



City Research Online

City, University of London Institutional Repository

Citation: Enisoglu, R. & Rakocevic, V. (2023). Low-Latency Internet Traffic Identification using Machine Learning with Trend-based Features. Paper presented at the 2023 International Wireless Communications and Mobile Computing (IWCMC), 19-23 Jun 2023, Marakesh, Morocco. doi: 10.1109/iwcmc58020.2023.10183084

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/36488/>

Link to published version: <https://doi.org/10.1109/iwcmc58020.2023.10183084>

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Low-Latency Internet Traffic Identification using Machine Learning with Trend-based Features

Ramazan Enisoglu, Veselin Rakocevic

City, University of London

London, United Kingdom

ramazan.enisoglu@city.ac.uk - veselin.rakocevic.1@city.ac.uk

Abstract—Identifying the type of network traffic has several advantages, such as detecting and preventing applications that violate an organization’s security policy or improving Quality of Service (QoS) and Quality of Experience (QoE) through traffic engineering. To enhance QoS support for Internet Service Providers (ISPs), a fine-grained classification scheme for network traffic is proposed in this paper. Statistical analysis and the throughput patterns of FTP, video conferencing, and video streaming traffic reveal that using new statistical features can be more effective at distinguishing the Internet traffic, especially from a QoS perspective, compared to the features commonly used in the literature, even for encrypted traffic. Machine Learning algorithms for classifying the low-latency traffic are trained using combinations of statistical features such as novel trend identification. Experiments are conducted to evaluate the proposed method using large-scale real network traffic data. Results show that our network can classify the single type of traffic with accuracy of over 97%, and identify the low-latency traffic in the traffic mix with accuracy of 87%.

Index Terms—Network traffic classification, k-NN, SVM, Machine Learning, Feature selection, Internet traffic mix, Statistical features, QoS, Low-Latency

I. INTRODUCTION

Internet traffic classification is a rapidly growing field, offering opportunities for intelligent management of network resources in the presence of the increasing volume of traffic requiring low-latency delivery. According to Ericsson’s forecast [1], video content is predicted to account for 80% of data usage by 2028. It is also estimated that 5 billion people will have 5G subscriptions by then. This simultaneous increase in the mobile network complexity and the performance expectations of the video applications creates challenges for network management and for ensuring QoS for multimedia services. Accurately identifying and categorizing network traffic is crucial for ISPs, mobile operators and regulators to address these challenges and optimally allocate network resources based on QoS requirements.

ISPs classify network traffic using a data-driven approach by extracting relevant features from traffic flows and using these features to identify the source application [2]. This is achieved through the use of modern Network Traffic Classification (NTC) methods that utilize advanced technologies such as Machine Learning (ML) and Artificial Intelligence (AI). The success of these NTC methods is dependent on the selection and use of representative features for discrimination, as well as the design of the feature extraction methods. These factors

play a crucial role in the overall performance and effectiveness of the NTC approach.

The research presented in this paper focuses on classifying Internet traffic into categories based on their latency requirements: low-latency applications like video conferencing and others like video streaming and file transfer that have more flexible latency requirements. This is motivated by the expectation that accurate real-time identification of the presence of low-latency traffic will enable better understanding of the necessary QoS levels and application of smart network resource management solutions, ultimately improving the end-user’s Quality of Experience (QoE). This paper introduces a low-latency Internet traffic classification scheme that utilizes machine learning algorithms trained with novel trend-related features. This scheme employs statistical analysis and data mining techniques to identify the most discriminating QoS-related features in the acquired Internet traffic data and applies the class. Our approach aims to identify low-latency traffic with a high level of granularity while preserving user privacy by avoiding access to packet payload content. Additionally, the method has minimal storage requirements, enabling near real-time classification.

Our main contributions are:

- The use of novel statistical features based on throughput trend to identify traffic patterns,
- Identification of low-latency application traffic,
- Development of a novel low-storage, near-real-time classification system using ML algorithms.

The experimental results show that this approach has an accuracy rate of over 97% for identifying single traffic types, including those requiring low-latency. In mixed traffic scenarios, it demonstrates an accuracy rate of up to 87% for identifying low-latency application traffics.

