Locally Weighted Classifiers for Detection of Neighbour Discovery Protocol DDoS and Replayed Attacks

Abeer Alsadhan[1], Abir Hussain[2], Panos Liatsis[3], Mohammed Alani[4], Hissam Tawfik[5], Phillip Kendrick[2], Hulya Francis[2]

[1]Department of Computer Science, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia.
aalsadhan@iau.edu.sa

[2]Department of Computer Science, Liverpool John Moores University, Liverpool, UK.

[3]Department of Electrical Engineering and Computer Science, Khalifa University, Abu Dhabi, United Arab Emirates.

[4]Department of Information Technology, Al Khawarizmi International College, Abu Dhabi, United Arab Emirates.

[5]School of computing, Leeds Beckett University, Leeds, UK.

ABSTRACT The Internet of Thing (IoT) requires more IP addresses than Internet Protocol version 4 can offer. To solve this problem, Internet Protocol version 6 was developed to expand the availability of address spaces. Moreover, it supports hierarchical address allocation methods, which can facilitate route aggregation, thus limiting expansion of routing tables. An important feature of the Internet Protocol version 6 (IPv6) suites is the Neighbour Discovery Protocol (NDP), which is geared towards substitution of the Address Resolution Protocol in router discovery, and function redirection in Internet Protocol version 4. However, NDP is vulnerable to Denial of Service (DoS) attacks. In this contribution, we present a novel detection method for Distributed Denial of Service (DDoS) attacks, launched using NDP in IPv6. The proposed system uses flow-based network representation, instead of packet-based. It exploits the advantages of Locally Weighted Learning techniques, with three different machine learning models as its base learners. Simulation studies demonstrate that the intrusion detection method does not suffer from overfitting issues, offers lower computation costs and complexity, while exhibiting high accuracy rates. In summary, the proposed system uses 6 features, extracted from our bespoke dataset and is capable of detecting DDoS attacks with 99% accuracy and replayed attacks with an accuracy of 91.17%, offering a marked improvement in detection performance over state-of-the-art approaches.

INDEX TERMS DDoS Attack, Neighbour Discovery Protocol, IPv6, Intrusion Detection System, Locally Weighted Learning.

I. INTRODUCTION

New discoveries in science led to major developments, and technology is constantly evolving to meet the demands of increasing numbers of users. The Internet is no exception in this transition period. Internet users appear to have exhausted the capacities of Internet Protocol Version 4 (IPv4), having taken advantage of its strength in providing reliable and interconnected computer networking with speeds comparable to local networks for information sharing and processing, online repository, multi-player network games, web-based applications, backups and other services [1]. The increasingly growing number of internet users led to the transitioning from IPv4 to IPv6, an easily implementable, interoperable and robust internet protocol in a 128-bit address format, providing up to 3.4×10^{38} unique IP addresses [2], with the capacity to overcome the depletion issues of IPv4 addresses [3]. The security of this network infrastructure against the attacks is crucial, as DDoS and or replay attacks exhaust the resources of the attacked network and are liable to cause great havoc to a network of devices [4]. Nowadays, the problem of securing an IPv6 network is becoming even more relevant, as attackers may exploit the vulnerabilities of the Neighbour Discovery Protocol (NDP) in order to launch DDoS and replay attacks [5].

This research aims at proposing a novel framework for an Intrusion Detection System (IDS), capable of securing IPv6 networks using advanced data mining techniques. The contributions of this work are as follows:

1) Construction of flow-based representation of packets as extraction of flow-based features for the development of IDS.

2) IDS structures based on locally weighted learning (LWL) for typical IPv6 networks, and the use of feature

selection techniques for best-selecting important features.

3) Three IDS methods using LWL -Bayesian networks, LWL-Decision Trees and the LWL-Naïve Bayes algorithms for detecting ICMPv6 DDoS and Replayed attack,

4) Simulation results verify that the proposed IDSs are capable of detecting Neighbour Discovery Protocol DDoS, and Replay attacks.

5) Moreover, the proposed IDSs are also effective in detecting any form of anomalies, when it deviates from normal flow-based packets.

The remaining part of this paper is organized as follows. A review of related works and associated background is presented in Section II. Section III describes the specifics of the proposed IDS. Experimental results and performance evaluation of the IDS methods are presented in Section IV. Finally, Section V summarizes the conclusions of this research and proposes directions for future work.

II. BACKGROUND

The term denial of service was originally coined by Gilgor before it was widely adopted [6]. This type of attack dating back to the 1980s accounts for more than one-third of all current networks attacks worldwide. A DoS attack exhausts the bandwidth and computational resources of its victim, as well as that of the users on the same network, by flooding it with packets [7]. DDoS attacks are a form of coordinated attacks, involving two or more computers targeting a victim [8].

There are two types of DoS attacks, namely, network level and application level. Network level attacks disable the connectivity of legitimate users by exhausting the network resources, e.g., bandwidth, router processing capacity etc. Application level attacks target the server resources, e.g. memory, CPU, and database/disc bandwidth [8]. DDoS attacks, on the other hand, are usually categorised into application level, state exhausting, and volumetric. Volumetric attacks represent 65% of DDoS attacks, making these the most common type of DDoS attack [9].

The standing difference between DDoS and DoS attacks is that the former is a large-scale DoS attack that uses severally distributed compromised machines from different internet regions to send traffic to a victim machine [8], while the latter is launched from a single machine to the victim machine. Examples of DDoS volumetric attacks are TCP flood attacks, DNS spoofing and UDP flood attacks. DDoS application level attacks usually inject malicious resources such as software for taking over control of a device or queries to its victim machine in order to exhaust its resources. Finally, state exhausting attacks include ping-of-death, which makes up 20% of all attacks related to DDoS [5].

A. Distributed Denial of Service Attack

DDoS attacks have been executed on computer networks and devices in recent times [10]. Some recent examples are the Oct 21st, 2016 attack on the East Coast, USA, which targeted a number of US banks, including Wells Fargo, JP Morgan, and the attack on the UK National Lottery in 2017 [5]. DDoS attacks often have different intentions and launch methods, which could be identified to detect the attack and potentially prevent it. DDoS attacks normally exhibit no obvious features, which can be used to identify such packets as malicious [11]. More so, the tools used in carrying out this form of attack are very easily obtainable, in fact a compromised machine used to send packets to victim might not be aware that it is being compromised, which makes the attacks quite frequent. Maintaining a simple structure of many-to-one features as depicted in Figure 1, DDoS attacks are complex to detect or resolve and can wreck a terrible impact on attacked network or device as attackers remain hidden through the usage of IP spoofing while executing this deadly attack [12].

Furthermore, the attacker's incentives play a role in this type of the attack [13]. The attacker's Incentives include intellectual challenges often carried out by younger hackers, and cyber warfare. These are mostly politically motivated, for financial gain, or serving ideological beliefs.

Considering that the victim machine has limited resources compared to the attackers', DDoS is a dangerous form of attack, which exploits the general delinquencies found in networks of computers and devices[4].

