DDoS Classification using Combined Techniques

Mohd Azahari Mohd Yusof[1], Noor Zuraidin Mohd Safar[2], Zubaile Abdullah[3], Firkhan Ali Hamid Ali[4], Khairul Amin Mohamad Sukri[5], Muhamad Hanif Jofri[6], Juliana Mohamed[7], Abdul Halim Omar[8], Ida Aryanie Bahrudin[9], Mohd Hatta Mohamed Ali @ Md Hani[10]

College of Computing, Informatics and Mathematics,
Universiti Teknologi MARA (UiTM) Cawangan Melaka Kampus Jasin, Malaysia[1]
Faculty of Computer Science & Information Technology (FSKTM),
Universiti Tun Hussein Onn Malaysia (UTHM), Parit Raja, Batu Pahat, Johor, Malaysia[2, 3, 4, 5]
ICT as Enabler (iCAN) Focus Group, Department of Information Technology, Center for Diploma Studies,
Universiti Tun Hussein Onn Malaysia (UTHM), Pagoh Higher Education Hub, 84600 Pagoh, Johor, Malaysia[6, 7, 8, 9, 10]

***Abstract*—Now-a-days, the attacker's favourite is to disrupt a network system. An attacker has the capability to generate various types of DDoS attacks simultaneously, including the Smurf attack, ICMP flood, UDP flood, and TCP SYN flood. This DDoS issue encouraged the design of a classification technique against DDoS attacks that enter a computer network environment. The technique is called Packet Threshold Algorithm (PTA) and is combined with several machine learning to classify incoming packets that have been captured and recorded. Apart from that, the combination of techniques can differentiate between normal packets and DDoS attacks. The performance of all techniques in the research achieved high detection accuracy while mitigating the issue of a high false positive rate. The four techniques focused in this research are PTA-SVM, PTA-NB, PTA-LR and PTA-KNN. Based on the results of detection accuracy and false positive rate for all the techniques involved, it proves the PTA-KNN technique is a more effective technique in the context of detection of incoming packets whether DDoS attacks or normal packets.**

Keywords*—DDoS; machine learning; accuracy; false positive rate

I. INTRODUCTION

The world now desperately needs an Internet to share resources with other users no matter where they are. It provides many facilities for users to perform daily activities including online games, social media and information search related to teaching and learning. Internet is available 24 hours a day to all users. However, the Internet is often threatened by several network attacks from attackers around the world and this includes DDoS attacks as said by study [1].

When a DDoS attack is launched by an attacker, the computer network or system is inaccessible at that time, even for users who have registered in the system. Typically, attackers apply botnets to perform DDoS attacks to get attacks with incredible speed. It can weaken the target server to serve all requests at that time. According to research [2], DDoS attacks can be categorized into three groups. These categories are volume-based attacks, followed by protocol attacks, and application layer attacks. Volume-based attacks are a category that involves attacks aiming to overwhelm network resources by flooding communication channels with a high volume of traffic. Volume-based attacks often utilize botnets, which are networks of compromised computers controlled by the attacker [3]. By leveraging thousands or millions of infected devices, the attacker generates a massive amount of network traffic, leading to system failures in the targeted infrastructure. The category of protocol attacks focuses on attacking network protocol layers. DDoS protocol attacks often exploit vulnerabilities within the communication protocols used in network infrastructure, such as TCP/IP [4]. Attackers may employ techniques like SYN floods, where they send an overwhelming number of SYN requests to the target server, causing an overload of requests and hindering the server's ability to serve legitimate users. Meanwhile, application layer attacks refer to targeting specific applications or services running on top of the network infrastructure [5]. Application layer DDoS attacks focus on exploiting vulnerabilities within the application's logic or resources it relies on. Attackers can generate various types of DDoS attacks from anywhere. An example of such an attack is the HTTP flood, where attackers overwhelm a web server by sending an abnormally high volume of HTTP requests. This flood of requests leads to a strain on server resources, causing a degradation in performance or even a complete service failure.

There are several types of DDoS attacks that can be generated by attackers from anywhere. These attacks encompass ICMP flood, UDP flood, Ping of Death, Slowloris, Zero-day attack, Smurf, and TCP SYN flood [6]. In order to protect against DDoS attacks, a robust and effective detection strategy is crucial.

