A Framework of Features Selection for IPv6 Network Attacks Detection

ZULKIFLEE M., MS AZMI, SSS AHMAD, S SAHIB, MKA GHANI
Faculty of Information and Communication Technology (FTMK)
Universiti Teknikal Malaysia Melaka (UTeM)
Durian Tunggal, Melaka
MALAYSIA
{zulkiflee, sanusi, sakinah, shahrinsahib, khanapi}@utem.edu.my

Abstract: - IPv6 technology introduced to solve problems in the previous IPv4 technology. New threats were discovered due to the exploitation of flaws in IPv6 architecture design. There is a demand to produce new intrusion detection technique for new threats of IPv6 network environment. In this paper, the method of features selection to produce the most significant feature is presented. The objective of this paper is to propose a framework to solve feature selection problem which in this paper the features of IPv6 packet will be the case. The fundamental method of feature selection was improvised to suit with this scenario. A data of IPv6 network attacks was produced by using an IPv6 testbed environment. The SVM and PSO were used in the process of determine the best features to detect IPv6 attacks. In the future, this framework can be applied in other domains which require features selection solution.

Key-Words: - : IPv6, IDS, Feature Selection, SVM, PSO

1. Introduction

IPv6 has been invented in year 1998 well-defined in RFC 2460. Nowadays, the number of IPv6 users has gradually increased. This is due to high demand of new IP addresses allocation which IPv4 cannot offer anymore. Users are still craving for unique IP addresses to be assigned to their nodes. The emergence of new technologies such as Internet of Things (IOT), cloud computing and wireless technology applications because the need of IP addresses are becoming more severe. Theoretically, IPv6 protocol is much better than IPv4 in several aspects [1, 2]. Hence, the migration in IPv6 network environment is inevitable eventually.

The implement of IPv6 is not a panacea for IPv6 security issues. The IPv6 security has becoming a major concern of the slow implementation of IPv6 [3-5]. The implement of IDS is an option to ease the security issues. IDS technology was adopted from the IDS in IPv4 environment. However, the detection techniques invented from IPv4 network environment are needed to be verified before being adapted in IPv6 network environment. Zagar [6] claimed that the IDS research in IPv6 is still in infant stage while it is have matured in IPv4 network environment. While, Hansman [7] emphasized the previous detection techniques produced were constructed by using an obsolete dataset. Therefore, further research is needed to identify the suitability of IPv4 detection mechanisms in IPv6 network environment [8-10].

A lack of IPv6 dataset has led slow blooming research in IPv6 security domain. A lot of contributing factors need to be considered to ensure the dataset produced is reliable. Currently, detection techniques produced for IPv6 network environment were not based on the same source [11-13]. Therefore, the results obtained by different researchers cannot be compared as the result gained from a test on different dataset. Thus, a generic dataset is needed to measure the previous techniques to identify its performance in more global perspective. The generated dataset will be the platform to compare different features performance on distinguishing between normal and attack packets.

In this paper, a framework of features selection method was proposed as the main objective. The methodology of constructing the best features to distinguish between normal and attack packets will be explicitly explained. The proposed method will be tested on a generic dataset produced by IPv6 testbed environment. In the following section, several related studies will be discussed. Next, the methodology of selecting features will be elaborated. Then, the process continues with an implementation of the selecting the best features to differentiate between normal and attack packets. Afterward, extensive discussion on the proposed

features selection method. Finally, the conclusion will be deliberated towards the end of this paper.

2. Related Work

2.1. IDS Framework

Based on several studies conducted by previous researchers [14, 15], multiple IDS frameworks have been proposed. These frameworks are supposed to lay out what an IDS is all about. The IDS framework presented in this work consists of the main components that characterize a system as an IDS. The frameworks from previous researchers have been analysed and synthesized. The following figure depicts the general IDS framework based on an analysis of the findings.