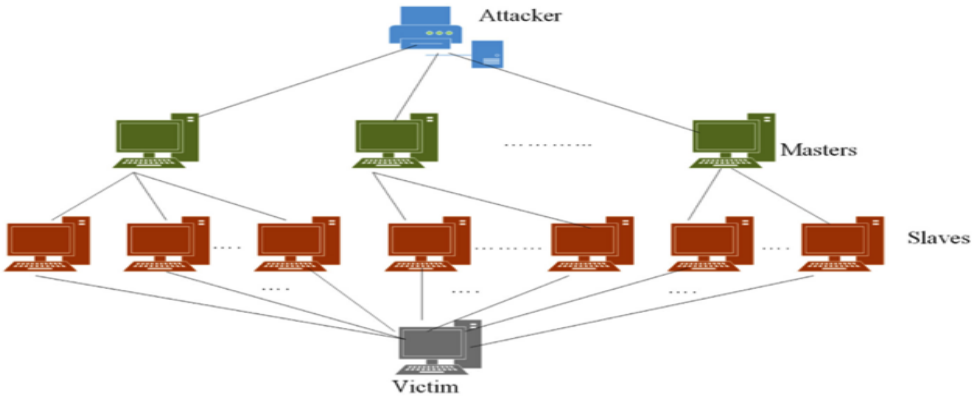

Figure 1: Architecture of DDoS attacks [14]

In a DDoS attack, the attacker creates master machines to control the slave machines to be used in executing the attack. Once the size of the compromised machines grows to a sufficient level, the attacker launches its bout on the victim by sending exhausting packets to the network or server resources, using spoofed IP addresses to cause loss of service [7]. A pictorial representation of DDoS attack architecture is seen in Figure 1.

B. Internet Protocol version 6 (IPv6)

IPv6 is the successor of IPv4 Internet protocol, offering a larger address space, and bringing about new features and improvements, which enable easy configuration, fast routing or packers and better quality of service [15]. IPv6 has unicast and multicast addresses, as IPv4. However, there is no distinct concept of a broadcast address in IPv6. Instead, a new type of address, i.e., anycast address, has been added to allow a message to be sent to any one member of a group of devices [16].

New features offered by this IP version include NDP and Internet Control Message Protocol version 6 (ICMPv6) which are being discussed in the next sub-sections.
However, IPv6 is vulnerable to various types of attacks, including Bypass, Port Scan and DoS, as depicted in Figure 2.

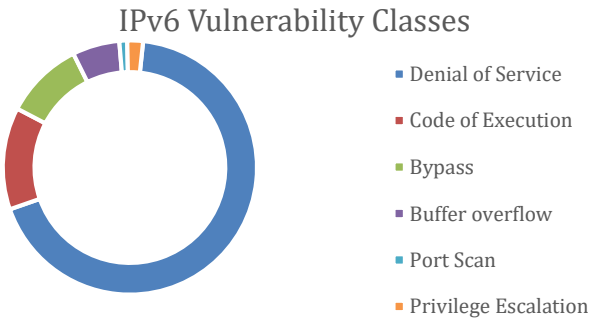

Figure 2: IPv6 vulnerability classes [14]

As reported by the National Vulnerability Database, DoS attacks are ranked as the top type of attack on networks of computers and devices, accounting for approximately 68% of all attacks.

1) INTERNET CONTROL MESSAGE PROTOCOL VERSION 6 (ICMPV6)

ICMPv6 is integrated as a core IPv6 protocol (while it is optional in IPv4) to facilitate communication between hosts. It is a multi-functional protocol that enables discovery of Path Maximum Transmission Unit (PMTU), address resolution, Neighbour Discovery and auto-configuration [14].

It is used for testing or diagnostic purposes and its messages are categorised as error messages (ranging from 1 to 127) and information messages (ranging from 128 to 255) [7]. It is vital in IPv6 network operations; however, it is susceptible to attacks such as Scan, DoS and Man in the Middle (MitM) [7], [21].

2) NEIGHBOUR DISCOVERY PROTOCOL (NDP)

The Neighbour Discovery Protocol is a network layer protocol that enables communication on local links. It carries out its functions solely depending on ICMPv6 by making use of five different types of ICMPv6 informational messages, sent by nodes and routers [12].

A typical task of NDP is to facilitate the auto generation of the IPv6 address for a node, often referred to as "Plug and Play" and also to enable the exchange of NDP messages between nodes [12].

3) IPV6 AND ICMPV6 SECURITY THREATS

As previously mentioned, IPv6, although designed to solve IPv4 issues, it inherits some security threats faced by IPv4. Moreover, IPv6, as with any new technology, brings its own set of security threats [9]. The exploitation of IPv6 implementation by cybercriminals during its initial phases offers opportunities for improvements and upgrades, in regards to security [10].

The most common security threats facing IPv6 include routing headers manipulation, reconnaissance attacks, fragment headers DoS attacks, the risks of tunnels, potential holes in dual stack approaches and ICMPv6 spoofing.

ICMPv6 lacks security awareness, and thus it is vulnerable to attacks [18]. ICMPv6 suffers from security flaws, such as the potential misuse of multicast addresses to perform reconnaissance attacks, and the potential to be used when executing a DoS attack [11]. Furthermore, the NDP part of ICMPv6 can be compromised and used to make victims unreachable.

4) NDP-BASED DDOS ATTACKS

DDoS attacks are carried out on IPv6 networks, mostly by using ICMPv6 in a number of possible scenarios. These include excessive transmission of ICMPv6 packets to a victim, transmission of error messages causing a drop in active sessions of established communication, and invalidation of legitimate addresses or disabling of interfaces when an infiltration occurs onto a link's maintenance messages [7].

To perform a DDoS attack using ICMP, attackers transmit "*ICMP_ECHO_REQUEST*" packets, using the broadcast IP address, to this victim's network. The machines on the victim's network reply to this request with the *"ICMP_ECHO_REPLY"* packets, and then the attacker may inundate the victim's network using an intermediary network. Consequently, this saturates the victim's network bandwidth, which leads to the unavailability of the network to its legitimate users [4].

Deploying a Network IDS (NIDS) at the default gateway of IPv6 network tends to curb DDoS attacks [17] as it captures packets passing through the network, and analyses them for the purpose of detecting DDoS attacks. The method of representation of the network traffic determines the performance of NIDS. The network traffic can either be represented as flow-based or packet-based [11].

C. Intrusion Detection Systems for NDP DDoS

IDS evolved in respect to addressing cybercrimes. In this age of rapid digital development and integration of digital services in the operations of several stakeholders (e.g., government, institutions, businesses and non-profit organisations), IDS is a vital part of security technology for protection of digital assets [10].

Being strategically located in a network in order to serve its purpose, IDS collects data from various systems and network sources for analysis with the aim of detecting threats or attacks. IDS offers network protection, higher accuracy, and the ability to detect previously unknown attacks using data mining techniques [19]. IDS can be categorized, based on its location for sourcing data, as Host-Based (HIDS) or Network-Based (NIDS). In addition, IDS can be categorized based on its detection mechanism as Signature-based (SIDS), Anomaly-based (AIDS), and Stateful Protocol Analysis [17], as shown in Figure 3. Anomaly-based IDS can be further be categorized as Rule-based IDS, and Artificial Intelligence-based [20]. Stateful Protocol Analysis enables IDS to detect and trace protocol states, thus providing important information for understanding and responding to an attack [8] ,[39].

