ITP-KNN: Encrypted Video Flow Identification Based on the Intermittent Traffic Pattern of Video and K-Nearest Neighbors Classification *

Youting Liu[1,2,3], Shu Li[1,2,3,4], Chengwei Zhang[1,2], Chao Zheng[1,2], Yong Sun[1,2], and Qingyun Liu[1,2,3]

[1] Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China
[2] National Engineering Laboratory of Information Security Technologies, Beijing, China
[3] School of Cyber Security, University of Chinese Academy of Sciences, Beijing, China
[4] Corresponding Author :lishu@iie.ac.cn

Abstract. As video dominates internet traffic, researchers tend to pay attention to video-related fields, such as video shaping, differentiated service, multimedia protocol tunneling detection. Some video-related fields, e.g., traffic measurement and the metrics for Quality of Experience, are based on video flow identification. However, video flow identification faces challenges. Firstly, the increasing adoption of Transport Layer Security makes payload-based methods no longer applicable. Secondly, traffic features differ when generated by different streaming protocols. This paper proposes a video flow identification method, called ITP-KNN, which utilizes the intermittent traffic pattern-related features (ITP) and the K-nearest neighbors (KNN) algorithm. The intermittent traffic pattern is caused by fragmented transmission, which is common among video streamings generated by different streaming protocols. Therefore, the intermittent traffic pattern is useful for overcoming the above challenges and then differentiating video traffic from not-video traffic. We develop a set of features to describe the intermittent traffic pattern. Preliminary results show the promise of ITP-KNN, yielding high identification recall and precision over a range of video content and encoding qualities.

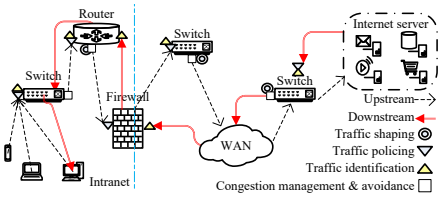
Keywords: Video streaming · Encrypted traffic · Traffic identification · Traffic pattern · Explainable machine learning · Feature selection.

1 Introduction

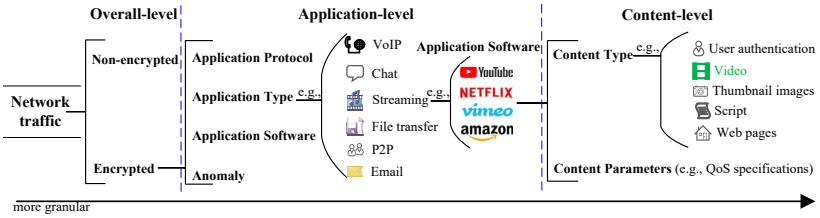
Video dominates the Internet: streaming video occupies 65% of traffic globally [28]. The increasing popularity of video has attracted the interest of different research groups. As a result, several video-related research directions have become hot topics, such as video traffic model [34, 24], the metrics for Quality of

* Supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDC02030600), Youth Innovation Promotion Association CAS and CAS Key Technology Talent Program.

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_21



(a) An example of Quality of Service (QoS) technology implementations .



(b) Multilevel categorization of traffic identification methods. We aim to solve a content-level traffic identification problem.

Fig. 1. What is video flow identification? QoS measures the ability of a network to provide differentiated service guarantees for diverse traffic following changing and complex network conditions [15]. Furthermore, traffic identification is a basic QoS technique, and ISPs can provide differentiated services based on traffic identification. Video flow identification is a content-level identification.

Experience (QoE) [10], video title identification [9, 29, 13], multimedia protocol tunneling [2].

Video flow identification (described in Figure 1) is needed. There is an assumption in much research [9, 29, 13, 32]: the flow is labeled as a video flow or not-video flow precisely. In other words, video flow identification is essential for practical applications of some video-related research, e.g., the metrics for internet video QoE. Besides, limiting or blocking video transmission also requires video flow identification. When streaming video transmission is limited, the resolution of the streaming video will switch to a lower one, in order not to interrupt the video playback. Thus, it is possible to limit video transmission to mitigate network congestion without significantly affected the Quality of Service. As an example, streaming-video services such as Netflix and YouTube began switching to standard definition to manage internet congestion during COVID-19 [26].

Two challenges of video flow identification are protocol diversity and traffic encryption.

