SPECIFICATION-BASED INTRUSION DETECTION SYSTEM FOR 802.11 NETWORKS USING INCREMENTAL DECISION TREE CLASSIFIER

by

VACHANA MANJUNATH HONNAMMA

Submitted in partial fulfilment of the requirements
for the degree of Master of Computer Science

at

Dalhousie University
Halifax, Nova Scotia
March 2018

To my Mom and Brother,

For their unconditional love and encouragement throughout my life.

To my friends for being a part of my life and making it as beautiful as it is.

To my supervisor for never losing his cool and always boosting my confidence and helping me in everything possible

And lastly

To all my well-wishers for standing by my side in every aspect of my success.

TABLE OF CONTENTS

LIST OF FIGURES..... vi
LIST OF TABLES..... vii
ABSTRACT..... viii
LIST OF ABBREVIATIONS..... ix
CHAPTER 1 - INTRODUCTION..... 1
1.1 Wireless Local Area Networks (WLANs) and Security..... 1
1.1.1 WLAN Goal 2
1.2 WLAN 802.11 Standards..... 3
1.2.1 Security Threats on WLAN..... 4
1.3 Intrusion Detection Systems for WLAN Security..... 5
1.3.1 Data Collection Regarding Suspects..... 6
1.3.2 Data Analysis..... 6
1.4 Intrusion Classification Techniques..... 7
1.4.1 Decision Trees and Neural Networks..... 8
1.4.3 Naive Bayesian and Fuzzy Sets..... 9
1.5 Scope and Motivation..... 9
1.6 Problem Statement..... 10
1.7 Aims and Objectives..... 11
1.8 Contributions..... 11
1.9 Organization of Thesis..... 12

CHAPTER 2 - BACKGROUND..... 13
2.1 Wireless Local Area Network (WLAN)..... 13
2.1.1 Different 802.11 WLAN standards..... 14
2.1.2 MAC 802.11 specific Intrusions on the WLAN..... 15
2.1.3 Effectiveness of IDS in WLANs Security..... 17
2.2 Types of Intrusion Detection Techniques..... 19
2.3 Conventional Data Collection Processes..... 20
2.3.1 Host and Network Based Data Collection..... 20

2.3.2 Direct and Indirect Monitoring for Data Collection…… 21
2.4 Data Analysis Techniques…….. 22
2.4.1 Pattern Matching Analysis Techniques…… 23
2.4.2 Data Mining Techniques…….. 24
2.5 Data Classification Methods in IDS Design……..... 26
2.5.1 Supervised Learning……..... 26
2.5.2 Semi Supervised and Unsupervised Learning…….. 27
SUMMARY……..... 29

CHAPTER 3 - LITERATURE SURVEY…… 30
3.1 Survey on MAC intrusions in WLAN…… 30
3.2 Common WLAN protection mechanisms, and its drawbacks…… 33
3.3 Review on IDS Techniques in WLAN……..... 34
3.3.1 Based on Anomaly Model…… 34
3.3.2 Based on Specification Model…….. 37
3.3.3 Based on Signature Model ……...……... 38
3.4 Feature Reduction Techniques ……..…… 40
3.4.1 Filtering Techniques…… 40
3.4.2 Wrapper Techniques…… 41
3.4.3 Hybrid Techniques……... 43
3.5 Data mining approaches for ……..…… 45
SUMMARY……..... 47

CHAPTER 4 - PROPOSED APPROACH ……..……..…… 48
4.1 Introduction…… 48
4.1.1 Importance of Feature Reduction in IDS…… 49
4.1.2 Decision Tree Algorithm…….. 50
4.2 Overview of the N-TBTDT…… 51
4.3 Optimal Feature Selection……..... 53
4.3.1 Normalized Information Gain Measurement for Feature Reduction……... 54
4.3.2 Chaotic PSO Based Attack Clustering…… 55

4.4 Improving Decision Tree Algorithm…….. 56
4.4.1 Tie Breaking Threshold Based Decision Tree Algorithm…… 57
4.5 Intrusion Detection and Categorization…….. 58
SUMMARY……. 58

