Journal of Information and Organizational Sciences
Volume 49, Number 2 (2025)
Journal homepage: jios.foi.hr

JIOS

DOI: 10.31341/jios.49.2.5
Open Access

UDC 004.056.55:004.738.5
Original Scientific Paper

Bootstrap Forest based method for Encrypted Network Traffic Analysis

Shobana Durairaju[1] and Aswani Kumar Cherukuri[1*]

[1]School of Computer Science Engineering and Information Systems, Vellore Institute of Technology, Vellore, Tamilnadu, India
[*]Correspondence: cherukuri@acm.org

PAPER INFO

Paper history:
Received 17 April 2025
Accepted 19 August 2025

Citation:
Durairaju, S. & Cherukuri, A. K. (2025). Bootstrap Forest based method for Encrypted Network Traffic Analysis. In Journal of Information and Organizational Sciences, vol. 49, no. 2, pp. 235-249

ABSTRACT

Encrypting communications and data over the Internet becomes essential in ensuring the privacy of communications and protecting the data from increasing threats. Hence, majority of Internet traffic and networked communications are encrypted now. However, encryption also provides a means for attackers to hide them behind encrypted communications and conduct malicious activities. Analyzing the unencrypted communications is relatively easy. The same task is highly challenging due to the presence of encryption in network communication. Conventional network analysis methods fail to analyze encrypted communications. There are methods like flow monitoring that are available to detect encrypted traffic and analyze traffic flow related features. By using traditional analysis methods, we could not achieve accurate detection and classification of encrypted network packets in various types of network traffic such as VoIP, Text, Audio, Video, VPN traffic. In our work, we have proposed the Bootstrap Forest model to analyze and classify encrypted network traffic. Bootstrap Forest model accurately classifies the encrypted network traffic using statistical and time-based features. The performance of the proposed model is evaluated and compared with the performance of other machine learning models under various performance metrics. The three publicly available datasets such as UNSW-NB15, ISCXTor 2016 and ISCXVPN 2016 datasets were used in our experimentations and evaluations. The experimental results show that our proposed model provides the best performance for classifying encrypted network traffic while comparing the F1 score with other methods.

Keywords: Encrypted Network Traffic, Benign Packets, Machine Learning, Network Packets, Network Traffic, Bootstrap Forest

1. Introduction

Network traffic analysis and classification is an important measure in network management. Classification of network traffic becomes essential in various aspects such as analyzing normal traffic, analyzing encrypted network traffic and further classifying normal and malicious traffic present in the network. Due to the major increase in encrypted network traffic, the availability of resources should be properly shared by the network and maintained by the network management to improve the quality of service. The data communication through encrypted IP packets have become as the standard in modern communication (Velan et al., 2015). Further, the amount of network traffic is rapidly increasing in each and every network. So, secured data communication plays a vital role in all the aspects of network communication. Even though the network is secured, malicious attackers are trying to analyze the network traffic and include their malicious

communications in the encrypted network. Identifying the malicious attacker communication involved in encrypted network traffic is still a major challenging task.

Various traditional methods are available for identifying malicious traffic and classifying encrypted network traffic. Some of them are Deep Packet Inspection (DPI), Port-based, Payload-based, Protocol-based etc. But still these methods have many limitations on their own. In general, both plaintext traffic information and encrypted data information are available in encrypted network traffic (Velan et al., 2015). Due to unavailability of data in the encrypted network traffic, some of the traditional methods, for example DPI and Pattern matching, could not provide accurate traffic classification. Classifying encrypted network traffic is still a major problem. Nowadays, various machine learning models like Support Vector Machine (SVM), Neural networks, Random Forest, K-Nearest Neighbor (K-NN), Decision Tree, Naive Bayes, etc., are used to improve the performance on classifying encrypted network traffic.

The main purpose of this work is to perform better analysis and improve the classification performance in encrypted network traffic using our proposal to apply Bootstrap Forest (BF) model. The proposed BF model works on the concept of bagging, which makes prediction based on multiple decision trees. Further, the proposal of using BF model could satisfy the research gap identified in the various machine learning models used in the classification of encrypted network traffic by concentrating the misclassification rate. To show the effectiveness of the proposed model, experiments were carried out on the three publicly available datasets such as UNSW-NB15 (Moustafa & Slay, 2015), ISCXTor 2016 (Xu et al., 2022) and ISCXVPN 2016 datasets (Aouedi et al., 2022), (Lotfollahi et al., 2020). To perform the analysis on the encrypted network traffic, we have done the literature review on the various machine learning models which are already used and compared their strengths and limitations with the proposed model. In the above mentioned datasets, statistical and time-based features were taken for the analysis and classification of encrypted network traffic. The experimental results of BF model and other machine learning models are compared and discussed.

The main objectives of our work can be summarized as follows:

- First, we have analysed the publicly available datasets and identified the statistical and time-based features to classify encrypted network traffic.
- Next, we have done a study on how these selected features can be used for the analysis and classification of encrypted network traffic.
- Further, we have proposed the BF model to improve the performance of encrypted network traffic analysis and classification.
- In this work, we have selected three publicly available datasets to evaluate the performance of some machine learning models and compared them with our proposed BF model.
- Then, the experimental results of the proposed BF model are compared with the results of other machine learning models in classifying encrypted network traffic.

This is the first work to analyse and classify encrypted network traffic using BF model. The rest of the paper is organized as follows. Section 2 defines the main sources of existing Encrypted Network Traffic Analysis (ENTA), various protocols and its structure used in encrypted network traffic, types of network traffic, existing classification techniques used for ENTA and followed by related work. Section 3 explains the proposed model used for the classification of encrypted network traffic. Section 4 presents the features and datasets used for the classification of encrypted network traffic, experimental setup and its analysis and results. Finally, section 5 summarizes our research and conclusions.

2. Background

ENTA: Encrypted network traffic can be analysed from the sources of information obtained from the protocols used in secure communication. They are Hypertext Transfer Protocol Secure (HTTPS), Internet Protocol Security (IPsec), Secure Shell (SSH) protocol and Secure Real-time Transport Protocol (SRTP). HTTPS uses Transport Layer Security (TLS) as well as its predecessor Secure Socket Layer (SSL). HTTPS can be used in all the online applications while transferring sensitive data such as online banking, email service, online purchase, etc. SSL is used to encrypt the user data in order to provide secure communication when data is transferred through the Internet. Through its secure communication, it provides data privacy, authentication and data integrity. In public network, IPsec is used to transfer the data in secured manner by providing its authentication and by encrypting the network packets that are communicated in the Internet Protocol (IP) Network. SRTP is used to control and provide security for real-time transport protocol sessions. For secure data communication, most of the networks use protocols such as HTTPS, TLS, SSH and IPsec (Cherukuri et al., 2024). IPsec provides its security in both transport and tunnel mode. In the transport mode, IP packet gets encrypted and in the tunnel mode, the entire IP packet along with the header details gets encrypted. We can also understand the security from the basic frame format of TLS and SSH protocol packet format. We can

gather the payload and header information from the protocol frame format for inspection (Elmaghraby et al., 2024). Here, Payloads can be verified from the information obtained from the payload of the packet. Packet Inspection follows the predefined patterns like regular expression as digital signatures for each protocol.