The research presents several significant contributions in the following manner:

- In this research, a DDoS attack classification algorithm called the Packet Threshold Algorithm (PTA) was developed to accurately distinguish incoming packets as either normal or malicious. It specifically targets TCP SYN flood, Smurf, UDP flood, Ping of Death, or normal packets. The PTA utilizes a packet threshold mechanism to differentiate and classify incoming traffic.
- To enhance detection capabilities and address the issue of false positives, the PTA was combined with various machine learning techniques, including Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Naïve Bayes (NB), and Logistic Regression (LR). The

objective of this combination was to mitigate the problem of incorrectly classifying DDoS packets as normal packets or vice versa, as experienced in previous techniques.

- In addition to integrating machine learning algorithms with the PTA to enhance the overall performance and accuracy of the detection system, this research also conducted a comprehensive evaluation of the effectiveness of each technique. The results were presented to identify the most efficient approach for detecting malicious packets within a network environment. Furthermore, this study explored potential enhancements and optimizations to further advance the state-of-the-art in DDoS attack detection.

Having a reliable and precise detection strategy is of utmost importance in safeguarding against DDoS attacks. The combined approach of the PTA and machine learning algorithms significantly enhances the system's capability to accurately differentiate and classify incoming packets. By reducing false positives, this strategy provides a more effective defense against DDoS attacks, ensuring the integrity and availability of network resources [7].

The DDoS detection problem is enhanced using machine learning models such as SVM, KNN, Naïve Bayes, and Logistic Regression, which are well-suited for handling classification jobs. Naïve Bayes is strong at probabilistic classification, SVM is good at separating data points, KNN is good at pattern recognition, and Logistic Regression is good for binary classification. By adjusting to a variety of packet behaviors, these models help distinguish between malicious and legitimate packets with accuracy.

There are, nevertheless, certain restrictions. Large datasets may be a problem for SVM and KNN, affecting computing efficiency. The independence between features assumed by Naïve Bayes may not hold true for complex packet dynamics. Non-linear correlations between features may be difficult for logistic regression to handle, which could reduce its accuracy for complex packet classifications. When selecting the best model for DDoS detection, these limitations must be considered.

This paper is divided into several sections. Related work is presented in Section II. Next, in Section III, the methodology is presented, and the evaluation of techniques is described in Section IV, followed by results and discussion in Section V. The final section, Section VI, provides a brief summary of this paper.

II. Related Work

Despite the substantial research efforts dedicated to countering DDoS attacks, the challenge of mitigating them endures. Researchers have introduced various techniques in their attempts to combat the actions of DDoS attackers. Table I provides a summary of the methods proposed by these researchers to address such attacks.

Starting with the first study conducted by study [8], this research addresses the pressing issue of distributed denial-of-service (DDoS) attacks within the context of 5G networks. It emphasizes the predominant focus of previous studies on radio access networks (RAN) and voice service networks, often overlooking the vulnerabilities inherent in core networks (CN). These core network components, including the Access and Mobility Management Function (AMF), Session Management Function (SMF), and User Plane Function (UPF), are pivotal in providing expansive 5G coverage but are susceptible to DDoS attacks. The study introduces a methodology and a threat detection system tailored to counter signalling DDoS attacks specifically targeting 5G standalone CNs. By leveraging fundamental machine learning classifiers and preprocessing techniques such as entropy-based analysis (EBA) and statistics-based analysis (SBA), the research demonstrates the effectiveness of proactive defense strategies against these attacks. Notably, the results underscore the RF classifier as the top performer, achieving an impressive average accuracy of 98.7%.

The second study, led by [9], underscores the critical role of the internet as a fundamental communication tool in contemporary society. In tandem with the internet's indispensability, the frequency and severity of cyber-attacks have escalated, with DDoS attacks ranking among the top five most impactful and costly cyber threats. DDoS attacks disrupt legitimate users' access to network resources, necessitating the development of swift and accurate detection methods to mitigate their considerable damage. The study adopts machine learning classification algorithms, including LR, DT, RF, Ada Boost, Gradient Boost, KNN, and NB to detect DDoS attacks using the CICDDoS2019 dataset, encompassing eleven distinct DDoS attack types characterized by 87 features. The research evaluates classifier performance through various metrics, revealing that AdaBoost and Gradient Boost excel in classification, while LR, KNN, and NB also exhibit strong performance. However, DT and RF classifiers demonstrate less effective classification results.