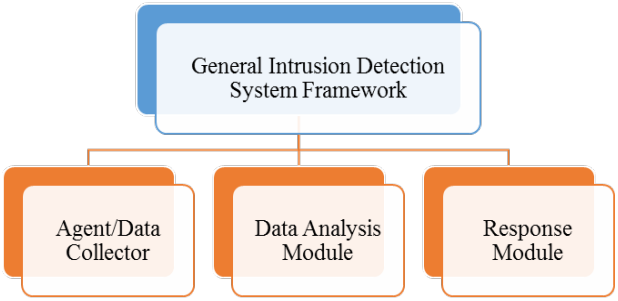

Fig. 1: General IDS Framework

Fig. 1 illustrates the general IDS framework. It shows that the framework consists of three main components, namely, the *Agent/Data Collector*, the *Data Analysis Module*, and the *Response Module*. The *Agent/Data Collector* is the module responsible for collecting the data that tends to be used for analysis purposes. The input varies and encompasses event logs, traffic logs, and system logs. The second module is the *Data Analysis Module*, which is responsible for analyzing the gathered data in order to detect whether any abnormal activities have occurred within the monitored boundary. If any abnormal activity detected, the system will raise an alarm for further action. Finally, the *Response Module* provides an appropriate action whenever an alarm is triggered. The response includes sending alerts to an email address or even to the network administrator's mobile phone. However, the main focus in this paper is only on the *Data Analysis* module. The function of this module is to analyse the input to decide whether the input contains possible threats or it is clean. In this particular module, the specific research area is to define the most significant features in order to produce an attack pattern policy more effectively.

2.2. IPv6 Attacks Classification

In IPv4 research domain, there is a dataset to construct detection techniques for IPv4 environment which known as KDD99 dataset. This KDD99 dataset is considered as a de facto standard dataset for IPv4 network environment as it was widely used by researchers [16-18]. Based on KDD99 dataset, the simulated attacks in classified into four categories, namely, Denial of Service Attack (DoS), User to Root Attack (U2R), Remote to Local Attack (R2L) and Probing Attack [19]. Besides that, the attack also can be classified based on the protocol used to launch its attacks. The main protocols used in IPv6 are TCP, UDP and ICMPv6 [20-22]. From other perspective, the attacks also can be classified based on the function or services offer in IPv6 environment such as neighbor advertisement and routing header [23, 24]. Hence, the IPv6 network attacks can be classified in several perspectives. Each classification is differentiated based on the coverage of selected network attack whether it covers the various attack categories, protocols used or services offer in IPv6 network environment.

2.3. Feature Selection Method

Feature selection is one of the main elements in the IDS framework. The selection of features used in an intrusion detection technique will eventually influence the impact of the proposed technique's performance. Based on Dash and Liu [25], the feature selection process can be found in the following figure. This task is crucial, as the final outcomes will directly impact the formation of a detection technique. The better the features selected are, the better the detection technique that will be produced. The selection of appropriate features will decrease data dimensionality and enhance the algorithm capability [26].

Fig. 2 shows the processes involved in selecting features. The process will begin with the reception of the original feature set from a complete dataset. The feature generation process will take place based on the received dataset. In this process, the weight age of each feature from the original dataset will be analyzed and ranked individually. Then, the best features will be proposed for the next step, which is called the feature evaluation. The evaluation process will then assess the quality of the selected features compared to the original feature set. Eventually, the relationship for each feature will be reported and then it is up to the end-user to evaluate whether the proposed features are the best

fit to represent the original dataset or vice versa. If the proposed features meet the predefined criteria, then the selected features will be used in further actions. Otherwise, the process of feature generation from the original dataset will be repeated until the predefined criteria are satisfied.

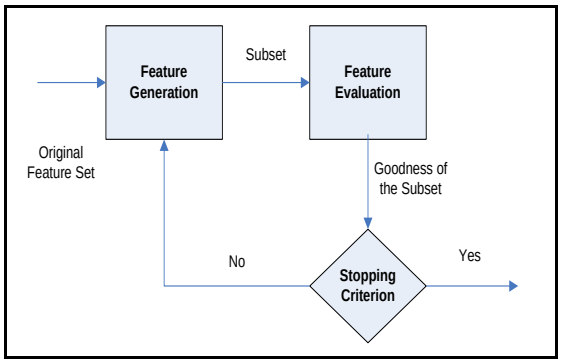

Fig. 2: Feature Selection Process [25]

2.4. Feature Significant Evaluation

For this particular task, a bio-inspired computing technique is used. According to Dressler [27], bio-inspired computing techniques such as Particle Swarm Optimization (PSO) are distinguished for solving issues that are related to the optimization process and pattern recognition. One of the main features offered by this technique is its diversity, which can improve a system's ability to react to unknown as well as unpredicted scenarios [28]. Based on these two authors, it is asserted that PSO is capable of identifying an optimized process and has a diversity feature that can identify unknown scenarios. Given its capabilities, this work utilizes the PSO technique to identify the best features to use when detecting network attacks in the IPv6 network environment.

This study uses the PSO algorithm produced by Moraglio [29]. They addressed the process of identifying the most significant features by adapting PSO and attribute space exploration. The technique was adapted in this study as a method to identify the most significant features to distinguish between normal and attack data within the IPv6 network environment. The algorithm used by the author can be found in the following table.