DPI, Shallow Packet Inspection (SPI), Medium Packet Inspection (MPI) are the three types of packet inspection techniques used to monitor the network packets. SPI is the basic network packet monitoring technique which is used to monitor and inspect the data packets based on source IP address, destination IP address, source port number and destination port number. It can inspect the packets only at Network Layer, Data Link Layer and Physical Layer in OSI and TCP/IP Layers. MPI can act as an intermediate node in the networks. It is also called middle-boxes and they are used to monitor and inspect the data packets only at presentation layer, session layer, transport layer, network layer, data link layer and physical layer. At the same time, MPI cannot be used in Application Layer and it can just act as a gateway between nodes and Internet Service Providers. DPI is the only technique which uses IP header and payload information to monitor and inspect the data packets in the networks. The major advantage of DPI is that we can monitor and inspect the real time data packets. Also, DPI can be used in all layers of OSI and TCP/IP models. Both SPI and MPI did not use the IP header and payload information to inspect the data packets in the networks while the same can be done using DPI. So, DPI is the best technique when compared to SPI and MPI and it can be used in all the TCP/ IP Layers. DPI is the best traditional packet inspection technique because it can check whether every packet is transmitted in the correct format or not.

2.1. Types of network traffic

During the data communication in Internet, many different types of network traffic namely, web browsing, email, chat, live streaming, file transfer, VoIP, VPN, Secure browsing, torrents, etc.. enters into the network at once. They can be classified based on their traffic pattern and their format. Voice traffic is the real time traffic. It is a very sensitive traffic type. The major factors to be considered in Voice traffic for lossless communications are delay time, jitter and packet loss. The protocol used in Voice traffic is UDP. The important application in voice traffic is VoIP in which we can make voice calls using Internet connection instead of traditional phone lines. Video traffic plays an important role in today's scenario. It can be of one among the two types such as stored video traffic and real time video traffic. Stored Video traffic like watching YouTube is not real time video traffic in which the time delay and packet loss does not affects its traffic. At the same time, Videoconferencing is a real-time video traffic in which time delay and packet loss play an important role. If the time delay and packet loss are high, this type of communications will become impossible. The protocol used in video traffic is also UDP. Next, data traffic is not a real time traffic and it is not a sensitive traffic type. The main applications in data traffic are emails, web pages, file transfer, etc. Time delay and packet loss are not the important factors in data traffic. The protocol used in data traffic is TCP. In data traffic, if delay and packet loss occurs, TCP will help in retransmission of data which does not affect the communication. Among all the other types, encrypted traffic plays an major role in the current environment. Here, network traffic is defined as the total amount of data that is moving across various connected devices (i.e., network) at a particular time. Network data is always encapsulated in network packets, which provide the payload in the network.

Encrypted network traffic plays an important role in secured communication. HTTPS, TLS/SSL plays a vital role in improving security in data transfer. It can be used in all online applications while transferring sensitive data such as online banking, email service, online purchase etc. We can classify the network traffic based on protocol-based, port-based, flow-based and machine learning-based methods. They are as follows:

Protocol-based classification method (Velan et al., 2015): We can classify network traffic based on information collected from the packet header. This is the traditional model used to identify and classify network traffic. The protocol format should be changed every time based on the protocols used in every packet. It is a very fast and simple method but susceptible to attack.

Port-based classification method (Boumhand et al., 2023): In this method, the network traffic can be classified based on the port information of the application type received from TCP/UDP. From that we can get the details of source and destination port numbers of the packets. By comparing the port numbers used by the protocols with the standard port numbers, we can classify the network easily. This method is very simple, but it is vulnerable to port detection attack.

Flow-based classification method (Azab et al., 2024): This method is used to classify the network traffic by analysing the flow-based features. Various flow-based features play a major role in the classification of encrypted network traffic. From this, we can manually select and derive the features to classify the network.

Payload-based classification method (Sheikh & Peng, 2022): This method analyses the network traffic based on the information available in the packet's payload. In addition, this method uses the pattern matching approach to classify the encrypted network traffic. So, we can consider it as DPI.

2.2. Techniques used for ENTA

Statistical based analysis (Velan et al., 2015; Papadogiannaki & Ioannidis, 2021): It is a method of analysing the pattern of encrypted network traffic data. Statistical based analysis helps us to identify the normal behaviour of the network traffic and the same is used to detect the anomalies present in the network traffic. It is a simple method to classify the encrypted network traffic.

Machine learning based analysis (Ahmad et al., 2021; Bagui et al., 2017): It trains the machine learning model based on the known input and predicts the future output. In this method, Random Forest, Decision Tree, Logistic Regression, SVM, Naive Bayes and K-NN classification algorithms can be used to classify the encrypted network traffic. By applying these methods, we can analyse and classify the network traffic to achieve accurate classification results without the knowledge of human experts.

Deep learning-based analysis: (Seydali et al., 2023; Lotfollahi et al., 2020) Deep Learning is a type of machine learning techniques which uses artificial neural networks to perform computations for complicated data. This method also works well in classification problems.

2.3. Related Work

Various researchers have used traditional methods like DPI and port-based methods for analysing and classifying encrypted network traffic. However, these methods produce less performance in classifying encrypted network traffic. Nowadays, protocol-based, payload-based, flow-based and machine learning based classification methods are widely used to identify the important features for classifying encrypted network traffic (Zhao et al., 2021). (Chen Y et al., 2024) used hierarchical features to avoid the issue of inaccurate feature generation and improved the efficiency. (Seydali et al., 2023) introduced a hybrid learning approach called as CBS which uses spatio-temporal and statistical features to achieve higher accuracy when compared to other models. (Velan et al., 2015) have done a deep survey on the methods of ENTA and classification. In their survey, they have used protocol structure and identified that the protocol structure gives the lot of information from the initial connection phase for ENTA and classification. Then, they have done survey on payload and feature-based classification methods and encryption protocols which are used for encrypted network traffic. Next, they have done the survey on feature-based classification methods and identified strength and weakness of feature-based classification for encrypted network traffic.

