3.1 Data Description

The traffic data we used is collected by our self-developed Traffic Monitoring System [14]. The system can monitor packets and aggregate them into records. Each record contains a lot of useful information about user's requests to the Internet. The dataset used ranges from February 21 to February 26,2017 (Monday - Sunday), covering about 223,800 users. Moreover, the size of datasets is about 4TB.

3.2 Data Filtering

A complete IQIYI video interactive process usually includes: select the video and click the play button on the web page, starting to request video resources, starting playing the advertisement, ending playing the advertisement, downloading the initial video cache, downloading the subsequent video cache while playing, starting buffering, ending buffering, channel conversion and playing the next video, etc.

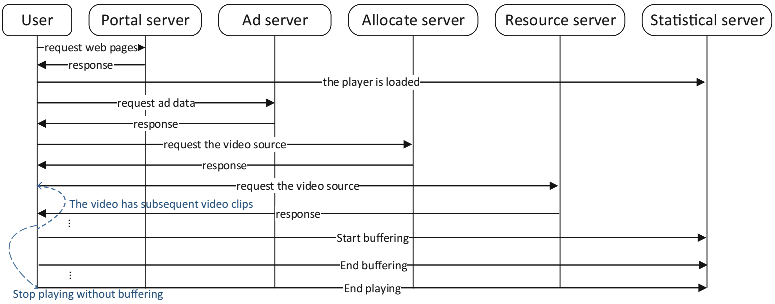

Fig. 1. IQIYI video interactive process.

As shown in the Fig. 1, a playing process needs several servers to participate in. By experimenting and reproducing the interactive process many times, we find some features of its domain names of different servers. For example, almost all IQIYI's domain names end with "qiyi.com" or "iqiyi.com". The statistical server, which collects the status information of the video during playback, its domain name is fixed to "msg.71.am". The allocation server is responsible for scheduling video resources and getting their specific locations to the client side, and its domain name is also fixed to "data.qiyi.com". All videos' URIs, including advertisement, start with "/videos/". We also find that there are many differences in the request URI when using different platforms to watch an IQIYI video, such as Windows OS and Android OS, which is worthy of our attention when conducting an experiment.

Based on the IQIYI's hosts we find above, we can easily filter the data about IQIYI from the massive data. But there are still some redundant records we don't need in calculating the quality metrics. We just need to extract key records about the video event during playing such as downloading data event. In this case, we use the regular expression to finish this work.

3.3 Video Metrics Calculation

In this part, we calculate the video quality metrics at a per-video level (or view).

Joining Time. When we decide to watch a video and click the playing button on the web, we can't see the video contents at once, which has an impact on user's experience. Based on the description of Fig. 1, the joining time T_j is defined as the time interval between the time when the user clicks the playing button and the time the video starts to play, including the time T_1 to init buffer and the time T_2 to play the advertisement. We define:

$$T_j = T_1 + T_2 \tag{1}$$

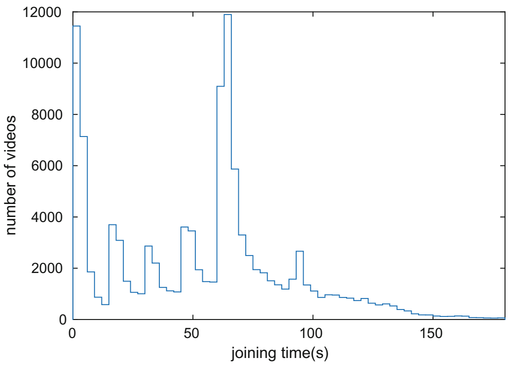

Fig. 2. The histogram of video joining time

Six peaks appear in Fig. 2, and the interval time between peaks is about 15s. Actually, the peak is caused by advertisement playing. In IQIYI, each advertisement's time length is fixed to fifteen seconds strictly, and the number of the advertisements being played is decided by the length of the video. A User, who chooses a video length at about 45 min, usually needs to watch four advertisements before the video starts to play. That means the user needs to wait at least 60 s.

Video Resolution. Video resolution is one of the most intuitive metrics. Through capturing the video list information the allocation server sends to the client, we know IQIYI provides five kinds of video resolutions. According to the status of the network, the player requests the proper resolution to download.