Fig. 3 shows the steps taken in order to determine the most significant features based on the prepared dataset. Most of the values were proposed by Moraglio et al. (2007). The algorithm has 20 iterations. The population size of a single iteration is set to 20. Each particle was evaluated in order to obtain its personal and then the global best scores. The objective function of this task was to obtain the maximum classification score to distinguish between normal and attack data. Selection for CfsSubsetEval was based on recommendations from previous studies [30, 31], which claimed that CfsSubsetEval was one of the best attribute selection evaluation tools.

```
1: while (i <= 20) do (iteration)
2:    for all particle i do
3:          initialize position X_i at random in the search space
4:    end for
5:    while (j <= 20) do (population size)
6:          for all particle i do
7:                Execute the objective function (CfsSubsetEval)
8:                set personal best x̂_i as best position found so far
                    by the particle
9:                set global best ĝ as best position found so far by
                    the whole swarm
10:         end for
11:         for all particle i do
12:               update position using a randomized convex
                   combination
                        X_i = CX ((X_i, 0.33), (ĝ, 0.33), (x̂_i, 0.34))
13:               mutate X_i(P(X_i) = 0.01)
14:         end for
15:   end while
16: end while
```

Fig. 3 : Geometric PSO Algorithm with CfsSubsetEval

2.5. Feature Evaluation

This task is meant to evaluate the performance of selected features in classifying different type of packet correctly. In this study, the selected features will be assessed on its capability to differentiate between normal and attack IPv6 packets. For this task, a data mining technique called *Support Vector Machine* (SVM) is used. SVM is a technique meant for two-class problems [32, 33]. In this case, the technique was used to differentiate between the normal and attack packets based on the chosen features. SVM is meant to draw a hyperplane to differentiate the data into two main classes. SVM has been successfully applied to several studies in order to solve various issues [34-36]. In this study, the SVM technique is used as a classifier to distinguish between normal and attack data. A representation of the SVM can be found in the following figure.

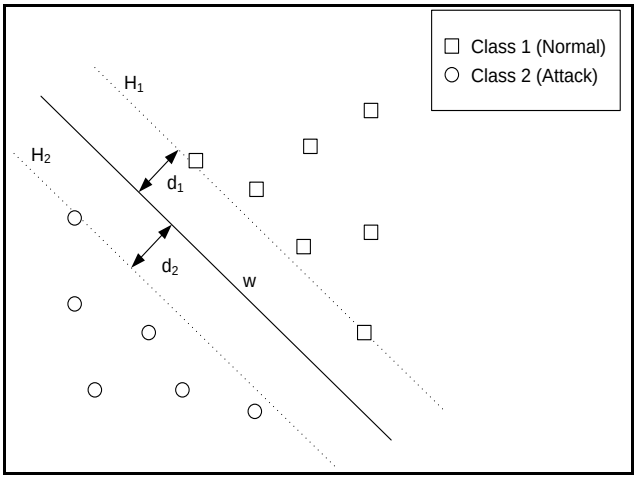

Fig. 4 SVM Hyperplane

Fig. 4 shows how SVM is used to draw a hyperplane that distinguishes between normal and attack data. In the figure, w is defined as normal to the hyperplane. Then the two planes, H1 and H2, reduced the margin of error from the original hyperplane as described by:

$$x_i.w + b = +1 \text{ for } H_1 \qquad \text{Eq. 1}$$
$$x_i.w + b = -1 \text{ for } H_2 \qquad \text{Eq. 2}$$

Eq. 1 and Eq. 2 show the acceptable margin from the original hyperplane. In these equation, the variable of $\frac{b}{\|w\|}$ defines the offset of the hyperplan from the origin along with the normal vector (w). Fig. 4, the margin from H_1 to the hyperplane is represented by d_1, while for H_2 it is d_2 where d_1 is equal to d_2, which is also known as SVM's margin [37]. The classifier can be defined as the following:

$$f(x) = sgn\left(\sum_{i=1}^{n} a_i y_i K\left(x_i \times x\right) + b\right) \qquad \text{Eq.3}$$

Eq. 3 shows the function to classify normal and attack data where sgn() represents the sign function, a_i is a Lagrange multiplier, $K\left(x_i \times x\right)$ is the kernel function where x_i is a training sample, and x is a sample to be classified. The Radial Basis Function (RBF) is the kernel function, as proposed by the LibSVM authors [38].

3. FRAMEWORK OF FEATURES SELECTION

Feature selection process was officially introduced in year 1997. Since then, the process was rapidly used and evolved to solve several issues related to the feature selection. In this paper, the framework of feature selection will be proposed. Fig. 5 shows the proposed framework of feature selection. In this framework, there are five phases needed to be performed. These phases were evolved from the original feature selection process produced by Dash and Lie (1997) [25]. As an example, this framework will be tested on a scenario where the scenario is to define the best features to differentiate between normal and attack packet based on IPv6 network environment.