(Moustafa & Slay, 2015) proposed a method to generate the features of UNSW-NB15 dataset. In this work, they have explained the different types of features used in other three different datasets and analysed those features to perform Network Intrusion Detection Systems. Further, they have compared those features with UNSW datasets features. They have identified and stated that UNSW datset can be used for analysing NIDS in the research community. (Lotfollahi et al., 2020) proposed a framework that can analyse both traffic characterization and application identification in encrypted network traffic using publicly available VPN-non-VPN dataset. They have done experiments using various machine learning models and compared the performance of the models using performance metrics. (Bagui et al., 2017) used time related features to classify encrypted network traffic in VPN and non-VPN datasets. The authors proposed a new framework using machine learning models, comparing the classification performance and identified the suitable model for encrypted network traffic classification using time related features. (Xu et al., 2022) proposed an encrypted traffic classification with Path signature using session packet length and done traffic path transformations to expose additional features. The authors applied a few machine learning models on six different datasets and compared the performance of the models to identify the suitable model. (Dener et al., 2023) proposed a model which combines feature selection, gradient recurrent unit's algorithms to achieve data balancing. The authors selected the features by using Random Forest algorithm and used some oversampling techniques to reduce the negative effect in imbalanced datasets. The authors have implemented various machine learning models and compared the performance with their proposed model.

(Hu et al., 2023) used spatial and temporal features using LSTM and CNN and further they have included automatic mapping and labelling of feature mechanisms. This mechanism helps to adopt network traffic datasets quickly and can handle both encrypted and unencrypted traffic. By using various machine learning models, they have achieved higher performance in classifying encrypted network traffic. (Elmaghraby et al., 2024) proposed a framework using features selected from packet length, time stamps or TLS information, encrypted payload information as input. The authors applied some ensemble method and achieved higher

performance in classifying encrypted network traffic. (Ahmad et al., 2021) proposed feature clusters method and used machine learning models. The authors achieved better performance by their proposed method using feature clusters and achieved better classification performance. (Aouedi et al., 2022) proposed an ensemble-based approach for ENTA and achieved better performance.

(Zhao et al., 2021) presents a survey on the various common features used in the classification of encrypted network traffic. (Azab et al., 2024) have done a review on network traffic classification based on various classification techniques, datasets and features used. They have also reviewed malicious behavior identification using various datasets, features and challenges in network traffic classification. (Chen Z et al., 2023) proposed a multi-flow encrypted network traffic classification method with various features. (Peng et al., 2024) have used both plaintext and encrypted text for classifying encrypted network traffic and they have done experiments on two publicly available datasets. (Lashkari et al., 2017) used time-based features to analyse Tor traffic flows and application type traffic. (Boumhand et al., 2023) proposed a framework which is used to detect multi-activity traces in network traffic classification. (Sheikh & Peng, 2022) have done a survey on encrypted network traffic classification techniques, procedures to perform analysis, datasets, various machine learning models and the challenges. (Papadogiannaki & Ioannidis, 2021) have done a survey on various traditional ENTA techniques, applications and challenges.

3. Proposed Methodology

In our work, we propose applying ensemble enabled BF model for ENTA and its classification. Considering the literature, on achieving some more additional benefits, we have proposed the BF model in place of Random Forest in our ENTA and its classification. JMP has introduced this BF model by implementing Random Forest algorithm as part of it with some additional features. Among the various additional features, column contribution is the major one. It helps to identify the set of most contributing important predictors or features in the dataset which could simplify the model by reducing the dimensionality and focus on the most relevant data. Further, BF provides another advantage in terms of model interpretability. Here, it provides a way to explicitly identify the most valuable features with which the model works better while the Random Forest model could not provide it explicitly. Further, column contribution provides the visual representation of the most contributing predictors or features along with its quantifiable measure to reduce the impurity. These are the key benefits we could achieve, while applying the BF model Instead of Random Forest model. Overall, the main difference between Random Forest model and BF model is in configuring and controlling, further it is not in algorithm.

BF model makes predictions by combining multiple decision trees. Each tree in Bootstrap samples is built with random samples of training data with replacements. It also trains the model by building multiple decision trees to improve the performance of the model with less error rate. In Random Forest, we need not mention how many predictors required to be used for implementation. In JMP (SAS Institute Inc., 2023), we can manually enter the number of predictors. Further, all of these predictors are sampled at each split in the decision tree. In addition, we can identify and check whether the selected predictors are more important for classification or not. It is decided based on the noticed misclassification rate. BF provides measures for calculating the error rate for out-of-bag observations to estimate predictive performance. This out-of-bag misclassification rate serves as a cross-validation as like the estimate of model performance. If, this value is higher in validation than training, we can choose the other predictors. The predictors can also be identified by column contributions in JMP. So, error rate can be easily identified and rectified using BF model. It makes final predictions by averaging the predictions of multiple decision trees. BF model uses bagging to reduce variance and improve classification accuracy. By using the BF model, we can easily identify and reduce the misclassification rate of the models. So, we can train the model and create a model to obtain the better performance in analyzing and classifying the encrypted network traffic. In this way, BF model helps us to make interactive analysis and reporting.

Figure 1. shows the proposed architecture of BF model for ENTA and classification. First, raw encrypted network traffic datasets are collected. Those data sets are then preprocessed to get extracted features. From those extracted features, further important features are selected. For experimental analysis, 70% of samples are taken for training and 30% of samples are taken for testing from the dataset. Then, the BF model is applied by entering BF specifications. The BF model performs bagging to improve the performance of the model by reducing bias. From the raw datasets, the BF model makes a final prediction by combining multiple decision trees obtained from every random subset of samples with replacements taken for training. It also trains the model by building multiple decision trees and makes predition by aggregating multiple decision trees. It also diplays misclassification rate and makes final predictions with lower misclassification rate to improve the performance of the model. Thus, the BF model improves performance.

Figure 2. shows the workflow of ENTA and classification using machine learning models. Workflow consists of following steps such as dataset collection, data representation and preprocessing with labeling. Then by applying Principal Component Analysis, important features can be extracted for analysis and classification. Here, we have used three datasets like UNSW-NB15, ISCXTor 2016 and ISCXVPN 2016 and their statistical and time-based features can be analysed to classify the encrypted network traffic. We have mainly focused on flow-based statistical features, session-based and time-based features in the preprocessed data. We have split each dataset into training and testing sets to implement various algorithms like SVM, BF and Neural networks. Here, R-square value and confusion matrix can be obtained from the experiments. The R-square value denotes the performance of the model and the other performance metrics like Precision, Recall and F1 score can be calculated from confusion matrix. While comparing the R-square value and F1 score, BF model could show the better performance with less misclassification rate.