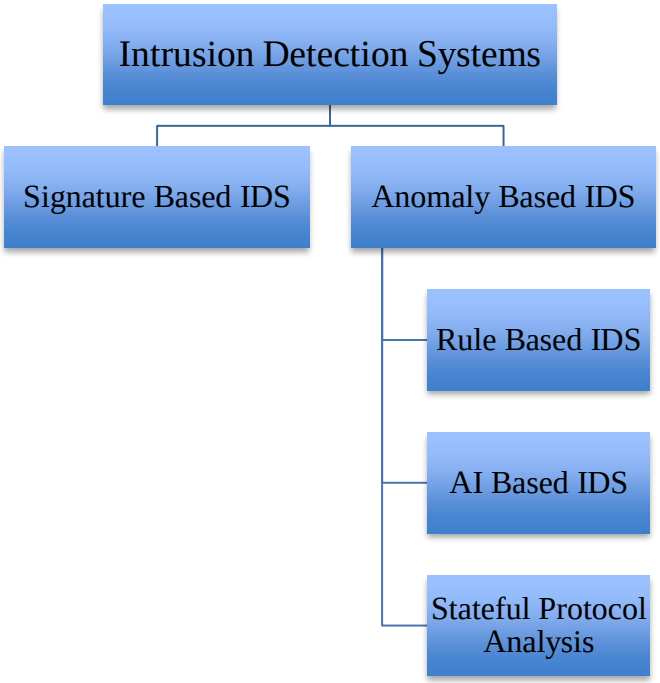

Figure 3: Categorization of IDSs based on their detection mechanisms [19].

III. REVIEW OF RELATED WORK

Since the early years of IPv6, and even before its live implementation, security researchers have studied various types of DDoS attacks. The works of [22] and [13] took a broad perspective in discussing DoS and DDoS attacks in IPv6. As early as 2007, Yang et al., in [15] discussed the vulnerabilities of NDP and the possibility of utilizing this protocol in launching DDoS attacks. This research explained that NDP could be used in three ways to perform DoS attacks, namely duplicate address detection, neighbor unreachability detection failure, and neighbor solicitation/advertisement spoofing. It further discussed solutions for other types of attacks using IPSec, however it did not elaborate on counter measure strategies for NDP attacks. A proposal for a framework for detecting DoS and DDoS in IPv6 was published in 2012, by Gao and Chen [25], mainly focusing on TCP-flood, UDP-flood, and ICMP-flood attacks detection.

Several publications surveyed security threats in IPv6, and in particularly, focusing on NDP, ICMPv6, and DDoS. Elejila et al., in 2016 examined intrusion detection systems designed to thwart ICMPv6-based DDoS attacks. [22]. In this work, they considered existing vulnerabilities in ICMPv6 and how these could be exploited to launch DDoS attacks. They proposed IDSs tailor made for IPv6, or extended from IPv4 IDS and evaluated these systems in terms of accuracy in detecting attacks.

Satrya and Yulianto proposed a method based on hybrid analysis to detect DDoS flooding attacks in IPv6 networks [26], The work introduced a prototype for detection of DDoS flooding attacks using source addresses, analytical methods and analysis of network flow. The proposed system reported 85% accuracy, and an average attack detection time of 2 minutes and 56 seconds. The research in [27] was in-line with the current issues facing IPv6 architecture

design. The authors developed a framework for selecting important features for detection of attacks in IPv6 networks. This made use of Particle Swarm Optimization (PSO) techniques for feature selection on the original dataset, which resulted in the selection of 5 features, referred to as ProFeat (2013), and made use of Support Vector Machines (SVM) for detection of DDoS attacks. The proposed model yielded an average accuracy of 99.95% over 3 datasets.

Another survey, published in 2016 by Anbar et al. in [23], reviewed the security vulnerabilities related to IPv6 NDP. The review also refers to some detection and prevention mechanisms of common attacks exploiting IPv6 NDP.

Zhang and Wang, in [28], discussed a wider range of attacks that are possible through IPv6 NDP. Their research explored state-of-the-art solutions, including SEND, use of IPSec AH and MAC in NDP. The work in [29] presented the prevention of DDoS attack in a cloud infrastructure using machine learning algorithms. The authors proposed an automated IDS, which focused on resource utilization instead of the usual packet monitoring, using Artificial Neural Networks to detect attacks and store the results in a database to be used for future referencing.

The v6IIDS framework was proposed in [12], which discusses– an intelligent ICMPv6 DDoS Flooding-attack Detection system. This was developed using the back-propagation algorithm for detection of ICMPv6 DDoS attacks, following feature selection using Principal Component Analysis (PCA) on the original dataset. Furthermore, the dataset features were first ranked using IGR in order to identify important and effective features. In summary, the efficiency of v6IIDS was evaluated and was shown to have accuracy of 98.3% using real datasets from the Nav6 laboratory

Ahmed et al., published a more recent survey in 2017. This work reviewed the NDP protocol specifications, and provided a detailed account of potential threats, and various countermeasures [24].

The work in [5] focused on developing a DDoS attack detection system from sources in the cloud . In their contribution, it was established that DoS attacks are often executed from virtual machines in the cloud and that most of the previous research efforts focused on analysing network traffic on the victim's side instead. Leveraging statistical information sourced from both virtual machines and the cloud server's hypervisor, the authors comparatively analysed the performance of nine machine learning algorithms, including K-means, Logistic Regression, Expectation Maximization, Support Vector Machines, Naïve Bayes, Decision Tree, and Random Forest. Their results showed an accuracy of 99.7% in detecting four types of DoS attacks.

An efficient statistical approach was used to detect Distributed Denial of Service attacks using features extracted from traffic and a dynamic threshold detection algorithm [30]. In essence, the detection mechanism operates by calculating four attributes based on the characteristics of DDoS within a specified time interval and when the computed value is higher than a present threshold, network traffic is flagged as an attack. The validation of the algorithm was carried out using two DARPA datasets and a simulated dataset. Overall, the proposed model achieved an accuracy of 99.5%, a true positive ratio (TPR) of 98% and a true negative ratio (TNR) of 99.6% on the DARPA 98 dataset, and 99.5% accuracy, 99.5% TPR and 99.5% TNR on the DARPA 2000 dataset. Using the synthetic dataset, the algorithm produced an overall accuracy, TPR and TNR of 99.5%, respectively.

Hybrid artificial intelligence detection models for covert channel attacks in IPv6 were implemented by Saleh et al., [31].

The hybridized framework used Fuzzy Logic and Genetic Algorithms, resulting to a robust model that optimally generated rules capable of detecting covert channels. The research resulted in the development of a detection model with a 97.7% detection rate, superior to state-of-the-art techniques.

