World Wide Web
https://doi.org/10.1007/s11280-022-01040-3

A novel feature-based framework enabling multi-type DDoS attacks detection

Lu Zhou[1] · **Ye Zhu**[1] · **Yong Xiang**[1] · **Tianrui Zong**[2]

Received: 18 August 2021 / Revised: 4 March 2022 / Accepted: 8 March 2022

Abstract
Distributed Denial of Service (DDoS) attacks are among the most severe threats in cyberspace. The existing methods are only designed to decide whether certain types of DDoS attacks are ongoing. As a result, they cannot detect other types of attacks, not to mention the even more challenging mixed DDoS attacks. In this paper, we comprehensively analyzed the characteristics of various types of DDoS attacks and innovatively proposed five new features from heterogeneous packets including entropy rate of IP source flow, entropy rate of flow, entropy of packet size, entropy rate of packet size, and number of ICMP destination unreachable packet to detect not only various types of DDoS attacks, but also the mixture of them. The experimental results show that the proposed fives features ranked at the top compared with other common features in terms of effectiveness. Besides, by using these features, our proposed framework outperforms the existing methods when detecting various DDoS attacks and mixed DDoS attacks. The detection accuracy improvements over the existing methods are between 21% and 53%.

Keywords DDoS attacks · Machine learning · Communication system security · Computer networks

1 Introduction

Distributed Denial of Service (DDoS) attacks are great threats to the availability of online services. A DDoS attack, which occurs when the attacker purposely sends large malicious packets from multiple sources to a victim, aims to make the server unavailable by

All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

This article belongs to the Topical Collection: *Special Issue on Decision Making in Heterogeneous Network Data Scenarios and Applications*
Guest Editors: Jianxin Li, Chengfei Liu, Ziyu Guan, and Yinghui Wu.

✉ Lu Zhou
zhoulu@deakin.edu.au

Extended author information available on the last page of the article

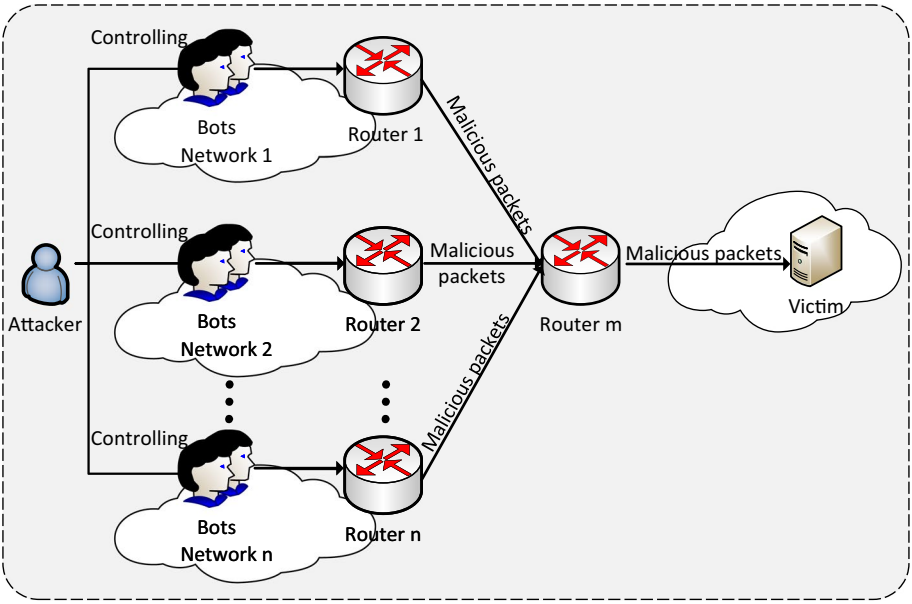

Fig. 1 A sample threat model for a DDoS attack

exhausting its vital resources such as bandwidth, CPU, and memory. DDoS attacks happen more and more frequently every year and a recent report shows that there are more than 23,000 attacks per day (i.e., 16 attacks in every minute) in 2018 [8]. Moreover, DDoS attacks are often organized by individual botnet families. Currently, these families tend to collaborate to arrange mixed DDoS attacks, where various types of attacks are launched on a victim at the same time [45].

Popular DDoS attacks include SYN flooding, DNS amplification, Low Rate, Pulsing attack, and Spoofing attacks [26]. An SYN flooding attack exploits the vulnerability of three-way handshaking in the TCP connection and opens massive half-open TCP connections by deliberately withholding the corresponding final ACK to exhaust the SYN-Queue resources [28]. In a DNS amplification attack, the source IP addresses of the requests sent to the DNS server are pointed to the victim's IP address by the attacker. Since the packet size of a received reply from the DNS server is much larger than the packet size of the request, the number of packets received by the victim will be multiple times more than the number of request packets [22]. A Low Rate DDoS attack aims to evade the traditional anomaly detection by reducing the rate and number of the attack packets below the thresholds so that the attack traffic can be concealed in benign traffic [32]. In a Pulsing attack, the attacker periodically sends a large number of malicious packets within a short time to form pulse-shaped traffic, which can overwhelm the buffer of the victim [40]. To conceal the identity of the attackers and compromised bots, an attacker can apply a spoofing strategy in the implementation phase where the attack packets contain modified information, especially fake source IP addresses, to make it difficult to distinguish and trace back the attack packets [21]. In this paper, we refer SYN flooding, DNS amplification, Low Rate, and Pulsing to as the attacks without using spoofing and refer Spoofing to as any attacks that are launched with spoofing strategy. Figure 1 illustrates an example of a DDoS attack scenario. In a DDoS attack, the attacker first targets a specific victim (e.g., government online

service). Then instead of directly connecting to the victim, the attacker remotely controls thousands of bots, which have been infected with malware, from different networks to send malicious packets to the victim. Since these malicious packets can exhaust the computational resources of the victim, the normal users cannot access to the victim.

1.1 Related work

One detection solution is the traditional anomaly detection. These methods extract a certain feature from the traffic and compare the value with that of normal traffic, and a DDoS attack is detected if the deviation exceeds a threshold. The authors in [47] proposed a generalized entropy method, which measures the probability distribution differences between the attack and legitimate flows, to detect the Low Rate DDoS attack. In [50], an adaptive correlation analysis method was proposed, which analyzes the correlation between the aggregated flows on the victim and the suspicious flows, to decide whether the Pulsing attack is ongoing. In [28], the authors proposed a normalized entropy method, which measures the abnormal changes of randomness concerning IP destination flows, to detect the SYN flooding attack. The authors in [51] proposed an expectation of packet size method, which measures the difference on the variance of packet size between the attack and normal traffic, to detect the Low Rate DDoS attack. An entropy variations method was proposed in [49], which measures the entropy variation between the normal and attack traffic, to detect and traceback DDoS attacks. A joint entropy-based detection and mitigation scheme was proposed in the Software-defined networking (SDN) [25]. An attack is detected when both bandwidth and joint entropy of the attack traffic exceed the thresholds. Wang et al. proposed a three-layers sketch-based structure to detect and mitigate DDoS attacks [46]. The detection theory is based on the divergence of probability distance for the incoming requests, because the attackers are usually persistently launching malicious requests. However, due to only a few features being considered, these methods are only effective on specific attacks and can hardly detect different types of attacks. Moreover, the detection performance is also limited when an attacker mixes various types of attacks. Recently, machine learning (ML) methods offer a new route for many classification and regression tasks. Apart from the usage in the prevention of cyber attacks on energy internet systems [31], predicting traffic flow for electric vehicles charging service [30], and network representation learning [17, 33, 48], they have also been applied to classify normal and attack traffic based on a set of features and thus decide whether a DDoS attack is ongoing [23, 43]. Zhu et al. proposed a privacy-preserving cross-domain detection for SDN [53]. They apply the KNN algorithm to detect a certain type of attack. The authors in [42] proposed a genetic algorithm combined with the KNN classifier to detect known and unknown DDoS attacks. Asmir et al. proposed a framework with feature engineering and machine learning-based detection [10]. Seven features are selected and tested on five methods, but most of the features (i.e., from_router, from_switch, to_router, and to_switch) could only be obtained in the simulation environment. Alsirhani et al. proposed a DDoS attack detection system that dynamically selects classification algorithms to detect different DDoS patterns [13]. Ali et al. [12] proposed a method based on multilevel autoencoder-based feature learning. Jia et al. proposed a Convolutional Neural Network (CNN) model to detect DDoS attacks [24]. However, their performance heavily relies on the extracted features that present the key patterns of different types of attacks. The selected features, such as ratio of ICMP, number of SYN, number of ACK, number of HTTP, variance of packet size, inbound to

outbound traffic ratio, entropy of IP source flow, and entropy of packet type, are ineffective to detect multiple types of attacks. Therefore, the existing methods cannot determine the most informative features for various types of DDoS attacks detection.