II. RELATED WORK

Recently, studies on the Internet traffic classification have addressed three issues: characterization types, such as VoIP, file transfers, and video identification of Internet applications (YouTube, Netflix, Skype), and user actions within applications (YouTube video resolution).

Three approaches exist for solving Internet traffic classification issues:

- Payload-based (DPI) [3], which is prone to privacy invasion and high computational costs, and challenges in dealing with encrypted traffic.
- Port-based methods [4] are less effective due to increased use of dynamic and default ports.
- Statistical and machine learning-based methods [4], using time and size features and supervised learning algorithms as classifiers.

This paper will focus on describing the most relevant works, specifically those that use statistical and machine learning methods.

There have been a variety of studies focused on the generation of features for network traffic analysis. Moore et al. [5] created a list of 248 descriptors based on bidirectional flow. Moore and Zuev [6] used a Naive Bayes Kernel estimator and their own descriptors to classify network traffic. Fahad et al. [7] evaluated 5 techniques for feature selection and proposed an integrated approach, resulting in improved performance and accuracy compared to other techniques [5].

Also, there have been numerous studies that have focused on utilizing only flow-based features, such as time and size related features, for network traffic analysis. Draper-Gil et al. [8] used features like flow bytes per second and inter-arrival time to classify Non-VPN and VPN traffic with 80% accuracy using C4.5 and kNN classifiers. Zhang et al. [9] used 20 simple flow-based features, modeling the correlation information in traffic flows of the same app using bag of words technique. Their new robust traffic classification scheme performed better than other machine learning methods.

Recently, new approaches have emerged in payload-based traffic classification. Wang et al. [10] used a neural network and normalized byte sequences for packet payloads. They improved classification results with 1-D Convolutional Neural Networks (CNNs) on the ISCX VPN-nonVPN traffic dataset [11]. Lotfollahi et al. [12] also used CNNs and auto-encoders for the same dataset and got good results. Aceto et al. [13] proposed a multimodal deep learning framework combining payload and protocol fields for classifying mobile encrypted traffic.

Payload-based methods have limitations in analyzing encrypted traffic such as VPN or Tor connections. Encrypted traffic is often protected by strong encryption algorithms with random initialization vectors, making it challenging for payload inspection. While Wang et al. demonstrated success in classifying VPN traffic using one-dimensional convolution neural networks [10], the reasons for their success may be more complex than the use of similar encryption methods and keys. Therefore, while the weakness of payload inspection for encrypted traffic highlights the need for alternative approaches, the factors contributing to the success of more sophisticated methods.

Recently, deep learning has been used in traffic classification research. Lopez-Martin et al. [14] used a combination of RNN and CNN to classify traffic based on 6 packet features and achieved accuracy of over 95% using port info and 84% without it. Chen et al. [15] transformed flow data into auto-

convolution representations and fed it into a neural network, but relied on other info like target IP. Zhang et al. [16] proposed an adaptive model update scheme. Iliyasu and Deng [17] proposed a semi-supervised approach using a GAN, which achieved high accuracy with few labeled samples, addressing challenges in creating ground truth for encrypted datasets.

As previously discussed, prior studies have utilized various combinations of classification techniques successfully, however, none of the works have focused on extracting throughput based pattern recognition features from network traffic for the purpose of performing NTC for applications that require low latency. This is the unique aspect of our proposed work.

III. METHODOLOGY

The proposed method for network traffic classification aims to efficiently identify low-latency requiring applications in mixed traffic environments by utilizing novel trend recognition features. The method consists of two primary stages: (1) extraction of significant features from network traffic and (2) classification using machine learning algorithms. The use of this two-step process ensures a more accurate and efficient identification of low-latency traffic, improving the overall network performance.

The methodology and process involved will be further explained in the following sections, including data collection, feature selection, scaling for pattern recognition, and presentation of experimental results.

A. Dataset

The data collection process was performed on a local WiFi network, and the set of traffic traces comprised FTP, video streaming, video conferencing, and mix of these. The traces were collected over a 100 Mbps Internet connection, and the throughput measurements of the active traffic in both directions was recorded and stored. In total, over 400,000 samples of throughput values (more than 20 hours of applications usage) were collected, with over 60,000 samples of each traffic type. The traffic classes and their sample sizes used for classification are shown in Table II.

