Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 8060333, 13 pages
https://doi.org/10.1155/2022/8060333

Research Article

Labelled Dataset on Distributed Denial-of-Service (DDoS) Attacks Based on Internet Control Message Protocol Version 6 (ICMPv6)

Selvakumar Manickam,[1] Adnan Hasan Bdair AIghuraibawi,[1,2] Rosni Abdullah,[3] Zaid Abdi Alkareem Alyasseri,[4,5] Karrar Hameed Abdulkareem,[6] Mazin Abed Mohammed,[7] and Ayman Alani[8]

[1]*National Advanced IPv6 Centre (NAv6), Universiti Sains Malaysia, Penang 11800, Malaysia*
[2]*Baghdad College of Economic Sciences University, Baghdad, Iraq*
[3]*School of Computer Sciences, Universiti Sains Malaysia, Penang 11800, Malaysia*
[4]*Information Technology Research and Development Center (ITRDC), University of Kufa, Najaf, Iraq*
[5]*ECE Dept. Faculty of Engineering, University of Kufa, Najaf, Iraq*
[6]*College of Agriculture, Al-Muthanna University, Samawah 66001, Iraq*
[7]*College of Computer Science and Information Technology, University of Anbar, 11, Ramadi, 55431 Anbar, Iraq*
[8]*Faculty Of Computing and Informatics, Universiti Malaysia Sabah, Jalan UMS, 88400 Kota Kinabalu Sabah, Malaysia*

Correspondence should be addressed to Adnan Hasan Bdair AIghuraibawi; adnan_hcan@yahoo.com

Received 22 December 2021; Accepted 21 February 2022; Published 18 April 2022

Academic Editor: Nawab Muhammad Faseeh Qureshi

The most dangerous attack against IPv6 networks today is a distributed denial-of-service (DDoS) attack using Internet Control Message Protocol version 6 (ICMPv6) messages. Many ICMPv6-DDoS attack detection mechanisms rely on self-created datasets because very few suitable ICMPv6-DDoS attack datasets are publicly available due to privacy and security concerns. When implemented in a real network, however, a detection system that relies on a dataset with incorrect packet or flow representation and contains unqualified features generates a large number of false alerts. The goal of this work is to create a comprehensive ICMPv6-DDoS attack dataset that can be used for tuning, benchmarking, and evaluating any detection systems designed to detect ICMPv6-DDoS attacks. The proposed datasets met the criteria for a good dataset, ensuring their usefulness to other researchers. A GNS3 network simulation tool is used to simulate an IPv6 network and generate ICMPv6 traffic for the dataset. The generated traffic contains both normal and abnormal ICMPv6 traffic, with the abnormal traffic containing ten different ICMPv6-DDoS attacks based on RA and NS message flooding. Five classifiers were chosen, varying in terms of type, classification performance, and the number of features used, and the results were as follows: decision tree 80%, support vector machine 78%, naïve Bayes 80%, *k*-nearest neighbours 81%, and neural networks 81%. The proposed dataset has been shown to accurately represent attack traffic in tests, with a high detection accuracy and a low false-positive rate.

1. Introduction

IPv6 introduced a slew of new features and benefits, including an autoconfiguration mechanism that provides automatic IPv6 address configuration for IPv6-enabled hosts. The Internet Control Message Protocol for IPv6 (ICMPv6) specification differs from that of its IPv4 counterparts in some ways. For example, the address resolution protocol had been replaced by the Neighbour Discovery Protocol (NDP), and some administrative controls had been modified. When an IPv6-enabled device is plugged into an IPv6 network, NDP generates a unique IPv6 address without the need for manual configuration [1, 2]. ICMPv6 is a newer implementation of ICMP that is intended for IPv6, and it uses the same strategy as ICMPv4 [3]. Nevertheless, it is an upgraded version that is critical in IPv6 networks [1], where it is often used for control reasons such as testing, diagnostics, fault isolation, and reporting erroneous operations [4, 5]. The modern design of

the ICMPv6 has security flaws that expose it to a variety of threats. Durdağı and Buldu (2010) discovered a variety of IPv6 threats, including DDoS and reconnaissance attacks. According to a study published in 2012 [6], DoS and DDoS attacks constituted a significant portion of attacks against IPv6 networks, as shown in Figure 1.

One method of overcoming DoS and DDoS attacks is through prevention, which attempts to avoid or minimize losses from such attacks. Attempts at source address spoofing or address stealing could be prevented through network policy enforcement, and because the ICMPv6 protocol is critical to the operation of the IPv6 network, it cannot be disabled [7, 8]. Despite the fact that numerous detection systems have been proposed to address this issue, it has yet to be completely resolved. Regrettably, detection systems have a low rate of detection accuracy. The main reason for this is that most detection systems used nonrepresentative datasets and did not use strong attack-related features [9].

ICMPv6 is a critical component and an important feature of IPv6, and every IPv6 node must be completely implemented according to RFC 4443 [10]. Figure 2 displays the ICMPv6 header format. Table 1 shows some instances of probable ICMPv6 characteristics. ICMPv6 detects faults discovered when processing packets [11, 12] and performs additional internet-layer functions such as diagnostics. ICMPv6 has two sorts of messages: error notification and information notification, which are used in router advertisement and neighbour solicitation [2] processes that define relationships between neighbouring nodes by using type and code fields to distinguish services, both of which are susceptible to denial-of-service (DoS), Man-in-the-Middle (MITM), and spoofing attacks [13, 14]. Fields with empty values in ICMPv6 packets can be used in the future [15].

The advantage of packet-based representation is that all the details in each record, such as time, header, and payload, can be obtained without preprocessing. As a result, the data is immediately available to the detection system, as opposed to the flow-based representation, which requires preprocessing before use. The choice of an appropriate dataset demands an understanding of the underlying mechanism that generates the data. In terms of anomaly detection, packet-based representation has nine benefits: (a) required for rapid implementation of countermeasures to prevent major malware propagation; (b) required to generalize solutions and acquire new knowledge that helps future enhancements and refinement; (c) required to validate the new detection system; (d) required to study attack behaviour; (e) for replication of studies, attempting to develop a particular detection system, must be able to replicate the detection system's application using the same dataset and receive consistent and trustworthy findings, (f) quicker and more cost-effective predictors, as compared to other systems; (g) to tune parameters, most detection systems have particular parameters that affect their performance and detection accuracy; and (h) to choose the feature set, detection accuracy is directly proportional to the strength of the specific features. Select the most capable set of features [12, 16, 17].

The absence of publicly available IPv6 datasets impedes advancement in the field of IPv6 network security. As a result, the majority of existing detection systems for IPv6 networks cannot be based on a single dataset, and some must rely on a self-created dataset [18–21]. However, numerous factors must be considered in order to create a reliable dataset. A new dataset must meet all requirements before it can be used. A dataset must include both normal and abnormal traffic data representing diverse scenarios, as well as all important and relevant features labelled [22, 23].

Additionally, several existing datasets are out of date. Therefore, novel attacks and new technological advantages, such as IPv6 and IPv6 datasets, are required for evaluating and testing detection systems, particularly those based on IPv6 anomaly detection. Another concern levelled at some existing datasets is the absence of the traffic generation component. Hence, the aggregated traffic's correctness cannot be determined. As indicated in [24, 25], "datasets in the intrusion detection domain have been widely criticized for their accuracy and capacity to reflect real-world settings."

When different datasets are used, comparing different detection systems is impossible [21]. Thus, a benchmark dataset is required to measure and compare the performance of existing technologies to those of other technologies. Since the dataset is used to evaluate and compare the performance of detection systems in distinguishing normal and abnormal traffic, it must be labelled to include ICMPv6-DDoS attacks with the most significant and relevant features for evaluation and comparison.