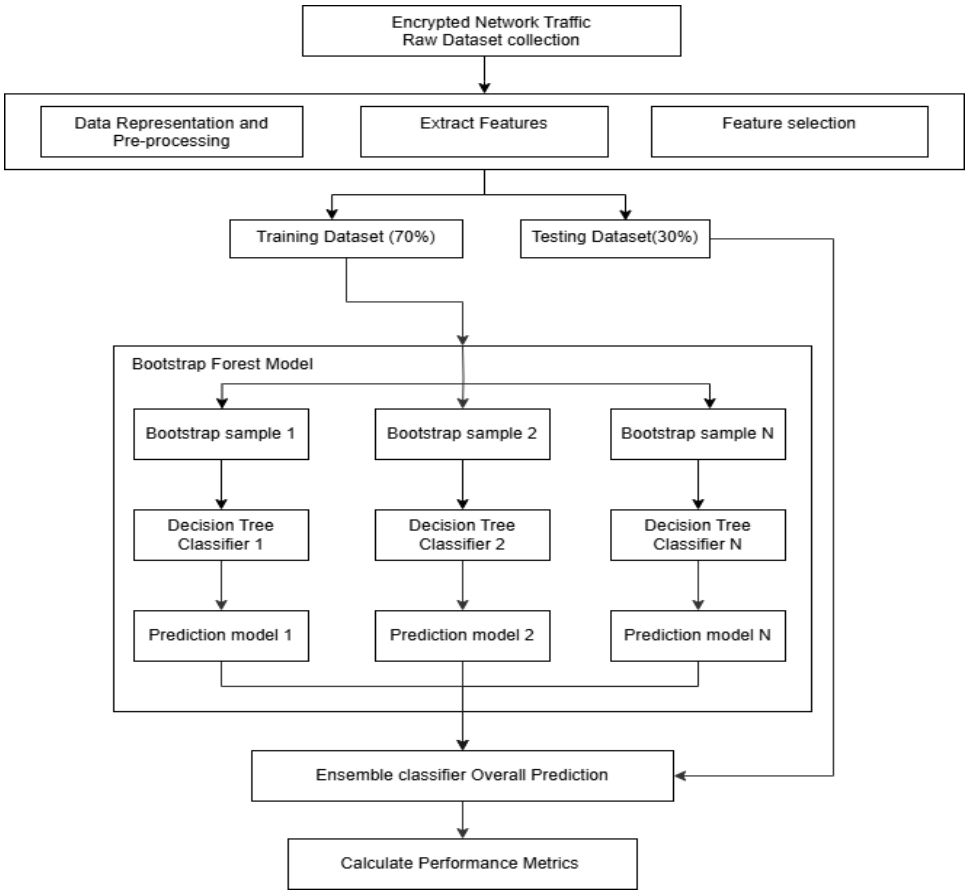

Figure 1. Proposed Architecture of Bootstrap Forest (BF) Model

4. Experimental Setup and its Analysis

4.1. Features used for ENTA and classification

Encrypted network traffic uses different types of features for its analysis and classification. They are as follows. Statistical features (Seydali et al., 2023; Azab et al., 2024): We can define the statistical features from our datasets. We can obtain the packet length by calculating the mean, average, std length of the packets. From these features we can analyze the original length which can be used for classification.

Temporal features (Hu et al., 2023): We can identify the temporal features by analyzing packet size, packet frequency and session duration of the packets. These time dependent features can be used to identify and classify encrypted network traffic.

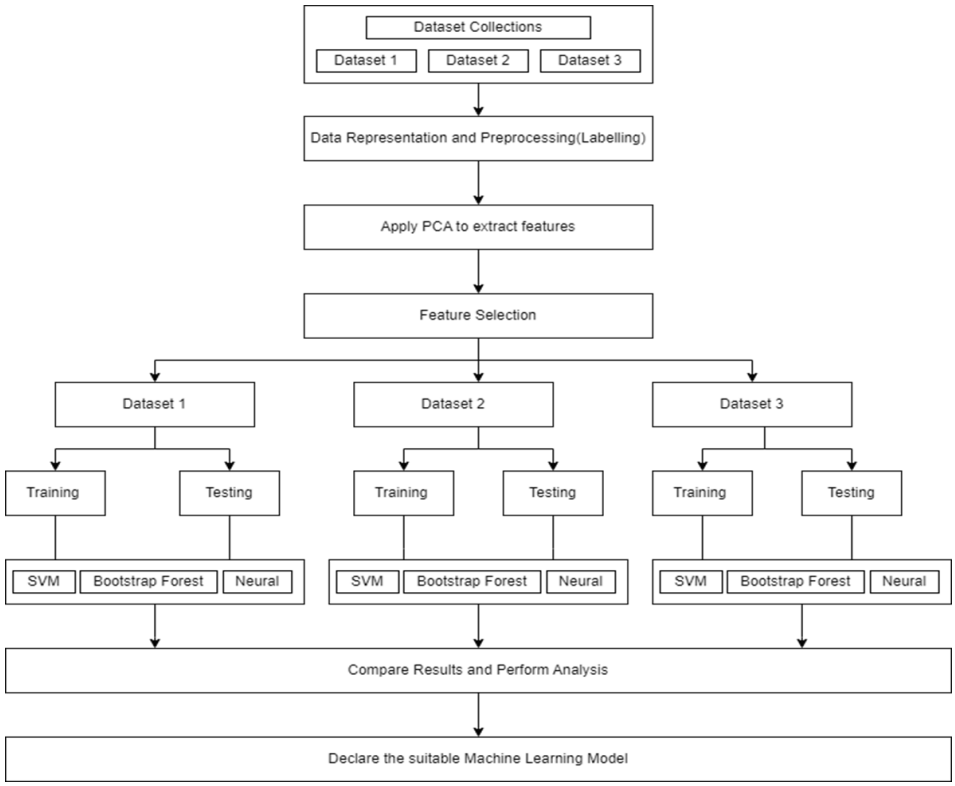

Figure 2. Workflow of ENTA using machine learning models

Residual Plain text features (Peng et al., 2024): The plain text features obtained from the payload of the packets can be used for the analysis of encrypted network traffic. The features like cipher suite, compression method and TLS information can be extracted from the packet payload and used for analysis of packets.
Spatial features (Hu et al., 2023): We can identify the spatial features from the particular session of network packets in bytes. It helps to capture the spatial relationship between network packets in bytes and that can be used to analyze and classify the encrypted network traffic.

Inter-Arrival Time sequence features (Seydali et al., 2023; Bagui et al., 2017; Lashkari et al., 2017): We can calculate the Inter-Arrival Time (IAT) from the flow of packets like IAT min, IAT max, IAT mean, IAT std. From the calculated IAT, we can analyze the packets by matching with the time sequence from the normal packets.

Length sequence features (Seydali et al., 2023; Chen Z et al., 2023): We can find the length sequence features such as min, max, mean and std length of the packets. By calculating these features, we can match with the normal packet values which can be used for analysis and classification.