1.2 Contributions

To accurately detect SYN flooding, DNS amplification, Low Rate, Pulsing, Spoofing, and mixed DDoS attacks, in this paper, we comprehensively analyzed the characteristics of various types of DDoS attacks and proposed five new features from heterogeneous packets including entropy rate of IP source flow, entropy rate of flow, entropy of packet size, entropy rate of packet size, and number of ICMP destination unreachable packet. The average detection accuracy of the proposed features on specific types of attacks and mixed attacks are 0.99. Compared to the existing methods, the detection accuracy improvements are between 21% and 53%.

Our main contributions are summarized as follows:

- We propose five new features that are entropy rate of IP source flow, entropy rate of flow, entropy of packet size, entropy rate of packet size, and number of ICMP destination unreachable packet. Since the value of each feature has the same trend of change under all five types of DDoS attacks, the proposed features are effective on DDoS attacks detection.
- We theoretically analyze the improvements of the proposed features over the existing features and evaluated their effectiveness on the real DDoS attack datasets.
- By using the five features, our proposed framework outperforms the existing methods when detecting all five types of DDoS attacks and mixed DDoS attacks. The detection accuracy improvements over the existing methods are between 21% and 53%.

The rest of the paper is organized as follows. Section 2 presents the details of the proposed framework. The feature performance analysis is demonstrated in Section 3. Section 4 evaluates the performance of the proposed framework using real data and compares it with the state-of-the-art methods. Finally, Section 5 concludes the paper.

2 Proposed framework

This section presents our proposed framework to detect multi-type DDoS attacks. Figure 2 shows the overall architecture of the proposed detection system. Firstly, we extract the proposed features from the samples. Secondly, the feature preprocessing step is conducted. Finally, the classification process is performed to detect if a victim is under different types of attacks.

2.1 Raw dataset splitting

Since all features need to be extracted from samples, prior to the proposed framework, it is necessary for us to measure over sliding windows of time and split the raw dataset into file fragments to collect samples. For a given raw pcap file L with time length T, the raw file is split as $L = \{l_i | i \in (1, N)\}$, where l_i denotes the ith file fragment in time interval

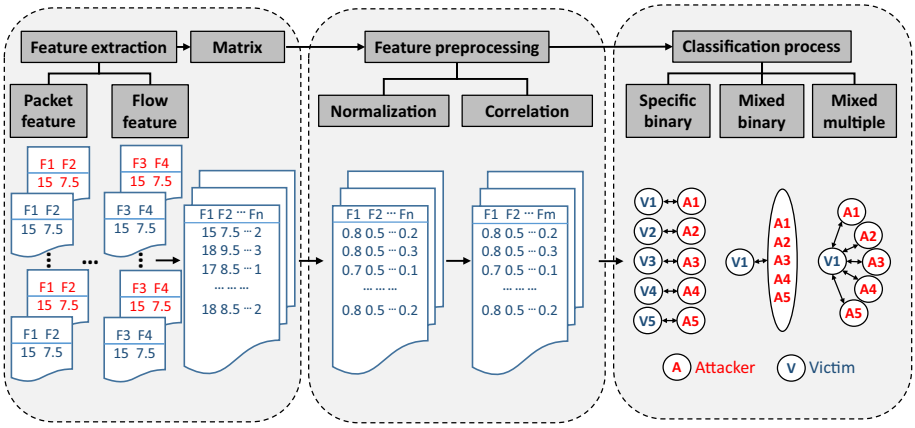

Fig. 2 The overview of the framework

$[T_i, T_{i+\triangle t}]$ according to window time $\triangle t$. Then we can have the number of file fragments N as follows:

$$N = \frac{T}{\triangle t}. \tag{1}$$

2.2 The proposed feature extraction

In this subsection, we analyze the characteristics of different DDoS attacks, introduce the proposed five new features including entropy rate of IP source flow, entropy rate of flow, entropy of packet size, entropy rate of packet size, and number of ICMP destination unreachable packet, demonstrate their improvements over the existing features, and explain their effectiveness against all five DDoS attacks.

2.2.1 Entropy rate of IP source flow

We consider the different distribution of flow between legitimate and attack traffic. Since in DDoS attacks, an attacker generates massive malicious packets and these packets present aggregated flow distribution, while the normal flows present more randomness distribution in the network [49]. The randomness of flow distribution can be measured by entropy measurement. That is, the higher randomness the source flow has, the greater the entropy is [47].

Following the previous work [47], we classify the packets with same IP source address into IP source flow. For a given data X with source flow distribution P_s, $P_s = \{P_i | i \in [1, n_s]\}$, where P_i is the probability of ith flow and n_s is the number of IP source flows. According to [47], the entropy of IP source flow $H_s(X)$ can be defined as:

$$H_s(X) = -\sum_{i=1}^{n_s} P_i \log(P_i). \tag{2}$$

From (2) it is clear that if $\exists i \in [1, n_s]\ P_i = 1$, $H_s(X)$ takes its minimum value 0; If $\forall i \in [1, n_s]\, P_i = \frac{1}{n_s}$, $H_s(X)$ reaches its maximum value $\log(n_s)$. To be more specific, when n_s is fixed, the aggregated distributed P_i makes $H_s(X)$ decrease, while the uniform distributed P_i makes $H_s(X)$ increase. Therefore, under an assumption that the number of IP source flow n_s is not significantly increased, such as the Low Rate DDoS attack, entropy of IP source flow will decrease when a DDoS attack is ongoing. However, one important character of most DDoS attacks is that the number of IP source flow n_s will increase dramatically. In this scenario, $H_s(X)$ will be monotonically increasing with n_s [20]. As a result, entropy of IP source flow will decrease under the Low Rate DDoS attack while increase under other DDoS attacks [52], which makes entropy of IP source flow incapable of effectively justifying whether a DDoS attack is happening or not.

To accurately detect a DDoS attack, we propose entropy rate of IP source flow and it is defined as:

$$H'_s(X) = \frac{-\sum_{i=1}^{n_s} P_i \log(P_i)}{n_s}, \tag{3}$$

where P_i is the probability of the ith IP source flow and $H'_s(X)$ is the entropy rate of IP source flow. According to [20], $H'_s(x)$ is monotonically decreasing with n_s. When n_s increases substantially during a DDoS attack, $H'_s(x)$ will be reduced by a considerable amount. As for the Low Rate DDoS attack, n_s is stable while the entropy, which is the numerator in (2), is reduced. As a result, $H'_s(x)$ will also decrease. Therefore, our proposed entropy rate of IP source flow will decrease under all DDoS attacks, which indicates the effectiveness of this feature in detecting various DDoS attacks.