The third study, conducted by [10], addresses the ongoing challenge of effectively managing DDoS attacks, which pose a significant threat to network security by inundating target networks with malicious traffic from multiple sources. Despite the availability of various conventional methods for detecting DDoS attacks, rapidly identifying these threats using feature selection algorithms remains a formidable task. In this study, a hybrid approach is introduced, incorporating feature selection techniques such as chi-square, Extra Tree, and ANOVA, in conjunction with four machine learning classifiers: FR, DT, KNN, and XGBoost. The primary goal is to enable early detection of DDoS attacks on IoT devices. To validate the proposed methodology, the research employs the CICDDoS2019 dataset, which encompasses a wide range of DDoS attacks, and conducts assessments in a cloud-based environment (Google Colab). The experimental results demonstrate the superior performance of the hybrid methodology, achieving an impressive 82.5% reduction in features and attaining 98.34% accuracy with ANOVA for XGBoost, thereby facilitating the early identification of DDoS attacks on IoT devices.

The fourth study, conducted by study [11], pioneers a comprehensive approach to address pressing security concerns in IoT networks, with a specific focus on the persistent threat

posed by DDoS attacks. Their innovative solution involves the integration of SDN with IoT to reinforce security measures and access control. Despite this integration, DDoS attacks continue to pose a formidable challenge. To tackle this issue head-on, the study introduces an advanced machine learning-based security framework. They meticulously craft a controlled testing environment for simulating DDoS attacks, capturing network logs, preprocessing them into a structured dataset, and employing a trio of robust algorithms, namely NB, DT, and SVM for network packet classification. Remarkably, their framework attains impressive accuracy rates, achieving 97.4% for NB, 96.1% for SVM, and an outstanding 98.1% for DT, unequivocally showcasing its effectiveness in mitigating DDoS threats while optimizing resource utilization and proficiently managing network traffic. This pioneering approach holds substantial promise for elevating the security posture of IoT networks.

TABLE I. PAST STUDY MACHINE LEARNING TECHNIQUE

Title of Paper / Year of Published	Machine Learning DDoS Detection Techniques			
	SVM	*NB*	*LR*	*KNN*
Machine Learning Based Signalling DDoS Detection System for 5G Stand Alone Core Network (2022)	✓	✓	×	×
Detection of DDoS Attacks Using Machine Learning Classification Algorithms (2022)	×	✓	✓	✓
Analysis of Machine Learning Classifiers for Early Detection of DDoS Attacks on IoT Devices (2022)	×	×	×	✓
Towards a Machine Learning-Based Framework for DDoS Attack Detection in Software-Defined IoT (SD-IoT) Networks (2023)	✓	✓	×	×
Detection of DDoS Attack in IoT Traffic using Ensemble Machine Learning Techniques (2023)	✓	✓	✓	×

In the final study conducted by study [12], the focus is on investigating DDoS attacks within the context of the IoT. The research utilizes machine learning classifiers, including both bagging, and boosting techniques, to categorize attack traffic, making use of the CICDDoS2019 dataset designed to simulate DDoS attacks on the UDP and TCP protocols commonly employed in IoT networks. To tackle data imbalance, the study employs an ensemble sampling approach that combines random under-sampling and ADASYN oversampling. Feature selection is carried out using two methods: the Pearson correlation coefficient and the Extra Tree classifier. The results reveal that RF performs the best with minimal training and prediction time, and Extra Trees for feature selection outperforms the Pearson correlation coefficient method in terms of overall time efficiency for most classifiers. However, it's noteworthy that when using the Pearson correlation coefficient for feature selection, RF remains the optimal choice for attack detection.

After conducting an extensive analysis of prior research in the field of DDoS detection using machine learning methods, it becomes evident that there is a pressing need to improve the process of feature selection in the datasets utilized. It is of paramount importance to minimize the occurrence of false positives in order to achieve a heightened level of detection precision. This revelation underscores the critical importance of carefully selecting relevant and efficient features for incorporation into DDoS detection and classification methodologies. By enhancing feature selection techniques, the potential for generating false positive alerts can be significantly reduced, resulting in outcomes that are more reliable and precise.

III. PROPOSED METHODOLOGY

This section introduces the research methodology, which is organized into four phases as illustrated in Fig. 1, and it outlines various research activities.