Protocol diversity challenges video flow identification as it makes traffic features differ. Protocol diversity results from the application of kinds of streaming protocols. Streaming protocols are designed to provide online video playback without completely downloading it first. There are various popular streaming protocols, e.g., Dynamic Adaptive Streaming over HTTP (DASH), HTTP Live Streaming (HLS), or even private protocol. Protocol diversity makes traffic fea-

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_21

tures differ. For example, packet size is different when different streaming protocols generate flows. The advanced method presented by Li et al. [18] requires detecting the upstream request, which depends on the packet size, so it is not universal enough to deal with other streaming protocols.

Video flow identification faces the other challenge, traffic encryption. *Sandvine*'s report states that more than 50% of global traffic is encrypted [27]. When traffic is encrypted, conventional payload-based methods, such as deep packet inspection, are no longer applicable. Fortunately, traffic analysis (TA) still work even when traffic is encrypted [23]. However, for encrypted traffic, fine-grained classification is a tough task[4], and video flow identification is a kind of fine-grained identification.

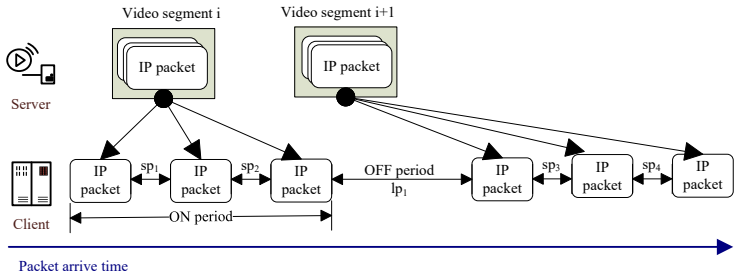


Fig. 2. Transmitting video in segments leads to the traffic showing an intermittent pattern — continuously arrive of packets during ON periods, and a suspend for transmission during OFF periods. The intermittent traffic pattern consists of lots of ON periods and OFF periods one-to-one. Furthermore, continuously arrive of packets means a set of small PAIs (e.g., sp_i), smaller than the PAI (e.g., lp_j) that reflects the suspend for transmission (OFF period). Therefore, the probability distribution image of PAIs will contain two peaks. One peak is influenced by ON periods, consisting of sp_i, and i value range is from 1 to n in increments of 1 while n is the number of small PAIs (sp_i). The other peak reflects OFF periods, consisting of lp_j, and j value range is from 1 to z in increments of 1 while z is the number of larger PAIs (lp_j).

In this paper, we propose a video flow identification method: ITP-KNN. ITP-KNN is a kind of machine learning-based traffic analysis, using intermittent traffic pattern-related features (ITP) as input and the K-nearest neighbors (KNN) algorithm as classifier. Regularly and fragmentarily transmission of video, which is caused by the application of streaming protocols, makes the traffic traces showing the intermittent pattern [21] (as shown in Figure 2). Thus, the intermittent traffic pattern helps differentiate video traffic from not-video traffic.

We summarize our key contributions as follows:

- We develop a set of features to depict video transmission patterns: the intermittent traffic pattern.
- We interpret why the features we developed can help to identify video flows.

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_21

– We implement and evaluate ML-based TA in video flow identification: we propose a video flow identification framework, and evaluate it on five public datasets.

2 RELATED WORK

2.1 Meachine Learning (ML)-based Traffic Analysis (TA)

TA has been proved applicable in encrypted traffic identification [22, 4]. McGregor et al. [20] first apply ML algorithms to identify traffic. After that, researchers began to pay attention to the application of ML algorithms in traffic identification [7].

Manual feature selection is required when using conventional ML algorithms [35]. Encrypted traffic identification lacks public datasets to evaluate different methods. Draper-Gil et al. [8, 17] released their datasets, and they used time-related features to realize VPN or Tor traffic identification. However, they did not explain why the time-related features can help identify traffic. Shi et al. [30] developed a set of network path-related features to identify the traffic source; the disadvantage is that models that use network path-related features need to retrain regularly.

In recent years, deep learning (DL) has been implemented in traffic identification with different aims [25]. Wang [37] showed that compared with conventional ML algorithms, DL had improved the result of traffic identification. However, their model is not explainable.

DL obviates the requirement of manually feature selection since the feature selection runs automatically through training [25]. It is a high cost for domain experts to select features for identifying various types of traffic. Therefore, the DL-based TA is more suitable for multiclass identification than conventional ML-based TA. Conventional ML-based TA is not inferior to the DL-based one; the problems that prevent the application of DL are that DL requires a large and representative dataset, and the result of DL lacks interpretability.

