Traffic Pattern Plot: Video Identification in Encrypted Network Traffic

Ali S. Kamal[1], Syed M. A. H. Bukhari[1] Muhammad U. S. Khan[1], Tahir Maqsood[1], and Muhammad A B Fayyaz[2]

[1] Department of Computer Science, COMSATS University Islamabad, Abbottabad, Pakistan
alisherkamal13101@gmail.com, {ammar,ushahid,tmaqsood}@cuiatd.edu.pk
[2] OTEHM, Manchester Metropolitan University Manchester, United Kingdom
m.fayyaz@mmu.ac.uk

Abstract. Most of the internet traffic is encrypted and it is a challenge to identify streaming videos in the internet traffic. In this paper, we present a methodology named Traffic Pattern Plot (TPP) to identify video streams in encrypted network traffic. The proposed methodology plots the video traffic flows and uses a Convolutional Neural Network (CNN) to detect the videos. The results show that the traffic pattern plot generated from 120 seconds of sniffing network traffic is enough to identify the video even in the encrypted network traffic with 94% accuracy.

Keywords: Traffic Pattern, VPN traffic classification, YouTube Video identification, image classification

1 Introduction

In recent years, video identification in internet traffic has become an important research area [4, 7, 17, 31]. It is helpful to profile clients based on their online activities. Similarly, profiling users is also helpful in tracking down militants and adversaries. However, with Virtual Private Networks (VPNs), clients further hide their identity, making it difficult for security agencies to identify militants. Previous works on video identification use handcrafted features used by traditional machine learning algorithms such as Naïve Bayes (NB) and Support Vector Machine (SVM) to identify videos in Internet traffic.

In recent years, the Convolutional Neural Network (CNN) has shown an advantage over traditional machine learning algorithms. CNN has shown its success in many areas, such as medical imaging, pattern recognition, human activity recognition [8, 9, 15], graph classification [10, 29], data mining [34, 35], and natural language processing [1, 11, 16, 27, 28, 30]. One of the key applications of CNN is image processing. The image processing by CNN is fairly mature, allowing it to run on devices with low computation power, such as mobile phones.

In this work, we present a method for video identification in encrypted internet traffic using the traffic flow as an image similar to the method presented in [13]. For this purpose, we capture the video traffic flow and plot the Traffic Pattern Plot (TPP). The CNN is trained on the TPPs and is used to identify the videos in the internet traffic. The method can detect the type of traffic, i.e., VPN or non-VPN, with a high accuracy of 98%. Moreover, different traffic flows of video qualities are detected with an accuracy of 88%, 94%, and 87% for auto quality, 360p, and 1080p videos.

Our main contribution is a generic method for video identification in encrypted internet traffic. The method uses the packet size per arrival time to plot the traffic. Moreover, rather than capturing or plotting the whole video, we just plotted the videos of 120 seconds. Moreover, instead of plotting the bi-directional traffic flow, we only consider the unidirectional flow of traffic. The same CNN model is used for all types of dataset in the paper, and no attempt is made to gain additional accuracy by tuning the hyperparameters of the CNN model for a specific type of traffic flow. Unlike the techniques [19,20,36] that rely on the payload content of the packet to classify the traffic flow, our proposed method relies on the size of the packet per arrival time. Moreover, the storage requirement of our method is quite minimal because the size of the Maximum Transmission Unit (MTU) is 1500, which requires only four bytes to store the value.

To summarize, the main contributions of this paper are as follows.

- A visualization technique of the video traffic pattern.
- A method that uses traffic patterns as images to identify videos in encrypted internet traffic.

The paper organization is as follows. Section II presents the literature review and Section III presents the details of the creation of TTPs. The method of video detection using the TPPs is discussed in Section IV. Section V presents details of the results and experiments performed on different datasets, and lastly, Section VI concludes the paper.

2 Related Work

For the last couple of years, two primary internet traffic problems, internet traffic categorization and internet traffic classifications, have been studied. There are a variety of methods for resolving these issues, such as statistics-based methods, Machine Learning, Deep Packet Inspection (DPI) [21, 24]. The DPI methods are computationally extensive and, due to encryption, cannot handle the majority of today's internet traffic. However, size and time-related characteristics are manually extracted in machine learning methods, and complicated patterns or supervised learning algorithms are used as classifiers. Moore et al. [22, 23] generated a list of 248 descriptors based on a bidirectional flow in 2005, including Round-Trip Time (RTT) statistics, frequencies, packets inter-arrival time statistics, size-based statistics. To categorize the internet traffic, they used a Naive Bayes Kernel estimator.