2.2.2 Entropy rate of flow

In the IP source flow, the flows are nondirectional, while in this subsection, we define the flow as directional flow. In the entropy rate of flow, we classify the packets with same source IP address and destination IP address into a flow. Therefore, we can measure the abnormality of the flow distribution on the destination side, especially on the targeted victim. An important feature of DDoS attacks is that the aggregated flows are purposely targeted on the victim, which is the attack goal [50]. According to [39], the ith flow F_i can be defined by the ith flow header D_i and the ith packet set S_i as $F_i = \{D_i, S_i\}$. D_i is used to identify the flow and can be defined as:

$$D_i = (IP_{src}, IP_{dst}), \tag{4}$$

where IP_{src} and IP_{dst} are the source and destination IP address, respectively. $F_{ab} = F_{\mathbf{ab}} + F_{\mathbf{ba}}$, where **ab** represents the packets from source IP a to destination IP b, and vice versa. Let n_f be the number of flows and P_f be the probability distribution of S_i, the entropy of flow $H_f(X)$ can be calculated using (2) with P_f. Similar to Section 2.2.1, when the number of flows is significant during a DDoS attack, the entropy of flow will increase. When the number of flow is stable and the distribution of P_f concentrates, such as in the Low Rate DDoS attack, the entropy of flow will decrease. Therefore, we propose entropy rate of flow to avoid the conflicting results as the entropy rate of flow will drop under all DDoS attacks. The entropy rate of flow $H'_f(X)$ is defined as:

$$H'_f(X) = \frac{H_f(X)}{n_f}. \tag{5}$$

2.2.3 Entropy of packet size and entropy rate of packet size

We consider the features that are related to the packet size and propose two features, which are entropy of packet size and entropy rate of packet size. Since a DDoS attack is the resource competition between attackers and defenders: a successful DDoS attack is that the resources of attackers outnumber those of defenders, and the key resources of attackers are the enormous packets [32]. To avoid being detected, a common strategy is that attackers send the packets with random sizes, while the sizes of normal packets are either small for protocol packets (e.g., SYN packets) or maximized for application data (e.g., HTTP packets) [51]. To distinguish the normal and attack traffic, we use entropy of packet size and entropy rate of packet size to measure abnormal distribution on the packet size.

According to the DIX Ethernet V2 protocol, the maximum packet size in the network is 1514 bytes [34]. Therefore, we define $NP_k = |\{j|PS_j = k\}|$, where PS_j is the size of the jth packet and NP_k is the total number of packets whose sizes are k, $k = 0, 1, \ldots, 1514$, and $|.|$ returns the total number of elements in a set. We also define $I = \{k|NP_k \neq 0\}$ and $n_d = |I|$. For $i \in I$, the probability for ith packet size PPS_i can be calculated as $PPS_i = \frac{NP_i}{\sum NP_i}$. Let us define the entropy of packet size $H_{PS}(X)$ as:

$$H_{PS}(X) = -\sum PPS_i \log(PPS_i). \tag{6}$$

Due to the randomness of packet sizes in the attack traffic, the value of n_d will grow during a DDoS attack and thus the entropy of packet size will increase. To measure the change of entropy of packet size, we propose entropy rate of packet size. The entropy rate of packet size $H'_{PS}(X)$ is defined as:

$$H'_{PS}(X) = \frac{H_{PS}(X)}{n_d}. \tag{7}$$

Since the entropy rate of packet size is monotonically decreasing with n_d, and n_d increases when the attack happens, the entropy rate of packet size will decrease under all DDoS attacks. Therefore, we use entropy of packet size and entropy rate of packet size as our distinctive features for DDoS attacks detection.

2.2.4 Number of ICMP destination unreachable packet

The number of ICMP destination unreachable packet measures the abnormal number of ICMP destination unreachable packets when a DDoS attack is ongoing. The destination unreachable packet is generated by the host or the inbound gateway when the host is unreachable or the service port is inactive to a request packet [1]. During a DDoS attack, since the destination host is overwhelmed by massive attack packets and thus unreachable for a request packet, the number of ICMP destination unreachable packet will increase dramatically. In addition, when an attacker performs port scanning in the early stage of a DDoS attack, a destination unreachable packet is generated if the service port is inactive.

Particularly, the number of these packets is significant when an attacker launches a Spoofing attack as the packet header (e.g., IP source address) has been modified. Therefore, we use the feature number of ICMP destination unreachable packet to measure the abnormality of DDoS attacks.

Note that the five proposed features are extracted from heterogeneous packets, that is, the feature number of ICMP destination unreachable packet is extracted from ICMP packets while the other four features are extracted from all types of packets.

2.3 Feature preprocessing

The extracted features are preprocessed using the Min-Max normalization and Pearson correlation coefficient, respectively. Note that the packet-based and flow-based features are extracted separately. Then by following [23, 43], and [37], they are further combined to construct the feature matrix in the classification step.

2.3.1 Feature normalization

To fairly compare the effectiveness of different features, we first normalize each feature set x into the range [0, 1] using the Min-Max normalization:

$$x_{norm} = \frac{x - \min(x)}{\max(x) - \min(x)}, \tag{8}$$

where $\max(x)$ and $\min(x)$ are the maximum and minimum values of the corresponding feature, respectively, and x_{norm} is the normalized feature.

2.3.2 Pearson correlation coefficient

To avoid using linear-related features, we select linearly independent features using the Pearson correlation coefficient (PCC) that can be defined as:

$$r_{UV} = \frac{\sum_{i=1}^{N}(U_i - \bar{U})(V_i - \bar{V})}{\sqrt{\sum_{i=1}^{N}(U_i - \bar{U})^2 \sum_{i=1}^{N}(V_i - \bar{V})^2}}, \tag{9}$$

where U_i and V_i are two different normalized features extracted from the ith file fragment, $\bar{U}$ and $\bar{V}$ are the mean values of feature U and feature V, respectively, and r_{UV} is the PCC between U and V. The PCC value ranges from -1 to 1, where -1 means the two features are negatively linearly related, 0 means the two features are irrelevant, and 1 means the two features are positively linearly related.

2.4 Classification tasks

After all features are extracted and selected, the classification process is conducted to classify normal and attack samples. To test the effectiveness of the features, six classification methods are applied, which are Decision Tree [41], Deep Learning [29], K Nearest Neighbor [19], Logistic Regression [27], Random Forest [16], and Support Vector

Table 1 The overview of the extracted features

Feature No.	Feature
F1	Entropy rate of IP source flow
F2	Entropy rate of flow
F3	Entropy of packet size
F4	Entropy rate of packet size
F5	Number of ICMP destination unreachable packet
F6	Number of packet
F7	Entropy of packet type
F8	Maximum of packet size
F9	Variance of packet size
F10	Number of ICMP packet
F11	Ratio of ICMP packet
F12	Number of HTTP packet
F13	Ratio of HTTP packet
F14	Number of DNS packet
F15	Ratio of DNS packet
F16	Number of SYN packet
F17	Ratio of SYN packet
F18	Number of ACK packet
F19	Ratio of ACK packet
F20	Number of flow
F21	Entropy of flow
F22	Entropy of IP source flow
F23	Entropy of IP destination flow
F24	Entropy of source port
F25	Entropy of destination port
F26	Mean inbound to outbound traffic ratio
F27	Entropy of inbound to outbound ratio

Machine [44]. For a fair comparison, we use the default parameters for all algorithms in sklearn [9].

Remark 1 Since one of our major points is to demonstrate that our features are effective by even using basic models with default parameters, we choose not to select advanced classification models or tune parameters of the models to achieve higher classification scores.

For each classification method, the classification process is performed on a victim who suffers specific types of attacks and mixed type attacks. In specific types of attacks classification, a victim suffers one type of attack at a time and we conduct binary classification between the normal and attack traffic. In the mixed type attacks classification, a victim is attacked by the mixed attacks. We use binary classification to detect whether the victim has been attacked and multi-class classification to identify the types of attacks.

3 Feature performance analysis

To show the excellent effectiveness of the proposed features, in this section, we compare our five features with the other common 22 features and rank all features in terms of detection performance. The details of all features are listed in Table 1, where our proposed features are named as F1-F5 and the other 22 common features are represented by F6-F27.

3.1 Feature extraction

Because various packet types are utilized by botnets, such as HTTP, ICMP, SYN, and DNS, which target multiple layers of the victim [45], the extracted 22 common features range from the data link layer to the application layer to monitor abnormal behaviors in those layers. Meanwhile, these features are privacy-preserving since they are not related to the content, which protects the application data. These features can be categorized into packet-based and flow-based features and extracted separately.