B. Features

After collecting samples of the different Internet traffic, we first extracted different features from them. Table I demonstrates the new dataset after feature extraction. In total we have 12 features and 5 of them are novel trend-based features. We used these features training the ML algorithms. The description of the features are given in Table III. While calculating average throughput ratio over time, 3sec chosen as representative number.

C. Slope (Trend)

Experiments analyzing traffic patterns revealed the importance of the changing throughput. To address this, a new feature was developed using the slope of a linear function to determine the trend in the rate. The slope is calculated as the ratio of the vertical change between two points to the

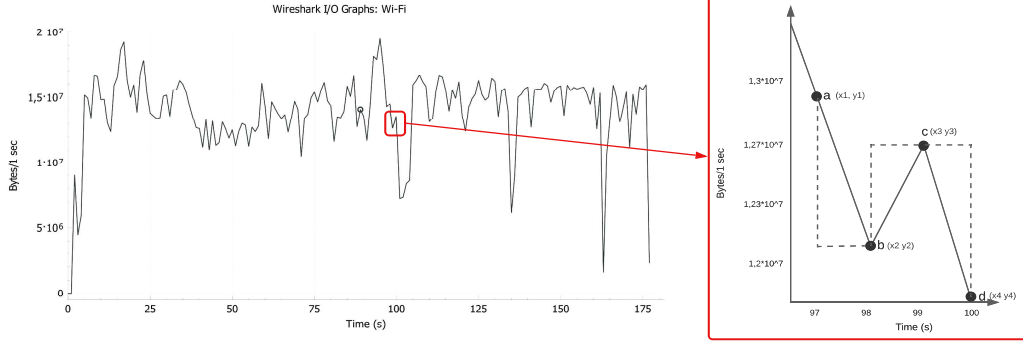


Fig. 1: Throughput of FTP traffic and demonstration of trend calculation

New Dataset with Features											
D	U	DU (A)	DA3 (B)	UA3 (C)	RAD3 (D)	RAU3 (E)	Rank of A	Rank of B	Rank of C	Rank of D	Rank of E
50350000	622020	74,92	38,318	0,45896	1,02230	1,02214	9	9	8	8	8
40970000	379980	107,82	37,482	0,44901	1,0060	1,00809	9	9	8	6	6
30120000	602670	49,97	37,256	0,44541	0,99404	1,027	8	9	8	5	8

TABLE I: Dataset v2

Traffic Line	Application	Total Samples
VoIP	Skype, Zoom, MS Teams	62177
Video Streaming	YouTube, Netflix, Prime	68987
File Transfer	FTPS, Torrent, Steam	72366
Mix Traffic	All	214987

TABLE II: Traffic lines, applications and sample sizes

ID	The Features
D	Throughput in downstream
U	Throughput in upstream
DU	D to U ratio
DA3	The average throughput in downstream over 3 sec
UA3	The average throughput in upstream over 3 sec
RAD3	Throughput rate (slope) in downstream
RAU3	Throughput rate (slope) in upstream

TABLE III: Features

Number	RDU	RDA3	RUA3	RRAD3	RRAU3
0	$\leq 0,5$	$\leq 0,5$	$\leq 0,0058$	$\leq 0,029$	$\leq 0,029$
1	0,9	0,75	0,0116	0,058	0,058
18	200	30	0,3	1,5	1,5
19	≥ 300	≥ 40	$\geq 0,4$	$\geq 3,93$	$\geq 3,93$

TABLE IV: Scaling the features. RDU is rank of DU

Traffic Mixes
1 FTP, 1 Video Streaming
1 FTP, 1 Video Conferencing
1 Video Conferencing, 1 Video Streaming
1 FTP, 2 Video Streaming
1 FTP, 3 Video Streaming
1 Video Conferencing, 2 Video Streaming
1 FTP, 1 Video Conferencing, 1 Video Streaming

TABLE V: Traffic Mixes

horizontal change between the same two points. In figure 1, the pattern of FTP traffic is shown on the left, and on the right, the slope is calculated for the time interval 97-100. This new feature provides a way to determine the overall trend of the throughput rate - a slope greater than 0 indicates an increase, while a slope lower than 0 indicates a decrease. This feature was applied to various types of internet traffic, both single and mixed, in order to generalize traffic patterns and provide a comprehensive understanding of the data and also used to create new feature sets III for training the ML algorithms.