In a closer relation to this work, a research carried out by Elejla et al., and published in 2018, used flow-based representation of packets [17]. The first similarity to our research is the nature of the dataset. Specifically, the authors used a flow-based representation of traffic and a new set of features for attack detection, which represent flow features and also differentiate between normal and ICPMv6-based DDoS attacks. The researchers utilised seven machine-learning algorithms for IDS model development. The models were developed using cross-validation (CV) and a test set and were evaluated using four metrics, namely True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). For CV, two of the models had similar accuracies of 85.67%, while for the test set (a simulation of a real-time scenarios), Random Forest produced the highest accuracy result of 85.83%. The work evaluated the proposed models against Packet-based IDSs and proved to be more effective, since the best Packet-based IDS model achieved an accuracy of 61.28% in comparison.

Another recent research work focused on the detection of low-rate DDoS attacks [10]. It revealed various statistics relating to DDoS attacks, including SYN, HTTP, ICMP, UDP, and TCP flood types. The contribution proposed a novel method for low-rate DDoS attack detection using a Long Short-Term Memory (LSTM) classifier.

Finally, HTTP flood attack detection in the application layer of a network is the aim of the research work conducted in [32]. The objective of this research was fast and early detection of DDoS, and proposed a Bio-Inspired IDS that detects DDoS anomalies in a network by focusing on the application layer of the network. Benchmarking on the CAIDA dataset, the work used the bat algorithm for early detection of App-DDoS, and reported an accuracy of 94.8%.

IV. RESEARCH METHODOLOGY

This research proposes an IDS model for detection of NDP DDoS attacks. This is intended as an end-to-end architecture, which analyzes network traffic and detects the occurrence of an attack. A feature extraction method for high dimensional data is used to appropriate encapsulate the characteristics of network traffic. Next, locally weighted learning, discussed in sub-section D, is used to classify NDP messages. The work of Elejla et al., [17] is used as a benchmark, by adopting their flow-based data representation, model development method and evaluation metrics.

A. Proposed Framework

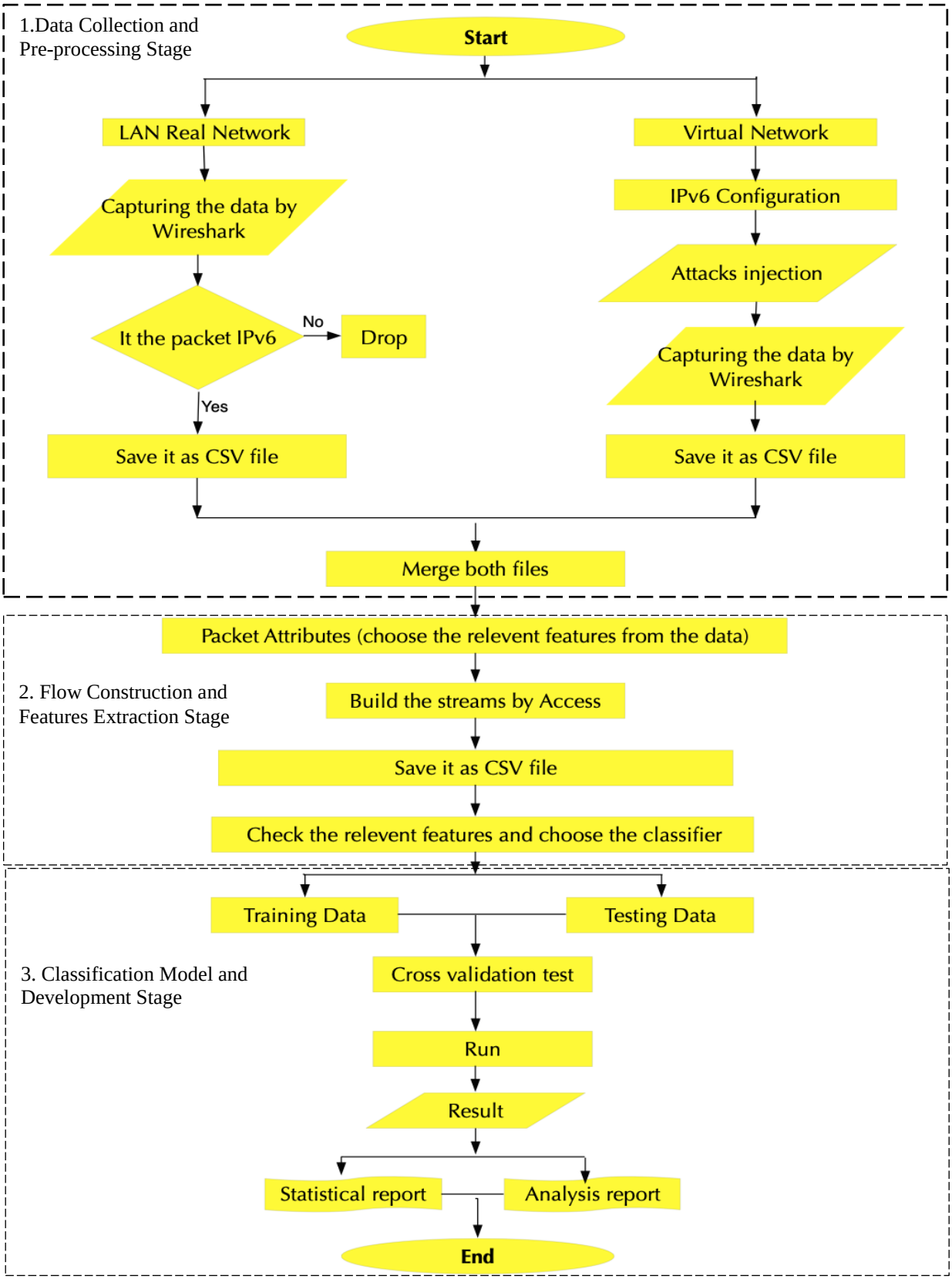

Figure 4: Flowchart of the proposed model.

As can be seen in Figure 4, the framework of the proposed IDS is structured into three main components: data collection and pre-processing flow construction and feature extraction, mining and classification. These components consist of several tasks that are completed in stages, and explained in the subsequent sub-sections.

B. Data Collection and Pre-processing stage

In order to generate the dataset, a real network was used to capture the traffic using the Wireshark tool [35]. Furthermore, a virtual test bed network was created using the Graphical Network Simulator-3 (GNS3) [36] and attached to the real network. The attacks were not performed on the real network due to the potential damage. Instead, the DDoS attacks were simulated on the virtual network (by disconnecting it from the real network), which has the same settings and configuration as the real network. Both malicious and normal traffic instances were captured to develop the required training, validation and testing datasets.
Let x represents a set of N number of packets P, and X_{IPv6} is the set of IPV6 packets. Algorithm 1 depicts the data capturing process.