The purpose of this study is to describe the process of traffic collection and labelling. The article offered a set of representative packet-based features for representing packet-based datasets. These datasets are critical for three reasons: (1) It provides a representative collection of features for the labelled packet-based dataset of ICMPv6-based DDoS attacks. (2) It significantly minimizes the amount of private or sensitive data in comparison to packet-based datasets: virtual machines are used. (3) Using a set of classifiers, it was demonstrated that the proposed dataset's format aided in achieving acceptable detection accuracies.

The following sections comprise the remainder of this paper. Section 2 discusses related works that have been published in existing IPv6 datasets. Section 3 discusses data collection and preprocessing for the purpose of creating, capturing, filtering, and labelling the proposed datasets, followed by Section 4 which describes the validation and evaluation processes. Finally, Section 5 is the paper's conclusion.

2. Related Work

A dataset is a collection of tables that are closely related and correspond to a specific experiment. It is required for evaluating, testing, comparing, and deploying detection systems [26]. The majority of the datasets are IPv4 traffic, and there are no public pure IPv6 traffic datasets. Then we create a dataset that only contains IPv6 traffic. The dataset is composed of two components: the application label and traffic data [27]. As a result, the KDDCUP99 dataset used for intrusion detection in IPv6 networks is frequently derived from the dataset's 10% test set, which serves as a dataset training set; the normal

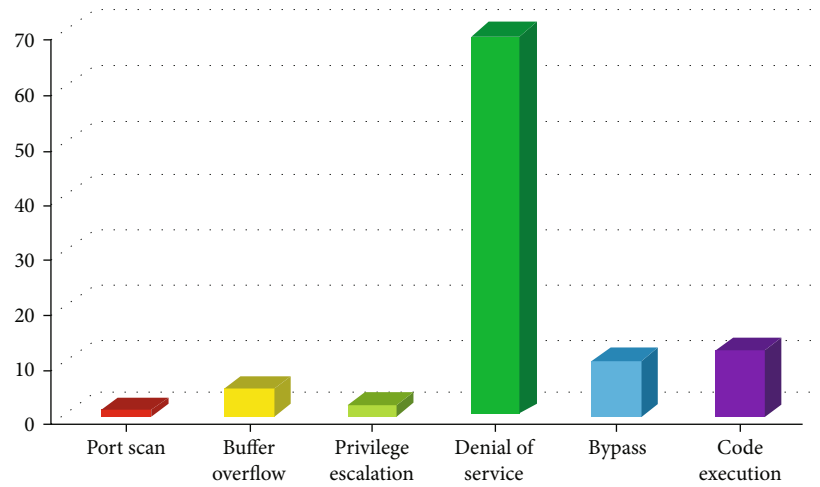

Figure 1: IPv6 vulnerability classes.

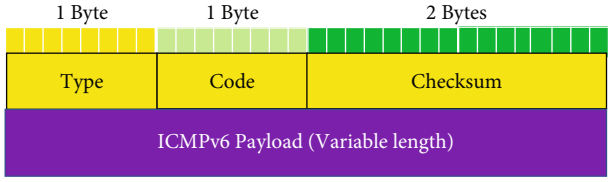

Figure 2: ICMPv6 header format.

Table 1: Identify features in IPv6 header fields.

No.	Features	Class
1	Traffic class	Normal or attack
2	Flow label	Normal or attack
3	Hop limit	Normal or attack
4	Payload length	Normal or attack
5	Source address	Normal or attack
6	Next header	Normal or attack
7	ICMPv6 type	Normal or attack
8	ICMPv6 code	Normal or attack
9	ICMPv6 payload	Normal or attack

type samples and attack type samples associated with attacks are chosen [28]. Furthermore, a dataset is a fundamental and essential component of artificial intelligence, machine learning, data mining, genetic algorithms, and many statistical techniques used to detect apparent relationships between cases in a pooled dataset. IPv4 detection solutions have been tested and benchmarked using a variety of datasets, including the DARPA intrusion detection evaluation dataset [29], NSL-KDD [30], KDD [31], CAIDA [32], and LBNL [33]. The Defense Advanced Research Projects Agency and the Air Force Research Laboratory created the DARPA dataset in 1998 as the first benchmark dataset for testing and evaluating IDS.

However, even the most realistic datasets rarely contain all of the attributes required to generate an accurate IDS solution. For example, in a realistic dataset, private and confidential information such as source and destination IP addresses would not be made available because releasing them could expose the organization to attacks. Other datasets are deficient in some statistical characteristics or are unable to meet the needs of current technologies [25, 34, 35].

Datasets are primarily used to assess a detection solution's ability to distinguish normal traffic from abnormal or attack traffic. Typically, researchers use one or more datasets depending on the needs of the solution. However, most existing datasets can only be used to evaluate and test IPv4 solutions because they do not include IPv6 network threats [36]. One of the issues that IPv6 security researchers face is the dearth of available IPv6 datasets [26]. Researchers are forced to create their own IPv6 dataset for the purpose of evaluating and testing IPv6 detection systems, which is typically inferior to a good standard (benchmark) dataset. As a result, several IPv6 datasets have been created by researchers for their own use solely for the purpose of achieving their stated research objectives.

A group working on MAWI [37] proposed a dataset not for security research but for studying IPv6 traffic behaviour. The MAWI dataset only contains normal daily-use IPv6 traffic, implying that there is no malicious traffic in the dataset. Barrera and Van Oorschot [38], for example, used the MAWI dataset in their research to visualize IPv6 network traffic and analyse the security aspects of the IPv6 protocol. As a consequence, because it lacks malicious traffic, the MAWI dataset is unsuitable for testing detection systems. Additionally, injecting malicious traffic into the existing MAWI dataset via new network traffic data is likely to create a heterogeneous and biased dataset, as the two types of traffic originate from two distinct networks.

The Center for Applied Internet Data Analysis (CAIDA) [32] suggested an IPv6 dataset that contained only specific types of IPv6 attacks. CAIDA is another name for the Ark IPv6 topology dataset. CAIDA was created by utilizing the Paris traceroute technique through the use of a scamper programmer. The dataset is unlabelled and uses a packet-based representation. The CAIDA dataset can be used to evaluate

an IDS that does not require labelled traffic, such as discovering interfaces of new IPv6 routers [39, 40]. However, because detection system testing and evaluation rely on detection accuracy rate and other measures to distinguish different class labels, the CAIDA dataset is insufficient for such purposes due to the unlabelled feature. Furthermore, the source and destination IPv6 addresses were removed from the dataset, which is critical to the operation of several detection systems.

Zulkiflee et al. [41] proposed a five-stage testbed for creating IPv6 datasets by employing a variety of attacks. To launch DoS and probing attacks, three attack tools from The Hacker's Choice (THC) Toolkit, FloodRouterv6, Smurf6, and Alive6, were used. The authors used a packet-based traffic representation with five attributes to model the IPv6 network traffic in the dataset. The five attributes are the source port number, the time interval between packets, the destination IP address, the source IP address, and the protocol. The authors' included attack scenarios met the requirements outlined in [21]. The dataset, however, does not include all possible ICMPv6 DoS attacks. Furthermore, this dataset has not been made publicly available for use by other researchers.