The IQIYI's statistical server plays a role as a monitoring station, which collects the status of the video player cyclically, and we call this detection signal as the heartbeat signal. In the request URI of the heartbeat signal, the parameter called defi is discovered to be the flag of a video's resolution.

IQIYI defines the resolution over "896 * 504" as high resolution. The pie in Fig. 3 reports that 73% of views are at the resolution "896 * 504". Due to the fact that an important factor limiting the resolution is the quality of the network, so we can infer that the resolution of "896 * 504" is the best compromise between the clarity of a video and quality of network when watching a video in the IQIYI platform.

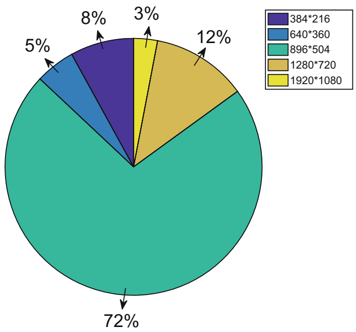

Fig. 3. The distribution of video's resolution

Buffering Ratio and Buffering Frequency. Buffering Ratio is represented as a percentage. It is the fraction of the total session time spent in buffering. The player goes into a buffering state when the video buffer becomes empty and moves out of buffering (back to playing state) when the buffer is replenished. It is worth to note that the buffering ratio doesn't take the number of buffering events into account. Although the buffering ratios of a video is same, the number of buffering events may vary. The author in [12] gives the conclusion that multiple buffering events have the different influence on user's engagement comparing to a single buffering event, causing more users to stop the playing process.

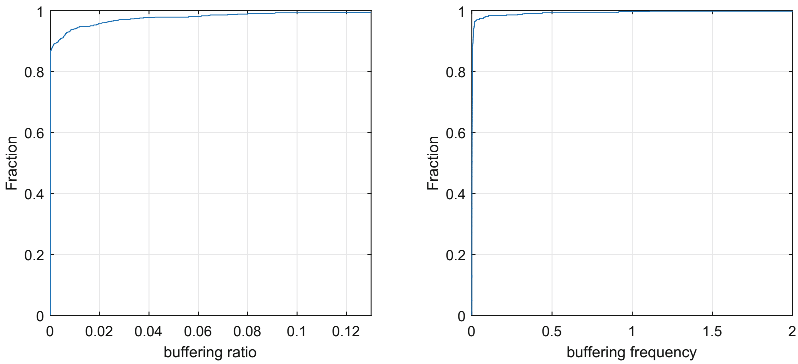

Fig. 4. The CDF of buffering ratio and buffering frequency

As expected in Fig. 4, most viewing sessions experience good quality, both having a very low buffering ratio and low buffering frequency. We note that the possibility of buffering during playback in IQIYI is small since we only find about thirty thousand buffering records from the traffic data collected in seven days.

4 Correlation Between Quality Metrics and Experience

This section aims to investigate the implicit relationships between the quality metrics and user's experience based on the results discovered in Sect. 3. The relationships between the quality metrics and the effective user experience can be extremely complex, and the quality metrics themselves have subtle interdependencies and have implicit tradeoffs. For example, some research has shown that although switching bit rate to adapt to the bandwidth condition can reduce buffering effectively, the high rates of bit rate switching annoy users to some extent.

4.1 Settings and Assumptions

Unlike other video providers such as YouTube that streams short videos, IQIYI typically provides TV series usually length at about 45 min. Figure 5 shows the distribution of the lengths of the videos being played in seven days. It is clear that users prefer to choose the TV series when using IQIYI platform. In order to eliminate the impact of user's bias on QoE, we filter the videos length at 40 min to 50 min to analyze in the following experiments. We also note that the impact of joining time can be ignored when the video length is fixed and has been playing, as explained in Sect. 3.

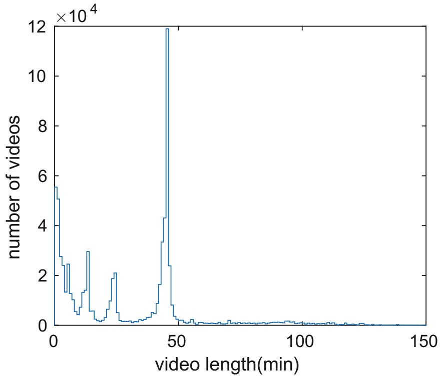

Fig. 5. The histogram of video length

Then, we point out how we distinguish the videos that are abandoned due to poor viewing experiences from the massive records. As other researchers [12], we also adopt the video abandonment to do this work. Besides, we take the dragging events into account for complementary. When a user drags the process bar during playback, we infer that the user has poor engagement, which also reflects user's experience. After making these assumptions, we introduce the FP-Growth algorithm.