3.1. Implementation

In this implementation section, the framework of feature selection will be executed to identify the best features in order to differentiate between normal and attack packets in IPv6 network environment. In this framework, five tasks are needed to be performed. However, in this case study the process of comparing result was discarded as there is not prior result to this study.

3.2. Problem Formulation

Problem Formulation is the first phase of feature selection. In this phase, the objective of this project is needed to be clarified. The background issues and the main problem are needed to be justified. In this case, the problem is to identify the best features of IPv6 packets which can be used to differentiate between normal and attack packets effectively.

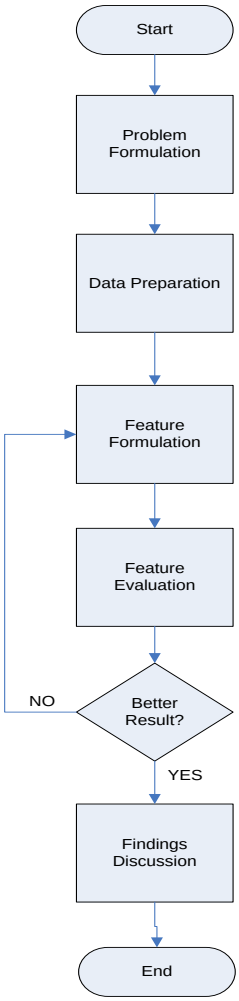

Fig. 5 The Framework of Features Selection

3.3. Data Preparation

Next, once the problem has been justified the data are needed to be prepared. As discussed in Section 3, the data for this scenario will be based on the protocol perspective. Therefore, the IPv6 main protocols which are TCP, UDP and ICMPv6 will be covered in this dataset. There two scenarios are needed to be simulated, namely, Normal Scenario and Attack Scenario. What is more, the Normal Scenario represents normal IPv6 packets transactions while the Attack Scenario signifies attack launch in IPv6 environment. The following table shows the attacks and the protocol used to launch each attacks

Table 1: The Coverage of the Selected Attacks

Attack Type	Targeted Victim	Protocol Used		
		TCP	UDP	ICMPv6
IPv6 Alive6	Node	X	X	
IPv6 FloodRouter	Network			X
IPv6 Smurf6	Node			X

Table 1 depicts the coverage of the selected attacks used in the attack traffic scenario in this study. There were three types of attacks used as stated in the table, namely, *Alive6*, *FloodRouter*, and *Smurf6*. For *Alive6*, the attack was meant for a dedicated node and uses the TCP and UDP protocols to launch its attack. Meanwhile, the *FloodRouter* attack was meant for the network victims and uses the ICMPv6 protocol to launch its attack. Finally, *Smurf6* is a type of attack that targets a dedicated node and uses the ICMPv6 protocol to launch its attacks. All the main protocols covered by these attacks are summarized in the table; the targeted victims were not only dedicated nodes but also network victims as well. The detail data collection processes was elaborated in the previous project [39].

3.4. Feature Formulation

Feature formulation will be the next stage. In this task, the list of available features will be identified. The best features to distinguish between normal and attack IPv6 packets will be determined. There are four processes involved, namely, 1) *Common Features*, 2) *Feature Justification*, 3) *Feature Significance Evaluation* and 4) *Final Feature Selection*. The process begins by identifying the common features used by the main IPv6 protocols: TCP, UDP, and ICMPv6. After these features are identified, the justification for each selected feature is analyzed to ensure the feature's suitability with the study's objective. Then, based on the findings, the features passing the justification process are evaluated using an attribute selection technique from data mining. The findings identify the significance of the parameters chosen for the datasets. Finally, after the evaluation, the best features are identified for further assessment.

3.4.1. Common Features

This is the first process that must be completed before continuing on to the next task. This task is to identify the common features among the main protocols used in the IPv6 environment, which are TCP, UDP, and ICMPv6. Although these protocols use the same IPv6 packet structure, each one conveys different features in its packets. All the features that can be extracted directly from the IPv6 packets are considered basic features in the work. During this task, a packet from each protocol was captured using the TCPDump tool; all the features contained in each packet were obtained for further analysis. All features were extracted from the packet header from each protocol. The generic features represent the common features among these main

protocols in IPv6. Based on the figure, there are seven common features: timestamp (time), hop limit (hlim), next payload (npayload), source IP address (SrcIP), destination IP address (DstIP), next length (nlength), and protocol used (protocol). These seven features were evaluated to determine the most significant features that are representative of the IPv6 packets for further analysis.