Najjar and Kadhum [31] proposed a five-stage IPv6 dataset creation process. The GNS3 was used to generate simulated network traffic. The network topology included six PCs, two routers, and a server. The generated datasets contain a normal IPv6 traffic behaviour representing SLAAC, DAD, DUD, and address resolution processes. THC Toolkit was used to generate abnormal or attack traffic. The abnormal behaviour data aids in the detection of new attacks and aids researchers in understanding the ICMPv6 protocol and messages. Since the dataset is labelled and includes both normal and attack traffic, it can assist researchers in evaluating detection systems. However, it only contains two DoS attack scenarios and is not publicly available to researchers. The benefit of using a dataset for machine learning is that it can identify and detect attacks without relying on attack signatures.

Saad et al. [34] suggested an IPv6 dataset for evaluating a detection system for DoS attacks utilizing ICMPv6 echo request messages. The dataset contains traffic from both normal and abnormal (attack) IPv6 networks traffic. One disadvantage of this dataset is that it does not contain complex situations (just DoS attacks), as its purpose is to evaluate the detection system's performance in detecting ICMPv6 echo request message flooding attacks. Further, this dataset is not available online for other researchers to use due to the privacy and confidentiality of the real network traffic's content. Saad et al. [34] proposed an intelligent approach for ICMPv6 flooding attack detection by an intelligent back-propagation neural network (v6IIDS) using ICMPv6 flooding echo request message dataset in IPv6 networks.

Omer et al. [17] proposed a five-stage testbed for creating an ICMPv6 DDoS dataset in a more recent flow-based dataset. The dataset source was normal network traffic in a real network in a Universiti Sains Malaysia (USM) laboratory, and it contains various attack scenarios. The requirements of a good dataset were met, making it useful to other researchers. The datasets are labelled, publicly accessible and available, represented, contain a set of 11 flow features for ICMPv6 DDoS attacks, and contain a few packets details. However, it cannot detect attacks that rely on packet's payload. In addition, preprocessing for flow construct is needed. Omer et al. [42] tested and assessed numerous detection methods against their dataset, demonstrating a strong attack traffic representation by attaining robust and high detection accuracies and also low false-positive rates. Table 2 summarizes the existing IPv6 datasets as well as their limitations.

3. The Proposed Datasets

In general, a detection system is used to analyse network traffic for suspicious behaviours in order to help researchers in resolving security issues in IPv6 networks. Due to the inability of current IPv6 datasets to match the criteria of detection systems for ICMPv6 DDoS attacks, as defined in Section 2, a new alternative dataset is required to serve as a reference for testing and evaluating detection systems for ICMPv6 DDoS attacks. Furthermore, because of its qualified and relevant feature set, this dataset can be used to investigate and study ICMPv6 DDoS attacks.

To be worthy of being used as a reference for testing and assessing detection systems by researchers aiming to protect the ICMPv6 protocol or IPv6 networks, this dataset must adhere to all of the conditions outlined in Section 2. This section outlines the process of preparing, developing, and generating traffic for the proposed dataset. During this process, a dataset is generated and collected in an IPv6 network containing regular traffic and tools to launch attacks on the IPv6 networks using ICMPv6 messages, such as NA, RS in RA, and NS messages. This data enables researchers to conduct studies and find solutions to security issues in a safe and nondestructive environment. For example, the advantage of a packet-based network representation is that it is possible to obtain all of the details in each record, such as its time, head, and payload. Furthermore, the data is immediately available for use by the detection system without any preprocessing. A flow-based detection system, on the other hand, cannot detect using head and payload data.

In the first and second stages, a network simulation tool (GNS3) is used to generate both normal and abnormal or attack traffic. Due to security and privacy concerns, a simulator is used to generate traffic rather than to capture real network traffic data. Sensitive organizational information, such as source and destination IP addresses, cannot be included because deploying them could expose the organization to attacks [25, 34, 35].

The next three stages are traffic capturing, traffic filtering, and dataset labelling. The steps for generation are as follows:

(1) Create a virtual network with computers connected via a switch and router

(2) Computer's exchange ICMPv6 RA and NS messages between themselves, and the network traffic is captured as normal data

Table 2: Comparison of existing IPv6 datasets.

No.	Datasets	Description	Disadvantages
1	MAWI [37]	The daily effects of normal IPv6 traffic. Intended for nonsecurity purposes. Publicly available online. Packet-based representation.	Cannot be used for detection systems evaluation. Only comprises normal IPv6 traffic. Traffic representation based on ten features.
2	CAIDA [32]	Includes a few IPv6 attacks type. The traffic is unlabelled. Publicly available online. Packet-based representation.	Cannot be used for detection systems evaluation. Source and destination IP addresses features are removed. Utilizes limited features.
3	Zulkiflee et al. [41]	Only three types of IPv6 attacks were included. THC Toolkit utilized to perform attacks. Packet-based representation.	Does not cover all potential ICMPv6 DDoS attacks. Not available online. Traffic representation based on six features.
4	Najjar and Kadhum [31]	Uses GNS3 tool to a create dataset on a virtual network. Comprises three IPv6 attack types. Normal and abnormal (attacks) traffic are labelled. Packet-based representation.	Does not cover all potential ICMPv6 DDoS attacks. Not available online. Traffic representation based on seven features.
5	Saad et al. [34]	Includes ICMPv6 echo request message DoS attack packets. Created based on a real network. Packet-based representation.	Does not cover all potential ICMPv6 DoS attacks. Not available online. Traffic representation based on eight features
5	Omer et al. [17]	The source of normal traffic is a real-life network. Achieved the requirements of a good dataset. Represented features for ICMPv6 DDoS attacks.	Contains a few package details. Unable to detect attacks that rely on package payload. Preprocessing for flow construct is needed. Only contain a set of 11 traffic flows.

(3) The Hacker's Choice IPv6 (THC-IPv6) Toolkit is used to generate attack traffic, and the network traffic is captured as abnormal or suspicious data

(4) The normal and abnormal data, which includes ICMPv6 (RA and NS) messages, is saved for use in the next step

The proposed dataset was created in four stages approach, as shown in Figure 3.

3.1. Traffic Capturing Stage. Using Wireshark [43], the capabilities of detection systems can be evaluated using a variety of testing techniques, including cross-validation and supplied set testing. The suggested packet-based datasets in this paper can be classified into training and testing datasets. The collected and filtered traffic data is then converted to a comma-separated value (CSV) format. A field called class is added to make labelling possible. Figure 2 depicts the implementation of the stage. The primary goal of this stage is to collect normal and abnormal traffic data, as well as a sufficient number of features, to evaluate the performance of the proposed approach. This step captures two types of datasets, one normal dataset in Stage 1 and one abnormal dataset in Stage 2.

Typically, IDS is used to analyse the network traffic in order to detect anomalies. A prerequisite is the provision of a dataset for the purpose of studying and investigating the behaviour of ICMPv6 DDoS attacks. It is then straightforward to define a subset of features for evaluating and testing the IDS's performance. All of these activities take place at the laboratory of the National Advanced IPv6 Centre (NAv6). The topology of the network is depicted in Figures 3 and 4.

(i) A router serves as a gateway to connect the network's devices

(ii) A switch connects the router and nodes

(iii) A monitoring machine captures network traffic through tools

(iv) The machines targeted for attacks are known as victim nodes

(v) Attack nodes generate abnormal traffic directed at the machines of their victims

(vi) The normal traffic is generated by nodes, which include servers, laptops, tablets, and PC

The dataset should meet the requirements for evaluating proposed IDSs in the detection of DDoS attacks based on ICMPv6 messages. The dataset's normal and abnormal traffic is generated in two stages, as discussed in the following subsections.

This step's primary purpose is to build a dataset with normal and abnormal (attack) traffic, as well as a good amount of attributes for evaluating the proposed approach's performance. This process generates two sorts of datasets: normal traffic and abnormal traffic comprising ICMPv6-based DDoS attack packets.