CHAPTER 5 - PERFORMANCE EVALUATION OF N-TBTDT …….. 59
5.1 Introduction to N-TBTDT……..... 59
5.1.1 Extracted Feature Set…… 60
5.2 AWID Dataset……..…. 61
5.3 Experimental Evaluation……..…. 61
5.3.1 Performance Metrics…… 62
5.4 Experimental Results……. 62
5.4.1 TDS Vs Detection Accuracy…… 62
5.4.2 TDS Vs False Positive Rate……. 64
5.4.3 TDS Vs Precision……. 65
5.4.4 TDS Vs F-Score……... 66
5.4.5 Number of Attacks Vs Detection Accuracy……. 67
5.4.6 Number of Attacks Vs Classification Accuracy……..... 69
SUMMARY……. 70

CHAPTER 6 - CONCLUSIONS AND FUTURE DIRECTIONS …… 72
6.1 Conclusions…… 72
6.2 Future Directions…… 73

REFERENCES……... 74

LIST OF FIGURES

Figure 4.1: Optimal Feature Selection Using Normalized Information Gain and Chaotic PSO Optimization Algorithm……..... 51
Figure 4.2: Decision Tree Classifier Based Intrusion Detection System…….. 52
Figure 5.1: TDS Vs Detection Accuracy…….. 63
Figure 5.2: TDS Vs False Positive Rate……..... 64
Figure 5.3: TDS Vs Precision……..... 65
Figure 5.4: TDS Vs F-Score……. 67
Figure 5.5: Number of Attacks Vs Detection Accuracy……..... 68
Figure 5.6: Number of Attacks Vs Classification Accuracy…… 69
Figure 5.7: Test results …… 70

LIST OF TABLES

Table 2.1: Different WLAN Standards…… 15
Table 2.2: Types of WLAN Intrusion…… 16
Table 2.3: Comparison of WLAN security protocols…… 18
Table 2.4: Comparison of Different Data Analysis Techniques…… 25
Table 2.5: Comparison of Different Classification Technique…… 27
Table 3.1: Different Types of WLAN Intrusions…… 31
Table 3.2: Features of WLAN Basic Security Schemes…… 33
Table 3.3: Compilation of IDS Techniques…… 35
Table 3.4: Comparison of various Types of IDSs…… 39
Table 3.5: Comparison of Various Feature Reduction Techniques…… 44
Table 5.1: Extracted Feature Set…… 60
Table 5.2: TDS Vs Detection Accuracy…… 62
Table 5.3: TDS Vs False Positive Rate…… 64
Table 5.4: TDS Vs Precision…… 65
Table 5.5: TDS Vs F-Score…… 66
Table 5.6: Number of Attacks Vs Detection Accuracy…… 67
Table 5.7: Number of Attacks Vs Classification Accuracy…… 69

ABSTRACT

Wireless Local Area Networks (WLANs) represent a popular wireless networking technology, and they provide high-speed internet connections based on the 802.11 standard. WLAN is subject to different kinds of intrusions, such as injection, impersonation, flooding etc. Several attempts in the development of efficient WLAN protection schemes have ended up with the inadequate security mechanisms that are vulnerable to various 802.11 standard intrusions. Intrusion Detection Systems (IDSs) play an inevitable role in WLAN security.

The objective of this thesis is to present a specification-based IDS technique, named as Normalized information gain and Tie Breaking Threshold-based Decision Tree (N-TBTDT) that utilizes various data mining techniques to improve the IDS performance significantly. The main contributions of the proposed IDS are feature reduction using normalized information gain, the chaotic Particle Swarm Optimization (PSO) for feature extraction and improved Very Fast Decision Tree (VFDT) for intrusion classification. The bias compensation factor-based tie-breaking threshold promises the efficient decision tree construction, rather than the random selection of tie-breaking threshold. This shows a significant improvement in the detection accuracy of N-TBTDT.

To evaluate the performance of N-TBTDT, two different scenarios are created. Firstly, the training dataset size is varied, and secondly, the number of attacks is varied from low to high. The N-TBTDT exploits different performance metrics such as detection accuracy, false positive rate, precision, F-Score, and classification accuracy in experimental evaluation. The experimental results show that the improved decision tree classifier accurately detects and classifies the 802.11 specific intrusions and, it attains 99.94% of detection accuracy. For a training dataset size of 177.2MB, the false positive rate of N-TBTDT decreases by 0.95%, when compared to existing system - Normalized gain based IDS for MAC Intrusions (NMI).