D. Scaling

In this research, we developed a custom scaling method optimized for the network. By analyzing individual traffic patterns, feature values were scaled between 0 and 19. 20

levels have been chosen to demonstrate the operation, this will be improved in the future. Table IV displays our scaling system, with feature names represented by rank(R) of DU, DA3, UA3, RAD3, and RAU3. The first column assigns specific ranges to each feature. Ranges calculated by analysing the patterns of each traffic types. This rule applies to all features, and the table constructed according to that.

E. Traffic Types

In this study, we examined different traffic types, including FTP, video streaming, video conferencing, and a combination. FTP traffic is characterised by increasing throughput over time, although this may be limited by the maximum network capacity in practice. Video streaming utilises network capacity for buffering, resulting in a very volatile throughput, while

video conferencing demands constant use of a few Mbps of network capacity.

In addition to single traffic flows, we also examined the combination of several traffic types to identify the presence of low-latency traffic among them. Table V shows the combination of traffic mixes that we analyzed.

F. Machine Learning Algorithms

In this paper we employed five machine learning algorithms: k-NN, SVM, Random Forest and C4.5 for traffic classification and k-Means for clustering. These algorithms have a proven track record of effectively handling complex data.

IV. CLASSIFICATION RESULTS AND ANALYSIS

In this section, we present the results of our classification model and analyze its performance. We begin by analyzing the impact of different features on the model's performance. We then report the classification accuracy of the model and compare it to the performance of other models or baselines. Finally, we provide a detailed analysis of the model's performance, including any insights into its strengths and weaknesses, as well as any potential areas for improvement.

A. Selection of Features

In this section, we evaluate the performance of the feature(s) with different ML algorithms. We used all the combinations of features during the evaluation process and results are shown in Table VI. Each row represents a different combination of features and the columns represent the accuracy scores of the different algorithms on those features.

Feature(s)	Acc (k-NN)	Acc (SVM)	Acc (k-Means)	Acc (C4.5)	Acc (RF)
1 (RDU)	83,13	83,58	79,76	78,54	81,21
2 (RDA3)	81,7	82,55	77,6	77,4	79,71
3 (RUA3)	79,46	68,76	74,12	66,13	69,65
4 (RRAD3)	49,31	52,51	48,43	49,23	51,12
5 (RRAU3)	50,09	51,71	50,02	49,62	48,60
1,2	91,23	90,85	87,76	86,43	88,92
4,5	52,9	53,15	50,98	50,78	51,03
1,2,3	94,91	93,7	90,13	89,76	90,11
1,2,3,4	95,71	95,03	93,45	93,78	93,98
1,2,3,4,5	96,47	95,9	94,1	94,21	94,38

TABLE VI: Impact of the features

To evaluate the impact of the features for classifying single traffic type, we started using one feature at a time. We observed that feature 1 (RDU) is a powerful discriminator, as it enables algorithms to classify individual traffics with an accuracy of around 83.5%. Similarly, features 2 and 3 also stood out as powerful features. On the other hand, features 4 and 5, which capture the slope between intervals, act as more supportive trend features. Their power can be seen when classifying traffic in a mix.

When using features 1, 2, and 3 together, kNN and SVM achieved an accuracy of 94.91% and 93.7% respectively when classifying individual traffics. Among all the ML algorithms, kNN performed the best in most cases. This is likely due to its ability to effectively classify data based on similarity to its

nearest neighbors. However, feature number 1 was found to be crucial for kNN and its removal resulted in a significant decrease in accuracy.

Finally, the algorithms were trained using 4 and 5 features together. When using 4 features, kNN achieved the highest accuracy of 95.71%, while k-means achieved the lowest accuracy of 93.45%. Similarly, when using all 5 features, kNN again achieved the highest accuracy of 96.47% while k-means achieved the lowest accuracy of 94.1%.

B. Tuning the Machine Learning Algorithms

For the kNN machine learning algorithm, we first utilized the 10-fold cross validation method to determine the optimal sample for the training set. After evaluating the performance of the algorithm across 10 distinct training sets, we identified the best set with a default value of $k = 5$. Subsequently, we conducted a hyper-parameter tuning procedure to examine the performance of the algorithm for different values of k . As a result of our analysis, we determined that the optimal value for k was 7. This value was used to achieve the highest classification accuracy on the test dataset. Similar tuning procedures are applied for SVM, k-means, C4.5 and RF algorithms.