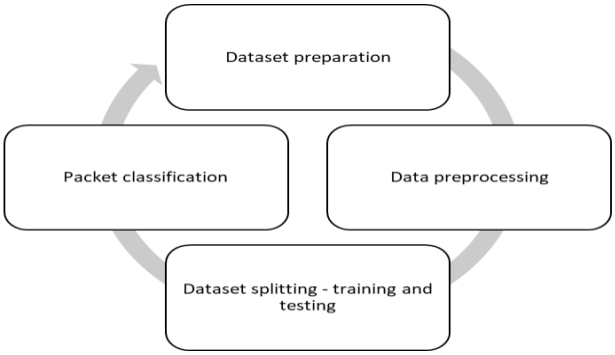

Fig. 1. Methodology of proposed DDoS detection.

A. Dataset Preparation

A dataset containing several types of DDoS attacks and normal packets is provided in the first phase, as shown in Fig. 2. The dataset is relevant to research activities as it records multiple incoming packets, which are the primary focus.

```
In [3]: import pandas as pd
        ddos_df = pd.read_csv('Desktop/ddos_dataset.csv')
        ddos_df.head()
```

Out[3]:

	Src_Addr	Dst_Addr	Pkt_ID	From_Node	To_Node	Pkt_Type	Packet_Type	Pkt_Size	FID
0	3.00	24.30	389693	21	23	tcp	1	55	4
1	15.00	24.15	201196	23	24	tcp	1	55	16
2	24.15	15.00	61905	23	22	udp	2	60	16
3	24.90	9.00	443135	23	21	tcp	1	55	10
4	24.80	8.00	157335	23	21	tcp	1	55	9

Fig. 2. Sample of DDoS dataset.

It includes various features such as source address, destination address, packet type, packet size, and packet class. For instance, the source address refers to the IP address of the sender generating the packet or traffic, while the destination address represents the IP address that receives the packets or traffic.

B. Data Preprocessing

The second research phase is data preprocessing. This phase is crucial in research work as it requires expertise to transform the data into a comprehensible format. Two activities were conducted in this phase: data cleaning and data reduction. Data cleaning is indeed the first activity in the research process, as presented in Fig. 3. This method is called identification of missing values, which is utilized in the research. It indicates that if there is a missing value, the output

will show a value 1, 2, and so on. This means that there are missing values or empty cells in the Src_Addrs, Pkt_ID, and From_Node columns in the dataset used.

```
In [1]: import pandas as pd
        ddos_df = pd.read_csv('Desktop/original_ddos_dataset.csv')
        ddos_df.isnull().sum()

Out[1]: Src_Addr        2
        Dst_Addr        0
        Pkt_ID          1
        From_Node       1
        To_Node         0
        Packet_Type     0
        Pkt_Size        0
```

Fig. 3. Identification of missing values.

The second activity involves data reduction, reducing the number of data samples by identifying and eliminating duplicate rows in the dataset, as presented in Fig. 4.

```
In [1]: import pandas as pd
        ddos_df = pd.read_csv('Desktop/original_ddos_dataset.csv')
        ddos_df.head(10).duplicated()

Out[1]: 0    False
        1    False
        2    False
        3     True
        4    False
        5    False
        6     True
        7    False
        8    False
        9    False
```

Fig. 4. Identification of duplicate data.

In this case, data duplication occurs in rows 3 and 6, which need removal to generate high-quality data and facilitate analysis. Both activities assist in obtaining complete, consistent, and high-quality data within the dataset.

C. Data Splitting

In the third phase of the research, known as data splitting, further investigation proceeds. The dataset, consisting of a total of 240,000 samples, is partitioned into two distinct sets: the training set and the testing set, as outlined in Table II.

The training set plays a crucial role in assessing the effectiveness of machine learning methods by utilizing data samples from the dataset. On the other hand, the testing set is employed to evaluate these methods. The train and test functions were formed to separate these two sets of data. The dataset was divided according to the data distribution outlined in Table II. For example, the data separation for 80: 20 ratios allocates 80% for the training set and the remaining 20% for the testing set.