3.1.1 Packet features

The packet features include basic statistic features and protocol features. Statistical features, such as the number of packets and entropy of the packet type, measure the basic volume of packets. Since 75.7% of the attacks are volumetric attacks [7] and these attacks mainly depend on massive packets, the feature number of packets is used to measure the volume of the attacks. Meanwhile, DDoS attacks are launched using certain packet types such as TCP, UDP, and ICMP [45]. To measure the abnormal distribution of packet types, the feature entropy of packet type is utilized as another basic packet feature. We also extract other statistical features including the maximum of packet size and variance of packet size to measure abnormal packets on packet size. Protocol features based on the changes in number and ratio of protocol packets. Since the main protocols utilized by DDoS botnets are ICMP and HTTP [45], we adopt the number of ICMP packet, the ratio of ICMP packet, the number of HTTP packet, and the ratio of HTTP packet as the protocol features. We have also included other DDoS attacks related protocol features such as the number of DNS packet, the ratio of DNS packet, the number of SYN packet, the ratio of SYN packet, the number of ACK packet, and the ratio of ACK packet.

3.1.2 Flow features

The flow features measure the distribution difference between legitimate and attack flows. In a network, the nondirected packets may be legitimate, while the aggregated flows that purposely target on the victim are malicious. According to [52], the number of flows, which is a basic flow feature that measures the volume of the aggregated flows, has a significant increase in most DDoS attacks, especially in the Spoofing attack. Therefore, we use the number of flows as one of the flow features. When a DDoS attack is ongoing, the traffic between the inbound and outbound presents imbalanced traffic ratio. To measure the imbalanced traffic ratio between the inbound and outbound traffic, the mean inbound to outbound traffic ratio and the entropy of inbound to outbound ratio are proposed to calculate the average imbalanced traffic ratio. We also extract other entropy-based flow features such as the entropy of flow, the

Table 2 The details of the datasets

Dataset		Duration (s)	Packet	PacketRate	Average bits/s
Normal	ISCX	29,135.87	11,709,971	401	2.864 Mbps
	LLS	6,616.45	347,987	56.4	78 kbps
Attack	SYN	300	3,368,576	11,228.6	66 Mbps
	DNS	300	9,319,873	31,064.8	39 M
	LowRate	300	166,448	554.8	998 K
	Pulsing	300	37,116	299.9	130K
	Spoofing	300	543,957	1,813.2	992K

entropy of IP source flow, the entropy of IP destination flow, the entropy of source port, and the entropy of destination port to measure the abnormal distribution on the port number.

3.2 Feature ranking

3.2.1 Datasets

We collect both normal and attack datasets to rank the features. Normal datasets help us evaluate the intrinsic properties of legitimate traffic and explore abnormal features that the attacks present. However, the normal traffic also varies, which can affect the detection results. Therefore, in this paper, we use two normal datasets to avoid overfitting. The details of the normal datasets are listed below.

- ISCX: This dataset is provided by the Information Security Centre of Excellence (ISCX) at the University of New Brunswick, which resembles the true real-world data to generate benchmark datasets for intrusion detection [6]. The datasets consist of normal and anomalous traffic. We choose the dataset on Monday, which has benign background traffic that resembles normal human activities. This dataset is also used by [11].
- LLS: This dataset is from the Lincoln Laboratory of MIT [2], which contains background traffic that is used to test intrusion detection evaluations. This dataset is also used by [47].

We collect five real-world sophisticated attack datasets regarding the attack datasets to explore the important features that are effective to specific attacks and evaluate the detection performance. The details of these datasets are listed as follows and demonstrated in Table 2.

- SYN flooding (SYN): This dataset is from the Impact [4], which is collected by the University of Southern California-Information Sciences Institute and contains an SYN flooding attack.
- DNS amplification (DNS): This dataset comes from the Impact [4] and contains a DNS amplification attack.

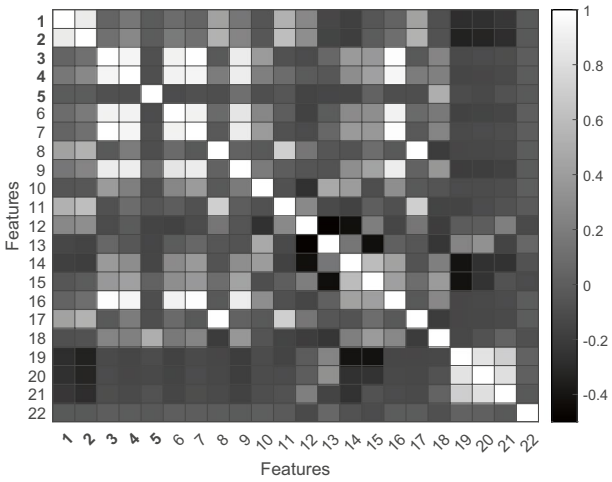

Fig. 3 The Pearson correlation coefficient of the features

- Pulsing (Puls.): This dataset is provided by the Impact [4]. This dataset is also used by [51].
- Low Rate (Low): This dataset is a Low Rate DDoS attack from the Center for Applied Internet Data Analysis (CAIDA) [3]. This dataset is also used by [47].
- Spoofing (Spoof.): This dataset is from the CAIDA [5]. The attack is implemented by modifying the source IP address of the packets to conceal the identity of attackers and compromised machines.

3.2.2 Performance ranking

Before ranking the features, we first preprocess all features using the steps introduced in Section 2.3. Figure 3 illustrates the Pearson correlation coefficient among the features F1-F22. Since the Pearson correlation coefficients between the F23 and F21, F24 and F22, F25 and F22, F26 and F19, and F27 and F22, are 0.9857, 0.9770, 0.9663, 0.9804, 0.9829, respectively, which means the features in each pair are highly linearly correlated. Therefore, we remove the redundant features F23-F27, which are entropy of IP destination flow, entropy of source port, entropy of destination port, mean inbound to out bound traffic ratio, and entropy of inbound to outbound ratio, to avoid over-fitting in the classification stage.

The area under the ROC (Receiver Operating Characteristics) curve (AUC) [15] is a common metric to rank the feature importance, especially in binary classification with imbalanced data distribution [18], which measures the likelihood that a randomly drawn positive example has a higher modeled probability than a randomly drawn negative example [15]. For a given dataset $\mathcal{D} = \{(\mathbf{x}_i, y_i) \in \mathbb{R}^d \times \{-1, +1\} | i \in [n]\}$, where $[n] = \{1, 2, ..., n\}$, denote the positive examples as $\mathbf{x}_i^+$ and the negative examples as $\mathbf{x}_j^-$. The AUC of feature f on $\mathcal{D}$ is:

$$\text{AUC} = \sum_{i=1}^{n_+} \sum_{j=1}^{n_-} \frac{\mathbb{II}[f(\mathbf{x}_i^+) > f(\mathbf{x}_j^-)] + \frac{1}{2}\mathbb{II}[f(\mathbf{x}_i^+) = f(\mathbf{x}_j^-)]}{n_+ n_-}, \tag{10}$$

where n_+ and n_- are the numbers of positive and negative examples, respectively and $\amalg[\pi]$ is the indicator function that returns 1 if π holds and 0 otherwise. The value of the AUC ranges from 0 to 1 and a higher AUC score of a feature infers that a binary classifier usually can achieve a better classification performance with that feature [18]. Therefore, in this paper we evaluate the feature effectiveness and rank the independent features according to the mean AUC scores.

We measure the AUC scores of the features on the five real attacks, which are SYN flooding, DNS amplification, Low Rate, Pulsing and Spoofing attacks. In a specific type of attack detection, the AUC score of a feature varies according to the background traffic. We apply cross-validation on the five attacks using the two normal traffic, which are ISCX and LLS, as the background traffic. In a nutshell, we have 10 detection pairs in this experiment. Since in a real DDoS attack scenario, the real traffic in the attack scenario contains both attack and background traffic. Therefore, for each detection pair, the attack traffic is obtained by mixing the attack dataset and the last 300 seconds of its corresponding normal dataset.

Remark 2 In all attack file fragments, a few file fragments may be inevitably normal. However, due to the missing labels in real datasets, we cannot precisely label each file fragment. For simplicity, we consider all file fragments split from the normal traffic as normal samples, and all file fragments divided from the merged attack traffic, which is a combination of background traffic and attack traffic, as attack samples, following the previous work [47].