2.2 Video Streaming

Li et al. [18] presented Silhouette to identify YouTube video flows. Silhouette depends on the packet size of the upstream request, so it is not universal enough to deal with other streaming protocols. Shi et al. [32] proposed a method that can identify the video source. However, the features adopted by them are strongly related to the network condition, so regular training is needed. Casas et al. [5] achieved an application-level identification; some features used by them are computational complex. They concluded that the flows' label in the ground-truth dataset should carry what content (e.g., video, Web pages or YouTube thumbnail images) the flow carries. Garcia et al. [11] improved Ground Truth techniques by applying unsupervised clustering methods on DPI-labeled traffic. However, the method is designed for offline analysis only.

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_21

Video title identification enjoys the most attention. Advanced researches choose to establish the video fingerprints database [9, 29, 13]. Video fingerprint is a type of information leakage caused by DASH and Variable Bit Rate. The common flaw is that building video fingerprint is hard to cope with the growing number of video titles. Besides, the fingerprint-based method can only identify the known video, and the best result for video outside the training set is to mark as unknown.

As early as 2016, Dubin et al. [9] first identified YouTube video titles, assuming that adversaries can directly observe encrypted video flows at the network layer, yet their detectors are susceptible to noise. Schuster et al. [29] extended the attack scenario to one where direct eavesdropping was not feasible. They used a neural network to identify video titles, and there were no false positives for video outside the training set. Recently, Gu et al. [13] presented a method to build video fingerprint by using differentiated bit rate, which eliminated the impact of the switching of video resolution.

3 Preliminaries

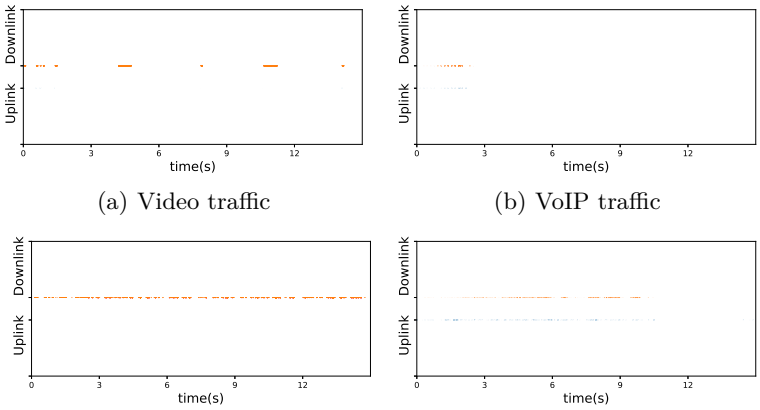
3.1 Streaming Protocols

As the popularity of Internet video streaming services has increased, streaming protocols have developed several generations to make the utmost of advanced networking techniques. Nowadays, HTTP-based adaptive streaming protocols become the most popular because of the advantages of easy to deploy and firewall penetration [21, 3, 18]. Video-related companies or institutions (e.g., Netflix, Adobe, Apple, and Microsoft) have all proposed HTTP-based adaptive streaming protocols[19].

HTTP-based adaptive streaming protocols are designed to provide smooth video playback [21]. With the adoption of HTTP-based adaptive streaming protocols, the video resolution can be dynamically adjusted according to the network condition. Moreover, HTTP-based adaptive streaming protocol can avoid the waste of bandwidth resources and reduce network congestion. For example, the customer may not watch the entire video. In this case, the video with such protocols is periodically delivered in segments if the customer continues watching.

Streaming protocols have the above-mentioned advantages due to the segmented transmission of video. As shown in Figure 2, video is divided into several video segments, and each video segment usually contains content that lasts for a few seconds. The client request one video segment at a time, and then the server will send the video segment. Therefore, the video resolution can be adjusted as soon as the network condition changes. Moreover, video transmission can be suspended immediately when the customer does not continue watching the video.

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_21



(a) Video traffic (b) VoIP traffic

(c) File transfer traffic (d) Browser traffic

Fig. 3. Packet arrival times of different traffic. We ignore the packet size, and all packets appear as points on the horizontal line. Compared with not-video traffic, video traffic shows an intermittent pattern (Downstream). Besides, intervals of upstream requests of video traffic are related to the duration of video segments, consistent with the traffic model proposed by Waldmann et al. [34].