3.4.2.Feature Justification

Next task is Feature Justification. The features used for the justification were gathered based on the common features from the three main protocols in IPv6: TCP, UDP, and ICMPv6. However, three features were discarded from the dataset: *hlim*, *npayload*, and *nlength*. In addition, two new features were taken into consideration: source (*SrcPort*) and destination port addresses (*DstPort*). Furthermore, another feature, timestamp (*time*), was replaced with the interval time between two consequent packets (*TimeIntvl*).

Based on the common features between the three main IPv6 protocols, three features were eliminated: *hlim, npayload,* and *nlength*. *hlim* was discarded because the network scenario in this study was limited to two networks that only represent an inside and outside network. The hops between packets were almost the same but this *hlim* value is not consistent in the real IPv6 network since packets originate from various places. Next, *npayload* and *nlength* were excluded because the packets generated in the testbed environment came from packet generator tools. They were more focused on the header information than the payload of the packets. Adding padding to the generated packets would not have resolved the issue because the payload of the packets would vary depending on their application. Therefore, taking these features into consideration would have probably yielded questionable results in the end.

On the other hand, two features were added: source port (*SrcPort*) and destination port addresses (*DstPort*), even though these features were not among the common features for the main IPv6 protocols. *SrcPort* and *DstPort* were included because the value of the port address could resolve the application used by the packet. This feature is also known as the "application address." For example, if a packet used value 80 in the *DstPort*, the packet would be probably an HTTP packet. Some of the latest detection techniques also incorporate *SrcPort* and *DstPort* as features in their techniques [40-42], because the port addresses can contribute to the network attack behavior pattern. Hence, the addition of the *SrcPort* and *DstPort* features in the analysis stage will be useful in assessing the significance of these features for intrusion detection.

Next, the time feature was replaced with a feature measuring the time between two consecutive packets (*TimeIntvl*). According to Onat [43] and Lei [44], time was a main element in their proposed detection techniques because it is useful in producing a real-time detection technique. However, Kline [45] showed that the use of a time series or a timestamp could cause a problem in detecting anomalous traffic accurately. Therefore, the decision to discard the time feature was not a good idea since it is useful in constructing the generated features such as number of connections; in such a case, packet frequency based on time and bandwidth usage based on time cannot be derived if they are needed. Hence, a feature derived from the timestamp, *TimeIntvl*, was introduced to the work. *TimeIntvl* is meant to improve the time feature by conveying more useful information rather than simply storing only the timestamp information. Although retaining the time feature could yield features in the future, the detection process must add another task to generate additional features instead of using the stored data to begin the detection analysis straight away. What is more, the process of generating derived features such as number of connections per time and number of packets per time will eventually tamper with the originality of the other features. Therefore, *TimeIntvl* was introduced to convey the useful information from a timestamp without compromising the value of the other features.

Table 2 lists the features selected for further analysis. There were three stages prior to making a decision regarding the selected features. The first stage involved listing the common features, where seven features were found*: time, hlim, npayload, nlength, SrcIP, DstIP,* and *Protocol*. In the next stage, three additional features were introduced based on some justifications. From this reasoning, the time feature was replaced with *TimeIntvl* and another two features were added to the future analysis list: *SrcPort* and *DstPort*. Finally, after all these considerations had been made, only six features were selected for the feature analysis process, namely, *TimeIntvl, SrcIP, DstIP, SrcPort, DstPort,* and *Protocol*. These features then went through the feature analysis process, which determined whether the selected features were significant enough to be stored in the database for further analysis. The process of identifying the importance of the selected features will be discussed in the upcoming section.

Table 2: Selected Features for Feature Analysis

List of Features	Phases		
	Common Features	Introduced Features	Selected Features
time	X		
hlim	X		
npayload	X		
nlength	X		
SrcIP	X		X
DstIP	X		X
Protocol	X		X
TimeIntvl		X	X
DstPort		X	X
SrcPort		X	X

3.4.3. Feature Significance Evaluation

In this third section, the process of identifying the significance of the selected features will be elaborated explicitly. In order to conduct this task, attribute selection in a data mining technique was used. *Particle Swarm Optimization* (PSO) was employed to evaluate the best features representing the dataset. At this stage, three datasets were evaluated in order to analyze the importance of each feature with regard to the dataset. The data prepared for the analysis are presented in the following table.

Table 3: Dataset for Features Analysis

Dataset	Data Identification	Dataset Size (records)
Data1 (*Alive6*)	D1A	142,640
Data1 (*FloodRouter*)	D1F	253,832
Data1 (*Smurf6*)	D1S	195,827

Table 3 provides details about the datasets used for feature analysis in order to identify their best representative features. These datasets can be considered generic datasets because the information touched all three main IPv6 protocols and the targeted victims consist of both nodes and the network. The five features identified from the datasets in the previous section were evaluated and the final findings classified the best features for use in creating a detection technique to recognize IPv6 network attacks.