Algorithm 1: Data capturing from LAN.
Read Data:
Let X be a set of packets P
$X = \{P_1, P_2, P_3, \ldots\ldots\ldots, P_N\}$
Let $P_N \in X_{IPv6}$ Where $X_{IPv6} \subset X$
Put $X_{IPv6} \rightarrow CSV_1$
where CSV_1 is a Comma Separated Value file for real data

Various software tools were used to build the virtual testbed. GNS3 version 1.3.2 was used to build the virtual test bed network and connect it to the real network. Oracle's virtual machine was used to install different types of OS. Different operating systems were considered for the purposes of, e.g., generating normal traffic (e.g., Windows XP SP2), performing attacks (e.g., Kali) or collecting traffic (e.g., Windows 7 SP1). The CISCO's router and switch were used to connect the OSes together and attach the virtual network to the real network. THC toolkit and SI6 tools were used to perform the attacks. Wireshark was used to collect traffic (normal and malicious). Algorithm 2 shows the virtual data attacks creation process.

Algorithm 2: Virtual data.
Let V represent a virtual network
Configure $V \Rightarrow V \in V_{IPv6}$
where V_{IPv6} is IPV6 virtual network
Let $t \in T$
where T represents a set of attacks,
$V_{Attack} = \{V/V \in V_{IPv6} \mid t \in T\}$
Put $V_{Attack} \rightarrow CSV_2$
where CSV_2 is a Comma Separated Value file for virtual data

C. Flow Construction and Feature Extraction Stage

By combining the traffic collected from the real and virtual networks, a dataset of IPv6 traffic was developed in Comma Separated Value (CSV) format. The traffic packets were filtered to contain instances of NDP traffic as depicted in Figure 5, which are the target of this research. The packets were filtered based on the type of IP (i.e. either version 4 or 6), followed by the ICMPv6Type (i.e. if it is Neighbour Advertisement (NA), Neighbour Solicitation (NS), Router Advertisement (RA), or Router Solicitation (RS)). Next, the CSV file was imported in Database and traffic packets were labelled into normal or attack, using prior information about attack injection.

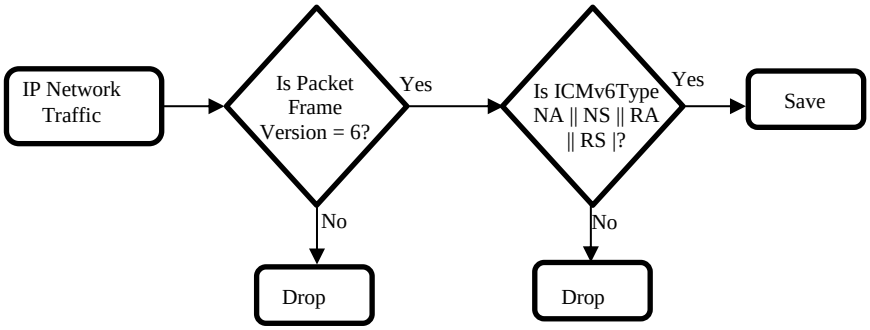

Figure 5: Flowchart of ICMPv6 packets filtration

The stages of feature extraction, malicious streams labelling and development of IDS using machine learning algorithms are presented in Algorithm 3.

Algorithm 3: Proposed Data analysis

Let $X \in \{CSV_1 + CSV_2\}$

Let F be a set of features

{ICMP_Type, Packet_Number, Byte_Number, Destination, SAT, Byte_Ratio}

Find $X_f \in F$

where X_f is a set of all extracted features from X.

Let R = {0, 1} where 0= Normal, 1= Attack

$\forall\ r \in R$, if $r = 0 \Rightarrow$ destination $\in \{destination_x$ to $destination_y\}$

If $r = 1 \Rightarrow$ destination $\in \{ destination_{fake} = destination_x$ to $destination_y$ & $Time_{fake} = Time_x$ to $Time_y\}$

Let ML be a set of machine learning algorithms

Pass Xf to $ML_{algorithm} \Leftrightarrow$

ML= 0, where X_f | ML= 1 where X_f

The required packet attributes were extracted from each packet in order to prepare it for the feature construction and streams building phase. These attributes are IPv6Src address, IPv6Dst address, NDPType, packet time, packet Length, packet Traffic_Class, packet hop_limit, packet Flow_Label, packet Next_Header, packet Checksum, packet Payload and Length. These attributes were used to finally create stream features such as IPv6Src addresss, IPv6Dst address, NDPType as well as packet features such as Flow_label, Packet Next_Header etc.

The next phase is building the streams based on their definition and extract the features that are mentioned previously. This phase is performed using MYSQL queries, applied on the traffic to convert the packets into streams with several features for each stream:

$$S6_{NDP} = (\text{IPv6Dst}, \text{NDP type}, T)$$

The dataset used for model development contains 16 features excluding the class labels (normal or attack). These include Source, Destination, ICMPv6Type, SAT, SH, ByteNumber, PacketsNumber, L_Diversity, TC_Diversity, and HL_Diversity, among others.

D. NDP-Based IDS Model using Machine Learning

The proposed intrusion detection model aims to detect ICMPv6-based DDoS attacks. A novel method was introduced in this study's model development, which makes use of locally weighted learning techniques. Specifically, the LWL method was adopted to build fast and efficient models using Bayesian Networks (BN), Decision Tree (DT), and Naïve Bayes (NB) algorithms, as the base learners. According to [33], LWL is a method that enables base algorithms to provide a local function for the prediction of a current point of interest, using a subset of the training data, instead of a global function that makes use of the complete training data. It is often referred to as lazy learning, as it delays the processing of training data until a query point requires a prediction.

LWL offers advantages in terms of the range of parameter values. In many cases, a global method cannot produce a parameter value that could provide a sufficiently good approximation. This is resolved in LWL methods that are typically non-parametric and execute current predictions based on a subset of data.

Furthermore, LWL method is most effective in reducing computation costs [33], since it utilises data points that are close to the query point which are assigned higher weights than those far away in the prediction task. An excellent overview of the LWL can be found in [33], [37].

Bayesian Networks are said to be a probabilistic graphical model that represents some set of variables and conditional dependencies inherent in them using a direct acyclic graph (DAG) [34]. Nodes represent random variables, hypotheses, or latent variables, while the edges between nodes depict the probabilistic dependencies among the random variables [38].

Decision Tree (DT) is a classification algorithm that uses inductive learning on a labelled training set to construct a model. It recursively divides data items into various classes using the information extracted from the training set attributes. This recursive process is terminated when all the data items in the current subsets are of the same class. The node of the tree specifies the attribute used for data item partitioning, while the edges between two nodes or a node and a leaf depict the possible values of an attribute.

Naïve Bayes make use of Bayes theorem of calculating the posterior probability based on prior probability, the probability of the observations and the likelihood that the observational data fits the hypothesis. Naïve Bayes is an intuitive and simple method based on Bayes' rule of conditional probability.

The algorithms were chosen based on their ability to support interpretation and easy to understand models. Moreover, the type of learning that each of the algorithms employs is different, thus, their combination could potentially result to a robust classification model. Prior to this research work, locally weighted method of learning had not been implemented for developing IDS capable of detecting ICMPv6-based DDoS and Replayed type of attacks.