Table 3 shows the AUC scores for all features under various attacks when ISCX and LLS are used as the background traffic. Since the AUC of a feature varies according to the detection pairs, we use the mean value of AUC over the 10 detection pairs to present the overall importance of the feature and rank the features according to the mean AUC. From Table 3, we can observe that the AUC scores of the proposed five features F1-F5 are highly ranked compared to other common features, which demonstrates the importance of these features in the classification.

4 Experimental results

4.1 Experiment settings

We use a 64-bit Windows 10 system with an Intel(R) Cores(TM) i7-8650U CPU @1.90 GHz and 16 GB RAM to deploy our framework. The datasets we used for experiments are the real public datasets shown in Section 3.2.1. The algorithms to parse the packet header, extract features from the datasets and conduct ML methods for classification are implemented in Python 3.7. In the following experiments, all raw files are split into file fragments using a two-seconds time window, which is a common choice for intrusion detection [38].

4.2 Evaluation metric

Since imbalanced data is the inherent characteristic in real DDoS attacks detection, in this paper we use the F1-score, which is a widely used metric for

Table 3 The AUC score of each feature for the binary classification

Feature	ISCX					LLS					Mean	Rank
	SYN	DNS	Low	Puls.	Spoof.	SYN	DNS	Low	Puls.	Spoof.		
F1	0.95	0.99	0.84	0.64	0.93	0.93	0.93	0.93	0.81	0.93	0.88	**4**
F2	0.95	0.99	0.87	0.59	0.88	0.94	0.94	0.93	0.78	0.93	0.88	**5**
F3	0.99	1.0	0.87	0.62	0.95	0.99	1.0	0.99	0.92	0.99	0.93	**1**
F4	0.99	0.99	0.85	0.58	0.94	0.96	0.96	0.96	0.86	0.96	0.90	**3**
F5	0.87	0.50	1.0	1.0	1.0	0.88	0.50	1.0	1.0	1.0	0.87	**6**
F6	0.99	1.0	0.86	0.59	0.95	0.99	1.0	0.99	0.90	0.99	0.92	2
F7	0.86	0.89	0.79	0.62	0.53	0.62	0.63	0.90	0.83	0.66	0.73	15
F8	0.50	0.51	0.56	0.58	0.60	0.87	1.0	0.71	0.69	0.71	0.67	20
F9	0.79	0.52	0.57	0.61	0.76	0.99	0.81	0.70	0.67	0.59	0.70	18
F10	0.87	0.50	1.0	1.0	1.0	0.88	0.50	1.0	1.0	1.0	0.87	7
F11	0.87	0.50	0.99	0.99	0.99	0.88	0.50	0.99	0.99	0.99	0.86	8
F12	0.99	0.53	0.53	0.53	0.53	1.0	0.50	0.50	0.50	0.50	0.61	22
F13	0.96	0.77	0.67	0.56	0.73	0.99	0.52	0.52	0.51	0.52	0.67	19
F14	0.91	0.55	0.55	0.55	0.55	0.99	0.83	0.83	0.83	0.83	0.74	14
F15	0.64	0.74	0.66	0.59	0.71	0.53	0.53	0.58	0.73	0.53	0.62	21
F16	1.0	0.60	0.79	0.66	0.60	1.0	0.50	0.87	0.74	0.50	0.72	16
F17	0.76	0.79	0.63	0.66	0.77	0.88	0.56	0.75	0.71	0.56	0.70	17
F18	0.99	0.59	0.62	0.58	0.74	0.99	0.70	0.86	0.85	0.97	0.78	13
F19	0.65	0.99	0.93	0.61	0.98	0.55	0.93	0.86	0.69	0.92	0.81	11
F20	1.0	0.51	0.70	0.60	1.0	0.99	0.80	0.99	0.98	0.99	0.85	9
F21	0.99	0.55	0.60	0.51	1.0	0.99	0.80	0.97	0.91	0.99	0.83	10
F22	1.0	0.51	0.57	0.65	0.78	0.99	0.81	0.98	0.65	0.99	0.79	12

imbalanced data [14], to evaluate the classification performance. Although the overall accuracy is another common evaluation metric, it is not suitable for imbalanced distribution [36]. Therefore, in this paper we only used F1-score to measure the detection performance for all of the methods. The F1-score can be calculated as:

$$\text{F1-score} = \frac{2}{C}\sum_{c=1}^{C}\frac{\text{Precision}_c \times \text{Recall}_c}{\text{Precision}_c + \text{Recall}_c}, \tag{11}$$

$$\begin{cases} \text{Precision}_c = \frac{\text{TP}}{\text{TP+FP}}; \\ \text{Recall}_c = \frac{\text{TP}}{\text{TP+FN}}, \end{cases} \tag{12}$$

where C is the total number of predicted classes, Precision_c is the Precision value of the cth class, Recall_c is the Recall value of the cth class, $c = 1, 2, \ldots, C$. TP is the true positive which means that the positive point is detected as positive, FP is the false positive which means that the negative point is detected as positive, and FN is the false negative which means that the positive point is detected as negative. From (11) and (12), we can clearly see that the F1-score is an unweighted average over the classification result of each class,

Table 4 The binary classification results for 10 detection pairs

Attack	Normal	DT	DL	KNN	LR	RF	SVM
SYN	ISCX	0.99	0.76	0.92	0.97	0.98	0.98
	LLS	0.99	0.87	0.98	0.99	0.99	0.99
DNS	ISCX	1.0	0.49	1.0	1.0	1.0	1.0
	LLS	0.99	0.48	0.99	0.99	0.99	0.99
LowRate	ISCX	0.99	0.99	0.99	1.0	0.99	0.99
	LLS	0.99	1.0	1.0	0.99	1.0	1.0
Pulsing	ISCX	1.0	0.99	1.0	1.0	1.0	0.99
	LLS	0.99	1.0	0.99	1.0	1.0	1.0
Spoofing	ISCX	1.0	0.98	1.0	1.0	1.0	0.99
	LLS	0.99	0.98	1.0	0.99	1.0	0.99

meaning that it is ideal for imbalanced data. The value of F1-score ranges from 0 to 1 and a higher F1-score value indicates a better result.

Due to the abrupt increasing of DDoS attack data, the selection of the splitting ratio for the training and testing set may cause overfitting or underfitting. Attackers can change the strength and peak time according to the attack ability and defense ability of the target. Therefore, using the fixed training and testing set may produce a high variance, while splitting the main attack features in the training set may cause overfitting. In contrast, the biased data that occur in the testing set may cause underfitting. A commonly used approach to assess the quality and stability of the data and classifiers is the K-fold cross-validation (CV). In K-fold CV, the samples are divided into K equal subsets and trained for k times. For each iteration, one subset is used for testing, $K-1$ groups are applied for training, and the mean value is used for overall evaluation. In this paper, all classification methods are applied with 10-fold CV and the F1 score is the averaged value of the 10-fold CV.

4.3 Classification results in different attack scenarios

To evaluate the performance of our proposed features, we design the experiments for two strategies that attackers may take to launch the five attacks: (i) launching a specific type of attack and (ii) launching mixed types of attacks. As mentioned in Section 2.4, the classifications are conducted using Decision Tree (DT), Deep Learning (DL), K Nearest Neighbor (KNN), Logistic Regression (LR), Random Forest (RF), and Support Vector Machine (SVM) classifiers.

4.3.1 Specific type of attacks

In the specific type of attacks, attackers launch one of five attacks at a time. We test our framework when a victim suffers a binary classification between the normal and the particular attack traffic. We continue to use the 10 detection pairs to evaluate the performance. For each detection pair, we label the normal traffic as 0 and the attack traffic as 1. Table 4 shows that most of classifiers can achieve high accuracy in the binary classification. For example, for DT, KNN, LR, RF, and SVM, the F1-scores under all attack scenarios are close to 1.