Table 4: PSO Analysis Result

Analyzed Features	Dataset			Analysis Finding
	D1A	D1F	D1S	
TimeIntvl	X	X	X	3
SrcIP	X	X	-	2
DstIP	-	-	-	0
SrcPort	-	-	X	1
DstPort	-	X	-	1
Protocol	X	-	X	2

Table 4 provides the findings from the PSO analysis conducted in the study. As seen on the table, three datasets were used in the PSO analysis. The first dataset results show that only three features substantially represented the dataset: TimeIntvl, SrcIP, and protocol. In the second dataset, again, only three features were significant enough to represent the dataset: TimeIntvl, SrcIP, and DstPort. Finally, the third dataset shows that only TimeIntvl, SrcPort, and protocol were noteworthy in representing the dataset. The main feature from these findings, TimeIntvl became the proposed feature; TimeIntvl was important to all the datasets. Meanwhile, the DstIP feature was the least significant. Therefore, DstIP was accordingly discarded from the final feature selection list.

3.4.4. Final Feature Selection

In this final stage, the final selected feature will be identified based on the result obtained from the feature significance evaluation. The following table shows the final selected features.

Table 5: Final Selected Features (ProFeat 2013)

Author (Year)	No. of Features	Features List
ProFeat (2013)	*5*	*TimeIntvl*
		SrcIP
		SrcPort
		DstPort
		Protocol

Table 5 shows the set of features selected as the enhanced features proposed by the work. After the feature analysis was conducted, it was concluded that only five features would be listed as enhanced features to be used in a future detection technique to identify IPv6 network attacks. These features are *TimeIntvl*, *SrcIP*, *SrcPort*, *DstPort*, and *Protocol*. They will be classified as the enhanced features proposed by the study and will be referred to as (*ProFeat 2013*) in the paper. Chapter 5 will explicate the process of implementing and evaluating the enhanced features proposed by this study.

3.5. Feature Evaluation

In this task, the features selected in the previous stage will be assessed. The assessment was based on the capability of the selected features to differentiate between different types of packets. In this study, the selected will be evaluated its performance on classifying normal or attack IPv6 packets correctly. The following table shows the dataset prepared for this particular task.

Table 6: Dataset for Features Analysis

Dataset	Data Identification	Dataset Size (records)
Data2 (*Alive6*)	D2A	161,699
Data2 (*FloodRouter*)	D2F	250,088
Data2 (*Smurf6*)	D2S	201,588

Table 6 provides details on the dataset prepared for the feature evaluation process. Similar to the previous tasks, three attacks were taken into this assessment task, namely, Alive6, FloodRouter and Smurf6. Each of these dataset is comprised by two minutes of IPv6 traffic packets where 1 minute of normal traffic and another 1 minute of a specific attacks. A data mining called SVM is used to perform this task. Each data were tested on a ten-fold cross validation mode to bias in the final outcomes. From this task, the result will determine the capability of the selected features on classifying different type of IPv6 packets. The following figure shows the accuracy score of the selected features of the tested dataset.

Table 7 provides the results that have been extracted from the feature evaluation process on the prepared dataset. The selected feature was represented by ProFeat 2013. From the accuracy score, it can be concluded that the selected features are capable of correctly classify the IPv6 packets with an outstanding score. The average accuracy score for this test is 99.95%. From this result, the features used by *ProFeat 2013* can be used to produce an outstanding detection technique to detect IPv6 network attacks in the future.

Table 7: Feature Evaluation Results

Tested Features	Accuracy Score (%)		
	Dataset		
	D2A	D2F	D2S
ProFeat 2013	99.94	99.94	99.98

Table 7 provides the results that have been extracted from the feature evaluation process on the prepared dataset. The selected feature was represented by ProFeat 2013. From the accuracy score, it can be concluded that the selected features are capable of correctly classify the IPv6 packets with an outstanding score. The average accuracy score for this test is 99.95%. From this result, the features used by *ProFeat 2013* can be used to produce an outstanding detection technique to detect IPv6 network attacks in the future.

3.6. Findings Discussion

This is the last phase this proposed framework. In this stage, the findings obtained throughout the study will be documented for future references. All results regardless either positive or negative results were compiled and the justification of each obtained result will be elaborated. This documentation will help future researcher to avoid replication process conducted in this study. Some research gaps also are needed to be defined to help this research continues.