Feature selection plays very important role to analyse the encrypted network traffic. When, we analyse the flow of data from encrypted network traffic, we can see temporal features, spatial features and statistical features. From the literature, to get high accuracy in classifying encrypted network traffic, we have taken statistical and time-based features to improve the prediction accuracy in ENTA and classification. The list of flow-based statistical features is shown in the Table 1 (Seydali et al., 2023). Out of 79 features, we first consider the features related to packet length. The most important features related to packet length considering encrypted network traffic classification are min packet length, packet length mean, backward packet length min, backward packet length mean, packet length std, average packet size and average forward segment size. From these features, we can analyse the length of the packet for classification. In general, the

packet size for Maximum Transmission Unit (MTU) can be 1500 bytes consisting of IP packet with IP header. So, we can classify the traffic by considering the size of packet. If the packet size is lesser or minimum than 1500 bytes, we can consider as benign packet. If the packet size is greater than or equal to 1500 bytes, we can consider as normal packet. If a packet size is lesser than 1500 bytes then there might be the possibility of

Sl. No	Feature Name	Feature Description	Minimum and Maximum value	How it is used for analysis
Statistical Features				
1	Min Packet Length	Lower limit on length of a packet	21-1500 bytes	If the size of the packet is less than MTU it is considered a malicious otherwise normal packet.
2	Packet Length Mean	Mean length of a packet		
3	Packet Length Std	Standard deviation length of a packet		
4	Backward Packet Length Min	Minimum size of the packet from receiver to sender		
5	Backward Packet Length Mean	Mean size of the packet from receiver to sender		
6	Average Packet Size	Average size of packet	1460 bytes	
7	Average Forward segment size	Average size of the segment from source to destination		
8	Flow IAT Mean	Mean time between two packets sent in the flow	Bot communication generates uniform sized small TCP/UDP packets	If the IAT exceeds the threshold value, then it is considered a malicious otherwise normal packet
9	Flow IAT Std	Standard deviation time between two packets sent in the flow		
10	Flow IAT Max	Maximum time between two packets sent in the flow		
11	Flow IAT Min	Minimum time between two packets sent from sender to receiver		
12	Min Packet Length	Lower limit on length of a packet	21 bytes	If the packet or header length exceeds the threshold value, then it is considered a malicious otherwise normal packet
13	Max Packet Length	Upper limit on length of a packet	65535 bytes	
14	Fwd Header Length	Total bytes used for packet header from source to destination	20-60 bytes	
15	Idle Max Time	Maximum time a flow was idle before becoming active	600 seconds or 10 minutes for TCP/UDP flows	If the idle time of the flow exceeds 600 sec, then it is considered a malicious otherwise normal packet

Table 1. The list of flow-based features used for analysis

Sl. No	Feature Name	Feature Description	Minimum and Maximum value	How it is used for analysis
Residual Plaintext Features				
16	Cipher suite	Cryptographic algorithm used	TLS AES 128 GCM SHA256 TLS AES 256 GCM SHA384 TLS AES 128 CCM SHA256 TLS AES 128 CCM 8 SHA256 TLS CHACHA20 POLY1305 SHA256 TLS AES 256 GCM SHA384 TLS CHACHA20 POLY1305 SHA256 TLS AES 128 CCM SHA256 TLS AES 128 CCM 8 SHA256	If the cipher suite information gathered matches the following standard five combinations, then it is considered as normal otherwise, malicious packet
17	Compression method	Compression method length	One byte length always and method set to null (0)	If the compression method value equals to null then it is considered as normal otherwise malicious packet
18	TLS extension information	Extension information	Every field has standard value	If the length of the extension field matches the standard length, then it is normal otherwise malicious packet

Table 1. The list of flow-based features used for analysis - continued

DDoS attack and if the attack has happened, the size of packets is always very small. Further, attacker always generates smaller packets or empty packets. So, from the features related to Packet Length, we can classify malicious or benign traffic.

Considering encrypted network traffic classification, other important features related to IAT (Bagui et al., 2017; Lashkari et al., 2017) are calculated as follows: The time between the two packets sent from the same source helps to identify the features such as Flow IAT Mean, Flow IAT Std, Flow IAT Max and Flow IAT Min. From these features, we can analyse the IAT for encrypted Network Traffic classification. By using the features related to IAT, we can classify Botnet traffic from normal traffic, because normal bot communication will indicate more uniform sized, small TCP/UDP packets. If the IAT exceeds the threshold value, then it is considered as malicious, otherwise normal packet. The features like Min and Max Packet Length, Fwd Header Length (Chen Z et al., 2023) are used to identify and classify encrypted network traffic. If the packet or header length exceeds the threshold value, then it is considered as malicious otherwise normal packet. The feature Idle Max Time is used to classify the normal or malicious TCP/UDP flow packets. If the Max Idle Time of the TCP/UDP flow exceeds 600 seconds or 10 minutes, then it is considered as malicious otherwise normal packet. Feature selection is one of the most important parts in encrypted network traffic classification because it plays a very important role in its performance. We have done a major study on the features used for different machine learning models and found that various features are extracted and used for classification namely, flow-based, payload-based, packet-based, statistical based features, etc. (Zhao et al., 2021; Seydali et al., 2023; Bagui et al., 2017; Hu et al., 2023). Similar studies are already available in the literature for identifying the features. However, our work is different from others in terms of how these features are useful for analysing and classifying the encrypted network traffic. In our work, we have identified the statistical and time-based features which are used to identify and classify the malicious and benign packets in the encrypted network

traffic. In addition, with the support of column contributions in JMP we can identify the features which has strong impact on our model performance. To analyse each feature, we have identified packet length and IAT. Further, to classify malicious and benign packet, we have compared the captured packet length and IAT with the standard values. The list of flow-based features are already shown in the Table 1. and they are helpful in analysing and classifying encrypted network traffic. By this work, we could provide a better understanding on feature selection to be used for machine learning based ENTA and classification. The first step in the analysis is to detect encrypted network traffic then to classify the network into various categories and subsequently identify any malicious traffic present in the network. Our main aim is to perform accurate classification of encrypted network traffic using machine learning algorithms.

4.2. Datasets used for encrypted network traffic classification

There are several datasets available in the literature for the purpose of ENTA and classification, such as UNSW-NB15 (Bhandari et al., 2023), NSL KDD, ASNM-CDX-2009, ISCXVPN 2016 (Dener et al., 2023), ISCXTor 2016 (Xu et al., 2022) etc. From the survey, we can understand that they have collected, analysed and used different traditional methods for classification. Since, they have faced different challenges while implementing traditional methods for real time ENTA and classification, we have focused towards achieving accurate classification of encrypted network traffic using machine learning models with minimum number of features and without involving experts. From the available datasets, we have taken the following datasets for analysis.

UNSW-NB15 Dataset (Moustafa & Slay, 2015): In this research, first we have taken UNSW-NB15 dataset (UNB, 2020). It contains more than 68,000 flows in total. From these flows, the network contains seven groups of traffic classification. Out of 68,000 flows, we have taken more than 3,300 samples for analysis and classification. In this dataset, there are seven different types of traffic flow like fuzzers, exploits, shellcode, reconnaissance, DoS, generic and normal were analysed and classified.