TABLE II. DATA SPLITTING (TRAINING:TESTING)

No.	Data Splitting Training:Testing	No. of Samples	
		Training	Testing
1	50:50	120,000	120,000
2	60:40	144,000	96,000
3	70:30	168,000	72,000
4	80:20	192,000	48,000

D. Packet Classification

Quality data has been selected, and this research continues with the final phase, which is packet classification. In this phase, a technique called Packet Threshold Algorithm (PTA) has been proposed. This PTA is able to identify incoming packets whether normal packets or DDoS attacks. PTA is combined with several machine learning techniques, SVM, KNN, NB and LR. In the research, the functioning of this PTA was analyzed, as shown in Fig. 5. First, the PTA will check incoming packets based on a predefined packet threshold, which involves packet size and packet type received by the server. If the received packet is TCP or UDP or ICMP and a size of less than 60 bytes per second, PTA will issue the incoming packet category is normal packet. If the server receives TCP packets larger than 60 bytes per second, PTA will issue the incoming packet category is TCP SYN flood. Meanwhile, if the server receives a packet size exceeding 60 bytes per second and carries UDP packets, the PTA will issue the incoming packet category is UDP flood. If the type of packet received by the server is an ICMP packet and the size exceeds 65,535 bytes per second, PTA will issue the incoming packet category is Ping of Death. Meanwhile, if the ICMP packet size is less than 65,535 bytes per second but exceeds 60 bytes per second, PTA will issue the incoming packet category is a Smurf attack. The PTA will act to drop all packets received by the server, for which the packet size exceeds 60 bytes per second and the PTA allows packet sizes less than 60 bytes per second to enter the network environment. Finally, PTA is combined with machine learning by involving several phases or activities including features selection, data splitting, construction and evaluation of the techniques involved.

Here is a summary of how PTA determines the category of incoming packets. Firstly, PTA utilizes a predefined packet threshold to evaluate incoming packets. Secondly, PTA examines the packet type and size to determine their respective categories, as described above. Finally, based on the determined category, PTA performs specific actions on the packet: dropping all packets received by the server that exceed 60 bytes per second and allowing packets with sizes less than 60 bytes per second to enter the network. By employing this approach, PTA can accurately classify incoming packets as normal or belonging to various types of DDoS attacks.

```
Step 1: Start
Step 2: Check incoming packets
    if (packet size < 60) and (packet type = TCP) or (packet type = UDP) or (packet type = ICMP) then
        packet class = Normal
    if (packet size ≥ 60) and (packet type = TCP) then
        packet class = TCP SYN flood
    if (packet size ≥ 60) and (packet type = UDP) then
        packet class = UDP flood
    if (packet size ≥ 65,535) and (packet type = ICMP) then
        packet class = Ping of Death
    if (packet size ≥ 60 and packet size < 65,535) and (packet type = ICMP) then
        packet class = Smurf
Step 3: Features selection (x for input and y for target)
Step 4: Split data into training and testing set
Step 5: Build technique
Step 6: Train and test the technique
Step 7: Evaluation (TP, FP, TN and FN)
Step 8: End
```

Fig. 5. Packet Threshold Algorithm (PTA).

IV. Evaluation of Techniques

During the evaluation phase, detection accuracy and false positive rate are employed as metrics to analyze the precise number of packets detected by PTA and the occurrence of erroneous detections. This encompasses cases where normal packets are wrongly identified as DDoS attacks and instances where DDoS attacks are mistakenly classified as normal packets. The calculation of detection accuracy and false positive rate follows a widely accepted standard formula.

$$Accuracy = \frac{TP+TN}{TP+TN+FN+FP} \times 100 \quad (1)$$

$$FPR = \frac{FP}{FP+TN} \times 100 \quad (2)$$

The formula explanation above can be summarized as follows:

- True Positive (TP): Instances where the model correctly detected DDoS attacks when they occurred.
- False Negative (FN): Instances where the model failed to detect DDoS attacks when they were happening.
- False Positive (FP): Instances where the model incorrectly flagged normal traffic as DDoS attacks.
- True Negative (TN): Instances where the model correctly identified normal traffic as not being DDoS attacks.

These four evaluations can be illustrated using the confusion matrix in Table III. The confusion matrix is a crucial tool in machine learning, providing a detailed breakdown of a model's performance by categorizing predictions into TP, FN, FP, and TN. This breakdown helps assess both accuracy and the model's ability to identify positive and negative cases accurately. It is a fundamental instrument for improving classification model effectiveness in various domains, including DDoS attack detection.