4. CONCLUSION

As a conclusion, this paper is proposed a framework of feature selection procedure in general. As a case study, an implementation of defining the best features to differentiate between normal and attack IPv6 was executed. In general, the framework proposed in this paper can also be adapted in other domain which requires features selection solution. What is more, the feature proposed by this case study also can be used to produce detection technique to detection network attack in IPv6 network environment. In the future, the dataset produced in this case study will be improvised to facilitate studies which related to producing future IPv6 network attacks detection technique.

References:

[1] Waddington, D.G. and F. Chang, *Realizing the transition to IPv6*. IEEE Communications Magazine, 2002. **40**(6): p. 138-147.

[2] Triulzi, A. *Intrusion Detection Systems and IPv6*. 2003 [cited 2013 19 Nov 2013]; 2007]. Available from: http://www.alchemistowl.org/arrigo/Papers/ SPI2003-IDS-and-IPv6.pdf.

[3] Avi, T., *IPv6: new technology, new threats.* Network Security, 2011(8): p. 13-15.

[4] Barker, K., *The security implications of IPv6.* Network Security, 2013. **2013**(6): p. 5-9.

[5] Tang, S., K.-Y. Wong, and K. Yeung, *Record path header for triangle routing attacks in IPv6 networks.* WSEAS TRANSACTIONS on COMMUNICATIONS, 2008. **7**(12): p. 1202-1211.

[6] Zagar, D., K.i. Grgic, and S. Rimac-Drlje, *Security aspects in IPv6 networks implementation and testing.* Computers & Electrical Engineering, 2007. **33**(5-6): p. 425-437.

[7] Hansman, S. and R. Hunt, *A taxonomy of network and computer attacks.* Computers & Security, 2005. **24**(1): p. 31-43.

[8] Bellovin, S.M., B. Cheswick, and A.D. Keromytis, *Worm propagation strategies in an IPv6 Internet.* LOGIN: The USENIX Magazine, 2006. **31**(1): p. 70-76.

[9] Zulkiflee, M., M.A. Faizal, I.O. Mohd Fairuz, A. Nur Azman, and S. Shahrin, *Behavioral Analysis on IPv4 Malware in both IPv4 and IPv6 Network Environment.* International Journal of Computer Science and Information Security (IJCSIS), 2011. **9**(2).

[10] Pochiraju, A., *Time To Take IPv6 Thoughtfully*, in *Siliconindia*. 2012, Siliconindia Inc. p. 38-39.

[11] Lim, J.-D., 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.* International Conference on Control, Automation and Systems, 2007. ICCAS'07. , 2007: p. 404-407.

[12] Yun, K. and M. Zhu Jian. *Research of Hybrid Intrusion Detection and Prevention System for IPv6 Network*. in *Internet Technology and Applications (iTAP), 2011 International Conference on*. 2011.

[13] Ferdous, B., G. Bansal, N. Kumar, S. Biswas, and S. Nandi, *Detection of neighbor discovery protocol based attacks in IPv6 network.* Networking Science, 2013. **2**(3-4): p. 91-113.

[14] Ma, Y., H. Lian, and F. Zhang Xiao. *Researches on the IPv6 Network safeguard linked system*. in *Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on*. 2010.

[15] Zheng, Z. *Intrusion Detection System of IPv6 Based on Protocol Analysis*. in *Multimedia Technology (ICMT), 2010 International Conference on*. 2010.

[16] Jiong, Z. and M. Zulkernine. *A hybrid network intrusion detection technique using random forests*. in *Availability, Reliability and Security, 2006. ARES 2006. The First International Conference on*. 2006.

[17] Lee, J.-H., J.-H. Lee, S.-G. Sohn, J.-H. Ryu, and T.-M. Chung. *Effective value of decision tree with KDD 99 intrusion detection datasets for intrusion detection system*. in *Advanced Communication Technology, 2008. ICACT 2008. 10th International Conference on*. 2008: IEEE.

[18] Olusola, A.A., A.S. Oladele, and D.O. Abosede. *Analysis of KDD'99 Intrusion Detection Dataset for Selection of Relevance Features*. in *World Congress on Engineering and Computer Science*. 2010. San Francisco, USA.

[19] Tavallaee, M., E. Bagheri, W. Lu, and A.-A. Ghorbani. *A detailed analysis of the KDD CUP 99 data set*. in *Proceedings of the Second IEEE Symposium on Computational Intelligence for Security and Defence Applications 2009*. 2009.

[20] Stallings, W., *IPv6: the new Internet protocol.* Communications Magazine, IEEE, 1996. **34**(7): p. 96-108.

[21] Deering, S. and R. Hinden, *RFC2460: Internet Protocol, Version 6 (IPv6) Specification.* RFC Editor United States, 1998.