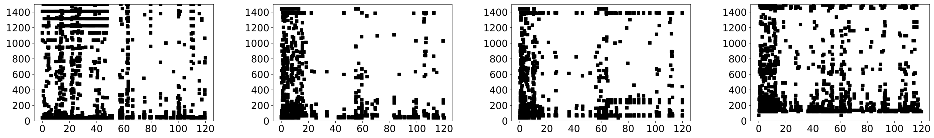
The work presented in [6] extended the work of Moore et al. [22, 23] by suggesting an integrated feature selection strategy by selecting five feature selection strategies to obtain an ideal set of features. Many works rely on flow-based features (such as time and size-related features). Gil et al. [3] classify the encrypted internet traffic using the C4.5 and K-nearest neighbor (KNN) classifier. The author leveraged the flow-based time-related features such as flow duration, flow bytes per second, flow packets per second, flow inter-arrival time, etc., and got the accuracy above 80%. The correlation information in traffic flows of the same application is described by utilizing the Bag of Words (BoF) technique presented by Zhang et al. [38,39]. The authors used 20 uni-directional flow-based statistical features. They presented a novel Resilient Traffic Classification (RTC) technique based on their BoF-based traffic classifier and demonstrated that it outperforms the most common machine learning methods. Neural networks have been used in several recent studies. Ertam et al. [5] employed a Genetic Algorithm (GA) to select features from a collection of twelve attributes. They apply Extreme Learning Machines (ELM) to a dataset containing seven classes of regular traffic, achieving a 95% accuracy.

Several papers have proposed a new technique based on the content of the payload from the packet. The content of each packet is converted to a sequence of bytes and given input to Artificial Neural Networks (ANN) for internet traffic classification is presented by Wang et al. [36]. Chen et al. [2] converted the network traffic data into an image of the flow parameter and input it to a neural network. They used information such as destination IP and did not fully detail their method to allow comparison. One of the first works to automatically select the feature is presented by Qin et al. [26]. The authors understand the relevance of relying on packets' Payload Size Distribution (PSD) probability in a bi-directional flow rather than manually generated features or handcrafted feature selection.

Video identification in encrypted Internet traffic has been studied in the last two decades [12, 14, 24, 25, 37]. The first and most related to our work that plot the traffic flow as an image and classify among different type of internet traffic is presented by Shapira et al. [32]. The authors used the ISCXVPN2016 dataset and extracted the traffic flow from each type of Internet traffic. The flow is plotted into images called FlowPics, and CNN is used to classify different internet traffic types. However, our work is different from the aforementioned work, as we classify the traffic flow of different videos. For this purpose, we captured the video traffic flow for 120 seconds with three different qualities and two traffic modes. The detail of the pattern plot is presented in Section III.

3 Traffic Pattern Plot

We setup an automated client that mocks the actual client playing the YouTube videos to evaluate the proposed method. For automation, we use selenium and WebDriver for Google Chrome. To capture the online stream of the videos, we use the Wireshark tool that captures the packets in the form of a Packet Capture file (PCAP). The capture is done for 120 seconds for each dataset. We created four types of datasets, including 1080p, 360p, auto-quality mode, and



(a) (b) (c) (d)

Fig. 1: Examples of Traffic pattern plots (a) 1080p video quality, (b) 360p video quality, (c) auto video quality, and (d) auto quality with VPN

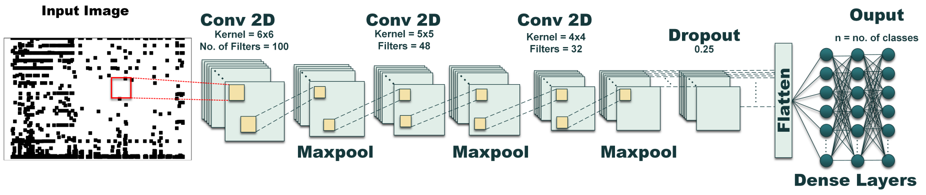


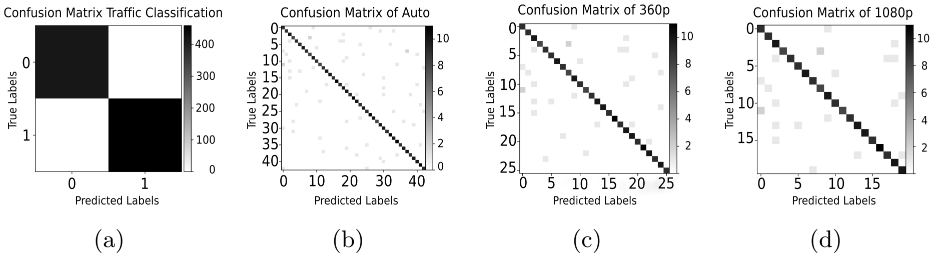
Fig. 2: The CNN model detail

auto-quality with VPN. For VPN, we use a desktop client, SurfShark, that uses the WireGuard protocol to further encrypt the internet traffic.