Table 5 The binary classification performance when a victim suffers five types of mixed attacks

	DT	DL	KNN	LR	RF	SVM
ISCX	1.0	0.92	1.0	1.0	1.0	1.0
LLS	0.99	0.98	1.0	1.0	1.0	1.0

4.3.2 Mixed attacks

In Section 4.3.1, we assume that there is only one attack ongoing. However, there is a trend that different botnets cooperate to launch mixed attacks aiming at the same victim [45]. Therefore, we also test our method in the mixed attack scenario. Assume that a victim is attacked by all five attacks at the same time, if we are only interested in detecting attacks regardless of their types, this detection task becomes a binary classification problem, where the attack traffic is generated by merging all attack datasets and the normal traffic, the normal traffic is labeled as 0, and the attack traffic is labeled as 1. Table 5 shows the classification results when ISCX and LLS are attacked by the five mixed attacks. Compared with the classification results in Table 4, the F1-scores of all classification methods are slightly higher since the feature abnormality increases in mixed attacks scenarios.

To evaluate the performance of identifying different attack types in the mixed attack stream, we label the two normal traffic as 0, and the five attack traffic as 1-5, respectively[1]. Table 7 shows the classification results. Most classifiers continue to show good performance. When the DT, KNN and RF classifiers are used, the mean F1-scores on the two normal datasets are close to 1. Therefore, all experiments show that the five proposed features combines with the framework can detect both specific and mixed attacks.

4.4 Comparison with existing methods

In this section, we select Random Forest as the classifier and compare our framework using the proposed five features with the following eight state-of-the-art algorithms. The window time $\triangle t$ for all methods are set to 2 for a fair comparison. The detailed descriptions and settings of these methods are listed as follows.

- RADAR [50]: This method detects SYN flooding by analyzing the SYN to ACK packet ratio. The threshold of the SYN/ACK ratio is set to 0.5 as indicated by the authors.
- Umbrella [35]: This method identifies malicious traffic by analyzing the packet loss rate and the number of packet. The threshold of the packet loss rate is set to 5% as indicated by the authors.
- GE [47]: This method measures the generalized entropy distance between legitimate and attack traffic. The order of generalized entropy α is set to 5 as indicated by the authors.
- Entropy [49]: This method measures the distance of entropy of IP source flow between legitimate and attack traffic.

[1] Both LR and SVM are originally designed for binary classification. Here, we use the multinomial loss fit for LR and one-vs-one strategy for SVM on multi-class classification as the default settings from the scikit-learn package in Python.

Table 6 The labels and ratios of the datasets. The ratio represents the percentage of the sample in the total datasets

Datasets		Binary class labelling		Multi-class labelling	
		Label	Ratio(%)	Label	Ratio(%)
Normal	ISCX	0	95.92	0	79.16
	LLS	0		1	16.75
Attack	SYN	1	4.07	2	0.81
	DNS	1		3	0.81
	LowRate	1		4	0.81
	Pulsing	1		5	0.81
	Spoofing	1		6	0.81

Table 7 The multi-class classification performance when a victim suffers five types of mixed attacks

	DT	DL	KNN	LR	RF	SVM
ISCX	0.97	0.62	0.93	0.65	0.98	0.68
LLS	0.98	0.67	0.94	0.82	0.98	0.75

- SAFETY [28]: This method measures the normalized entropy distance between legitimate and attack traffic.
- MLP [37]: This method selects three features, which are number of packets, entropy of IP source flow, and average of inter-arrival time, and chooses the best method MLP (i.e., DL) from four methods.
- SKM-HFS [23]: This method selects 6 features from 9 features using Hybrid feature selection method and uses Semi-supervised K-means algorithm to detect attacks.
- Fuzzy [43]: This method selects 8 features from 23 features based on chi-square and selects the best method Fuzzy c-means from 6 methods.

Entropy-based methods, i.e., GE, Entropy, and SAFETY, identify a file fragment as an attack when its entropy value is smaller than that of the normal benchmark or a threshold based on the assumption that the number of flows is stable. However, this classification criterion should be opposite on some datasets when the number of flows increases [52], e.g., in the LLS-Spoofing detection pair task. Thus, we manually reverse their classification labels if the mean entropy of attack samples is greater than that of normal samples. Note that the entropy-based methods do not provide the entropy thresholds in their parameter settings, and thus we have to search for the optimal thresholds by ourselves. For a fair comparison, we search for the best threshold for each entropy-based method on each dataset and report the highest F1-score as the best performance for each method. The search range is in 100 entropy values between minimum and maximum entropy values. For our proposed framework, the five new features, which are entropy of packet size, entropy rate of packet size, entropy rate of IP source flow, entropy rate of flow, and number of ICMP destination unreachable packet, are selected for all datasets. We conduct comparison experiments in two scenarios: a victim suffers a specific attack and five mixed attacks.

Table 8 The experimental comparison results on five types of DDoS attacks

Method	ISCX					LLS				
	SYN	DNS	Low	Puls	Spoof.	SYN	DNS	Low	Puls.	Spoof.
RADAR	0.49	0.49	0.51	0.49	0.49	0.47	0.47	0.49	0.47	0.47
Umbrella	0.49	0.49	0.49	0.49	0.48	0.48	0.48	0.48	0.48	0.48
GE	*0.96	0.50	*0.50	0.66	*0.50	*0.98	*0.57	*0.71	*0.55	*0.53
Entropy	*0.92	*0.50	0.50	*0.51	*0.99	*0.99	*0.57	*0.77	*0.70	*0.99
SAFETY	*0.50	0.54	0.55	*0.52	*0.91	*0.71	0.66	0.60	*0.53	*1.0
MLP	0.49	0.49	0.49	0.49	0.49	0.48	0.48	0.48	0.48	0.48
SKM-HFS	0.27	0.40	0.40	0.40	0.27	0.99	0.33	0.30	0.41	0.99
Fuzzy	0.52	0.63	0.49	0.47	0.47	0.55	0.94	0.52	0.43	0.42
Our	**0.98**	**1.0**	**0.99**	**1.0**	**1.0**	**0.99**	**0.99**	**1.0**	**1.0**	**1.0**

The bold entries is used to highlight the results of our method

4.4.1 Detection performance on specific type of attacks

We conduct binary classification to compare the detection performance under the specific type of attacks for all methods using the binary class labeling shown in Table 6. Table 8 shows the F1-scores of our proposed framework and the compared methods for 10 detection pairs[2]. The performances of RADAR, Umbrella, GE, Entropy, and SAFETY are limited, as these methods only consider few features and use fixed thresholds. RADAR is effective on simulation data [50], but it is limited in the real DDoS attack detection. The reasons are that it only applies one feature (SYN/ACK ratio) to detect SYN flooding, and the proposed threshold of 0.5 may not hold in real data. This is because the ACK packet responds to not only the SYN packet, but also other protocol packets such as PSH and FIN packets, which can significantly decrease the SYN/ACK ratio. In fact, from our observation, for ISCX and LLS, the mean SYN/ACK ratio for each attack is much smaller than the threshold. As a consequence, the mean F1-score of RADAR is only 0.48 for 10 detection pairs.

Although Umbrella is effective on simulation data [35], the detection performance is also limited in the real detection. This is because the packet loss rates of the five attacks are lower than the proposed 5%. For instance, the highest packet loss rate is 3% under the Spoofing attack, which is still less than the threshold. Therefore, the average detection rate of Umbrella over all datasets is only 0.48.

The three entropy-based methods have unstable performance over all datasets. As shown in Table 8, Entropy is only effective on SYN flooding and Spoofing attacks but vulnerable for DNS, Low rate, and Pulsing attacks. Similarly, GE can only detect SYN flooding attacks and SAFETY is only effective on Spoofing attacks. This is because each of these methods only considers one entropy feature and thus can only detect certain types of attacks. Besides, these methods are only effective on specific attack samples. For example, in [47], only one attack sample is used to evaluate the performance of GE against the Low Rate DDoS attack. However, in this paper we test the effectiveness of each method using

[2] The scores with * indicates that the class labels are reversed because the attack entropy is greater than the normal entropy.