[22] Zagar, D. and K. Grgic. *IPv6 Security Threats and Possible Solutions*. in *Automation Congress, 2006. WAC '06. World*. 2006.

[23] Jeong, J.P., *IPv6 Router Advertisement Option for DNS Configuration.* IETF, 2007.

[24] Nakibly, G. and F. Templin, *Routing Loop Attack using IPv6 Automatic Tunnels: Problem Statement and Proposed Mitigations.* RFC 6324, 2011.

[25] Dash, M. and H. Liu, *Feature selection for classification.* Intelligent data analysis, 1997. **1**(1-4): p. 131-156.

[26] Faezipour, M., M. Nourani, and S. Addepalli. *A Behavioral Analysis Engine for Network Traffic*. in *Consumer Communications and*

Networking Conference (CCNC), 2010 7th IEEE. 2010.

[27] Dressler, F., *Self-organization in sensor and actor networks*. 2008: John Wiley & Sons.

[28] De Castro, L.N., *Fundamentals of natural computing: basic concepts, algorithms, and applications*. 2006: CRC Press.

[29] Moraglio, A., C. Di Chio, and R. Poli, *Geometric particle swarm optimisation*, in *Genetic Programming*. 2007, Springer. p. 125-136.

[30] Ye, P., Y. Kim, and O. Buzek, *LING773-Negotiations Project Report.* 2011.

[31] Persada, A.G., N.A. Setiawan, and H.A. Nugroho. *Comparative study of attribute reduction on arrhythmia classification dataset*. in *International Conference on Information Technology and Electrical Engineering (ICITEE), 2013* 2013: IEEE.

[32] Vapnik, V., *The nature of statistical learning theory*. 2nd ed. 2000, New York: Springer.

[33] Boser, B.E., I.M. Guyon, and V.N. Vapnik. *A training algorithm for optimal margin classifiers*. in *Proceedings of the 5th Annual Workshop on Computational learning theory*. 1992: ACM.

[34] Nivre, J., J. Hall, J. Nilsson, A. Chanev, G.l. Eryigit, S. KÃbler, S. Marinov, and E. Marsi, *MaltParser: A language-independent system for data-driven dependency parsing.* Natural Language Engineering, 2007. **13**(2): p. 95-135.

[35] Hanke, M., Y.O. Halchenko, P.B. Sederberg, S.J. Hanson, J.V. Haxby, and S. Pollmann, *PyMVPA: A python toolbox for multivariate pattern analysis of fMRI data.* Neuroinformatics, 2009. **7**(1): p. 37-53.

[36] Dorff, K.C., N. Chambwe, M. Srdanovic, and F. Campagne, *BDVal: reproducible large-scale predictive model development and validation in high-throughput datasets.* Bioinformatics, 2010. **26**(19): p. 2472-2473.

[37] Fletcher, T., *Support vector machines explained.* Tutorial paper., Mar, 2009.

[38] Chih-Chung, C. and L. Chih-Jen, *LIBSVM: A library for support vector machines.* ACM Transaction Intelligent Systems Technology, 2011. **2**(3): p. 1-27.

[39] Zulkiflee, M., 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), 2014. **9**(8): p. 1434-1441.

[40] Zhao, D., I. Traore, A. Ghorbani, B. Sayed, S. Saad, and W. Lu, *Peer to Peer Botnet Detection Based on Flow Intervals*, in *Information Security and Privacy Research*. 2012, Springer. p. 87-102.

[41] Stiawan, D., M. Idris, and A.H. Abdullah, *The Prevention Threat of Behavior-based Signature using Pitcher Flow Architecture.* IJCSNS International Journal of Computer Science and Network Security, 2010. **10**(4): p. 289-294.

[42] Ramaki, A.A., R.E. Atani, R.K.I. Abadi, and M. Tavaghoe, *Enhancement Intrusion Detection using Alert Correlation in Co-operative Intrusion Detection Systems.* Journal of Basic and Applied Scientific Research, 2013. **3**(6): p. 272-279.

[43] Onat, I. and A. Miri. *A real-time node-based traffic anomaly detection algorithm for wireless sensor networks*. in *Systems Communications, 2005. Proceedings*. 2005.

[44] Lei, L., J. Xiaolong, M. Geyong, and X. Li. *Real-Time Diagnosis of Network Anomaly Based on Statistical Traffic Analysis*. in *Trust, Security and Privacy in Computing and Communications (TrustCom), 2012 IEEE 11th International Conference on*. 2012.

[45] Kline, J., N. Sangnam, P. Barford, D. Plonka, and A. Ron. *Traffic Anomaly Detection at Fine Time Scales with Bayes Nets*. in *Internet Monitoring and Protection, 2008. ICIMP '08. The Third International Conference on*. 2008.