3.2 The intermittent traffic pattern

As described in Section 3.1 and Figure 2, video traffic exhibits the intermittent pattern because of the segmented transmission. To be transmitted over the network, one video segment will be divided into several IP packets. Consisting of many ON periods and OFF periods, the intermittent traffic pattern seem like a faucet: when the faucet is turned on (ON periods), the water flooding, and when the faucet is turned off (OFF periods), transmission suspend. During the ON period, a video segment is transmitted, resulting in the continuous arrival of packets; during the OFF period, video transmission will not be performed. Through Packet Arrival Interval (PAI), we can observe the intermittent traffic pattern of video more intuitive, as shown in Figure 2. Moerover, the intermittent traffic pattern can be observed in most HTTP video streaming [21], such as video streaming generated by different video service providers, e.g., Youtube [30] and Netflix [1].

We use PAI to describe the intermittent traffic pattern, as shown in Figure 3. There is a gap between two video segments because one video segment is transmitted after the sever received a request from the client. As a result, the arrival interval of two successive packets that belong to the same video segment is smaller than the interval between two successive video segment transmission. Therefore, PAI obtained in ON periods is generally smaller than PAI obtained in OFF periods. As a consequence, the distribution of the PAI of video traffic is characteristic. The probability distribution image of PAIs will contain two peaks. The highest peak consist of smaller PAIs (ON period-related), and the

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_21

other one consist of PAIs that are related to the interval between two successive video segment transmission.

3.3 State-of-the-art method

Silhouette Silhouette, presented by Li et al., is a method to detect YouTube videos from network traffic dataflow. The method has several steps. First, two traffic features (i.e., average downstream payload size, data rate) are extracted. Second, Application Data Units (ADUs) are detected based on two thresholds (i.e., segment length threshold for video ADU, packet length threshold for upstream request). Third, three thresholds are used to determine a flow as a YouTube video. Thresholds are tuned from observing hundreds of YouTube video sessions.

Within encrypted traffic classification, Silhouette is a rare exception, which achieves high recall and none false positives without machine learning algorithms. Therefore, we choose it as one of the state-of-the-art methods.

Machine Learning-based methods

Candidate methods Since it was applied for traffic classification in 2004 [20], Machine Learning (ML)-based traffic analysis has been attracting much interest [22, 4]. Different researchers use different methodological datasets to evaluate their methods [33]. As a result, their methods are not directly comparable [33], increasing the difficulty of selecting a method as the state-of-the-art method. Below, we describe three types of traffic features that are often used in traffic classification.

- **Time-related features.** Time-related features are a set of time-related features (e.g., packets per second, bytes per second, flow duration) that first proposed by draper et al. [8].
- **Network path-related features.** Network path-related features refer to the distribution of Packet Arrival Intervals (PAIs). The distribution of PAIs is proved related to the network path [30].
- **Raw traffic.** The first few packets of the flow are observed enough for traffic identification [25]. Besides, wang et al. use the first 784 bytes of the flow with Convolutional Neural Network (CNN) algorithms, and finally get a better result than conventional ML algorithms in encrypted traffic identification [36, 35]. Therefore, we use the first 784 bytes of the flow as a type of traffic features, marked as raw traffic.

We choose KNN and CNN as two candidate algorithms, the reasons list as follows:

- KNN. KNN, representatives for conventional machine learning algorithms, has been proved effective in encrypted traffic classification [22, 30, 9]. Moreover, it behaves better than other algorithms (e.g., SVM, J48) in some tasks [30, 8] (e.g., video source identification).

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_21

– CNN. CNN, a typical algorithm of deep learning, is good at classifying sequential data (network traffic can be seen as sequential data) [14, 12]. Besides, wang et al. find CNN better than conventional ML algorithms in application-level traffic identification [36, 35]. Moreover, CNN was proved to have an excellent result in video title identification [29].

Table 1. Candidate methods vs. our method (ITP-KNN).

Mark	Feature	Algorithm
Draper-Gil et al.'s method	Time-related features	KNN
Shi et al.'s method	Network path-related features	KNN
Wang et al.'s method	Raw traffic	CNN
ITP-KNN (our method)	Intermittent traffic pattern Pattern-realted features	KNN

Finally, we get three Candidate methods, list in Table. 1 .

Table 2. The F_1 score of candidate methods (%).

	K=1	K=3	K=5	K=7	K=9
Draper-Gil et al.'s method	96.47	96.96	96.95	97.05	97.00
Shi et al.'s method	97.62	97.95	98.24	98.39	98.49
Wang et al.'s method	4.27 (recall=43.8%,precision=2.24%)				

Final battle We select 94500 flows randomly from five public datasets [10, 9, 18, 8, 17]. We use these flows as a dataset to evaluate candidate methods. 10-fold cross-validation was used when tested the performance, and all the features are extracted at the end of the flow. Results are listed in Table 2. Shi et al.'s method performs best; therefore, we choose Shi et al.'s method as one of the state-of-the-art methods to judge our methods.