The PCAP files in our dataset represent each corresponding stream of a specific video. After downloading the data, the PCAP files are then preprocessed using Scapy (a Python library for packet manipulation of internet traffic). ACK, SYN, FIN, DNS, Ethernet header, and no payload packets are removed during the pre-processing step. The payload packet size that is the actual data of the video stream is extracted against the arrival time. Our objective is to create images of each PCAP file of the corresponding video. For this purpose, the payload packets are plotted into a scatter plot, traffic pattern plot upon arrival time. Each video stream file is plotted in a separate TPP. For uniformity, all the packet arrival time is normalized to 120 seconds as each PCAP of the video stream is captured at different time. After plotting the packet size, the TPPs are used to train the CNN model. Examples of the TPP are shown in Figure 1. We train a sequential convolutional neural network (SCNN), as presented in [17,18] on our datasets. The model diagram is presented in Figure 2.

4 Experiments and Results

This section presents the results of our experiments. For this purpose, we created a dataset containing the TTPs of different traffic flows, as mentioned in Section 3. We present a video classification method that uses image processing that has not been discussed in the previous literature. Therefore, we compare our results with the techniques presented in [17,18] and VGG16 [33] for image classification. The entire setup is deployed on Google Colab to train the SCNN model. The



(a) (b) (c) (d)

Fig. 3: Confusion Matrix of (a) Traffic Classification, (b) auto quality dataset, (c) 360p dataset, and (d) 1080p dataset

SCNN model is tuned by varying different parameter settings. Various settings are applied by adding more dense layers to the model, increasing the dropout to 0.25, and changing the filter size of the first layer to 100.

4.1 Creating separate dataset for different problems

To create our dataset, we capture the traffic stream in three different qualities 1080p, 360p, and auto quality. We use two different traffic modes to capture the traffic, i.e., VPN and Non-VPN, for auto quality. For auto quality traffic mode, we use 43 class samples of YouTube videos and capture each video traffic 55 times, making a total of 2356 traffic streams of videos. The same procedure is repeated for auto quality with using a VPN. We also created a balanced dataset with two classes, VPN and Non-VPN, containing 4712 streams. To create a 1080p traffic dataset, we selected 20 YouTube videos and captured 55 streams of each video, making a total of 1100 traffic stream files. The same process is repeated for creating 360p video files. However, the number of videos in the 360p dataset is 26, and each video is captured 50 times. We split the dataset for each type of traffic into a train and a test set. For this purpose, we split the data into 80-20, 80% of each dataset is used for training the model and the remaining 20% of the data is used for testing purpose. We use the basic accuracy metric to evaluate the performance of our model. In this scenario, the definition of accuracy is defined as provided by [17].

4.2 Results and Discussion

In this section we present the results of our proposed technique on the classification of traffic and different quality dataset. The results show that our proposed technique outperforms its counterpart with a high accuracy of 98%. In the traffic classification problem, our proposed technique performs 10% better than the BPS model. The accuracy of BPS model and VGG16 in the traffic classification problem is 3% and 89% respectively. Similarly, our proposed technique also outperforms the other techniques in different qualities dataset. The accuracy of our proposed technique in auto, 360p, and 1080p quality are 88%, 94%, and 87%

respectively. However, the accuracy of VGG16 is less than the TPP model. The accuracy of VGG16 for auto, 360p, and 1080p quality is 5%, 11%, and 17%, respectively. Similarly, the accuracy of the BPS model is 86%, 82%, and 88% for auto, 360p, and 1080p video quality. The confusion matrix of the aforementioned results is presented in Figure 3.