ISCXTor 2016 dataset (Xu et al., 2022): The second dataset which we have taken for analysis is ISCXTor 2016 dataset (UNB, 2023). The dataset contains more than 14,000 flows in total. Out of 14,000 flows, we have taken more than 8,000 samples for our analysis. In this dataset, there are six different types of network traffic flows like Audio, Browsing, File Transfer, Video, VOIP and P2P are analysed and classified.

ISCXVPN 2016 dataset (Dener et al., 2023): The third dataset which we have taken for analysis is ISCXVPN 2016 dataset (UNB, 2024). The dataset contains more than 14,600 flows in total. Out of 14,600 flows, we have taken more than 14,000 samples for our analysis. In this dataset, there are seven different types of network traffic flows such as chat, file transfer, mail, streaming, VOIP, P2P, browsing etc. are analysed and classified. From the samples, we have taken flow-based features for accurate classification of encrypted network traffic.

4.3. Results and Analysis

The proposed work was implemented and evaluated using JMP Pro, Version 18.0 (SAS Institute Inc., 2023) SAS Institute Inc., Cary, NC, USA a statistical analysis and machine learning tool.

Evaluation: To perform analysis and classification of encrypted network traffic, we have chosen the UNSW dataset (Ahmad et al., 2021) with 3350 samples for multiclass classification. For analysis and classification, statistical features such as dur, sbytes, dbytes, sttl, dttl, sload, dload, spkts, dpkts, smeansz and dmeansz were used. In order to perform the classification, various machine learning models, SVM, BF and Neural networks were used. Among these, BF model provides the best R-square and F1 score value which is fitted for the UNSW-NB15 dataset. The BF specification for implementing the datasets was given as follows: Number of terms sampled per split was 4, Maximum splits per tree was 30, Maximum number of terms was 10 and Random seed control for reproducibility was 123. Next, we have taken 14508 samples from ISCXTor 2016 dataset. For the analysis and classification of encrypted network traffic, the features (Baldini, 2020) such as flow duration, flow bytes, flow pckts, fwd IAT min, bwd IAT min, active mean, active max and active min were used. The same machine learning models as SVM, BF and Neural networks were used for classifying the encrypted network traffic. Among these, BF model provides the better R-square and F1 score value that is fitted for the ISCXTor 2016 dataset.

Then, we have taken 14651 samples from the ISCXVPN 2016 dataset (Dener et al., 2023). For analysis and classification of encrypted network traffic, the features like duration, total biat, mean biat, flow bytes, min flow IAT, max flow IAT and std active were used. The same machine learning models like SVM, BF and Neural networks were used for classifying the encrypted network traffic. Among these, BF model provides the best R-square and F1 score value that is fitted for the ISCXVPN 2016 dataset.

4.4. Model comparison

The SVM, BF, Neural networks models were compared using performance metrics such as Precision, Recall and F1 score. Table 2. provides the evaluation metrics obtained by applying various machine learning models. We could not compare the models by calculating accuracy for an imbalanced datasets. So, we have used the performance metrics such as Precision, Recall and F1-score to analyze and classify the encrypted network traffic. Here, the above mentioned performance metrics are calculated as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{F1 score} = \frac{2 * (\text{Precision} * \text{Recall})}{(\text{Precision} + \text{Recall})}$$

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{TN} + \text{FP} + \text{TP} + \text{FN}}$$

where TP,FP,TN and FN denotes true positive, false positive, true negative and false negative respectively.

The R-square value helps us to identify the suitable model for the given datasets. From Table 2, while comparing the R-square value of all the models for the three datasets, BF model provides better performance with the lowest misclassification rate. Further, we have calculated the performance metrics to compare the F1 score of all the models are given in the Table 2. This F1 Score is obtained from the confusion matrix for our three datasets which suggests the performance of the model. Here, the F1 score of the BF model is higher. It shows that, this model provides better performance compared to other models in classifying encrypted network traffic.

Algorithm	R-square	Precision	Recall	F1 score	Misclassification Rate	
					Training	Validation
UNSW NB15 dataset						
Bootstrap Forest (BF)	0.9421	69.930	66.457	68.150	0.0690	0.0418
Neural	0.8895	49.000	46.468	47.701	0.7085	0.7403
SVM	0.8890	57.018	48.320	52.316	0.1321	0.0776
ISCXTor 2016 dataset						
Bootstrap Forest (BF)	0.9332	78.732	79.329	79.029	0.2119	0.1950
Neural	0.8765	58.019	61.521	59.719	0.3000	0.3078
SVM	0.8189	69.618	60.897	64.948	0.3340	0.3349
ISCXVPN 2016 dataset						
Bootstrap Forest (BF)	0.9003	70.035	66.480	68.211	0.2549	0.2567
Neural	0.8035	50.958	36.384	42.455	0.3612	0.3686
SVM	0.8060	63.939	49.950	56.085	0.3278	0.3457

Table 2. Evaluation Metrics obtained by applying various Machine Learning Algorithms

The Receiver Operating Characteristic (ROC) curve and the Precision Recall curve help us to compare the performance of the machine learning models for an imbalanced datasets. The metric Area Under the Curve (AUC) is also useful to compare the performance of the models (i.e) AUC-ROC and AUC Precision Recall (AUC-PR). ROC curve is a graphical representation which shows the performance of the classification models. Precision is defined as the ratio of true positives to the total number of true positives and true negatives. It shows how many predicted responses are correctly classified as true positives. Recall is defined as the ratio of true positives to total number of true positives and false negatives. It shows that how many actual positive responses are correctly classified as positives. Particularly, it helps to identify the performance of the model

which has minority classes of samples present in the datasets taken for experiments. The performance of the machine learning models obtained from the experiments is shown in Figures 3, 4 and 5. Figures 3(A), 4(A) and 5(A) represents the ROC curves obtained for UNSW-NB15, ISCXTor 2016 and ISCXVPN 2016 datasets using machine learning models respectively. Here, the ROC curve lies above the diagonal and it displays the efficiency of the model.