Table 9 The comparison results when a victim suffers five types of attacks

	RADAR	Umbrella	GE	Entropy	SAFETY	MLP	SKM	Fuzzy	Our
ISCX	0.49	0.49	*0.98	*0.99	*0.50	0.49	0.27	0.56	**1.0**
LLS	0.47	0.48	*0.98	*0.99	*0.74	0.48	0.45	0.62	**1.0**

The bold entries is used to highlight the results of our method

large samples, which will lead to significant performance degradation for the entropy-based methods.

MLP performs poorly for two reasons. First, this method has not evaluated the correlation between features and adopted two linearly highly correlated features, which can cause overfitting. Second, the selected features such as entropy of IP source flow is not effective on detecting various DDoS attacks, as analyzed in Section 2.2.1. Similarly, SKM-HFS and Fuzzy also used ineffective features including entropy of IP source flow, ratio of ICMP packet, and mean inbound to outbound traffic ratio, which leads to their limited performance. Even the number of features utilized for these two methods is more than ours, the detection performance is still limited in this task.

However, our framework can accurately detect all types of DDoS attacks. As shown in Table 8, our proposed framework achieved the highest F1-score for each attack and outperforms the existing methods by a large margin. The reasons are twofold. First, compared to the existing methods, the proposed features in our method are the most effective features. The AUC scores of the proposed five features rank at the top under these detection tasks as shown in Table 3, while the features used by the comparison methods are lower ranked features, i.e., the entropy of IP source flow used in SKM-HFS, Entropy, and MLP methods ranks at 12, and ratio of ICMP packet used in Fuzzy method ranks at 8. Second, only a few features are used in the traditional methods in RADAR, Umbrella, GE, Entropy, and SAFETY, which further limits their detection performance. To be more specific, only one feature is applied in each traditional method while five features are used in our method. Therefore, the traditional methods are only effective on certain attacks while our method is effective on all types of attacks. All of the F1-scores of our proposed framework are no less than 0.98, which demonstrates the effectiveness of our proposed features.

4.4.2 Detection performance on mixed attacks

The effectiveness of all methods under mixed attacks is evaluated by the detection performance, where the attack traffic is generated by merging all attack datasets and the normal traffic, the normal traffic is labeled as 0, and the attack traffic is labeled as 1. Table 9 shows the F1-scores of all methods when the five attacks target the ISCX and LLS. The performance of RADAR and Umbrella under mixed attacks is much worse than their performance for the specific type attacks detection. On the contrast, the F1-scores of the entropy-based methods slightly increase. The reason is that the number of flows in the mixed attacks increases that makes the entropy distances between the attack traffic and normal traffic increase. Since the entropy-based methods utilize the entropy distances to detect attacks, a larger entropy distance will lead to better performance. Meanwhile, the detection results of MLP, SKM-HFS and Fuzzy are still unsatisfactory in this task. For our proposed framework, as shown in Table 9, it performs much better than the eight comparison methods.

Table 10 The average CPU seconds in the detection

	ISCX		LLS	
	Avg.	Total	Avg.	Total
RADAR	1.0e-06	0.014	1.2e-06	0.0038
Umbrella	7.6e-07	0.011	8.4e-07	0.0027
GE	2.4e-05	0.3532	1.6e-05	0.0517
Entropy	1.3e-05	0.1913	3.5e-05	0.1131
SAFETY	1.8e-05	0.2649	1.4e-05	0.0452
MLP	7.7e-07	0.0113	9.3e-07	0.0030
SKM-HFS	4.4e-07	0.0065	9.2e-07	0.0029
Fuzzy	2.7e-07	0.0038	3.3e-07	0.0009
Our	5.2e-07	0.0074	1.1e-06	0.0035

To sum up, our proposed framework that utilizes the proposed five features outperforms the compared methods against various DDoS attacks and mixed DDoS attacks. The detection accuracy improvements over the existing methods are between 21% and 53%.

4.4.3 Running time

In this subsection, we numerically analyze the latency of the methods in RADAR, Umbrella, GE, Entropy, SAFETY, MLP, SKM-HFS, Fuzzy, and our proposed method based on experiments. Although the ML based methods, which are MLP, SKM-HFS, and Fuzzy, and our proposed method, require the training phase, the training can be completed offline. Therefore, to achieve a fair comparison, we only calculate the time consumption in the testing phase for the real detection for all methods. Table 10 shows the running time of each method on ISCX and LLS. The window time is 2 seconds [38] and the values are averaged over all five DDoS attacks. From Table 10, we observe that even though the proposed method is not the most efficient one, the testing time is very short, with the average testing time of 5.2e-07 second for each sample. Therefore, our framework can be used for real-time detection.

5 Conclusion

In this paper, we propose five new features including entropy of packet size, entropy rate of packet size, entropy rate of IP source flow, entropy rate of flow, and number of ICMP destination unreachable packet to detect SYN flooding, DNS amplification, Low Rate, Pulsing, Spoofing, and mixed DDoS attacks. Based on the characteristics of these attacks, we comprehensively analyze the effectiveness on the proposed features and their improvements over the existing features. The experimental results show that our features outperform the other common features, as they ranked at the top in the performance list. When evaluating on the real-world attack traffic, by using the proposed features, our framework has satisfactory detection performance regardless of the types of DDoS attacks. The average F_1 scores of the proposed features with Random Forest classifier are 0.99 on specific types of attacks detection and mixed type attacks detection. The proposed framework can improve up to 53% on detection accuracy compared to the existing methods. In the future, we will

develop more effective DDoS attack detection methods based on the optimal features and extract more representative patterns for real-world practice.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions The authors have no relevant financial or non-financial interests to disclose. The authors have no financial or proprietary interests in any material discussed in this article.

Declarations

Conflicts of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

1. RFC 1812 (1995). http://www.networksorcery.com/enp/rfc/rfc1812.txt. Accessed 15 Aug 2021
2. MIT lincoln laboratory data sets (2000). https://www.ll.mit.edu/r-d/datasets/2000-darpa-intrusion-detection-scenario-specific-datasets. Accessed 15 Aug 2021
3. CAIDA DDoS attack 2007 (2007). https://www.caida.org/data/passive/ddos-20070804_dataset.xml. Accessed 15 Aug 2021
4. Information marketplace marketplace for policy and analysis of cyber-risk & trust (2009). http://www.impactcybertrust.org. Accessed 15 Aug 2021
5. CAIDA UCSD network telescope traffic samples (2012). https://www.caida.org/home. Accessed 15 Aug 2021
6. Information security centre of excellence (2017). https://www.unb.ca/cic/datasets/ids-2017.html. Accessed 15 Aug 2021
7. The 13th worldwide infrastructure security report (2019). http://itnewsafrica.com/pressoffices/arbor/index.htm. Accessed 15 Aug 2021
8. NETSCOUT threat intelligence report 2018 (2019). http://itnewsafrica.com/pressoffices/arbor/index.htm. Accessed 15 Aug 2021
9. Scikit-learn (2021). https://scikit-learn.org/stable/. Accessed 15 Aug 2021
10. Aamir, M., Zaidi, S.M.A.: DDoS attack detection with feature engineering and machine learning: the framework and performance evaluation. Int. J. Inf. Security **18**(6), 761–785 (2019)
11. Ahmed, M.E., Ullah, S., Kim, H.: Statistical application fingerprinting for DDoS attack mitigation. IEEE Trans. Inf. Forensics Security **14**(6), 1471–1484 (2019)
12. Ali, S., Li, Y.: Learning multilevel auto-encoders for DDoS attack detection in smart grid network. IEEE Access **7**, 108647–108659 (2019)
13. Alsirhani, A., Sampalli, S., Bodorik, P.: DDoS detection system: Using a set of classification algorithms controlled by fuzzy logic system in apache spark. IEEE Trans. Netw. Service Manag. **16**(3), 936–949 (2019)
14. Aurelio, Y.S., de Almeida, G.M., de Castro, C.L., Braga, A.P.: Learning from imbalanced data sets with weighted cross-entropy function. Neural Process. Lett. **50**(2), 1937–1949 (2019)
15. Bradley, A.: The Use of the Area Under the ROC Curve in the Evaluation of Machine Learning Algorithms. Pattern Recognit. **30**(7), 1145–1159 (1997)
16. Breiman, L.: Random forests. Mach. Learn. **45**(1), 5–32 (2001)