5 Conclusions

This paper presents a video identification technique in encrypted internet traffic using image classification. For this purpose, we captured the video traffic stream, extracted the packet size per arrival time, and plotted the packet size as Traffic Pattern Plot. The images are used to train the CNN model on different types of datasets including, the Traffic dataset, 1080p, 360p, and auto quality dataset. The proposed technique outperforms the other classification techniques. Moreover, the cost of storage, resources, and memory of our technique is low. The results have shown that videos can be identified in the encrypted internet traffic with high accuracy of 98%.

6 Acknowledgment

This material is based on work supported by the National Cyber Security Center (NCCS), Pakistan, and Higher Education Commission (HEC) under grant RF-NCCS-023.

References

1. Arshad, N., Bakar, A., Soroya, S.H., Safder, I., Haider, S., Hassan, S.U., Aljohani, N.R., Alelyani, S., Nawaz, R.: Extracting scientific trends by mining topics from call for papers. Library Hi Tech (2019)
2. Chen, Z., He, K., Li, J., Geng, Y.: Seq2img: A sequence-to-image based approach towards ip traffic classification using convolutional neural networks. In: 2017 IEEE International conference on big data (big data), pp. 1271–1276. IEEE (2017)
3. 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)
4. Dvir, A., Marnerides, A.K., Dubin, R., Golan, N.: Clustering the unknown - the youtube case. In: 2019 International Conference on Computing, Networking and Communications (ICNC), pp. 402–407 (2019). DOI 10.1109/ICCNC.2019.8685364
5. Ertam, F., Avcı, E.: A new approach for internet traffic classification: Ga-wk-elm. Measurement **95**, 135–142 (2017)
6. Fahad, A., Tari, Z., Khalil, I., Habib, I., Alnuweiri, H.: Toward an efficient and scalable feature selection approach for internet traffic classification. Computer Networks **57**(9), 2040–2057 (2013)

7. 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)
8. Hassan, H., Bashir, A.K., Ahmad, M., Menon, V.G., Afridi, I.U., Nawaz, R., Luo, B.: Real-time image dehazing by superpixels segmentation and guidance filter. Journal of Real-Time Image Processing **18**(5), 1555–1575 (2021)
9. Hassan, S.U., Saleem, A., Soroya, S.H., Safder, I., Iqbal, S., Jamil, S., Bukhari, F., Aljohani, N.R., Nawaz, R.: Sentiment analysis of tweets through altmetrics: A machine learning approach. Journal of Information Science **47**(6), 712–726 (2021)
10. Hassan, S.U., Shabbir, M., Iqbal, S., Said, A., Kamiran, F., Nawaz, R., Saif, U.: Leveraging deep learning and sna approaches for smart city policing in the developing world. International Journal of Information Management **56**, 102,045 (2021)
11. Iqbal, S., Hassan, S.U., Aljohani, N.R., Alelyani, S., Nawaz, R., Bornmann, L.: A decade of in-text citation analysis based on natural language processing and machine learning techniques: An overview of empirical studies. Scientometrics **126**(8), 6551–6599 (2021)
12. Khalife, J., Hajjar, A., Díaz-Verdejo, J.: Performance of opendpi in identifying sampled network traffic. Journal of Networks **8**(1), 71 (2013)
13. Khan, M., Baig, D., Khan, U.S., Karim, A.: Malware classification framework using convolutional neural network. In: 2020 International Conference on Cyber Warfare and Security (ICCWS), pp. 1–7 (2020). DOI 10.1109/ICCWS48432.2020.9292384
14. Khan, M.U., Bukhari, S.M., Maqsood, T., Fayyaz, M.A., Dancey, D., Nawaz, R.: Scnn-attack: A side-channel attack to identify youtube videos in a vpn and non-vpn network traffic. Electronics **11**(3), 350 (2022)
15. Khan, M.U.S., Abbas, A., Ali, M., Jawad, M., Khan, S.U.: Convolutional neural networks as means to identify apposite sensor combination for human activity recognition. In: 2018 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE), pp. 45–50 (2018)
16. Khan, M.U.S., Abbas, A., Rehman, A., Nawaz, R.: Hateclassify: A service framework for hate speech identification on social media. IEEE Internet Computing **25**(1), 40–49 (2021). DOI 10.1109/MIC.2020.3037034
17. Khan, M.U.S., Bukhari, S.M.A.H., Ali, S., Maqsood, T.: Isp can identify youtube videos that you just watched. In: 18th International Conference on Frontiers of Information Technology (FIT). IEEE (2021)
18. Khan, M.U.S., Bukhari, S.M.A.H., Maqsood, T., Fayyaz, M.A.B., Dancey, D., Nawaz, R.: Scnn-attack: A side-channel attack to identify youtube videos in a vpn and non-vpn network traffic. Electronics **11**(3) (2022). DOI 10.3390/electronics11030350. URL https://www.mdpi.com/2079-9292/11/3/350
19. Lopez-Martin, M., Carro, B., Sanchez-Esguevillas, A., Lloret, J.: Network traffic classifier with convolutional and recurrent neural networks for internet of things. IEEE Access **5**, 18,042–18,050 (2017)
20. Lotfollahi, M., Siavoshani, M.J., Zade, R.S.H., Saberian, M.: Deep packet: A novel approach for encrypted traffic classification using deep learning. Soft Computing **24**(3), 1999–2012 (2020)
21. Mohammad, S., Khan, M.U., Ali, M., Liu, L., Shardlow, M., Nawaz, R.: Bot detection using a single post on social media. In: 2019 Third World Conference on Smart Trends in Systems Security and Sustainablity (WorldS4), pp. 215–220. IEEE (2019)
22. Moore, A., Zuev, D., Crogan, M.: Discriminators for use in flow-based classification. Tech. rep. (2013)