4 Methods

4.1 Overview

We compare our method with two state-of-the-art methods. Therefore, we implemented three methods, as shown in Figure 4. The three methods are described as following. First, our method, we develop a set of intermittent traffic pattern-related (ITP) features and use the K-nearest neighbors (KNN) algorithm. Thus, our method is called ITP-KNN. Second, we choose Silhouette as one of the state-of-the-art methods, since Silhouette is a training-free method with high recall and none false alarm[18]. Third, we evaluated three kinds of encrypted traffic identification methods. According to the evaluation results, we choose Shi et

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_21

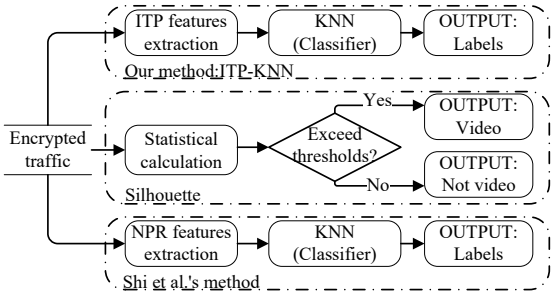


Fig. 4. Overview: our method vs. state-of-the-art methods. We develop a set of intermittent traffic pattern-related (ITP) features. We compare our method (ITP-KNN), with two state-of-the-art methods (Silhouette, and Shi et al.'s method). Firstly, our method ITP-KNN uses K-nearest neighbors (KNN) algorithm and ITP features to identify video flows in encrypted traffic. Secondly, Silhouette is a training-free method that identifies YouTube video flows based on some thresholds and heuristics [18]. Thirdly, Shi et al.'s method uses machine learning algorithms and a set of network path-related (NPR) features. As a result, it needs to retrain regularly [30].

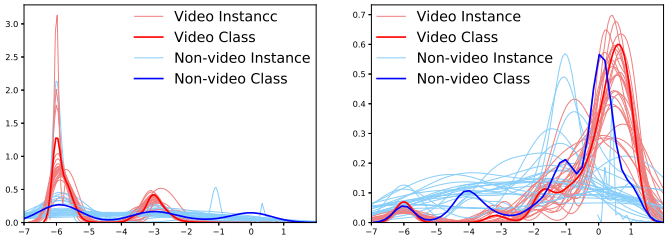
al.'s method (marked as Shi et al.'s method) as the other of the state-of-the-art methods. Shi et al.'s method achieves an application-level traffic identification with extremely high recall [30].

- **ITP-KNN**. The core of this part is feature extraction. We develop a set of features to highlight the intermittent traffic pattern and identify video flows based on these features. ITP-KNN is described in Section 4.2.
- **Silhouette** (detailed in Section 3.3). Silhouette is a training-free light-weight method. The experiment results show that Silhouette can identify YouTube video flows in encrypted traffic with high recall (QUIC-based streaming: 99%) and a zero false-positive rate [18].
- **Shi et al.'s method** (detailed in Section 3.3). Shi et al. [30] proposed a method that can identify video traffic sources (in other words, application software) at a client-side firewall. Furthermore, their method achieves the best average true positive rate at 94% while using the Nearest Neighbor algorithm [6]. For the sake of fairness, we modify the algorithm used in this method: we change the Nearest Neighbor algorithm to the KNN algorithm.

4.2 Our framework: KNN that uses ITP features as input

Feature Set: ITP features As described in Section 3.2, PAI can be used to describe the intermittent pattern of video traffic. There is a series of small PAIs (continuous packets arrival, ON period) followed by a large PAI (the duration of OFF period). Therefore, we can observe differences between video class and non-video class in the distribution of PAI. We randomly selected 50 video instances and 50 non-video instances and drawn their probability density function in Figure 5.

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_21



(a) Downstream (b) Upstream

Fig. 5. Distribution of log(PAI). Non-video instances include file transfer. There are two peaks in video class (downstream), which conform to our inference that video traffic shows intermittent pattern. Moreover, the log(PAI) corresponding to the highest peak in video class (upstream) is between 0 and 1, which is related to the duration of a video segment (1 to 10 seconds).