TABLE III. Confusion Matrix

		Predicted DDoS	
		DDoS	Normal
Actual DDoS	DDoS	TP	FN
	Normal	FP	TN

V. Result and Discussion

In this section, the experimental results for the various techniques employed are presented. Starting with an evaluation of the effectiveness of the proposed method for DDoS attack detection, followed by a comparative analysis with previously utilized techniques.

A. Performance Comparison of PTA with Machine Learning Techniques

This section presents the performance results for four combinations of PTA techniques with machine learning based on data splitting between training and testing sets, as shown in Table III. Upon analyzing the performance of each technique using a 50:50 data splitting, it becomes evident that the PTA-KNN technique attains the highest detection accuracy of 99.86%. It is closely followed by the PTA-SVM technique, which also achieves a detection accuracy of 99.86%. The PTA-LR technique achieves a detection accuracy of 99.12%, whereas the PTA-NB technique reaches a detection accuracy of 98.70%.

Shifting focus to the 60:40 data splitting, the PTA-KNN technique once again emerges as the frontrunner, achieving the highest detection accuracy of 99.86%. Remarkably, the PTA-KNN technique surpasses the detection accuracies achieved by the PTA-SVM, PTA-LR, and PTA-NB techniques, which are 99.66%, 99.17%, and 98.72% respectively. For the 70:30 data splitting, the PTA-KNN technique continues to outperform the other techniques with a detection accuracy of 99.84%. The PTA-SVM, PTA-LR, and PTA-NB techniques achieve respective detection accuracies of 99.65%, 99.16%, and 98.69%. Table IV shows the performance comparison of PTA with machine learning techniques. When considering the 80:20 data splitting, the PTA-KNN technique showcases an impressive detection accuracy of 99.83%, surpassing the PTA-SVM technique that achieves a detection accuracy of 99.63%. Furthermore, the PTA-LR technique demonstrates an impressive detection accuracy of 99.17%, whereas the PTA-NB technique achieves a slightly lower accuracy of 98.68%. Through meticulous examination, it can be deduced that the PTA-KNN technique showcases a remarkable efficacy in identifying incoming packets, regardless of their nature as DDoS attacks or normal packets. Observing the statistical outcomes presented in Fig. 6, which depict the effectiveness of the PTA-KNN technique in the research. This effectiveness stems from its utilization of packet type and size as key criteria. This conclusion is further supported by the exceptional detection accuracies achieved across various data splitting ratios: 99.86% for 50:50, 99.86% for 60:40, 99.84% for 70:30, and 99.83% for 80:20.

TABLE IV. Performance Comparison of PTA with Machine Learning Techniques

Technique	Data Splitting (Training:Testing)	Detection Accuracy	False Positive Rate
PTA-NB	50:50	98.70%	1.10%
	60:40	98.72%	1.08%
	70:30	98.69%	1.10%
	80:20	98.68%	1.08%
PTA-KNN	50:50	99.86%	0.01%
	60:40	99.86%	0.01%
	70:30	99.84%	0.01%
	80:20	99.83%	0.02%
PTA-SVM	50:50	99.66%	0.01%
	60:40	99.66%	0.01%
	70:30	99.65%	0.01%
	80:20	99.63%	0.02%
PTA-LR	50:50	99.12%	0.26%
	60:40	99.17%	0.25%
	70:30	99.16%	0.27%
	80:20	99.17%	0.26%

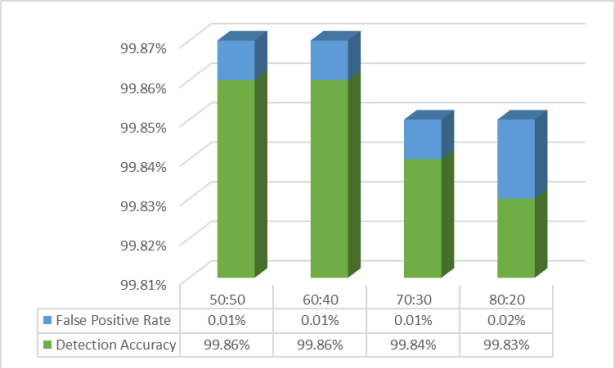

Fig. 6. Statistical outcomes of PTA-KNN technique.

