Combination between DASH and SVM machine learning in the field of video delivery

Jian Chen*, Chunming Yuan*
Beijing Advanced Innovation Center for Future Internet Technology, Beijing University of Technology, Beijing, China

*Corresponding author e-mail: s201641002@emails.bjut.edu.cn, s201641003@emails.bjut.edu.cn

Abstract. With the development of network and streaming media technology, network video traffic is growing rapidly. In order to better control and manage network traffic and guarantee the quality of service of network video, it is necessary to classify network video services effectively. In traffic identification and classification, feature analysis and acquisition of better features are the key points to achieve efficient classification. Starting with the characteristics of packet size distribution, rate, IP alternation, byte number ratio between downstream and upstream, number of sub-stream fragments and average packet arrival time interval, this paper uses Support Vector Machine (SVM) to verify the classification effect of this feature, and achieves a high classification accuracy.

1. Introduction
Video streaming is becoming one of the most popular services for mobile users. In fact, mobile video traffic, which accounted for 55% of total mobile traffic in 2015, will account for more than 70% by 2020. This significant growth has been accompanied by the widespread adoption of Dynamic Adaptive Streaming over HTTP on the standard DASH.

It is a popular standard for video streaming on the Internet, allowing for improved user experience in the presence of variable network conditions. DASH aims to avoid play interruption, which is known to be detrimental to perceived video quality [4]. The DASH server provides various quality levels (that is, bit rates) for the same content. When the network deteriorates the video client can switch to a lower video quality level (and therefore lower bit rate) to reduce the risk of playback disruption. Similarly, improved conditions allow for switching to higher levels of quality, ensuring optimal resource management. Information about the available video representation can be found in the media rendering description (MPD) file, which is downloaded by the client from the DASH server. The MPD file stores the bit rate, timing, structure and format/codec information or each representation of video and its fragments.

We propose a new architecture for adaptive HTTP video streaming in the context of MEC (section 3), introducing an adaptive algorithm to run on an MEC server with the aim of reducing network congestion while providing mobile users with improved execution quality (QoE). The main idea is to dynamically control the video representation of the client based on the current network state, thus driving the video adaptation mechanism of the client. To achieve this, we leverage (I) awareness of network-level information, which can be retrieved via the MNO exposed MEC API, and specific features of the (I I) DASH standard. Our approach demonstrates a use case for a mutually beneficial cp-mno
collaboration on the MEC infrastructure: we show (section iv) that fairness between mobile users is guaranteed, that overall perceived QoE is improved, and that better network resource utilization is improved by responding to congestion. Staleming. In addition, by actively incorporating MNO into the video transmission chain, our solution allows compensation for the loss of average MNO revenue per user after the introduction of a fixed rate. It is important to note that the proposed mechanism is standards-compliant and transparently coexists with the client-side bit rate adaptive algorithm. In fact, the final quality level selection decision is still made by the customer, which corresponds to the current DASH client standard specification.

2. Feature selection and simulation experiment

This paper mainly classifies and identifies the data streams of HTTP streaming media, online network television (Sopcast), HTTP download, Thunder download and QQ video.

2.1. Data reprocessing

Firstly, we use Wireshark to capture the package to get the experimental data. Considering the changing characteristics of the network and in order to capture more services, the time of each stream is set to 30 minutes. For HTTP HD video and HTTP standard definition video services, only the first 30 minutes of the video service are captured at a time. The time span of data acquisition is from November 2013 to February 2014, and data acquisition is carried out in three time periods: 9:00-12:00, 13:00-17:00 and 18:00-22:00. The training sample set and the test sample set contain data in three periods. The selection of data in the same period is random. After obtaining the original data, the flow characteristic parameters are extracted by tools, and then the flow characteristic parameters are analyzed and selected. Then the selected features are used to train the SVM model. The test sample set is classified and the output results are analyzed.

| Number | Name     | Training samples | Test samples |
|--------|----------|------------------|--------------|
| 1      | High     | 30               | 30           |
| 2      | Standard | 30               | 30           |
| 3      | Sopcast  | 30               | 15           |
| 4      | HTTP     | 30               | 15           |
| 5      | Thunder  | 30               | 20           |
| 6      | QQ       | 30               | 20           |

2.2. Criteria for measuring the effectiveness of classification

TP denotes the number of samples correctly classified by class j, FN denotes the number of samples classified by class j, FP denotes the number of samples classified by other classes, and the accuracy rate is the proportion of the number of samples correctly predicted in the predicted samples. The recall rate is correctly predicted as the proportion of the sample number of a class in the actual sample number of that class. The accuracy rate and recall rate reflect the correctness and completeness of classification respectively. The definitions of accuracy and recall rate are shown in Formula (1) and Formula (2), respectively. Formula (3) the overall recall rate defined by formula (3) shows the classification accuracy for the whole system as a whole.