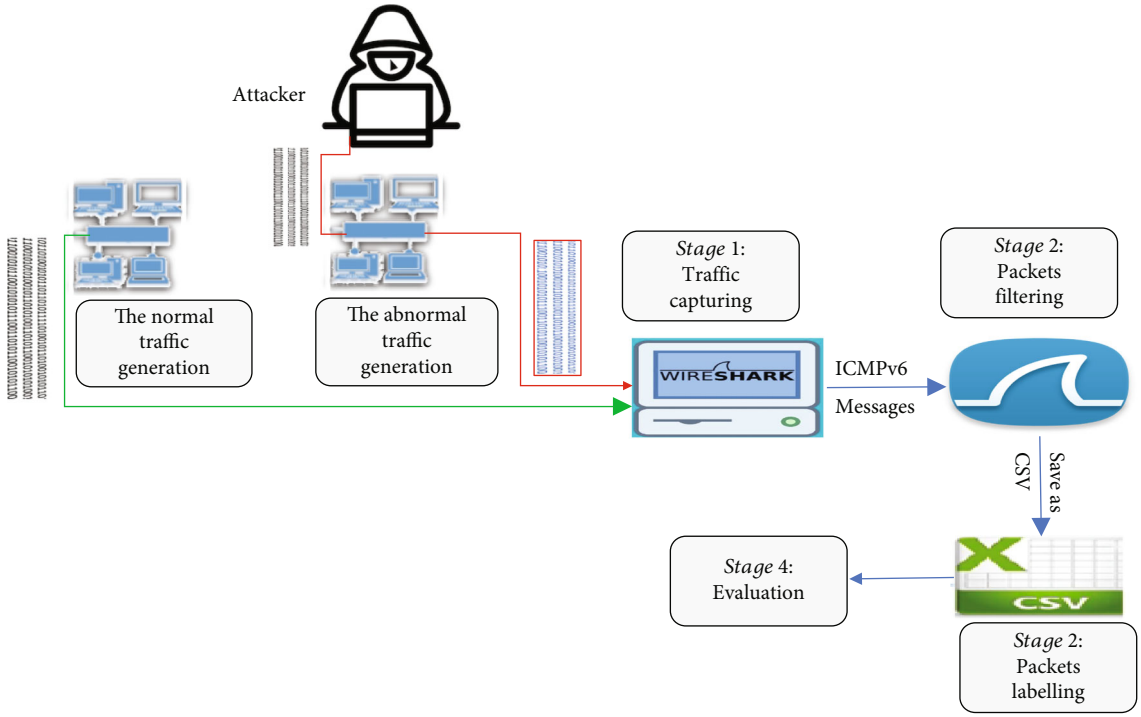

Figure 3: The proposed method.

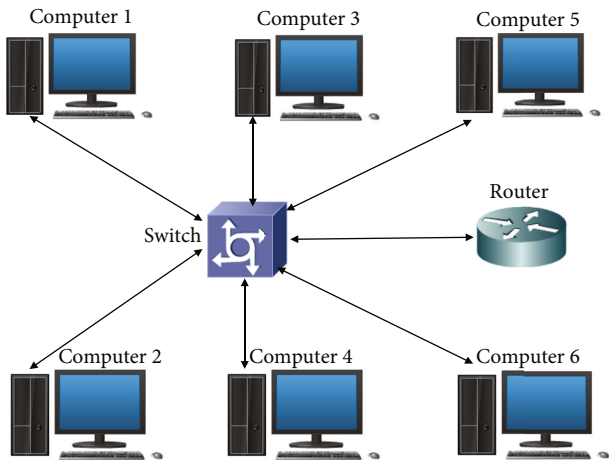

Figure 4: Virtual network topology for normal ICMPv6 traffic generation.

3.1.1. Generating Normal Traffic. There are numerous ways to generate IPv6 traffic, including the use of commercial traffic generator software. PACKETH, OSTINATO, and D-ITG are some examples of packet generator software for generating IPv6 traffic with user-defined features [17]. One limitation of these tools is their inability to simulate the behaviour of real-world devices. Furthermore, these tools are unable to mimic human behaviour, such as streaming, retyping, and typing speed. As a result, relying on these tools to generate a suitable dataset for our needs is impractical. Another way to generate IPv6 traffic is to set up a real IPv6 network with real computers and network devices, but this comes at a high cost. As stated in Section 2, one of the characteristics of a good dataset is that it must include normal traffic representation; the traffic represented in the dataset must appear and behave as realistically as possible. For the reasons stated above, creating a dataset using the actual production IPv6 network without jeopardizing the institution's privacy and security is not a viable option. As a result, using a network simulator like GNS3 [44] is a middle ground between the two extreme methods. GNS3 is used in the datasets to generate normal and abnormal traffic. Under typical working settings, the RA and NS messages are created by six hosts linked to a switch and a router, as displayed in Figure 4.

The Wireshark tool is listening in promiscuous mode on the connection between the switch and the router in the third stage. It captures, decodes, filters, and stores filtered data in a specific format, namely Packet CAPture (PCAP) files.

3.1.2. Generating Attack Traffic. The dataset used to evaluate a detection system must include both normal and abnormal traffic. The presence of different attack scenarios in the datasets is the second requirement for good datasets. Nevertheless, due to the negative impact on network performance and user devices, ICMPv6 DDoS attacks cannot be carried out in a real network. Since such abnormal traffic from ICMPv6-DDoS attacks cannot be captured in an actual network environment, a virtual network is used instead. Figure 5 shows the virtual network topology.

THC-IPv6 Toolkit [45] is used to trigger ICMPv6-DDoS attacks in order to generate abnormal or attack traffic. The attacks are directed at various victims, with several ICMPv6 messages originating from several attackers at different time intervals. The captured normal traffic is also replayed at high

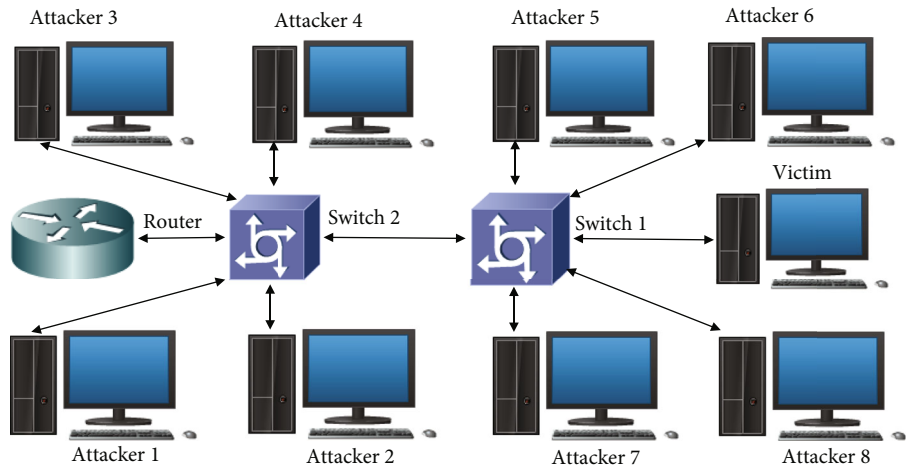

Figure 5: Virtual network topology for abnormal ICMPv6 traffic generation.

speed to flood the network in order to simulate flooding attacks. The variety of potential ICMPv6-DDoS attacks ensures the dependability of a detection model. Table 3 shows the different DDoS attack scenarios that were carried out.