4.2 FP-Growth Algorithm

In order to obtain the relationship between the video metrics and user experience, we implement the FP-Growth algorithm to mine [16] the implicit association rules.

The FP-growth algorithm is proposed by Han et al., and aims at mining frequent patterns without candidate generation. The FP stands for the frequent pattern. Given a dataset of transactions, the first step of FP-growth is to calculate item frequencies and identify some frequent items. The second step of FP-growth uses a suffix tree (FP-tree) structure to encode transactions without generating candidate sets explicitly, which is usually expensive to generate. After the second step, the frequent itemsets can be extracted [17]. The Spark.mllib provides a parallel implementation of FP-growth.

4.3 Analysis of Association Rules

To implement the RDD-Based FP-Growth API, we need to provide two parameters called *Support* and *Confidence*. *Support* is an indication of how frequently the itemset appears in the dataset, and *Confidence* is an indication of how often the rule has been found to be true. As related works show buffering events have the largest impact on user's experience, we concentrate on finding the rules related to buffering events. Based the analysis results discovered in Sect. 3, we have realized that the possibility of buffering during playback in IQIYI is small, and this indicates a small supportance value of rules including buffering events. Through extensive experiments, we find 0.17% is the threshold of *Support*. We set *Confidence* at a high level, 90%, to enhance accuracy. To verify the correctness, we calculate *lift* of the rules we have found. The *lift* is defined as:

$$lift\,(X \Rightarrow Y) = \frac{supp\,(X \cup Y)}{supp\,(X) \times supp\,(Y)} \tag{2}$$

If the rule had a lift of 1, it would imply that the probability of occurrence of the antecedent and that of the consequent are independent of each other. When two events are independent of each other, no rule can be drawn involving those two events [18].

We use eight letters to represent different video events or quality metrics shown in Table 1. Then, we get some preliminary results shown in Table 2. Finally, we calculate these rule's *lift* to verify correctness.

From Table 2, we see that:

Line 1: When users watch a high-resolution video and drag the process bar during playing, 94.6% of them tend to switch the channel or stop playing this video. We infer that the dragging behavior can be a signal that user is losing interest or meets some terrible video events, such as buffering when watching a high-resolution video.

Line 2: When users meet a buffering event in IQIYI, there is a great possibility that they have chosen a high-definition video.

Table 1. Definition of the symbols

Symbols	Definition	Attributes
A	Suffering buffering events	Video quality
B	Playing fluently	Video quality
C	Users drag the process bar	Video event
D	User don't drag the process bar	Video event
E	Switch channels or click to stop	Video event
F	Keep playing session	Video event
G	High-resolution videos	Video quality
H	Low-resolution videos	Video quality

Table 2. Mining results

Rules	*Confidence*	*lift*
$[C, G] \Rightarrow E$	94.62%	1.034
$[A] \Rightarrow G$	92.10%	1.032
$[A, G] \Rightarrow E$	96.77%	1.072
$[B, G] \Rightarrow D$	95.91%	1.047

Line 3: Encountering some buffering events while watching a high-resolution videos can have a great impact on the user's experience. We note that if the extent of buffering is severe, the player of IQIYI will stop the playing process automatically.
Line 4: It indicates that if users lose interest in the video contents or have a bad experience, they will drag the process bar or just stop the playing process. Watching a high-resolution video fluently means a good user experience, and 95.1% of users won't drag the process bar during playing.

5 Conclusion

In this paper, we investigate the video service quality of IQIYI in detail. We analyze the interactive processes when watching an IQIYI video in detail and find some features of IQIYI's hosts and request URIs. We distinguish the video events from large scale traffic records using Spark and calculate the quality metrics of IQIYI. Finally, we try to find out the implicit association rules between quality metrics and user's experience and get some findings. Our analysis provides a better understanding in quality metrics and user's experience. Further, improvements may be achieved by taking more metrics into account and investigating user's experience in more dimensions.