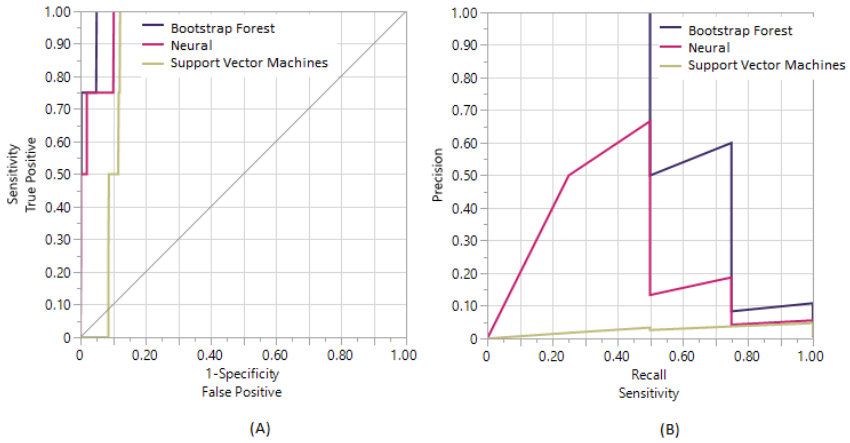

Figure 3. (A) ROC Curve for UNSW-NB15 dataset using machine learning models and (B) Precision Recall Curve for UNSW-NB15 dataset using machine learning models

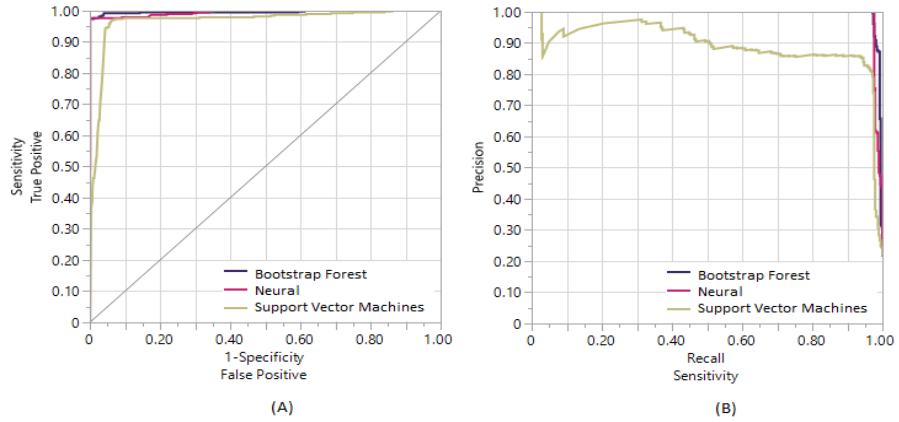

Figure 4. (A) ROC Curve for ISCXTor 2016 dataset using machine learning models and (B) Precision Recall Curve for ISCXTor 2016 dataset using machine learning models

The curve passes through the top left corner of the graph indicates the perfect model suitable for the given datasets. The goodness of the fit for the model are analyzed through the area under the curve.The ROC curve produced by the proposed model is very closer to the top left corner, when compared to ROC curve of all other models. It indicates the proposed model produces better performance compared to all other models. Here, for each dataset, ROC curve of BF model performs better when compared to all other models to classify the encrypted network traffic.

Figures 3(B), 4(B) and 5(B) represent the performance of Precision Recall curves obtained for UNSW-NB15, ISCXTor 2016 and ISCXVPN 2016 datasets using machine learning models respectively. It also helps us to predict the minority classes present in the samples taken for experimental analysis. The Precision Recall curve produced by the proposed model is more closer to the top right corner, when compared to Precision Recall curve of all other models. Further, Table 3. provides the metrics to compare the various machine learning models. From the Table 3. it is observed that the metrics AUC-ROC and AUC-PR of the proposed

model is higher when compared to all other models. It indicates that the proposed model produces better performance when compared to all other models. Here in each dataset, Precision Recall curve of the BF performs better compared to all other models to classify encrypted network traffic.

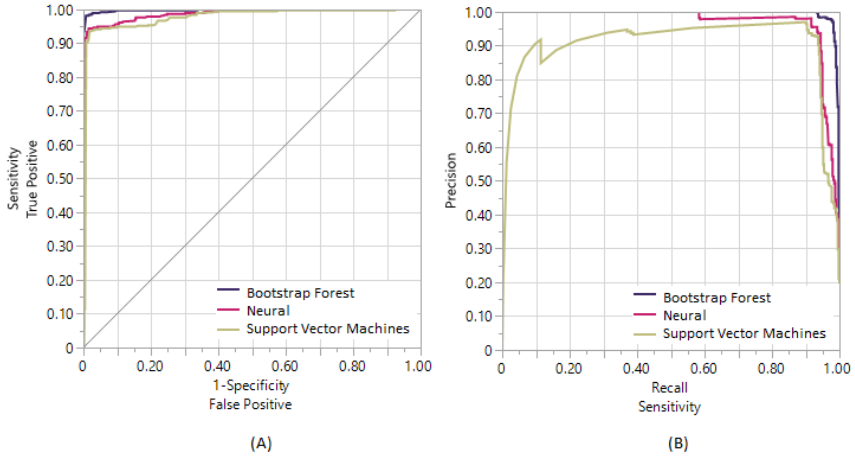

Figure 5. (A) ROC Curve for ISCXVPN 2016 dataset using machine learning models and (B) Precision Recall Curve for ISCXVPN 2016 dataset using machine learning models

Machine learning models	Dataset 1 UNSW-NB15		Dataset 2 ISCXTor 2016		Dataset 3 ISCXVPN 2016	
	AUC-ROC	AUC-PR	AUC-ROC	AUC-PR	AUC-ROC	AUC-PR
Bootstrap Forest (BF)	0.9869	0.6614	0.9960	0.9938	0.9982	0.9953
Neural	0.9688	0.2607	0.9946	0.9896	0.9879	0.9700
SVM	0.8971	0.0265	0.9682	0.8948	0.9776	0.8972

Table 3. Model Comparison using AUC-ROC and AUC-PR

We have done ENTA and classification by using an ensemble method called BF which is the main contribution of our work. The experimental results show that BF model provides better performance and lower misclassification rate when compared to other machine learning models. While focusing the limitations of our BF model, though it is robust ensemble method, they are not immune to the effects of imbalanced datasets, model interpretability and computational cost. In future works, we try to address these limitations.

5. Conclusion

The encrypted network traffic is growing tremendously every day in internet. This traffic needs to be analyzed to classify malicious traffic from the regular benign network traffic present in the network. Our work is to analyze and classify the encrypted network traffic by applying the various statistical and time-based features in different machine learning models. Similar studies are already available in the literature. Our proposed BF model is different from them in terms of analyzing and classifying encrypted network traffic. Moreover, this is the first analysis of encrypted network traffic classification using the BF model in the literature. By introducing bagging methodology, the proposed model is able to provide better performance with the lowest misclassification rate. In addition, we have analyzed features related to packet length and IAT and provided the description about how these features are useful for analyzing and classifying encrypted network traffic.