Moerover, video traffic is asymmetric, especially for the amount of the transmitted data. Downstream (client-to-server direction) of the video traffic and upstream (server-to-client direction) of the video traffic both show an intermittent pattern, but, they have differences, compared as follows:

- **Downstream**. During the ON period (the duration of ON period is recorded as T_{on}), a video segment is transmitted; thus, packets transmission bursts. The OFF periods are essentially the pauses between the two segments being transmitted.
- **Upstream**. A request contains few packets, and the interval between two upstream requests (recorded as T_{req}) is related to T_{on}; therefore, the data amount and transmission rate of upstream are smaller than those of downstream. (Waldmann et al. [34] found that the distribution of T_{req} is centered precisely around the duration of a video segment (T_{seg}), and T_{on} is close to but slightly less than T_{seg}, which can be observed in Figure 3 and 5 as well).

Table 3. Description of intermittent traffic pattern-related (ITP) features.

Feature category		Direction	Description
Quantized PAIs		Upstream	23 bin, the first one covering 0 to 5µs, and each subsequent bin has an upper limit two times larger than the previous one
		Downstream	23 bin, the first one covering 0 to 5µs, and each subsequent bin has an upper limit two times larger than the previous one
Summary statistics	Simple descriptive statistics	Upstream	maximum, minimum, and mean over the PAI time series; the two largest values and their bin number of the Quantized PAIs
		Downstream	maximum, minimum, and mean over the PAI time series; the two largest values and their bin number of the Quantized PAIs
	Higher-order statistics	Upstream	the skew, kurtosis, and coefficient of variation of the PAI time series
		Downstream	the skew, kurtosis, and coefficient of variation of the PAI time series
	Rates	Upstream	bytes per seconds and packets per second
		Downstream	bytes per seconds and packets per second

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_21

We use 70-dimensional features, listed in Table 3. We extract ITP features from each 15 seconds network trace and divide these features into two categories, as follows:

- **Quantized PAIs**. We use 23 bin with the first one covering 0 to 5μs, and each subsequent bin has an upper limit two times larger than the previous one. The last bin covers 10 seconds to 20 seconds. ITP-KNN processes upstream packets and downstream packets separately and concatenates the resulting features together. Therefore, Quantized PAIs are 46 dimensions.
- **Summary statistics**. Encrypted traffic provides two main sources of data: a time series of PAIs, and a time series of packet lengths [2]. This kind of feature comprises a sequence of summary statistics computed over the network traces of encrypted traffic, which is a set of features that prevalent for the problem of traffic identification [22, 8, 16]. As for the selection of summary statistics, we compute multiple descriptive statistics for upstream/ downstream traffic individually. This feature set includes simple descriptive statistics over the PAI time series - such as maximum, minimum, and mean - as well as higher-order statistics like the skew or kurtosis of the time series.

Selected Classifiers We describe the K-nearest neighbors (KNN) algorithm we have chosen for conducting our experiments as follows:

KNN is a supervised learning algorithm. Many researchers have utilized KNN to realize an application-level traffic classification and proved its better performance [22, 30, 9]. KNN performs even better than a deep learning algorithm, deep neural network, in video source identification [30, 31].

5 Evaluation and Discussion

In the experiments, we use five public datasets [10, 9, 18, 8, 17], together with some traffic captured by ourselves using Wireshark. In totally, we collected more than 450 thousand flows (about 350G). All the traffic was encrypted (HTTPS, QUIC, VPN, Tor, Shadowsocks).

From the above traffic traces, we randomly select 20010 video flows and 282650 not-video flows as our dataset. We use these 302660 flows to evaluate our method, i.e., ML-based TA that uses PR features (marked as ITP-KNN). We compare our method, ITP-KNN, with two state-of-the-art methods, Silhouette (training-free method) and Shi et al.'s method (performed best among ML-based methods in our evaluation).

Video flow identification requires high timely ability, which means the response time of the identification method should be as short as possible. For example, ISPs need to identify video flows in time in order to provide differentiated service. The response time of the identification method, affected mainly by the feature extraction time [8], shows how timely the method is. Therefore, we set the feature extraction time (flow duration) at 15 seconds (shorter than Shi et al.'s method).

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_21

Table 4. Our method (ITP-KNN) vs. two state-of-the-art methods (%).

	Silhouette			Shi et al.'s method			ITP-KNN		
	R	P	F_1	R	P	F_1	R	P	F_1
K=1	84.65	99.63	91.53	92.25	92.07	92.16	97.46	97.32	97.39
K=3				93.95	94.00	93.97	97.36	97.76	97.56
K=5				94.60	94.13	94.36	97.36	97.90	97.63
K=7				94.75	94.00	94.37	97.55	97.95	97.75
K=9				95.00	94.06	94.53	97.82	98.18	98.00

We used 10-fold cross-validation in the evaluation. Table 4 depicts the recall (R), precision (P), and F_1 score (F_1) obtained by using our method (ITP-KNN), and two state-of-the-art methods. Next, we present our main findings.