23. Moore, A.W., Zuev, D.: Internet traffic classification using bayesian analysis techniques. In: Proceedings of the 2005 ACM SIGMETRICS international conference on Measurement and modeling of computer systems, pp. 50–60 (2005)
24. Nguyen, T.T., Armitage, G.: A survey of techniques for internet traffic classification using machine learning. IEEE communications surveys & tutorials **10**(4), 56–76 (2008)
25. Nguyen, T.T., Armitage, G., Branch, P., Zander, S.: Timely and continuous machine-learning-based classification for interactive ip traffic. IEEE/ACM Transactions On Networking **20**(6), 1880–1894 (2012)
26. Qin, T., Wang, L., Liu, Z., Guan, X.: Robust application identification methods for p2p and voip traffic classification in backbone networks. Knowledge-Based Systems **82**, 152–162 (2015)
27. Safder, I., Hassan, S.U., Visvizi, A., Noraset, T., Nawaz, R., Tuarob, S.: Deep learning-based extraction of algorithmic metadata in full-text scholarly documents. Information processing & management **57**(6), 102,269 (2020)
28. Safder, I., Mahmood, Z., Sarwar, R., Hassan, S.U., Zaman, F., Nawab, R.M.A., Bukhari, F., Abbasi, R.A., Alelyani, S., Aljohani, N.R., et al.: Sentiment analysis for urdu online reviews using deep learning models. Expert Systems p. e12751 (2021)
29. Said, A., Hassan, S.U., Tuarob, S., Nawaz, R., Shabbir, M.: Dgsd: Distributed graph representation via graph statistical properties. Future Generation Computer Systems **119**, 166–175 (2021)
30. Sarwar, R., Zia, A., Nawaz, R., Fayoumi, A., Aljohani, N.R., Hassan, S.U.: Webometrics: evolution of social media presence of universities. Scientometrics **126**(2), 951–967 (2021)
31. 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)
32. Shapira, T., Shavitt, Y.: Flowpic: Encrypted internet traffic classification is as easy as image recognition. In: IEEE INFOCOM 2019-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), pp. 680–687. IEEE (2019)
33. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
34. Waheed, H., Anas, M., Hassan, S.U., Aljohani, N.R., Alelyani, S., Edifor, E.E., Nawaz, R.: Balancing sequential data to predict students at-risk using adversarial networks. Computers & Electrical Engineering **93**, 107,274 (2021)
35. Waheed, H., Hassan, S.U., Aljohani, N.R., Hardman, J., Alelyani, S., Nawaz, R.: Predicting academic performance of students from vle big data using deep learning models. Computers in Human behavior **104**, 106,189 (2020)
36. 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)
37. Zhang, J., Chen, C., Xiang, Y., Zhou, W., Xiang, Y.: Internet traffic classification by aggregating correlated naive bayes predictions. IEEE transactions on information forensics and security **8**(1), 5–15 (2012)
38. Zhang, J., Chen, X., Xiang, Y., Zhou, W., Wu, J.: Robust network traffic classification. IEEE/ACM transactions on networking **23**(4), 1257–1270 (2014)
39. Zhang, J., Xiang, Y., Wang, Y., Zhou, W., Xiang, Y., Guan, Y.: Network traffic classification using correlation information. IEEE Transactions on Parallel and Distributed systems **24**(1), 104–117 (2012)