LIST OF ABBREVIATIONS

AES	Advanced Encryption Standard
AP	Access Point
BMSL	Behavioral Monitoring Specification Language
BSS	Basic Service Set
DoS	Denial of Service
EFSAs	Extended Finite State Automata
GA	Genetic Algorithm
GFR	Gradually Feature Removal
HMM	Hidden Markov model
IDS	Intrusion Detection System
MABDIDS	Multi-agent Based Distributed WLAN IDS
MIMO	Multiple-Input-Multiple-Output antennas
NBC	Normalized Bias Compensation
NICs	Network Interface Cards
NIG	Normalized Information Gain
NMI	Novel anomaly based IDS for MAC Intrusions
NN	Neural Networks
N-TBTDT	Normalized information gain and Tie Breaking threshold based Decision Tree
PCA	Principal Component Analysis
PSO	Particle Swarm Optimization
SVM	Support Vector Machine
TDS	Training Data Size
TKIP	Temporal Key Integrity Protocol
UNII	Unlicensed National Information Infrastructure
VFDT	Very Fast Decision Tree
WEKA	Waikato Environment for Knowledge Analysis

WEP	Wired Equivalent Privacy
WiFi	Wireless Fidelity
WLANs	Wireless Local Area Networks
WPA	Wi-Fi Protected Access

CHAPTER 1 INTRODUCTION

1.1 Wireless Local Area Networks (WLANs) and Security

The tremendous growth of handheld devices proliferates the use of Wireless Local Area Networks (WLANs). The IEEE 802.11 is a series of wireless networking standards to define both radio standards and medium access control specifications and to implement the WLAN [1]. The 802.11 popularly called as Wireless Fidelity (WiFi) specifies an over-the-air interface between a wireless device and an access point (AP) or between two devices. These standards aim at providing a wireless Ethernet capability. The emergence of IEEE 802.11 mainly contributes to the popularity of WLAN deployment. The basic design of 802.11 emphasizes the convenience of WLAN deployment, rather than security. Moreover, the WLAN enables the users to access the information across many sectors in time and cost-effective manner. The ease of installation, convenience, no wiring, and mobility support tends to exploit the wireless LANs everywhere from home to large enterprise corporate networks. Recently, the growing popularity of WLAN and strong demand for internet-enabled devices emphasize the importance of having a secured network.

There are two types of WLANs namely ad hoc and infrastructure networks. In the ad-hoc mode, the network consists of wireless devices only without having any centralized architecture. The wireless devices can communicate with each other in the network using Network Interface Cards (NICs). The ad hoc mode is mostly suitable for small organizations. In the infrastructure-based WLANs, each node has a Basic Service Set (BSS) to connect to the AP directly. The infrastructure mode WLAN establishes the devices to communicate through an AP. Many WLANs support the infrastructure mode. Mostly, a group of wireless nodes within a geographic area, such as an office, an enterprise, a home, a restaurant, and the public zone deploy Wi-Fi networks in infrastructure mode. WiFi is an extension to the existing wired local area networks and allows the devices to network access under mobility. The advantages of WLAN are as given below.

User Mobility: The devices connect to the network resources such as internet using an air medium. This enables a group of devices to move around within a coverage area and still be connected to the network.

Rapid Installation: The installation time of WLAN is small. It is not required for wiring every workstation and every room. Without additional wiring or reconfigure the network, the WLAN makes the movement of the connected device very easy to the workstation.

Cost Reduction: Absence of wires and cables brings down the WLAN cost significantly. It also reduces the cost of trenching, drilling, and other methods which are needed for wired connections. The primary factor accomplishes the cost reduction is wireless routers.

Flexibility: The WLAN allows the system administrators to install it easily for temporary usages such as a conference or meeting without spending time in implementing cables and other ancillaries.

Scalability: To meet specific application and installation needs, the WLAN network topologies can be easily configured and scaled from small peer-to-peer networks to vast enterprise networks.

1.1.1 WLAN Goal

The WLAN consists a set of network nodes located in a limited geographical area, where each mobile node is capable of radio communication with an access point. Due to the flexibility in deployment and device mobility support, the WLAN is vulnerable to security intrusions [2-5]. Most of the WLAN equipment simultaneously supports three security schemes. Those are Wired Equivalent Privacy (WEP), Wi-Fi Protected Access (WPA) and WPA2. The security scheme can be differentiated from each other based on the encryption algorithms and the security key complexity. The security mechanisms in WLAN are evaluated by applying security detection mechanisms and key strength. The WEP protected WLAN provides security using an optional authentication process. To ensure the data confidentiality, the WEP exploits a key shared between the AP and only one WLAN device, named as key-mapping key. Both secret keys and key-mapping keys are subsequently utilized for protecting the communication between two nodes in

WLAN. However, the WEP has critical security flaws. Some of the security intruders utilize the advantage of a vulnerability on the key scheduling algorithm and break the WEP protection on WLAN [6-9].