C. Classification Reports for Individual Traffics

Figure 2 demonstrates the performance results of the different algorithms, trained with 5 features, classifying the three individual traffic type. Through the analysis of the results, although all algorithms achieved over 94% accuracy of classifying File Transfer, Video Streaming and Video Conferencing traffic, kNN performs 2% better than k-means, c4.5 and RF algorithms. On the other hand, identifying file transfers using SVM has a little higher accuracy score than kNN, but given this is the only area where SVM outperforms kNN, we will only use kNN for the remaining tasks.

As we have decided to proceed with the kNN algorithm, we test another performance metric for our network. In the following experiment, we compared some of the well-known scaling methods for features to our own. We evaluated our

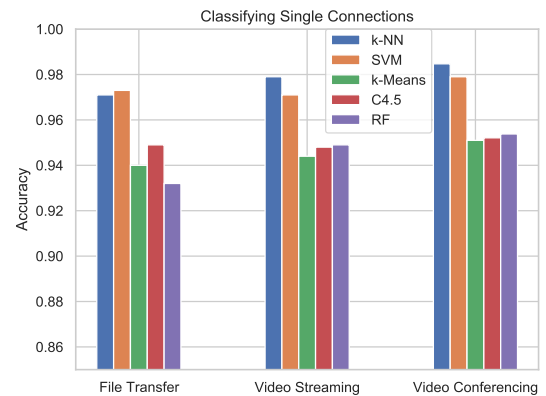


Fig. 2: Accuracy results of classifying single traffics

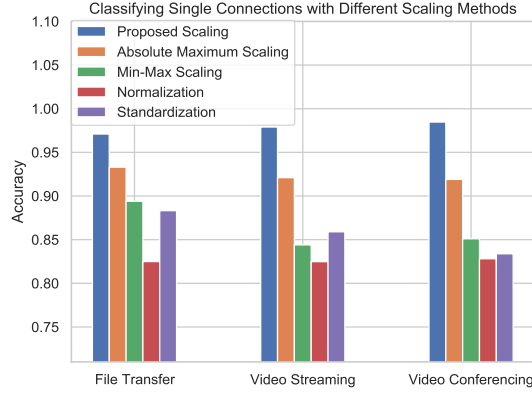


Fig. 3: Classifying single traffics with different scaling methods

network using five distinct scaling techniques, including the proposed one, absolute max, min-max, normalization, and standardization. Figure 3 demonstrates that our method outperforms other scaling methods by more than 2% for abs max. and 14% for normalization method. This is primarily due to the fact that we manually analyzed the patterns of individual traffic when developing our novel scaling system. Therefore, our method is superior to the alternatives.

In the following figure 4, we compare the accuracy results of our solution to several published proposals. We provided an expanded explanation of these proposals in section II. While each proposal has its own strengths and weaknesses, our network performed slightly better than those shown in Figure 4. In proposal [8], they tested their network's classification of encrypted traffic, resulting in an 89% accuracy rate. In contrast, proposals [15] and [14] use deep learning tools to classify traffic, achieving accuracy of more than 97%.

In the next experiment, we successfully attempted to classify not only individual traffic types, but also the exact traffic

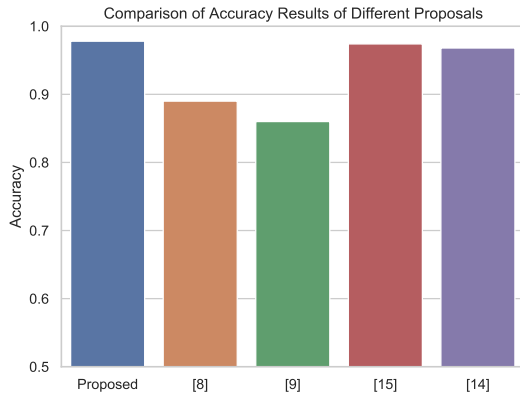


Fig. 4: Classification accuracy comparison of different approaches for single traffics