Acknowledgment. This work is supported in part by the National Natural Science Foundation of China (61671078, 61701031), Director Funds of Beijing Key Laboratory of Network System Architecture and Convergence (2017BKL-NSAC-ZJ-06), and 111 Project of China (B08004, B17007). This work is conducted on the platform of Center for Data Science of Beijing University of Posts and Telecommunications.

References

1. CNNIC: The 40th China statistical report on internet development. http://www.cnnic.net.cn/hlwfzyj/hlwxzbg/hlwtjbg/201708/P020170807351923262153.pdf. Accessed July 2017
2. China Economic Net: Video market tripartite confrontation. IQIYI occupies 51% of the market share. http://www.ce.cn/cysc/tech/gd2012/201702/13/t20170213_20160241.shtml. Accessed Feb 2017
3. Shunmuga Krishnan, S., Sitaraman, R.K.: Video stream quality impacts viewer behavior: inferring causality using quasi-experimental designs. IEEE/ACM Trans. Netw. **21**(6), 2001–2014 (2013)
4. Yu, F., Chen, H., Xie, L., Li, J.: Impact of end-user playout buffer dynamics on http progressive video QoE in wireless networks, vol. 136, no. 5, pp. 1996–2001. IEEE (2014)
5. Balachandran, A., Sekar, V., Akella, A., Seshan, S., Stoica, I., Zhang, H.: A quest for an internet video quality-of-experience metric. In: ACM Workshop on Hot Topics, Networks, pp. 97–102 (2012)
6. Dobrian, F., Awan, A., Zhan, J., Zhang, H.: Understanding the impact of video quality on user engagement. In: ACM SIGCOMM 2011 Conference, pp. 362–373 (2011)
7. Staelens, N., Pinson, M., Corriveau, P., De Turck, F., Demeester, P.: Measuring video quality in the network: from quality of service to user experience. In: 9th International Workshop on Video Processing and Consumer Electronics (VPQM 2015), pp. 5–6 (2015)
8. Seitz, N.: ITU-T QoS standards for IP-based networks. IEEE Commun. Mag. **41**, 82–89 (2003)
9. Moldovan, A.N., Ghergulescu, I., Muntean, C.H.: A novel methodology for mapping objective video quality metrics to the subjective MOS scale (2014)
10. BT ITU-R: Methodology for the subjective assessment of the quality of television pictures. EBU Technical Review (2015)
11. Hoßfeld, T., Seufert, M., Hirth, M., Zinner, T., Tran-Gia, P., Schatz, R.: Quantification of Youtube QoE via crowdsourcing. In: IEEE International Symposium on Multimedia, pp. 494–499 (2012)
12. Nam, H., Kim, K.H., Schulzrinne, H.: QoE matters more than QoS: why people stop watching cat videos. In: IEEE INFOCOM 2016 - the IEEE International Conference on Computer Communications, pp. 1–9 (2016)
13. Hossfeld, T., Egger, S., Schatz, R., Fiedler, M., Masuch, K., Lorentzen, C.: Initial delay vs. interruptions: between the devil and the deep blue sea. In: Fourth International Workshop on Quality of Multimedia Experience, pp. 1–6 (2012)
14. Rodriguez, D.Z., Abrahao, J., Begazo, D.C., Rosa, R.L., Bressan, G.: Quality metric to assess video streaming service over TCP considering temporal location of pauses. IEEE Trans. Consum. Electron. **58**(3), 985–992 (2012)
15. Balachandran, A., Sekar, V., Akella, A., Seshan, S., Stoica, I., Zhang, H.: Developing a predictive model of quality of experience for internet video. ACM SIGCOMM Comput. Commun. Rev. **43**(4), 339–350 (2013)

16. Verhein, F.: Frequent pattern growth (FP-growth) algorithm. University of Sydney, January 2018
17. Apache: Docs of frequent pattern mining RDD-based API. https://spark.apache.org/docs/latest/mllib-frequent-pattern-mining.html. Accessed July 2016
18. Wikipedia: Association rule learning. https://en.wikipedia.org/wiki/Association_rule_learning. Accessed Aug 2017