Referring to Table V, it is noteworthy that the detection accuracies presented therein exceed the performance of alternative techniques, thus emphasizing the superiority of the PTA-KNN technique. The detection accuracy percentages for the PTA-KNN technique are determined based on the number of successfully detected incoming packets. For the 50:50 data splitting, 119,827 incoming packets were accurately detected, while 173 packets were misclassified. In the case of the 60:40 data splitting, the PTA-KNN technique successfully identified 95,863 incoming packets as valid, with 137 packets misclassified. Similarly, for the 70:30 data splitting, the technique detected 71,882 incoming packets correctly, but there were 118 misclassified packets. Lastly, for the 80:20 data splitting, the PTA-KNN technique successfully detected 47,920 incoming packets, with 80 packets being misclassified.

TABLE V. DETECTION RESULTS FOR DIFFERENT DATA SPLITTING RATIOS AND PACKET TYPES USING COMBINATION TECHNIQUES

Technique	Data Splitting (Training:Testing)	No. of Incoming Packet Detected				
		Normal	Ping of Death	Smurf	TCP SYN Flood	UDP Flood
PTA-NB	50:50	106,499	354	733	208	10,649
	60:40	85,233	287	591	166	8,495
	70:30	63,943	206	455	128	6,324
	80:20	42,643	133	299	83	4,209
PTA-KNN	50:50	107,165	354	952	229	11,127
	60:40	85,766	287	765	182	8,863
	70:30	64,361	206	581	141	6,593
	80:20	42,914	133	393	88	4,392
PTA-SVM	50:50	107,168	354	731	216	11,123
	60:40	85,769	287	589	173	8,859
	70:30	64,363	206	454	135	6,590
	80:20	42,916	133	298	85	4,389
PTA-LR	50:50	107,168	354	529	212	10,680
	60:40	85,769	287	418	170	8,557
	70:30	64,363	206	321	131	6,376
	80:20	42,916	133	220	84	4,249

B. Performance Comparison between Proposed DDoS Detection Technique and Previous Techniques

This section presents a performance comparison between the proposed DDoS detection technique and existing methods, as displayed in Table VI. Within the provided table, which demonstrates performance comparisons in terms of detection accuracy for various techniques across different years of publication, it becomes evident that the highest and lowest accuracies vary significantly among the diverse techniques and algorithms employed. Notably, the proposed technique stands out with the highest overall accuracy of 99.86%, achieved using the KNN algorithm. However, it is essential to emphasize that the lowest accuracy values are somewhat dispersed. For instance, in the case of Park et al., the lowest accuracy values for LR and KNN are denoted as NA, indicating a lack of available data. In contrast, for other techniques, such as Gaur and Kumar, the lowest accuracy is attributed solely to the KNN algorithm, which attains an accuracy of 91.39%.

TABLE VI. PERFORMANCE COMPARISON BETWEEN PROPOSED DDOS DETECTION TECHNIQUE AND PREVIOUS TECHNIQUES

Technique/Year of Published	Performance Comparison in Terms of Detection Accuracy			
	SVM	NB	LR	KNN
Park et al. (2022)	98.76%	87.61%	NA	NA
Dasari and Devarakonda (2022)	NA	99.58%	99.58%	99.55%
Gaur and Kumar (2022)	NA	NA	NA	91.39%
Bhayo et al. (2023)	96.10%	97.40%	NA	NA
Pandey and Mishra (2023)	96.24%	98.23%	89.76%	NA
Proposed Technique (2023)	99.66%	98.72%	99.17%	99.86%

Overall, the proposed technique appears to exhibit the highest accuracy across most algorithms, rendering it a promising approach for detection. Nevertheless, it is crucial to consider other factors, such as computational complexity and practical applicability, when selecting a technique for a specific problem.

VI. CONCLUSION

The team has extensively researched the capabilities of the PTA technique in detecting both DDoS attacks and normal packets. This involves utilizing a predefined packet threshold that considers factors such as packet size and the specific packet types that attackers may generate. By integrating the PTA technique with diverse machine learning approaches, findings reveal that the PTA-KNN technique surpasses PTA-NB, PTA-SVM, and PTA-LR techniques in terms of detection accuracy and false positive rate percentage.