Fig. 5: Accuracy results of identifying exact traffic mixes

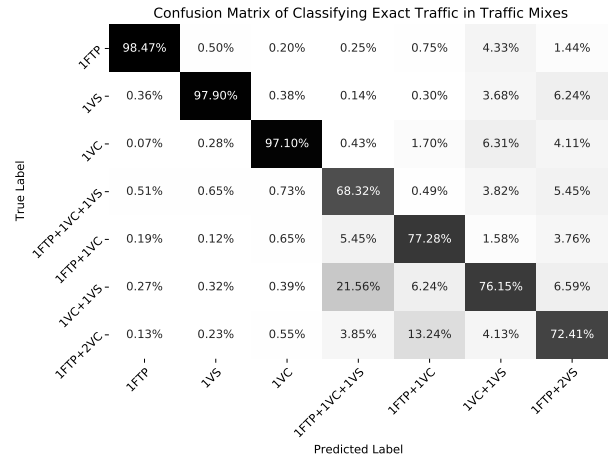


Fig. 6: Confusion matrix of classifying exact traffic mixes

mixes. Our results demonstrate the potential for identifying exact traffic mixes, despite the complexity of the task. To the best of our knowledge, our study is the first to address the exact traffic classification in traffic mixes. Figure 5 shows that we achieved an accuracy of around 84% in identifying the 1FTP+3VS traffic mixes. Furthermore, when multiple connections were made simultaneously, our network demonstrated its ability to identify the traffic mixes with a high degree of accuracy. The results of this experiment are further supported by the findings in the confusion matrix shown in Figure 6. The results highlight the promising performance of our network in classifying complex traffic mixes, with individual traffic types being classified with over 97% accuracy. This experiment demonstrates the potential for future research in this field and the promising outcomes that can be achieved with continued efforts.

In the final experiment, instead of identifying the exact content of a traffic mix, we decided to determine whether there is any low-latency application traffic in the traffic mix or not. Results of the experiments are shown in Figure 7. Overall, the

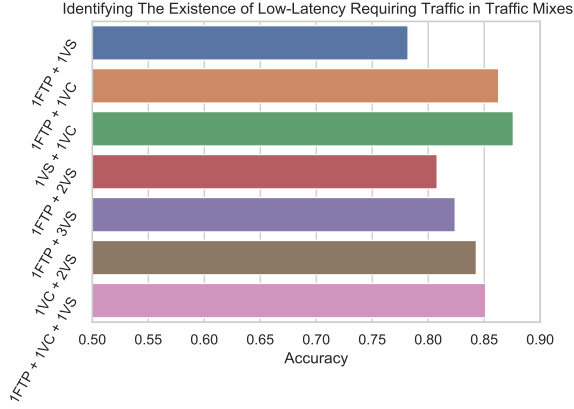


Fig. 7: Accuracy result of identifying the existence of low-latency requiring traffic in the traffic mix

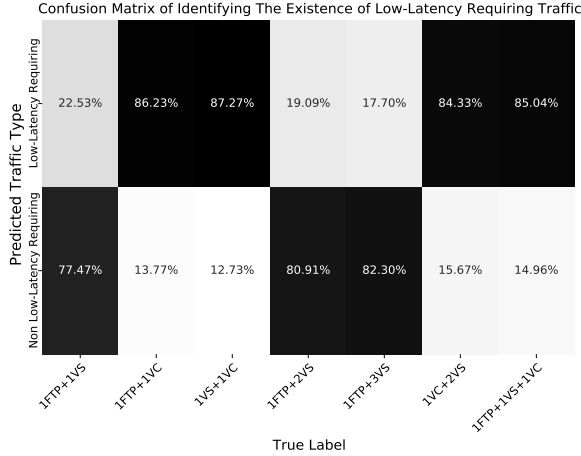


Fig. 8: Confusion matrix of identifying the existence of low-latency traffic

performance of this classification are much better compared to previous experiment. We obtained between 78% to 88% of accuracy when identifying low-latency application traffic in the traffic mix. Also, in Figure 8, we demonstrated the confusion matrix of this classification. We categorized video conferencing as a low-latency-requiring traffic type. This is due to the fact that video conferencing applications require a constant connection with low latency to maintain a high QoE. Our approach achieved an accuracy rate of 87.27% in predicting the mixed 1VS+1VC as low-latency-requiring traffic. Despite the intricate patterns present in FTP traffic, our method still showed favorable prediction outcomes.