V. RESULTS

In this section, the performance of the proposed flow-based ICMPv6 IDS model is presented.

The main goal is the development of an intrusion detection model that has strong predictive capabilities, does not suffer from overfitting, has low computational costs, and requires less time for model development and attack detection.

The representation of network traffic is usually packet-based, however, in light of recent research, a flow-based method of representation was implemented, which is found to provide improved patterns for detection of ICMPv6 attacks. Using flow-based data led to the engineering of considerable number of features. However, in this study, the selected features used to develop the models were handpicked based on domain expertise, resulting in the reduction of the number of features, by discarding those that are redundant or irrelevant and also those that may cause overfitting after running different experiments using the features, leading to lower computation costs and less time for model development.

Consequently, IDS models for ICMPv6 attacks are developed with the data serving as input for locally weighted learning of the classifiers used as base learners. As previously stated, these are the Bayesian Network (BN), Decision Tree (DT), and Naïve Bayes (NB).

The performance of the models was evaluated using a variety of metrics including accuracy, true positive ratio, true negative ratio, kappa statistics (a metric scored between 0 – 1, and resulting value greater than 0 indicate that the classifier is doing better than chance), f-measure, as well as the time taken to build the model. Model training was conducted using 10-fold cross validation. The Waikato Environment for Knowledge Analysis (WEKA) software was also used for generating advanced analyses.

The dataset used in this research was network traffic represented using a flow-based method. It consists of seven features including the class attributes, namely: ICMPv6Type, PacketsNumber, Destination, ByteNumber, SAT, BytesRatio, and MClass. The class attribute of MClass has three labels, i.e., normal, replayed and DDoS. There are a total of 123,436 instances, of which normal traffic accounts for 74,517 instances, with 238 instances for replayed traffic and 48,681 instances for DDoS traffic.

A. Locally Weighted Learning Model Results and Discussion

The Locally Weighted Learning algorithm was deployed using linear weighting kernels and 3 neighbours with the aforementioned classifiers serving as the base learners.

BN was used as the first base learner for the LWL algorithm in order to learn from the flow-based dataset and build its IDS model. The developed model achieved a very high overall accuracy of 96.48%. In total, the model correctly classified 119,098 of 123,440 instances and resulted into 4,342 incorrectly classified instances.

Table 1 depicts the evaluation metrics and the corresponding values for the LWL-BN model.

Model / Metrics	LWL – BN
Detection Rate	96.4825%
True Positive Ratio	0.9421
False Positive Ratio	0.0004
True Negative Ratio	0.999
False Negative Ratio	0.0579
Kappa Statistic	0.9282
F-Measure	0.967
Time Taken	0.2 sec

Table 1: Performance evaluation of LWL-BN model.

Table 1 shows that the LWL-BN model performs well on the dataset with a detection rate of approximately 97% and a kappa value of 0.9282, emphasising the strong predictive capability of the model.
A further probe into the results of the model performance can be used to discover the predictive strength of each class value. This is achieved by inspecting the values of the confusion matrix as depicted in Table 2. Table 3 shows the TPR, FPR and Sensitivity values for the replayed and DDoS labels.

Classes	Normal	Replayed	DDoS
Normal	70204	785	3582
Replayed	18	217	3
DDoS	5	0	48676

Table 2: Confusion matrix of the LWL-BN model

Attack Labels / Metrics	Replayed	DDoS
True Positive Ratio	0.912	0.9998

False Positive Ratio	0.006	0.047
Sensitivity	0.912	0.9998

Table 3: Performance evaluation of the LWL-BN model for the attack classes

From Table 2, it can be seen that the model achieved a very high detection rate of 99.98% for DDoS attacks by correctly classifying 48,676 instances, 94.2% detection rate for normal traffic and 91.17% for detecting the replayed type of attack. More so, from Table 3, the TP, FP and sensitivity values for the replayed and DDoS classes further demonstrate the detection capability of the LWL-BN model.

The Decision Tree algorithm was also used as a base learner for the LWL algorithm to learn from the flow-based dataset and build its IDS model. An evaluation of the performance model was then conducted. The Decision Tree model achieved a high overall accuracy of 93%, having correctly classified 114,868 of 123,440 instances but had a significant number of incorrectly classified instances, i.e., 8572 patterns.
Table 4 depicts the evaluation metrics and the corresponding values for the LWL-DT's model.

Model / **Metrics**	**LWL – DT**
Detection Rate	93.0557%
True Positive Ratio	0.885
False Positive Ratio	0.013
True Negative Ratio	0.999
False Negative Ratio	0.1145
Kappa Statistic	0.8614
F-Measure	0.937
Time Taken	0.2 sec

Table 4: Performance evaluation of the LWL-DT model.

It is evident from Table 4 that the LWL-DT model also fits the dataset well, with a detection rate of 93%, having a true positive ratio of 0.931 for the normal class and a kappa value of 0.8614, emphasising the strong predictive capability of the model. Further investigation into the result of this model's performance is used to discover the predictive strength for each class. This is achieved by considering the values of the confusion matrix of this model as depicted in Table 5 and also the TP, FP and Sensitivity for each attack are seen in Table 6.

Classes	**Normal**	**Replayed**	**DDoS**
Normal	65982	784	7751
Replayed	24	208	6
DDoS	611	60	48005

Table 5: Confusion matrix of the LWL-DT model

Attack Labels / **Metrics**	**Replayed**	**DDoS**
True Positive Ratio	0.874	0.986
False Positive Ratio	0.007	0.104

Sensitivity	0.874	0.986

Table 6: Performance evaluation of the LWL-DT model for the attack classes

As seen in Table 4, the confusion matrix reveals the predictive accuracies for each class. This model has 98.61% detection rate in DDoS attack types, 87.39% for detecting replayed attack types and 88.5% detection rate for normal packets. In Table 6, the LWL-DT model has a low false positive ratio of 0.007 for replayed attacks, and a 0.104 false positive ratio for DDoS attacks.

Lastly, the Naïve Bayes was used as a base learner for the LWL algorithm. The performance of the model was evaluated using the same procedure as the previous two base learners. The developed model achieved an overall detection rate of 96.024%. Table 7 depicts the evaluation metrics and their corresponding values for the LWL-NB model.

Model / Metrics	**LWL – NB**
Detection Rate	96.024
True Positive Ratio	0.935
False Positive Ratio	0.01
True Negative Ratio	0.99
False Negative Ratio	0.065
Kappa Statistic	0.9193
F-Measure	0.964
Time Taken	0.1 sec

Table 7: Performance evaluation of the LWL-NB model.

Table 7 shows that the LWL-NB model learnt the dataset successfully, having a 96% detection rate, with 0.01 false positive ratio, and a kappa value of 0.9193, which also indicates the strong predictive abilities of the developed model. A further examination into the results of this model's performance reveals the predictive strength for each class value in order to ascertain how best the model detects a certain type of attack. This is achieved by considering the values of the confusion matrix, as revealed in Table 8. Table 9 provides insight into the performance of the model with respect to each attack type.