Heuse [46] established The Hacker's Choice IPv6 (THC-IPv6) Toolkit to assess and test the security of IPv6 networks. It is capable of DDoS and MITM attacks [41, 42, 47, 48]. The malicious traffic used in this study, which consisted of ICMPv6 DDoS attacks, was generated with the help of THC-IPv6 Toolkit, which is built into the Kali Linux OS and makes it simple to launch attacks

Normal traffic is created by nodes, whereas abnormal traffic such as RA and NS messages is generated using THC-IPv6 Toolkit. Wireshark is used to capture the network traffic. In the first stage, it captures normal traffic for hours. In the second stage, the Flood-Router26 and Flood-Solicitate6 commands are used to generate abnormal traffic in order to capture the attack behaviours. The experiment is carried out in a virtual network, as shown in Figure 4, which was created using the GNS3 simulation tool. The tools included in THC-IPv6 Toolkit are used to initiate ICMPv6 DDoS attacks via the machines of eight attackers. Cross-validation tests are conducted to compare the proposed technique against currently available IDSs. The supplied set test includes two datasets, one for model testing and one for classifier training.

These messages can be abused for a variety of attacks, including the ICMPv6 flooding attack. An ICMPv6 flooding attack is carried out by sending a huge number of ICMPv6 messages to a victim, which can be a single node (PC or router) identified by its address. It can also target the entire IPV6 network by sending packets to an IPv6 multicast address. THC toolkit consists of various ICMPv6 flooding attacks that are launched from a machine running Kali OS. Since it is intended for intrusion detection and network security, the dataset is unprofitable if no attack scenarios are run. As a result, four distinct ICMPv6 flooding attacks were carried out in the network using THC Toolkit. All attacks were captured and logged using Wireshark from the monitor device.

THC Toolkit commands are used to launch attacks on the nodes of the victims for a set period of time. Table 4 shows the different types of attacks that can be carried out with the corresponding THC commands, as well as the target of each attack. Capturing normal and abnormal data is accomplished by sending and receiving ICMPv6 RA and NS messages between nodes via a switch or router, capturing natural data. THC-IPV6 Toolkit is used to generate DDoS attacks using ICMPv6 RA and NS messages. The Wireshark tool is used to capture the attack traffic. These captured data are considered abnormal or suspicious. Both normal and abnormal ICMPv6 (RA and NS) data are saved for filtering in the following step.

3.2. Packet Filtering Stage. ICMPv6 packet filtering is the stage in which data captured from the network is filtered by removing irrelevant data. Reduced data volume aids in the development of an effective detection system. Furthermore, filtering is a complex process in and of itself. Furthermore, irrelevant data in the dataset must be filtered out because it increases processing time and reduces the detection system's detection accuracy. Figure 4 shows a flowchart of the two stages of the ICMPv6 packet filtering stage.

The first step discards non-IPv6 packets with packet frame version values less than 6; other IPv6 packets may pass. The second step then discards non-ICMPv6 packets whose next header value is not equal to 58. Therefore, only ICMPv6 packets may pass through the packet filtering stage. To view only ICMPv6 packets, the captured traffic is filtered with the Wireshark filter ("IP. Version==6 && ipv6.nxt==58"). Figure 6 shows a sample of the captured ICMPv6 packet.

The captured data is not yet labelled, and the source IP address has not yet been identified as normal or abnormal. The process of identifying the packets is completed in the following step, packet labelling.

3.3. Packet Labelling Stage. After filtering, labelling the dataset is critical for the classifier or detection system to distinguish between normal and abnormal data during the training and testing phases. Following the conversion of the data to a CSV-formatted file, the dataset is labelled. In a new "class" field within the dataset, each category in the

Table 3: ICMPv6-based DDoS attack scenarios.

No.	Attack metric	Description
1	Sent packets	Any ICMPv6 packet
2	Source address	Either a real IPv6 address or a random (fake) IPv6 address
3	Destination address (victim)	Either a single device or group of devices (multicast address)
4	Attacking time	Different time distributed among the whole collection time
5	Attack type	Either ICMPv6 type or normal ICMPv6 traffic that is captured and resent in a short time period (replay attack)

Table 4: Details of attacks performed.

Attack name	THC command	ICMPv6 flooding attack packet	Target
Attack 1	flood_router26 eth0 "interface"	Router advertisement	All nodes
Attack 2	flood_router26 eth0 interface R	Router advertisement	All nodes
Attack 3	flood_router26 interface -R -a	Router advertisement	All nodes
Attack 4	flood_router26 interface -R -H	Router advertisement	All nodes
Attack 5	flood_router26 eth0 -H	Router advertisement	All nodes
Attack 6	flood_router26 eth0 -H -F	Router advertisement	All nodes
Attack 7	flood_router26 eth0 -H –s	Router advertisement	All nodes
Attack 8	flood_router26 eth0 –m	Router advertisement	All nodes
Attack 9	flood_router26 eth0 -G	Router advertisement	All nodes
Attack 10	flood_solicitate6 eth0 dest ip address	Neighbour solicitation	All nodes

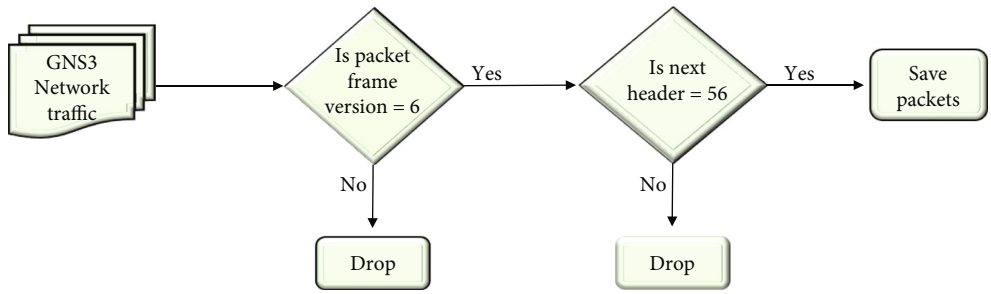

Figure 6: A flowchart of the filtering stage of ICMPv6 packets.

dataset registry is labelled as normal or attack data. To detect ICMPv6 flooding attacks, the IDS uses this label to differentiate between normal and abnormal data in the dataset. Each CSV file record is given a class label. The value of the class label is one of the packet's tuple-keys. It also means that the label value and the packet's tuple-keys value are the same. This ensures that the packet belongs to the same class, i.e., normal or attack. Table 5 shows the conversion of some captured traffic data into a CSV file format, as well as the addition of a "class" label in all records.

To evaluate the approaches, the dataset employs ICMPv6 messages as a benchmark dataset. Aside from the category label (normal or attack), the dataset contains twenty-one features. These characteristics are extracted by exchanging ICMPv6 messages between IPv6 network nodes, as shown in Figures 3 and 4. The reference dataset is used to generate a training set and a testing set. Each feature contains 16,881 dataset records containing ten different types of attacks. The characteristics of ICMPv6 packets are signified in a variety of formats. The analysis of various types of features needs extra processing time and hardware resources. To address these issues, the transformation process began by transforming the textual features to a numerical format, preparing them for use as input in the subsequent step (normalization). Excel functions are used to convert the format of dataset features from text values to numeric values.

The ICMPv6 message data contains approximately 13,504 training records, and the remainder is for testing. The ICMPv6 message dataset contains 21 fields of features, of which two are dropped: the package number and its time, which are considered redundant data for our needs. The label field, as described in the previous stage, is added. These

Table 5: Sample of captured datasets.