V. CONCLUSION AND FUTURE WORK

This paper presents a fine-grained classification scheme for network traffic to enhance QoS support for ISPs and Internet providers. By analyzing the throughput patterns of FTP, video conferencing, and video streaming traffic, it was found that new statistical features are more effective at distinguishing

Internet traffic than commonly used features in the literature. Machine Learning algorithms were trained using combinations of these features to classify low-latency requiring applications. Results of experiments using large-scale real network traffic data show that the proposed method achieved high accuracy in classifying single traffic and identifying low-latency requiring traffic in the traffic mix. This method can be used by ISPs and Internet providers to improve QoS and QoE through traffic engineering, as well as detect and prevent applications that violate an organization's security policy.

REFERENCES

- [1] "Ericsson mobility report, november 2022." <https://www.ericsson.com/en/reports-and-papers/mobility-report/reports>. Accessed: 2022-12-30.
- [2] G. D'Angelo and F. Palmieri, "Network traffic classification using deep convolutional recurrent autoencoder neural networks for spatial-temporal features extraction," *Journal of Network and Computer Applications*, vol. 173, p. 102890, 2021.
- [3] M. Finsterbusch, C. Richter, E. Rocha, J.-A. Muller, and K. Hanssger, "A survey of payload-based traffic classification approaches," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 2, pp. 1135–1156, 2013.
- [4] T. T. Nguyen and G. Armitage, "A survey of techniques for internet traffic classification using machine learning," *IEEE communications surveys & tutorials*, vol. 10, no. 4, pp. 56–76, 2008.
- [5] A. Moore, D. Zuev, and M. Crogan, "Discriminators for use in flow-based classification," *IEEE Communications Surveys & Tutorials*, 2013.
- [6] A. W. Moore and D. Zuev, "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.
- [7] A. Fahad, Z. Tari, I. Khalil, I. Habib, and H. Alnuweiri, "Toward an efficient and scalable feature selection approach for internet traffic classification," *Computer Networks*, vol. 57, no. 9, pp. 2040–2057, 2013.
- [8] G. Draper-Gil, A. H. Lashkari, M. S. I. Mamun, and A. A. Ghorbani, "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] J. Zhang, X. Chen, Y. Xiang, W. Zhou, and J. Wu, "Robust network traffic classification," *IEEE/ACM transactions on networking*, vol. 23, no. 4, pp. 1257–1270, 2014.
- [10] W. Wang, M. Zhu, J. Wang, X. Zeng, and Z. Yang, "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.
- [11] G. Draper-Gil, A. H. Lashkari, M. S. I. Mamun, and A. A. Ghorbani, "Characterization of encrypted and vpn traffic using time-related features," in *Proceedings of the 2nd International Conference on Information Systems Security and Privacy - ICISSP*, pp. 407–414, INSTICC, SciTePress, 2016.
- [12] M. Lotfollahi, M. Jafari Siavoshani, R. Shirali Hossein Zade, and M. Saberian, "Deep packet: A novel approach for encrypted traffic classification using deep learning," *Soft Computing*, vol. 24, no. 3, pp. 1999–2012, 2020.
- [13] G. Aceto, D. Ciunzo, A. Montieri, and A. Pescapé, "Toward effective mobile encrypted traffic classification through deep learning," *Neuro-computing*, vol. 409, pp. 306–315, 2020.
- [14] M. Lopez-Martin, B. Carro, A. Sanchez-Esguevillas, and J. Lloret, "Network traffic classifier with convolutional and recurrent neural networks for internet of things," *IEEE access*, vol. 5, pp. 18042–18050, 2017.
- [15] Z. Chen, K. He, J. Li, and Y. Geng, "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.
- [16] J. Zhang, F. Li, H. Wu, and F. Ye, "Autonomous model update scheme for deep learning based network traffic classifiers," in *2019 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, IEEE, 2019.
- [17] A. S. Iliyasu and H. Deng, "Semi-supervised encrypted traffic classification with deep convolutional generative adversarial networks," *IEEE Access*, vol. 8, pp. 118–126, 2019.