References

Ahmad, M., Riaz, Q., Zeeshan, M., Tahir, H., Haider, S.A., Khan, M.S. (2021). Intrusion detection in internet of things using supervised machine learning based on application and transport layer features using unsw-nb15 data-set. EURASIP Journal on Wireless Communications and Networking, 2021, 1–23, https://doi.org/10.1186/s13638-021-01893-8

Aouedi, O., Piamrat, K., Parrein, B. (2022). Ensemble-based deep learning model for network traffic classification. IEEE Transactions on Network and Service Management, 19 (4), 4124–4135, https://doi.org/10.1109/TNSM.2022.3193748

Azab, A., Khasawneh, M., Alrabaee, S., Choo, K.-K.R., Sarsour, M. (2024). Network traffic classification: Techniques, datasets, and challenges. Digital Communications and Networks, 10 (3), 676–692, https://doi.org/https://doi.org/10.1016/ j.dcan.2022.09.009

Bagui, S., Fang, X., Kalaimannan, E., Bagui, S.C., Sheehan, J. (2017). Comparison of machine-learning algorithms for classification of vpn network traffic flow using time-related features. Journal of Cyber Security Technology, 1 (2), 108–126, https://doi.org/10.1080/23742917.2017.1321891

Baldini, G. (2020). Analysis of encrypted traffic with time-based features and time frequency analysis. 2020 global internet of things summit (giots) (pp. 1–5).

Bhandari, A., Cherukuri, A.K., Ikram, S.T. (2023). Analysis of encrypted network traffic using machine learning models. Big data analytics and intelligent systems for cyber threat intelligence (Vol. 1, pp. 71–86). River Publishers, Denmark.

Boumhand, A., Singh, K., Hadjadj-Aoul, Y., Liewig, M., Viho, C. (2023). Network traffic classification for detecting multi-activity situations. 2023 ieee symposium on computers and communications (iscc) (pp. 681–687).

Chen, Y., Yang, J., Cui, S., Dong, C., Jiang, B., Liu, Y., Lu, Z. (2024). Unveiling encrypted traffic types through hierarchical network characteristics. Computers & Security, 138, 103645, https://doi.org/10.1016/j.cose.2023.103645 Retrieved from https://doi.org/10.1016/j.cose.2023.103645

Chen, Z., Cheng, G., Wei, Z., Niu, D., Fu, N. (2023). Classify traffic rather than flow: Versatile multi-flow encrypted traffic classification with flow clustering. IEEE Transactions on Network and Service Management, 21 (2), 1446–1466, https://doi.org/10.1109/TNSM.2023.3322861

Cherukuri, A.K., Ikram, S.T., Li, G., Liu, X. (2024). Classification of encrypted network traffic. Encrypted network traffic analysis (pp. 47–59). Springer,Cham, Switzerland.

Dener, M., Al, S., Ok, G. (2023). Rfse-gru: Data balanced classification model for mobile encrypted traffic in big data environment. IEEE Access, 11, 21831– 21847, https://doi.org/10.1109/ACCESS.2023.3251745

Elmaghraby, R.T., Aziem, N.M.A., Sobh, M.A., Bahaa-Eldin, A.M. (2024). Encrypted network traffic classification based on machine learning. Ain Shams Engineering Journal, 15 (2), 102361, https://doi.org/https://doi.org/10.1016/j.asej.2023. .102361

Hu, F., Zhang, S., Lin, X., Wu, L., Liao, N., Song, Y. (2023). Network traffic classification model based on attention mechanism and spatiotemporal features. *EURASIP Journal on Information Security*, *2023* (1), 6, https://doi.org/10.1186/s13635-023-00141-4

Lashkari, A.H., Gil, G.D., Mamun, M.S.I., Ghorbani, A.A. (2017). Characterization of tor traffic using time-based features. *International conference on information systems security and privacy* (Vol. 2, pp. 253–262).

Lotfollahi, M., Jafari Siavoshani, M., Shirali Hossein Zade, R., Saberian, M. (2020). Deep packet: A novel approach for encrypted traffic classification using deep learning. *Soft Computing*, *24* (3), 1999–2012, https://doi.org/10.1007/s00500-019-04030-2 Retrieved from https://doi.org/10.1007/s00500-019-04030-2

Moustafa, N., & Slay, J. (2015). Unsw-nb15: a comprehensive data set for network intrusion detection systems (unsw-nb15 network data set). *2015 military communications and information systems conference (milcis)* (pp. 1–6).

Papadogiannaki, E., & Ioannidis, S. (2021). A survey on encrypted network traffic analysis applications, techniques, and countermeasures. *ACM Computing Surveys (CSUR)*, *54* (6), 1–35, https://doi.org/https://doi.org/10.1145/ 345790

Peng, W., Cui, L., Cai, W., Ding, Z., Hao, Z., Yun, X. (2024). Efficiently and effectively: A two-stage approach to balance plaintext and encrypted text for traffic classification. *arXiv preprint arXiv:2407.19687,*, https:// doi.org/10.48550/arXiv.2407.19687

Seydali, M., Khunjush, F., Akbari, B., Dogani, J. (2023). Cbs: A deep learning approach for encrypted traffic classification with mixed spatio-temporal and statistical features. *IEEE Access*, *11,* 141674-141702, https://doi.org/10.1109/ACCESS.2023.3343189 Retrieved from https://doi.org/10.1109/ACCESS.2023.3343189

Sheikh, M.S., & Peng, Y. (2022). Procedures, criteria, and machine learning techniques for network traffic classification: a survey. *IEEE Access*, *10,* 61135–61158, https://doi.org/10.1109/ACCESS.2022.3181135

UNB (2020). UNSW-NB15 dataset:available online at https://www. unsw. adfa. edu. au/unswcanberra-cyber/cybersecurity/adfa-nb15-datasets. Accessed on, June 10,

UNB (2023). ISCXTor2016 dataset:available online at https://www.unb.ca/cic/datasets/tor.html. Accessed on, September 9,

UNB (2024). ISCXVPN2016 dataset:available online at https://www.unb.ca/cic/datasets/vpn.html. Accessed on, January 20,

Velan, P., ˇCerm´ak, M., ˇCeleda, P., Draˇsar, M. (2015). A survey of methods for encrypted traffic classification and analysis. International Journal of Network Management, 25 (5), 355–374, https://doi.org/10.1002/nem.1901 Retrieved from https://doi.org/10.1002/nem.1901

Xu, S.-J., Geng, G.-G., Jin, X.-B., Liu, D.-J., Weng, J. (2022). Seeing traffic paths: Encrypted traffic classification with path signature features. IEEE Transactions on Information Forensics and Security, 17, 2166–2181, https://doi.org/ 10.1109/TIFS.2022.3179955

Zhao, J., Jing, X., Yan, Z., Pedrycz, W. (2021). Network traffic classification for data fusion: A survey. Information Fusion, 72, 22–47, https://doi.org/10.1016/j.inffus.2021.02.009

SAS Institute Inc. (2023). *JMP® Pro (Version 18)*. Cary, NC: SAS Institute Inc.