No.	Time	Source	Destination	Protocol	Length	ICMPv6 type	ICMPv6 checksum	Class
1	0	1	ff02::2	ICMPv6	70	Router solicitation	0 × 7510	Normal
2	0.383949	2	ff02::2	ICMPv6	70	Router solicitation	0 × a235	Normal
3	0.607769	3	ff02::2	ICMPv6	70	Router solicitation	0 × dd85	Normal
4	1.0959	4	ff02::2	ICMPv6	70	Router solicitation	0 × 8477	Attack
5	1.135772	5	ff02::2	ICMPv6	70	Router solicitation	0 × 685c	Attack
6	1.679853	6	ff02::2	ICMPv6	70	Router solicitation	0 × 9007	Attack
7	2.175709	7	ff02::2	ICMPv6	70	Router solicitation	0 × a9c4	Normal
414	1458.569	414	fe80::f437:Ccfff:fe59:45d7	ICMPv6	86	Neighbour solicitation	0 × 7398	Normal
415	1458.57	415	fe80::7b:63ff:fe48:1798	ICMPv6	78	Neighbour advertisement	0 × b0fc	Attack
416	1460.649	416	fe80::4403:Bbdff:fed8:6aa0	ICMPv6	86	Neighbour solicitation	0 × ad71	Attack
417	140.1955	417	ff02::1	ICMPv6	1486	Router advertisement	0 × 00ec	Normal

Table 6: Features of IPv6 packet dataset.

No.	Features	Description
1	No.	Refer to the number of packets
2	Time	Indicates when the attack occurred
3	IP source	Source IP address utilized in the session
4	IP destination	Destination IP address utilized in the session
5	Protocol type	Protocol type (e.g., ICMPv6)
6	Packet size (length)	Number of ICMPv6 data bytes sent by source IP address
7	ICMPv6 code	ICMPv6 code
8	ICMPv6 type	ICMPv6 packet type (e.g., RA message)
9	ICMPv6-checksum	ICMPv6 checksum field in the ICMP header
10	ICMPv6-TA	ICMPv6 insects to receive the target address.
11	ICMPv6-option	ICMPv6 option type
12	ICMPv6-CHL	ICMPv6 insects to receive the current hop limit
13	ICMPv6-AF	ICMPv6 insects to receive the flag
14	ICMPv6-RL	ICMPv6 insects to receive the router time
15	ICMPv6-RH	ICMPv6 insects to receive reachable time
16	ICMPv6-RT	ICMPv6 insects to receive retrains time
17	Flow label	Variation in the label of packets
18	Next header	Variation in the next header of packets
19	Hope limit	Variation in the hop limit of packets
20	Payload length	Variation in the payload length of packets
21	Traffic class	Variation in the traffic class of packets

Table 7: Detection accuracy for five classifiers applied to datasets.

No.	Classifier	Accuracy
1	Decision tree	80.79%
2	Support vector machine (SVM)	78.78%
3	Naïve Bayes	80.54%
4	*k*-nearest neighbours (KNN)	81.57%
5	Neural networks	81.57%

multiattribute fields are converted to a single numeric format. (Table 6).Table 6 represents the details of the features that were extracted from the generated dataset

3.4. The datasets Evaluation Stage. In this section, we used learning-based anomaly detection by selecting five classification methods for the evaluation task. We adapted the classifier using the cross-validation test method, which involves training the classifiers on a part of the dataset (training) and then testing these trained models on the remaining dataset (test). Classifiers are utilized without any parameter optimization or tuning because the purpose of the experiment is to

Table 8: The comparison between the proposed datasets and existing IPv6 datasets.

No.	Dataset	Network configurations	Labelled	Attack scenarios	Number of features	Online
1	MAWI [37]	Unavailable	No	Normal traffic only	10	Available
2	CAIDA [32]	Unavailable	No	Nondiverse	Limited	Available
3	Zulkiflee et al. [41]	Available	Yes	Three IPv6 attack scenarios	6	Unavailable
4	Najjar and Kadhum [31]	Available	No	Two DDoS attack scenarios	7	Unavailable
5	Saad et al. [34]	Available	Yes	Echo request	8	Unavailable
6	Omer et al. [17]	Available	Yes	Diverse	11	Available
7	Proposed dataset	Available	Yes	Diverse	21	Will be available

demonstrate that the datasets are dependable and trustworthy for evaluating ICMPv6 DDoS detection methods. Furthermore, this section seeks to show that the proposed features are capable of distinguishing between ICMPv6 DDoS behaviours and normal behaviours.

The proposed datasets were evaluated using the evaluation scales, specifically the detection accuracy and false-positive rates. Along with the proposed features, these scales are used to estimate the efficiency of the proposed package-based representation. The detection accuracy rate (DAR) indicates an attacker's ability to classify attacks correctly. Equation (1) is used to determine DAR [49].

$$\text{Detection accuracy rate (DAR)} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100\%, \tag{1}$$

where FP is the sum of the attack samples that the classifier incorrectly predicted, namely, the false-positive; FN is the sum of the normal samples that the classifier incorrectly predicted, namely, the false-negative; TP is the sum of the attack samples that the classifier correctly predicted, namely, the true-positive; and TN is the sum of the normal samples that the classifier correctly predicted, namely, the true-negative.

The classifiers used in the experiment are commonly available in apply fast.ai [50], are an open-source Python-based library that uses PyTorch [51], and were established in 2016 by Jeremy Howard and Rachel Thomas with the goal of democratizing artificial intelligence and deep learning. fast.ai offers practitioners high-level components that can rapidly and easily produce state-of-the-art results in typical machine learning domains, as well as researchers components that can be mixed and coupled to construct novel methods. They were picked to prevent having to repeat the experiment. These classifiers were chosen as variables because they varied in terms of kind, classification performance, and quantity of features employed. Selector naïve Bayes and support vector machine (SVM) are classifiers applied to datasets with features. The testing method is used to calculate evaluation scales for applying workbooks to datasets. Tables 7 represents the evaluation metric and is the detection accuracy rate for applying classifiers to datasets with features using the test approach; the same classifiers were applied to the datasets.

Table 7 displays the accuracy achieved by classifiers employing 21 features. The cross-validation approach yielded detection accuracies for the classifiers evaluated. The findings also demonstrate the effectiveness of classifiers in detecting attacks. In the comparison, the accuracy of identifying attacks launched via RA and NS messages was nearly the same with no change.

The detection accuracies for each classification algorithm are shown in Tables 7 and 8 using the ICMPv6 DDoS dataset. Classification algorithms yielded a range of evaluation metrics. Nevertheless, some of them attained higher values than others. First, KNN and neural network algorithms outperformed all other algorithms in the majority of evaluation metrics. Second, the SVM achieved significantly lower values for the evaluation metric.

According to Table 7, the proposed datasets with the twenty-one features achieved high detection accuracies (up to 81.57%) for all classifiers used. Furthermore, it achieved robust and consistent evaluation rating scales in the test, confirming that the representation and features used are appropriate for such attack detection. However, the obtained results have still not reached high levels, implying that further development is conceivable. Based on the results above, interested researchers can improve them further by enhancing the classifiers or by utilizing feature selection technology to minimize the number of characteristics through the use of optimisation algorithms. Additionally, because these classifiers were implemented using default settings, fine-tuning the parameters may improve the outcome.

These findings demonstrated that the proposed datasets met the criteria for good datasets. Furthermore, by combining the datasets, this dataset can be used to evaluate and improve any proposed classification technique. The proposed datasets can also assist researchers in understanding the differences between abnormal (ICMPv6 DDoS attacks) and normal traffic in order to develop a more effective intrusion detection system.

4. Comparison with State-of-the-Art Datasets

In general, IPv6 datasets are created to meet specific research goals and needs. As such, as discussed in Section 2, many failed to completely meet the requirements of other researchers. Furthermore, many datasets were not publicly available for use by other researchers. Even though the CAIDA dataset is freely available to researchers online, it has been heavily modified due to privacy concerns. The protocol, source IPv6 address, and destination IPv6 address were all removed from the data, limiting the dataset's usefulness to other researchers. Therefore, the purpose of this paper is to

create an alternative reference dataset for detecting ICMPv6-DDoS attacks.