Classes	**Normal**	**Replayed**	**DDoS**
Normal	69648	1341	3528
Replayed	29	206	3
DDoS	443	5	48233

Table 8: Confusion matrix of the LWL-NB model.

Attack Labels / Metrics	**Replayed**	**DDoS**
True Positive Ratio	0.87	0.985
False Positive Ratio	0.011	0.047
Sensitivity	0.87	0.985

Table 9: Performance evaluation of the LWL-NB model.

Table 8 depicts the confusion matrix of the LWL-NB's model. This model achieved a high detection rate of 99.08% in detecting DDoS attacks, 93.47% accuracy in correctly classifying normal traffic and 86.55% in correctly classifying the 'replayed" type of attack.

B. Comparative Analysis of the Models

Table 10 depicts the detection rate, true positive ratio and false positive ratio of all developed models.

Models / Metrics	LWL – BN	LWL – DT	LWL – NB
Detection Rate	96.4825	93.0557	96.024
True Positive Ratio	0.9421	0.885	0.935
False Positive Ratio	0.0004	0.013	0.01

Table 10: Comparative analysis of each model's overall performances.

The LWL-BN model achieved the highest detection rate of 96.48% with the lowest false positive value of 0.0004%. The LWL-NB model is the next best model, with an accuracy rate of 96.024% and a false positive value of 001, while the LWL-DT model had the lowest overall detection rate of 93%. The LWL-DT model also had the highest false positive ratio of 0.013%.

A closely related work of [17], reported an overall accuracy of 85.3% at its best, having also used flow-based representation of packets but different sets of features and dataset.

Comparatively, all developed IDSs are capable of detecting DDoS and replayed attacks based on NDP-based network traffic as well in the detection of anomalies. They all demonstrated strong predictive abilities, however, the LWL-BN model proved to have the most favourable overall performance.

Taking into further consideration the performance of the models for each class of attack, the accuracy of each of the models for each attack type is considered in Table 11.

Model	Class Accuracy (%)		
	Normal	**Replayed**	**DDoS**
LWL-BN	94.21	91.17	99.98
LWL-DT	88.56	87.39	98.62
LWL-NB	93.47	86.55	99.08

Table 11: Comparative accuracies per attack type.

As depicted in Table 11, and also considering the DDoS attack type, all models are capable of detecting this attack with a very high rate, i.e., a minimum of 98.62%. As such, any of the models can be deployed for the purpose of detecting a DDoS attack.

Furthermore, the accuracy of the model in classifying normal packets varied with LWL-DT achieving the lowest detection rate of 88.56%, while the LWL-NB could identify a normal packet better than the LWL-DT with a detection rate of 93.47%. In this category of attack, the LWL-BN model triumphed with the highest detection rate of 94.21%.

Also, the LWL-BN model outclassed other models in the detection of replayed attacks, with an accuracy of 91.17%. The LWL-DT model outperformed the LWL-NB model with a detection rate of 87.39%, while LWL-NB had the lowest rate of 86.55%.

A pictorial representation of all model accuracy per class is depicted in Figure 9.

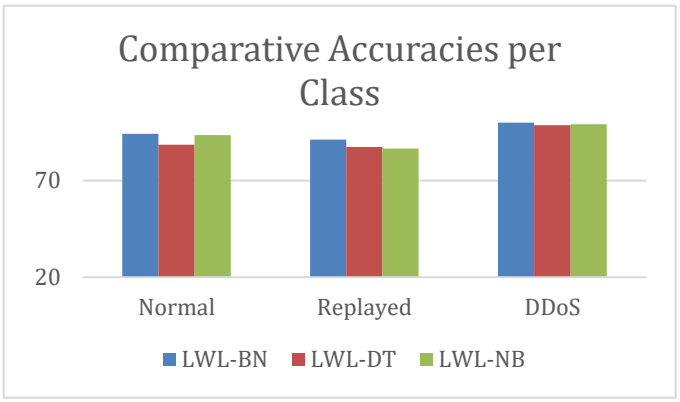

Figure 9: Comparative model accuracy per class

In conclusion, it is safe to infer that all deployed models are capable of classifying a DDoS attack, however, the LWL-BN model is more effective in detecting all attack types, followed by the LWL-NB model.

VI. CONCLUSION

Research into Intrusion Detection System for detecting DDoS attacks in ICMPv6 is in its infancy, compared with IPv4. Machine learning algorithms have been widely applied in developing such kinds of IDS. This paper presented IDS for detection of DDoS and replayed attacks, while successfully identifying normal packets in NDP type network traffic by making use of Locally Weighted Learning techniques. Three IDS models were proposed, namely, LWL-BN, LWL-DT, and LWL-NB, using the Bayesian Network, Decision Tree, Naïve Bayes, respectively, as base learners. Overall, the LWL-BN IDS surpassed the other IDS after evaluations were conducted using both the overall accuracy and class-type accuracies among the evaluation metrics. With respect to the developed IDS presented in this contribution for NDP, on average, it is capable of detecting DDoS attacks with 99% accuracy and replayed attacks with 91.17% accuracy, while normal traffic is detected with 94% accuracy, outperforming previous results of 85.3% accuracy [17], which have been used as our benchmark. Further research is needed in order to uncover better ways to achieve the same accuracy yielded by this flow-based IDS, using various techniques such as hybridising algorithms, ensemble methods, and alternative methods for feature selection. In addition, the real-time implementation and evaluation of the IDS presented in this study is being considered as imminent future work.

REFERENCES

[1] T. A. Duarte, “IPv4 to IPv6 Transition: Security Challenges,” 2013.

[2] B. Stockebrand, *IPv6 in practice: a Unixer’s guide to the next generation Internet. Springer Science & Business Media.* 2006.

[3] F. Beck *et al.*, “Automatic IPv4 to IPv6 Transition D1.2 - Network representation and pre-requisites,” *HAL Id inria-00407632*, 2009.

[4] T. Mahjabin, Y. Xiao, G. Sun, and W. Jiang, “A survey of distributed denial-of-service attack, prevention, and mitigation techniques,” *Int. J. Distrib. Sens. Networks*, vol. 13, no. 12, 2017.

[5] Z. He, T. Zhang, and R. B. Lee, “Machine Learning Based DDoS Attack Detection from Source Side in Cloud,” *2017 IEEE 4th Int. Conf. Cyber Secur. Cloud Comput.*, pp. 114–120, 2017.

[6] A. R. GAWANDE, “DDoS Detection and Mitigation using Machine Learning,” 2018.

[7] R. M. A. Saad, S. Ramadass, and S. Manickam, “A Study on Detecting ICMPv6 Flooding Attack based on IDS,” *Aust. J. Basic Appl. Sci.*, vol. 7, no. 2, pp. 175–181, 2013.

[8] M. Zekri, S. El Kafhali, N. Aboutabit, and Y. Saadi, “DDoS attack detection using machine learning techniques in cloud computing environments,” *2017 3rd Int. Conf. Cloud Comput. Technol. Appl.*, no. February 2018, pp. 1–7, 2017.