Accuracy rate and recall rate reflect the effect of classification from different aspects, which need to be considered comprehensively. F-measure weighted harmonic average of accuracy rate and recall rate is used as the final analysis parameters.

\[
\text{recall} = \frac{TP}{TP + FN}
\]  

(1)
3. Feature selection and experimental results

QQ videos differ significantly from the other five types of services in packet size distribution. Sopcast and Thunderbolt have similar distribution, but they differ greatly from HTTP download, HD and SD. HTTP download, HD and SD have similar distribution, whether from Hellinger distance of probability distribution or from Hellinger distance of probability distribution. From the point of view of information entropy, they are very similar, so it is difficult to separate them from each other in terms of packet size distribution. From the point of view of speed, the overall trend of HD is higher than the rate of standard definition. Through the previous analysis, we can see that due to the variability of the network environment, the stability of this feature is poor. The characteristics of Sopcast and QQ video are similar in the ratio of downstream and upstream bytes. Thunderbolt has a large span. The ratio of lower bound to Sopcast and QQ video crosses, but the upper bound is smaller than the ratio of SD, and the ratio of HD is larger than that of SD. The ratio of HTTP download is higher than that of HD. Large, some are similar to HD ratios. In terms of the number of sub-streaming fragments, QQ video and HTTP download are mainly based on the communication between two IP. The time interval between two valid packets may exceed the set threshold due to control overhead and other reasons, which leads to the emergence of fragments. However, from the experimental data, the number of sub-streaming fragments of both is very large. Less. The number of sub-stream fragments of Xunlei is larger than that of other services. Sopcast, HD and SD sub-stream fragments are hard to distinguish, but they are more downloaded than QQ and HTTP on the whole. It can be seen that the ratio of downstream bytes to upstream bytes can be used to identify HD and SDQ services, but QQ and Spotcast can be recognized as a whole. It is difficult to distinguish between them. In addition, the number of sub-streams is a feature that can better separate them, and Xunlei can also be separated accordingly. Therefore, the ratio of downstream bytes to upstream bytes and the number of sub-stream fragments are selected to train the model and classify the samples.

In order to verify the correctness of the analysis, the two features are selected to train the SVM model and identify the samples. Before discussing the classification results, the determination of sub-stream fragments is explained. In actual data, there will be protocols for inserting control packages such as LLMNR NBNS. This paper defines them as control overhead. It can be seen from the figure that the average control overhead time is less than 95% and less than 0.5s. Therefore, when sub stream fragments are defined, when sub stream fragments are encountered and the control overhead time is not more than 0.5 seconds, the sub stream fragments are considered not broken, otherwise the sub stream fragments are considered to have ended and a new sub stream fragment needs to be calculated.
It can be seen from the analysis of the ratio of the number of sub-stream fragments to the number of bytes that the number of sub-stream fragments in some services is much larger than the ratio of bytes. The direct use of sub-stream fragments in model training will have a great impact on the ratio of bytes. From the classification effect, the ratio of bytes in this case is greater than that in single case. The classification effect of each feature is even worse. This also shows that when the number of sub-streams is dominant in the training model, it is difficult to achieve better classification results: the role of two features in the training model needs to be adjusted. In order to compare the classification results with single feature using the ratio of bytes, the classification results with the ratio of the number of sub-stream segments and the number of bytes are significantly improved, which shows that the number of sub-stream segments plays a role in the model after the number of sub-stream segments is logarithmic.