In the research, potential areas for future enhancement have also been identified based on findings. One possible direction for improvement involves exploring adaptive thresholding techniques that dynamically adjust the packet threshold based on network conditions and attack patterns. Additionally, investigating the integration of anomaly detection algorithms and deep learning models could enhance the PTA technique's ability to detect emerging and sophisticated DDoS attacks.

These avenues for future research aim to further enhance the effectiveness and resilience of the PTA technique in combatting evolving cyber threats.

ACKNOWLEDGMENT

This work was supported by the Universiti Tun Hussein Onn Malaysia (UTHM) through Tier1 (vot Q508).

REFERENCES

[1] X. Wang, Y. Li, H. J. Khasraghi, and C. Trumbach, "The Mediating Role of Security Anxiety in Internet Threat Avoidance Behavior," Computer Security, vol. 134, pp. 1–14, Nov. 2023, doi: 10.1016/j.cose.2023.103429.

[2] R. M. A. Haseeb-ur-rehman et al., "High-Speed Network DDoS Attack Detection: A Survey," Sensors, vol. 23, no. 15. Multidisciplinary Digital Publishing Institute (MDPI), pp. 1–25, Aug. 01, 2023. doi: 10.3390/s23156850.

[3] Y. Li and Q. Liu, "A Comprehensive Review Study of Cyber-Attacks and Cyber Security: Emerging Trends and Recent Developments," Energy Reports, vol. 7, pp. 8176–8186, Nov. 2021, doi: 10.1016/j.egyr.2021.08.126.

[4] N. Abosata, S. Al-Rubaye, G. Inalhan, and C. Emmanouilidis, "Internet of Things for System Integrity: A Comprehensive Survey on Security, Attacks and Countermeasures for Industrial Applications," Sensors, vol. 21, no. 11. MDPI AG, pp. 1–29, Jun. 01, 2021. doi: 10.3390/s21113654.

[5] N. Tripathi and N. Hubballi, "Application Layer Denial-of-Service Attacks and Defense Mechanisms: A Survey," ACM Computing Surveys, vol. 54, no. 4. Association for Computing Machinery, pp. 1–33, Jul. 01, 2021. doi: 10.1145/3448291.

[6] L. Zhou, Y. Zhu, Y. Xiang, and T. Zong, "A Novel Feature-Based Framework Enabling Multi-Type DDoS Attacks Detection," World Wide Web, vol. 26, no. 1, pp. 163–185, Jan. 2023, doi: 10.1007/s11280-022-01040-3.

[7] F. M. Salem, H. Youssef, I. Ali, and A. Haggag, "A Variable-Trust Threshold-Based Approach for DDoS Attack Mitigation in Software Defined Networks," PLoS One, vol. 17, no. 8, pp. 1–19, Aug. 2022, doi: 10.1371/journal.pone.0273681.

[8] S. Park, B. Cho, D. Kim, and I. You, "Machine Learning Based Signaling DDoS Detection System for 5G Stand Alone Core Network," Applied Sciences (Switzerland), vol. 12, no. 23, pp. 1–27, Dec. 2022, doi: 10.3390/app122312456.

[9] K. B. Dasari and N. Devarakonda, "Detection of DDoS Attacks Using Machine Learning Classification Algorithms," International Journal of Computer Network and Information Security, vol. 14, no. 6, pp. 89–97, Dec. 2022, doi: 10.5815/ijcnis.2022.06.07.

[10] V. Gaur and R. Kumar, "Analysis of Machine Learning Classifiers for Early Detection of DDoS Attacks on IoT Devices," Arabian Journal for Science and Engineering, vol. 47, no. 2, pp. 1353–1374, Feb. 2022, doi: 10.1007/s13369-021-05947-3.

[11] J. Bhayo, S. A. Shah, S. Hameed, A. Ahmed, J. Nasir, and D. Draheim, "Towards a Machine Learning-Based Framework for DDoS Attack Detection in Software-Defined IoT (SD-IoT) Networks," Engineering Applications of Artificial Intelligence, vol. 123, no. 1, pp. 1–17, Aug. 2023, doi: 10.1016/j.engappai.2023.106432.

[12] N. Pandey and P. K. Mishra, "Detection of DDoS Attack in IoT Traffic using Ensemble Machine Learning Techniques," Networks and Heterogeneous Media, vol. 18, no. 4, pp. 1393–1408, 2023, doi: 10.3934/nhm.2023061.