A qualitative comparison was made in terms of various criteria as follows:

(i) Provides network configuration information: the term "network configuration" refers to the architecture, devices, and machines used to construct the IPv6 network and generate network traffic. This information is critical for various types of security research because it enables a better understanding of the dataset by revealing how network traffic is generated and other pertinent information

(ii) Dataset labelling: labelling allows a dataset to be used to improve or evaluate the performance of detection systems. Labelling is the deterministic addition of a category classification to each dataset's record that describes its source (normal or attacker). Unlabelled datasets, on the other hand, cannot be used to assess the accuracy of a detection system

(iii) Diverse attack scenarios: the detection system's testing against a dataset including a variety of attack scenarios ensured that the detection system faced a robust dataset with attacks that varied in terms of a variety of parameters, such as time, packets transmitted, source, frequency, and destination

(iv) Number of features: a dataset must contain features that are specific, important, and relevant enough to allow a detection system to accurately distinguish between normal and abnormal data. The more features there are, the more diverse the data that can be extracted from them

(v) Available online: making the dataset public allows other researchers to use it to evaluate their work and compare it to other existing work. The datasets proposed are made available and published on a website

The proposed datasets and the existing IPv6 datasets are compared qualitatively in Table 8. As displayed in Table 2, the work presented in this article attempted to address issues identified in other datasets. The majority of the existing datasets have not yet been classified. Labels are required for evaluating various detection systems. However, most of datasets are unclassified. Another issue with existing datasets is that they lack a variety of attack scenarios. For example, the dataset created by Najjar and Kazim includes two DDoS attack scenarios. Thus, it can only be used by detection systems that are designed to detect those two types of attacks. However, most detection systems are aimed at detecting as many attack types as possible, necessitating the use of datasets containing a variety of attack scenarios.

Additionally, all existing datasets are incapable of assisting in the generation of our flow-based represented datasets. MAWI and CAIDA datasets, for example, contain only normal traffic. As a result, they cannot be used to develop or evaluate mechanism-based attack detection. Saad et al.'s dataset is not publicly accessible. It is also an unlabelled dataset, so it is ineligible for our purposes. Although the Najjar and Kadhum dataset is labelled, it is not publicly available and lacks diversity in ICMPv6 DDoS attack scenarios.

5. Conclusion

ICMPv6 is a necessary and essential IPv6 protocol for any node in an IPv6 network to function properly. Nevertheless, ICMPv6 is vulnerable to a variety of attacks, including DoS, DDoS, and MITM attacks. Among the most dangerous attacks against IPv6 networks is the ICMPv6-DDoS attack. Many researchers have proposed various detection systems for ICMPV6-DDoS attacks. However, due to a lack of publicly available ICMPv6 datasets, these systems must rely on self-created datasets, which are insufficient to detect these attacks because they contain unqualified features that lead to misclassification.

There are also IPv6 datasets that were created for nonsecurity purposes, have limited attack scenario types, and are not deployed for others to use, or have unlabelled traffic. As a consequence, they could not be used to develop, train, or test detection systems for ICMPv6-DDoS attacks.

The proposed datasets met the requirements for a good dataset by being generated from traffic that contained the majority of possible ICMPv6-DDoS attack scenarios, contained traffic with accurate and complete labels, and had a convergent ratio between normal and abnormal (attack) representations using a representative set of features. Additionally, the representations and features used were evaluated using a variety of default parameterized classifiers. On the basis of detection accuracy and false-positive rates, the outcome demonstrates acceptable results and robustness.

The dataset's limitations do not stem from their generation from realistic traffic. Another barrier is the use of security testing tools in production networks, whether local or global networks, such as the Internet, due to the security of private information that can be identified by IP address users. The dataset that is being proposed is limited to ICMPv6 flooding attacks. Subsequently, in future work, we intend to broaden its diversity and make it more reliable against a variety of IPv6 attacks in traffic.

The ICMPv6 protocol is vulnerable to a variety of attacks other than ICMPv6 flooding attacks. Therefore, these attacks will be added to the dataset in the future. Furthermore, we intend to update and expand the network in order to increase the number of computers engaging in abnormal activity as well as the proportion of malicious traffic. Furthermore, the number of computers performing routine tasks will be increased. Including these activities is aimed at adding more reality to the dataset traffic and is required for labelling and validating the dataset.

Data Availability

The dataset was established at the University of Science Malaysia and is available by sending an email to the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] A. Conta, S. Deering, and M. Gupta, *Internet Control Message Protocol (icmpv6) for the Internet Protocol Version 6 (ipv6) Specification*, Request for Comments 4443 [online], 2006.

[2] T. Narten, E. Nordmark, and W. Simpson, *Neighbor Discovery for IP Version 6 (IPv6)*, 2007.

[3] R. M. A. Saad, S. Manickam, and S. Ramadass, "Utilizing data mining approches in the detection of intrusion in IPv6 network: review & analysis," *International Journal on Network Security*, vol. 4, no. 1, p. 35, 2013.

[4] D. M. Convery Sean, "IPv6 and IPv4 threat comparison and best-practice evaluation (v1. 0) 1 introduction," pp. 1–43, 2004, http://e-edu.pgisliven.com.

[5] J. Postel, *Internet protocol*, Internet Engineering Task Force, 1981.

[6] E. Durdağı and A. Buldu, "IPV4/IPV6 security and threat comparisons," *Procedia - Social and Behavioral Sciences*, vol. 2, no. 2, pp. 5285–5291, 2010.

[7] A. H. Bdair, R. Abdullah, S. Manickam, and A. K. Al-Ani, "Brief of intrusion detection systems in detecting ICMPv6 attacks," in *Computational Science and Technology*, vol. 603, Springer, Singapore, 2020.

[8] D. Security, *Chair for Network and Data Security*, 2017.

[9] O. E. Elejla, B. Belaton, M. Anbar, and A. Alnajjar, "Intrusion detection systems of ICMPv6-based DDoS attacks," *Neural Computing and Applications*, vol. 30, pp. 1–12, 2018.

[10] N. B. Lucena, G. Lewandowski, and S. J. Chapin, "Covert channels in IPv6," in *Privacy Enhancing Technologies*, vol. 3856, pp. 147–166, Springer, Berlin, Heidelberg, 2006.

[11] A. R. Choudhary, "In-depth analysis of IPv6 security posture," in *2009 5th International Conference on Collaborative Computing: Networking, Applications and Worksharing*, Washington, DC, USA, 2009.

[12] S. Cabuk, C. E. Brodley, and C. Shields, "IP covert timing channels: design and detection," in *Proceedings of the ACM Conference on Computer and Communications Security*, pp. 178–187, Washington DC USA, 2004.

[13] A. Carp, A. Soare, and R. Rughiniş, "Practical analysis of IPv6 security auditing methods," in *Proceedings -9th RoEduNet IEEE International Conference*, pp. 36–41, Sibiu, Romania, 2010.

[14] C. E. Martin and J. H. Dunn, "Internet Protocol Version 6 (IPV6) protocol security assessment," in *Proceedings-IEEE Military Communications Conference MILCOM*, Orlando, FL, USA, 2007.

[15] L. Frikha, Z. Trabelsi, and S. Tabbane, "Simulation, optimisation and integration of covert channels, intrusion detection and packet filtering systems," in *2009 Global Information Infrastructure Symposium*, pp. 1–4, Hammemet, Tunisia, 2009.