[9] A. Salih, X. Ma, and E. Peytchev, “Detection and Classification of Covert Channels in IPv6 Using Enhanced

Machine Learning," 2015.
[10] A. A. Chistokhodova and I. D. Sidorov, "Novel Method For Low-Rate Ddos Attack Detection," *Int. Conf. Inf. Technol. Bus. Ind.*, pp. 1–6, 2018.
[11] O. Elejla, B. Belaton, M. Anbar, and A. Alnajjar, "A Reference Dataset for ICMPv6 Flooding Attacks," *J. Eng. Appl. Sci.*, vol. 11, no. 3, pp. 476–481, 2016.
[12] R. M. A. Saad, M. Anbar, S. Manickam, and E. Alomari, "An intelligent ICMPv6 DDoS flooding-attack detection framework (V6IIDS) using back-propagation neural network," *IETE Tech. Rev. (Institution Electron. Telecommun. Eng. India)*, vol. 33, no. 3, pp. 244–255, 2016.
[13] M. Šimon, L. Huraj, and M. Hosťovecký, "IPv6 Network DDoS Attack with P2P Grid," *Creat. Intelligent, Technol. Data Sci.*, no. September, pp. 407–415, 2015.
[14] [O. E. Elejla, M. Anbar, and B. Belaton, "ICMPv6-Based DoS and DDoS Attacks and Defense Mechanisms: Review," *IETE Tech. Rev. (Institution Electron. Telecommun. Eng. India)*, vol. 34, no. 4, pp. 390–407, 2016.
[15] X. Yang, T. Ma, and Y. Shi, "Typical DoS / DDoS threats under IPv6," *2007 Int. Multi-Conference Comput. Glob. Inf. Technol. (ICCGI'07), IEEE.*, vol. 55, no. 55, pp. 1–6, 2007.
[16] Charles M. Kozierok. 2017, TCP/IP Guide available at: http://www.tcpipguide.com/
[17] O. E. Elejla, M. Anbar, B. Belaton, and B. O. Alijla, "Flow-based IDS for ICMPv6-based DDoS Attacks Detection," *Arab. J. Sci. Eng.*, no. March, 2018.
[18] S. Praptodiyono, I. Hasbullah, R. Murugesan, and S. Ramadass, "Survey of Internet Protocol Version 6 Link Local Communication Security Vulnerability and Mitigation Methods," *IETE Tech. Rev.*, vol. 30, no. 1, p. 64, 2013.
[19] A. O. Balogun, A. M. Balogun, V. E. Adeyemo, and P. O. Sadiku, "A Network Intrusion Detection System : Enhanced Classification via Clustering," *Comput. Inf. Syst. Dev. Informatics Allied Res. J.*, vol. 6, no. 4, pp. 53–58, 2015.
[20] [M. Moradi and M. Zulkernine, "A neural network based system for intrusion detection and classification of attacks," *Proc. 2004 IEEE Int. Conf. Adv. Intell. Syst. Appl.*, 2004.
[21] Y. Sun, C. Zhang, S. Meng, and K. Lu, "Modified Deterministic Packet Marking for DDoS Attack Traceback in IPv6 Network," 2011.
[22] O. E. Elejla, B. Belaton, M. Anbar, and A. Alnajjar, "Intrusion Detection Systems of ICMPv6-based DDoS attacks," *Neural Comput. Appl.*, no. December, pp. 1–12, 2016.
[23] M. Anbar, R. Abdullah, and R. M. A. Saad, "Review of Security Vulnerabilities in the IPv6 Neighbor Discovery Protocol Review of Security Vulnerabilities in the IPv6 Neighbor Discovery Protocol," *Inf. Sci. Appl.*, no. June, pp. 603–612, 2016.
[24] A. S. A. M. S. Ahmed, R. Hassan, and N. E. Othman, "IPv6 Neighbor Discovery Protocol Specifications , Threats and Countermeasures : A Survey," *IEEE Access*, vol. 5, no. January 2018, pp. 18187–18210, 2017.
[25] J. Gao and Y. Chen, "Detecting DOS / DDOS Attacks Under Ipv6," *Proc. 2012 Int. Conf. Cybern. Informatics. Springer, New York, NY.*, pp. 847–855, 2014.
[26] G. B. Satrya and F. A. Yulianto, "The Detection of DDOS Flooding Attack using Hybrid Analysis in IPv6 Networks," *3rd Int. Conf. Inf. Commun. Technol. (ICoICT). IEEE*, no. May, pp. 1–15, 2015.
[27] M. Zulkiflee, M. S. Azmi, S. S. S. Ahmad, S. Sahib, and M. K. A. Ghani, "A Framework of Features Selection for IPv6 Network Attacks Detection," vol. 14, no. January 2015, pp. 399–408, 2015.
[28] T. Zhang and Z. Wang, "Research on IPv6 Neighbor Discovery Protocol (NDP) Security J) Neighbor discovery message Neighbor discovery options," *2nd IEEE Int. Conf. Comput. Commun. (ICCC), IEEE.*, pp. 2032–2035, 2016.
[29] A. Jaiswal, P. C. Murthy, and M. BR, "Prevent DDOS Attack in Cloud Using Machine Learning," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 6, no. 6, pp. 575–579, 2016.
[30] S. Indraneel, V. Praveen, and K. Vuppala, "HTTP Flood attack Detection in Application Layer using Machine learning metrics and Bio inspired Bat algorithm," *Appl. Comput. Informatics*, 2017.
[31] A. Salih, X. Ma, and E. Peytchev, "Implementation of Hybrid Artificial Intelligence Technique to Detect Covert Channels Attack in New Generation Internet Protocol IPv6," pp. 173–190, 2017.
[32] J. David and C. Thomas, "Efficient DDoS Flood Attack Detection us- ing Dynamic Thresholding on Flow-Based Network Traffic," *Comput. Secur.*, 2019.
[33] P. Englert, "Locally Weighted Learning," Semin. Cl. Auton. Learn. Syst., vol. 1, no. 1, pp. 1–9, 2012.
[34] F. Faltin and R. Kenett, "Bayesian Networks," 2007.
[35] Wireshark User's Guide :for Wireshark 1.7 by Ulf Lamping, Richard Sharpe, Ed.Warnicke Copyright © 2017; Source: https://www.wireshark.org/about.html.
[36] GNS3. 2017, source: https://www.gns3.com.
[37] G. Atkeson, S. A. Schaal and Andrew W, Moore, Locally Weighted Learning, AI Review, 11, pp. 75-113, 1997.
[38] MacDermott, A., Baker, T. and Shi, Q., February. Iot forensics: Challenges for the ioa era. In 2018 9th IFIP

International Conference on New Technologies, Mobility and Security (NTMS) 2018 (pp. 1-5). IEEE.
[39] Karam Y, Baker T, Taleb-Bendiab A. Security support for intention driven elastic cloud computing. In 2012 Sixth UKSim/AMSS European Symposium on Computer Modeling and Simulation 2012 Nov 14 (pp. 67-73). IEEE.