1. The heuristic method that does not use machine learning algorithms (Silhouette) possesses a limited capability for identifying video flows. This finding is supported by the fact that Silhouette attains a recall of 84.65% (far lower than ML-based methods). We regard recall as the most important metric for the reason that high recall means that a method identified most of the video flows. For example, ISPs need to provide differentiated service guarantees, which require the ability to identify video flows as much as possible.

2. The application-level traffic identification method shows promising results for the identification of video flows (content-level). Results in Table 4 show that Shi et al.'s method performs well in this task. To sum up, these results were very encouraging, which means the present traffic identification methods can be applied to a more fine-grained identification task. However, when the identification task requires high recall and precision, the present traffic identification methods need to be improved.

3. Fine-grained traffic identification task demand for in-depth analysis of traffic. The results in Table 4 suggest this finding: our method (ITP-KNN) behaves best overall in the specific fine-grained traffic identification task (identification of video flows). Our method is based on an in-depth analysis of video traffic: we develop a set of features to stress out the unique transport pattern of video, the intermittent traffic pattern.

6 Conclusion

Experimental evaluations prove the advantages of our method; compared with baseline, our method is better overrall. The results show that our method can identify video flows from encrypted traffic.

Our framework is based on one assume: most online video traffic shows the intermittent pattern. Though theoretically universal, it has a flaw that it is only applicable to identify video traffic that generated by present streaming protocols. Though better than the baseline in universality, our method still has a margin of improvement. In future work, we will apply our method to the actual engineering to realize video flows identification in high-speed and large-scale network traffic.

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_21

References

1. Ameigeiras, P., Ramos-Munoz, J.J., Navarro-Ortiz, J., Lopez-Soler, J.M.: Analysis and modelling of youtube traffic. Transactions on Emerging Telecommunications Technologies **23**(4), 360–377 (2012)
2. Barradas, D., Santos, N., Rodrigues, L.: Effective detection of multimedia protocol tunneling using machine learning. In: 27th {USENIX} Security Symposium ({USENIX} Security 18). pp. 169–185 (2018)
3. Bitmovin: Bitmovin video developer report 2019. `https://go.bitmovin.com/video-developer-report-2019` (2019), accessed January 1, 2020
4. Cao, Z., Xiong, G., Zhao, Y., Li, Z., Guo, L.: A survey on encrypted traffic classification. In: International Conference on Applications and Techniques in Information Security. pp. 73–81. Springer (2014)
5. Casas, P., Mazel, J., Owezarski, P.: Minetrac: Mining flows for unsupervised analysis & semi-supervised classification. In: Proceedings of the 23rd International Teletraffic Congress. pp. 87–94. International Teletraffic Congress (2011)
6. Cover, T., Hart, P.: Nearest neighbor pattern classification. IEEE transactions on information theory **13**(1), 21–27 (1967)
7. Dainotti, A., Pescape, A., Claffy, K.C.: Issues and future directions in traffic classification. IEEE network **26**(1), 35–40 (2012)
8. Draper-Gil, G., Lashkari, A.H., Mamun, M.S.I., Ghorbani, A.A.: Characterization of encrypted and vpn traffic using time-related. In: Proceedings of the 2nd international conference on information systems security and privacy (ICISSP). pp. 407–414 (2016)
9. Dubin, R., Dvir, A., Pele, O., Hadar, O.: I know what you saw last minute—encrypted http adaptive video streaming title classification. IEEE Transactions on Information Forensics and Security **12**(12), 3039–3049 (2017)
10. Dubin, R., Dvir, A., Pele, O., Hadar, O., Richman, I., Trabelsi, O.: Real time video quality representation classification of encrypted http adaptive video streaming-the case of safari. arXiv preprint arXiv:1602.00489 (2016)
11. Garcia, J., Brunstrom, A.: Clustering-based separation of media transfers in dpi-classified cellular video and voip traffic. In: 2018 IEEE Wireless Communications and Networking Conference (WCNC). pp. 1–6. IEEE (2018)
12. Goodfellow, I., Bengio, Y., Courville, A.: Deep learning. MIT press (2016)
13. Gu, J., Wang, J., Yu, Z., Shen, K.: Walls have ears: Traffic-based side-channel attack in video streaming. In: IEEE INFOCOM 2018-IEEE Conference on Computer Communications. pp. 1538–1546. IEEE (2018)
14. Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., et al.: Recent advances in convolutional neural networks. Pattern Recognition **77**, 354–377 (2018)
15. H3C_Technologies: H3c s5120-si series ethernet switches configuration guide-release. `http://www.h3c.com.hk`, accessed December 10, 2019
16. Hayes, J., Danezis, G.: k-fingerprinting: A robust scalable website fingerprinting technique. In: 25th {USENIX} Security Symposium ({USENIX} Security 16). pp. 1187–1203 (2016)
17. Lashkari, A.H., Draper-Gil, G., Mamun, M.S.I., Ghorbani, A.A.: Characterization of tor traffic using time based features. In: ICISSP. pp. 253–262 (2017)
18. Li, F., Chung, J.W., Claypool, M.: Silhouette: Identifying youtube video flows from encrypted traffic. In: Proceedings of the 28th ACM SIGMM Workshop on Network and Operating Systems Support for Digital Audio and Video. pp. 19–24. ACM (2018)