[16] J. Weber, *Master Thesis IPv6 Security Test Laboratory, [M.S. thesis]*, Computational Engineering Department, Ruhr-University Bochum, Germany, 2013.

[17] O. E. Elejla, M. Anbar, B. Belaton, and S. Hamouda, "Labeled flow-based dataset of ICMPv6-based DDoS attacks," *Neural Computing and Applications*, vol. 31, no. 8, pp. 3629–3646, 2019.

[18] O. E. Elejla, *Flow-Representation Approach For ICMPV6-Based Ddos Attacks Detection, [Ph.D. thesis]*, School Comput. Sci., Univ. Sci. Malaysia, Penang, Malaysia, 2018.

[19] K. Yun and Z. J. Mei, "Research of hybrid intrusion detection and prevention system for IPv6 network," in *International Conference on Internet Technology and Applications, IEEE*, Wuhan, China, 2011.

[20] J. D. Lim, Y. H. Kim, B. H. Jung, K. Y. Kim, J. N. Kim, and C. H. Lee, "Implementation of multi-thread based intrusion prevention system for IPv6," in *ICCAS 2007- International Conference on Control, Automation and Systems*, pp. 404–407, Seoul, 2007.

[21] M. Zulkiflee, M. Ahmad, S. Sahib, and M. Ghani, "A framework of features selection for ipv6 network attacks detection," *WSEAS Transactions on Communications*, vol. 14, no. 46, pp. 399–408, 2015.

[22] A. Sperotto, G. Schaffrath, R. Sadre, C. Morariu, A. Pras, and B. Stiller, "An overview of IP flow-based intrusion detection," *IEEE Communication Surveys and Tutorials*, vol. 12, no. 3, pp. 343–356, 2010.

[23] P. Winter, E. Hermann, and M. Zeilinger, "Inductive intrusion detection in flow-based network data using one-class support vector machines," in *2011 4th IFIP International Conference on New Technologies, Mobility and Security*, Paris, France, 2011.

[24] F. Najjar and M. M. Kadhum, "Reliable behavioral dataset for IPv6 neighbor discovery protocol investigation," in *2015 5th International Conference on IT Convergence and Security (ICITCS)*, Kuala Lumpur, Malaysia, 2015.

[25] C. Brown, A. Cowperthwaite, A. Hijazi, and A. Somayaji, "Analysis of the 1999 DARPA/Lincoln Laboratory IDS Evaluation Data with NetADHICT," in *IEEE Symposium on Computational Intelligence for Security and Defense Applications*, Ottawa, ON, Canada, 2009.

[26] M. Tavallaee, N. Stakhanova, and A. A. Ghorbani, "Toward credible evaluation of anomaly-based intrusion-detection methods," *IEEE Transactions on Systems, Man, and Cybernetics Part C: Applications and Reviews*, vol. 40, no. 5, pp. 516–524, 2010.

[27] P. Ying, W. Shihui, and X. Xing, "Classification method of IPv6 traffic based on convolutional neural network," *Journal of Physics Conference Series*, vol. 1883, no. 1, p. 012088, 2021.

[28] Y. Alharbi, A. Alferaidi, K. Yadav, G. Dhiman, and S. Kautish, "Denial-of-service attack detection over ipv6 network based on KNN algorithm," *Wireless Communications and Mobile Computing*, vol. 2021, 6 pages, 2021.

[29] R. Lippmann, J. W. Haines, D. J. Fried, J. Korba, and K. Das, "The 1999 DARPA off-line intrusion detection evaluation," *Computer Networks*, vol. 34, no. 4, pp. 579–595, 2000.

[30] U. O. California and Kdd cup, "Datasets|Research|Canadian Institute for Cybersecurity|UNB," 1999, January 2019. https://www.unb.ca/cic/datasets/index.html.

[31] S. D. Hettich and S. Bay, *KDD-CUP-99 Task Description*, pp. 1–4, 1999.

[32] CAIDA, "CAIDA: Center for Applied Internet Data Analysis," 2014, November 2018, http://www.caida.org/home/.

[33] L. B. N. Laboratory, "LBNL/ICSI enterprise tracing project-project overview," 2004, December 2020, http://www.icir.org/enterprise-tracing/.

[34] R. M. A. Saad, S. Manickam, E. Alomari, M. Anbar, and P. Singh, "Design & deployment of testbed based on ICMPv6 flooding attack," *Journal of Theoretical and Applied Information Technology*, vol. 64, no. 3, pp. 795–801, 2014.

[35] A. Shiravi, H. Shiravi, M. Tavallaee, and A. A. Ghorbani, "Toward developing a systematic approach to generate benchmark datasets for intrusion detection," *Computers & Security*, vol. 31, no. 3, pp. 357–374, 2012.

[36] O. E. Elejla, A. B. Jantan, and A. A. Ahmed, "Three layers approach for network scanning detection," *Journal of Theoretical and Applied Information Technology*, vol. 70, no. 2, pp. 251–264, 2014.

[37] MAWI Working Group, "MAWI Working Group traffic archive," 2012, https://mawi.wide.ad.jp/mawi/.

[38] D. Barrera and P. C. Van Oorschot, "Security visualization tools and IPv6 addresses," in *6th International Workshop on Visualization for Cyber Security 2009*, pp. 21–26, Atlantic City, NJ, USA, 2009.

[39] M. Fomenkov and K. Claffy, *Internet measurement data management challenges*, 2011.

[40] M. D. Gray, *Discovery of IPv6 Router Interface Addresses via Heuristic Methods, [M.S. thesis]*, Naval Postgraduate School, 2015.

[41] M. Zulkiflee, N. Haniza, S. Shahrin, and M. K. A. Ghani, "A framework of IPv6 network attack dataset construction by using testbed environment," *International Review on Computers and Software (IRECOS)*, vol. 9, no. 8, p. 1434, 2014.

[42] O. E. Elejla, M. Anbar, B. Belaton, and B. O. Alijla, "Flow-based IDS for ICMPV6-based DDoS attacks detection," *Arabian Journal for Science and Engineering*, vol. 43, no. 12, pp. 7757–7775, 2018.

[43] Wireshark Foundation, "Wireshark · Go deep," *Abgerufen am*, vol. 2009, p. 27.05, 2010, 2010. https://www.wireshark.org/.

[44] A. Grossman, J. Marsili, B. Goudjil, and C. Eromenko, "Gns3 graphical network simulator," 2013, July 2019, https://www.gns3.com.

[45] M. Heuse, "THC-IPv6-Attack-Toolkit - aldeid," 2013, July 2019, https://www.aldeid.com/wiki/THC-IPv6-Attack-Toolkit.

[46] M. Heuse, *THC IPv6 attack tool kit*, Github, 2013, February 2016, https://github.com/vanhauser-thc/thc-ipv6.

[47] C. E. Caicedo, J. B. D. Joshi, and S. R. Tuladhar, "IPv6 security challenges," *Computer*, vol. 42, no. 2, pp. 36–42, 2009.

[48] V. Alangar and A. Swaminathan, "I Pv6 security: issue of anonymity," *Journal of Engineering and Computer Science*, vol. 2, no. 8, pp. 2486–2493, 2013.

[49] H. Tang and Z. Cao, "Machine learning-based intrusion," *The Journal of Computer Information Systems*, vol. 5, pp. 1825–1831, 2009.

[50] Fast.ai, "fast.ai | making neural nets uncool again," 2020, February 2021, https://www.fast.ai/.

[51] T. Doleck, D. J. Lemay, R. B. Basnet, and P. Bazelais, "Predictive analytics in education: a comparison of deep learning frameworks," *Education and Information Technologies*, vol. 25, no. 3, pp. 1951–1963, 2020.