Subsequently, WPA and WPA2 are developed as a quick alternative to the WEP. The WPA security scheme ensures independent authentication to every user and the implementation of the WPA in 802.11x solves the issues in WEP. It provides an efficient encryption mechanism such as Temporal Key Integrity Protocol (TKIP) and the Advanced Encryption Standard (AES). These encryption techniques effectively solve the issues associated with mutual authentication. Another version of 802.11i security protocol named as WPA2 highly contributes to the security of WLAN users. Both the security schemes such as WPA and WPA2 provide authentication and integrity to the users. Even though, both the security schemes such as WPA and WPA2 allow the users to access the firmware of AP directly, it is vulnerable to the password cracking attacks. The next version of 802.11w has been proposed using the robust security frames to ensure the authentication, integrity, and confidentiality. However, intelligent attackers launch disassociation attacks into the network by hacking the wall frames of 802.11w.

1.2 WLAN 802.11 Standards

Several WLAN solutions are available with varying levels of standardization and interoperability. The extensively deployed 802.11 standard has many extensions, and a few of them are under development. The 802.11 is the initially developed security standard for WLANs. It provides a data rate of 2 Mbps to the mobile nodes in WLAN. After that, 802.11a and 802.11b are developed from the initial 802.11 standards. The 802.11a standard provides 5 GHz band with the data rate of 54 Mbit/s. Compared to 802.11a, the 802.11b standard offers a high data rate of 11 Mbit/s with 2.4 GHz band. Both the standards follow the same media access method that is used in the 802.11 standards. The third amendment of 802.11 is 802.11g, which operates in the band of 2.4 GHz. The 802.11g exploits the same Orthogonal Frequency-Division Multiplexing (OFDM) based transmission scheme used in the 802.11a. The data rate of 802.11g is also

54 Mbit/s. The 802.11n is developed by introducing the Multiple-Input-Multiple-Output antennas (MIMO) in the previous 802.11 standards [10]. The 802.11n operates between 2.4 GHz and 5 GHz bands with a data rate of 600 Mbit/s. The 802.11ac is an evolutionary improvement to 802.11n. Compared to others, the 802.11ac delivers higher levels of performance that are commensurate with Gigabit Ethernet networking. The WEP is an optional encryption mechanism for 802.11 standards to ensure data confidentiality to the WLAN users. However, the WEP is ineffective, especially in de-authentication and plaintext attacks, due to the inherent nature of the wireless medium [11].

1.2.1 Security Threats on WLANs

The WiFi networks are mostly used in the office, an enterprise, a home, a restaurant, and the public zone. Due to the inherent broadcast nature of the wireless medium, the WLAN is highly vulnerable to the security attacks, compared to the wired LAN environment [5]. The intrusions on WLAN are traffic analysis and eavesdropping. One of these types is used by the intruder to violate the confidentiality and the integrity. In traffic analysis, an intruder attempts to hack three types of information. The first parameter is related to the identification of activities performed in the network. Secondly, the intruders try to identify the identification and physical location of the access point. Finally, an intruder gathers the information about the size and the number of packets over time by observing the network traffic. In eavesdropping, an intruder eavesdrops the complete or partial communication during data forwarding. By overhearing the network traffic, the intruders inject malicious behaviors into the network. Several security standards are proposed against the traffic analysis and eavesdropping intruders. However, these standards lack in identifying the advanced threats efficiently, since they snoop the keys that are employed by the 802.11 standard protocols.

By eavesdropping and analyzing the traffic information over time, the intrusions against the 802.11 security mechanisms are classified into four types such as Denial of Service, key retrieving, key stream retrieving, and Man in the Middle.

Denial of Service (DoS) Intrusion: An intruder targets the AP or the clients by continuously flooding the forged 802.11 management messages. This tends the AP to provide no frequent services to the legitimate users and even the network to crash.

Key retrieving Intrusion: The key retrieving is the passive intrusion, and it monitors the network activities for revealing the secret key that is used for packet encryption and decryption. The primary objective of the keystream intrusion is to threaten the confidentiality of the data.