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_21

19. Martin, J., Fu, Y., Wourms, N., Shaw, T.: Characterizing netflix bandwidth consumption. In: 2013 IEEE 10th Consumer Communications and Networking Conference (CCNC). pp. 230–235. IEEE (2013)
20. McGregor, A., Hall, M., Lorier, P., Brunskill, J.: Flow clustering using machine learning techniques. In: International workshop on passive and active network measurement. pp. 205–214. Springer (2004)
21. Mustafa, I.B., Uddin, M., Nadeem, T.: Understanding the intermittent traffic pattern of http video streaming over wireless networks. In: 2016 14th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt). pp. 1–8. IEEE (2016)
22. Nguyen, T.T., Armitage, G.J.: A survey of techniques for internet traffic classification using machine learning. IEEE Communications Surveys and Tutorials **10**(1-4), 56–76 (2008)
23. Raymond, J.F.: Traffic analysis: Protocols, attacks, design issues, and open problems. In: Designing Privacy Enhancing Technologies. pp. 10–29. Springer (2001)
24. Reed, A., Aikat, J.: Modeling, identifying, and simulating dynamic adaptive streaming over http. In: 2013 21st IEEE International Conference on Network Protocols (ICNP). pp. 1–2. IEEE (2013)
25. Rezaei, S., Liu, X.: Deep learning for encrypted traffic classification: An overview. IEEE communications magazine **57**(5), 76–81 (2019)
26. Sandvine: Covid-19 global internet trends. `https://www.sandvine.com/covid-19-trends`, accessed April 3, 2020
27. Sandvine: The global internet phenomena report september 2019 (2019)
28. Sandvine: The mobile internet phenomena report february 2020 (2020)
29. Schuster, R., Shmatikov, V., Tromer, E.: Beauty and the burst: Remote identification of encrypted video streams. In: 26th {USENIX} Security Symposium ({USENIX} Security 17). pp. 1357–1374 (2017)
30. Shi, Y., Biswas, S.: Protocol-independent identification of encrypted video traffic sources using traffic analysis. In: 2016 IEEE International Conference on Communications (ICC). pp. 1–6. IEEE (2016)
31. Shi, Y., Biswas, S.: A deep-learning enabled traffic analysis engine for video source identification. In: 2019 11th International Conference on Communication Systems & Networks (COMSNETS). pp. 15–21. IEEE (2019)
32. Shi, Y., Ross, A., Biswas, S.: Source identification of encrypted video traffic in the presence of heterogeneous network traffic. Computer Communications **129**, 101–110 (2018)
33. Velan, P., Čermák, M., Čeleda, P., Drašar, M.: A survey of methods for encrypted traffic classification and analysis. International Journal of Network Management **25**(5), 355–374 (2015)
34. Waldmann, S., Miller, K., Wolisz, A.: Traffic model for http-based adaptive streaming. In: 2017 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). pp. 683–688. IEEE (2017)
35. Wang, W., Zhu, M., Wang, J., Zeng, X., Yang, Z.: End-to-end encrypted traffic classification with one-dimensional convolution neural networks. In: 2017 IEEE International Conference on Intelligence and Security Informatics (ISI). pp. 43–48. IEEE (2017)
36. Wang, W., Zhu, M., Zeng, X., Ye, X., Sheng, Y.: Malware traffic classification using convolutional neural network for representation learning. In: 2017 International Conference on Information Networking (ICOIN). pp. 712–717. IEEE (2017)
37. Wang, Z.: The applications of deep learning on traffic identification. BlackHat USA **24** (2015)

ICCS Camera Ready Version 2020
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-50417-5_21