17. Chen, J., Zhong, M., Li, J., Wang, D., Qian, T., Tu, H.: Effective deep attributed network representation learning with topology adapted smoothing. IEEE Trans. Cybern., 1–12 (2021)
18. Chen, X., Wasikowski, M.: FAST: a roc-based feature selection metric for small samples and imbalanced data classification problems. In: SIGKDD, pp. 124–132. ACM (2008)
19. Cover, T., Hart, P.: Nearest neighbor pattern classification. IEEE Trans. Inf. Theory **13**(1), 21–27 (1967)
20. Cover, T.M., Thomas, J.A.: Elements of information theory. John Wiley & Sons (2012)
21. Duan, Z., Yuan, X., Chandrashekar, J.: Controlling IP spoofing through interdomain packet filters. IEEE Trans. Dependable Secure Comput. **5**(1), 22–36 (2008)
22. Georgios, K., Tassos, M., Dimitris, G., Stefanos, G.: Detecting DNS amplification attacks. In: CRITIS, pp. 185–196. Springer (2008)
23. Gu, Y., Li, K., Guo, Z., Wang, Y.: Semi-supervised k-means DDoS detection method using hybrid feature selection algorithm. IEEE Access **7**, 64351–64365 (2019)
24. Jia, Y., Zhong, F., Alrawais, A., Gong, B., Cheng, X.: Flowguard: An intelligent edge defense mechanism against IoT DDoS attacks. IEEE Internet Things J. **7**(10), 9552–9562 (2020)
25. Kalkan, K., Altay, L., Gür, G., Alagöz, F.: JESS: Joint entropy-based DDoS defense scheme in SDN. IEEE J. Sel. Areas Commun. **36**(10), 2358–2372 (2018)
26. Kalkan, K., Gür, G., Alagöz, F.: Filtering-based defense mechanisms against DDoS attacks: A survey. IEEE Syst. J. **11**(4), 2761–2773 (2017)
27. Kleinbaum, D.G., Dietz, K., Gail, M., Klein, M., Klein, M.: Logistic regression (2002)
28. Kumar, P., Tripathi, M., Nehra, A., Conti, M., Lal, C.: SAFETY: Early detection and mitigation of TCP SYN flood utilizing entropy in SDN. IEEE Trans. Netw. Service Manag. **15**(4), 1545–1559 (2018)
29. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**(7553), 436–444 (2015)
30. Li, C., Dong, Z., Chen, G., Zhou, B., Zhang, J., Yu, X.: Data-driven planning of electric vehicle charging infrastructure: A case study of Sydney. Australia. IEEE Trans. Smart Grid **12**(4), 3289–3304 (2021)
31. Li, C., Dong, Z., Yang, J., Chen, G., Meng, K., Hill, D.: AI-powered energy internet towards carbon neutrality: challenges and opportunities. TechRxiv (2021)
32. Li, Z., Jin, H., Zou, D., Yuan, B.: Exploring new opportunities to defeat low-rate DDoS attack in container-based cloud environment. IEEE Trans. Parallel Distrib. Syst. **31**(3), 695–706 (2020)
33. Li, Z., Wang, X., Li, J., Zhang, Q.: Deep attributed network representation learning of complex coupling and interaction. Knowledge Based Syst. **212**, 106618 (2021)
34. Liu, F., Wu, X., Li, W., Liu, X.: The packet size distribution patterns of the typical internet applications. In: IC-NIDC, pp. 325–332. IEEE (2012)
35. Liu, Z., Cao, Y., Zhu, M., Ge, W.: Umbrella: Enabling ISPs to offer readily deployable and privacy-preserving DDoS prevention services. IEEE Trans. Inf. Forensics Security **14**(4), 1098–1108 (2019)
36. Menardi, G., Torelli, N.: Training and assessing classification rules with imbalanced data. Data Min. Knowl. Discovery **28**(1), 92–122 (2014)
37. de Miranda Rios, V., Inácio, P.R., Magoni, D., Freire, M.M.: Detection of reduction-of-quality DDoS attacks using fuzzy logic and machine learning algorithms. Comput. Netw. **186**, 107792 (2021)
38. Mishra, P., Varadharajan, V., Tupakula, U., Pilli, E.S.: A detailed investigation and analysis of using machine learning techniques for intrusion detection. IEEE Commun. Surveys Tuts **21**(1), 686–728 (2019)
39. Pacheco, F., Exposito, E., Gineste, M., Baudoin, C., Aguilar, J.: Towards the deployment of machine learning solutions in network traffic classification: A systematic survey. IEEE Commun. Surveys Tuts. **21**(2), 1988–2014 (2019)
40. Rasti, R., Murthy, M., Weaver, N., Paxson, V.: Temporal lensing and its application in pulsing denial-of-service attacks. In: SP, pp. 187–198. IEEE (2015)
41. Safavian, S.R., Landgrebe, D.: A survey of decision tree classifier methodology. IEEE Trans. Syst., Man, Cybern. Syst. **21**(3), 660–674 (1991)
42. Su, M.: Real-time anomaly detection systems for denial-of-service attacks by weighted k-nearest-neighbor classifiers. Expert Syst. Appl. **38**(4), 3492–3498 (2011)
43. Suresh, M., Anitha, R.: Evaluating machine learning algorithms for detecting DDoS attacks. In: CNSA, pp. 441–452 (2011)
44. Suykens, J.A., Vandewalle, J.: Least squares support vector machine classifiers. Neural Process. Lett. **9**(3), 293–300 (1999)
45. Wang, A., Chang, W., Chen, S., Mohaisen, A.: Delving into internet DDoS attacks by botnets: Characterization and analysis. IEEE/ACM Trans. Netw. **26**(6), 2843–2855 (2018)
46. Wang, C., Miu, T.T.N., Luo, X., Wang, J.: Skyshield: A sketch-based defense system against application layer DDoS attacks. IEEE Trans. Inf. Forensics Security **13**(3), 559–573 (2018)
47. Xiang, Y., Li, K., Zhou, W.: Low-rate DDoS attacks detection and traceback by using new information metrics. IEEE Trans. Inf. Forensics Security **6**(2), 426–437 (2011)
48. Yang, Y., Guan, Z., Li, J., Zhao, W., Cui, J., Wang, Q.: Interpretable and efficient heterogeneous graph convolutional network. IEEE Trans. Knowl. Data Eng., 1–1 (2021)

49. Yu, S., Zhou, W., Doss, R., Jia, W.: Traceback of DDoS attacks using entropy variations. IEEE Trans. Parallel Distrib. Syst. **22**(3), 412–425 (2011)
50. Zheng, J., Li, Q., Gu, G., Cao, J., Yau, D.K.Y., Wu, J.: Realtime DDoS defense using COTS SDN switches via adaptive correlation analysis. IEEE Trans. Inf. Forensics Security **13**(7), 1838–1853 (2018)
51. Zhou, L., Liao, M., Yuan, C., Zhang, H.: Low-rate DDoS attack detection using expectation of packet size. Security Commun. Netw. **2017** (2017)
52. Zhou, L., Sood, K., Xiang, Y.: ERM: An accurate approach to detect DDoS attacks using entropy rate measurement. IEEE Commun. Lett **23**(10), 1700–1703 (2019)
53. Zhu, L., Tang, X., Shen, M., Du, X., Guizani, M.: Privacy-preserving DDoS attack detection using cross-domain traffic in software defined networks. IEEE J. Sel. Areas Commun. **36**(3), 628–643 (2018)

Authors and Affiliations

Lu Zhou[1] · Ye Zhu[1] · Yong Xiang[1] · Tianrui Zong[2]

Ye Zhu
ye.zhu@deakin.edu.au

Yong Xiang
yong.xiang@deakin.edu.au

Tianrui Zong
zongziqin@hotmail.com

[1] School of Information Technology, Deakin University, 3125 Victoria, Australia

[2] CNPIEC KEXIN LTD, 100020 Beijing, China