### Table 2. Classification effect combining byte number ratio and number of sub-stream fragments

| Name    | Accuracy | Recall   | F-measure |
|---------|----------|----------|-----------|
| High    | 1        | 0.0400   | 0.0769    |
| Standard| 1        | 0.1333   | 0.2353    |
| Sopcast | 1        | 0.7333   | 0.3188    |
| HTTP    | 1        | 0.0667   | 0.6333    |
| Thunder | 0.8169   | 1        | 0.1250    |
| QQ      | 1        | 0.2000   | 0.3333    |

When classifying network traffic, two characteristics, average packet size and average packet arrival time interval, are used to classify the six types of traffic in this paper. It can be seen that the classification effect is worse than that using the ratio of bytes and the number of sub-stream fragments. From the analysis of packet size and average packet arrival time interval in Chapter 3, it can be seen that the distribution of packet size characteristics of HD, SD and HTTP downloads is similar. In average packet arrival time interval characteristics, QQ video and Spotcast are similar, while Xunlei and HTTP downloads are similar; although the average packet arrival time interval of HD is similar. On the whole, it is lower than the standard clearance value, but due to the influence of network environment factors, this is between the two.

Relationships do not necessarily apply to data in different time periods, and they also have an impact on classification. Therefore, it is difficult to obtain better classification results when classifying the six types of services in this paper by using average packet size and average packet arrival time interval.
4. Conclusion
In recent years, with the rapid development of network video service, it occupies more and more proportion of the total Internet traffic. The transmission of streaming video has the characteristics of large data volume and high real-time, which puts forward higher requirements for network management: fast and efficient identification and classification of video service. Feature selection is the premise and basis of classification, which has an important impact on the final classification effect. Moreover, understanding and mastering the characteristics of these video services also plays an important role in the development of related network technology. This paper mainly analyses and identifies six video services, including HD and SD HTTP streaming media, HTTP download, QQ video, online network TV live broadcast (Sopcast) and Xunlei.

Firstly, the package size distribution of various services is analyzed and compared. It is found that the package size distribution of all kinds of services is relatively stable from the overall trend. From the CDF chart of package size and Hellinger distance measure of probability distribution, HTTP download, HD, and demarcation can be seen. Distribution similarity is very high, Sopcast and Xunlei services have similarities, and the distance between services is the largest between QQ video and other types of services. In the same time period, the average rate of HD data streams is higher than that of SD. Because of the variability of network state, the average rate of HD data streams in different time periods may be lower than that of SD data streams. The average packet arrival time interval of HD is smaller than that of SD data streams in general trend, but it is affected by network ring. The data in different time periods may not conform to the influence of environment. The ratio of the number of bytes in the downlink to the number of bytes in the upstream is higher than that in the standard, and the stability is better. At the same time, the number of sub-stream fragments is also analyzed. It is found that the number of sub-stream fragments in Xunlei service is more than that in other services.

The ratio of downstream bytes to upstream bytes and the number of sub-streams are selected as training features. The real network data captured in the campus network environment are used for classification experiments. The experimental results are analyzed. At the same time, the experimental results are compared with those based on the classification of packet size and arrival time interval.

References
[1] YAO CHING BANG. Personalized guidance and ubiquitous learning in intelligent library with multi-agent [C]. The 2nd International Conference on Computer and Automation Engineering, ICCAE 2010, 1: 578 - 582.
[2] KIM A, JEONG S H. A QoS/QoE control architecture for multimedia communications [C]. International Conference on Information Networking, 2012: 346 - 349.
[3] Sen S, Wang J. Analyzing Peer-to-Peer Traffic Across Large Networks. IEEE/ACM Transactions on Networking, 2004, 12 (2): 219 - 232k/.
[4] Karagiannis T, Papagiannaki K, Faloutsos M. BLINC: multilevel traffic classification in the dark [C]// ACM SIGCOMM Computer Communication Review. ACM, 2005, 35 (4): 229 - 240.
[5] Moore A W, Zuev D. Internet traffic classification using bayesian analysis techniques[C]// ACM SIGMETRICS Performance Evaluation Review. ACM, 2005, 33 (1): 50 - 60.
[6] Zhibin Y, Kil G B, Kim S. Traffic classification based on visualization[C]// Networked Embedded Systems for Enterprise Applications (NESEA), 2011 IEEE 2nd International Conference on. IEEE, 2011: 1 - 6.
[7] Ibrahim H A H, Mohd Nor S, Ahmed A. Internet Traffic Classification Algorithm Based on Hybrid Classifiers to Identify Online Games Traffic [J]. Jurnal Teknologi, 2013, 64 (3): 55 - 60.