Keystream Retrieving Intrusion: An intruder hacks the knowledge of the keystream and uses this knowledge to inject forged packets into the network.

Man-in-the-Middle Intrusion: It is a form of active eavesdropping, and this intruder creates independent connections with the two legitimate users. An intruder relays the message between the two communicating parties to inject malicious activities. The Man-in-the-Middle intrusion entirely controls the communication between two legitimate users. However, the users believe that they directly communicate with each other over a private connection securely.

1.3 Intrusion Detection Systems for WLAN Security

According to the security vulnerabilities, the 802.11 security protocols are inadequate to provide security in WLANs. Therefore, an external IDS of protection in Wi-Fi is essential to improve the growth of WLAN technology. An Intrusion Detection System (IDS) is a software to detect unauthorized access to a network. The IDS is device monitors the network continuously for malicious activity detection. The conventional IDS techniques widely apply the data mining techniques to extract the useful knowledge from a large number of network data. The classification and clustering algorithms are employed to deal with the unknown intrusions effectively. Most of the intrusion detection techniques exploit the data mining metrics such as false positive rate, false negative rate, and detection rate. A false alarm is a misclassification of a legitimate node as an intruder by the IDS. The complement of the false positive rate is referred as specificity. The false negative occurs when an IDS classifies an intruder in the legitimate class. In contrast, the true positive rate increases when an IDS classifies the intruders correctly.

1.3.1 Data Collection

The data collection phase enables the device to collect MAC features from the data sets, and it uses a set of optimal features to train the classifier to detect the intrusions regarding the MAC 802.11. The defense systems exploit the classification techniques to detect and classify the intrusions. In intrusion classification, all the features are not likely to be profitable to the IDS and the extraneous features have false correlations. This reduces the efficiency of IDS operations, and moreover, it results in high runtime and computational complexity of the IDS. To significantly improve the accuracy of detection, the conventional IDS techniques decide an optimal set of features, instead of considering all the extracted features. The feature reduction is the process of selecting a subset of relevant features to enhance intrusion detection accuracy. The correlation analysis process mines the correlation relationship between the features and constructs the optimal feature set. This process takes the entire MAC features as input and deviates the intrusion packets from normal using a simple distance-based heuristic measurement [12] [13]. However, this process results in high computational time and the storage requirement. The irrelevant and most relevant or redundant features cause noise in the classifier learning process, and this noise increases the false positive rate in the testing process. Recent techniques combine the filter and wrapper model to decide the set of optimal features. The filtering model estimates the information gain of each feature. According to the information gain, it ranks all the features, and consider the top-ranked features as an optimal set.

1.3.2 Data Analysis

The data analysis is a process of analyzing the set of optimal features and categorizing the packets under the specific classes. For data analysis, different intrusion detection techniques are employed. The IDS techniques are classified into the anomaly, specification, and signature-based IDSs [14]. The anomaly IDS techniques detect the features that are out of the usual activities [15]. The legitimate activities of a node are

stored in the history of the regular node activities. The principal advantage of the anomaly IDS is that it eliminates the need for specifying all known intrusions. One significant disadvantage of the anomaly IDS is the susceptibility to false positives. The specification-based IDS techniques identify the abnormal performance of the nodes, by defining the legitimate behavior of nodes as specifications. The main advantage of the specification-based IDS is a low false negative rate. The negative point is that this approach only reacts to known intrusions; but not identifies the unknown intrusions efficiently. Another main drawback in specification-based IDS is the effort required for defining the normal activities.

1.4 Intrusion Classification Techniques

The classification techniques exploit a set of labeled data instances to train a classifier and classify the test instances into predefined classes using a learning model. The anomaly-based IDS using a classification model operates in two phases. The training phase reduces the features into the optimal set and trains the classifier with optimal features. The testing phase classifies a test instance into legitimate or malicious with the support of trained classifier. The classification-based anomaly IDS technique exploits either one class or multi-class classifier. The one class classification techniques assume that all the instances taken for classifier training have only one class label. This kind of techniques decides the discriminative boundary around the total number of legitimate instances using one class classifier. If the test instance falls under the decided boundary region, the instance is classified under malicious class. In contrast, the multi-class classification-based anomaly IDS techniques assume that the training data includes multi-class instances. Such techniques enable a classifier to distinguish the features under legitimate and rest of malicious classes. Many intelligent classification techniques are proposed, and the most common existing classifiers are decision trees, Neural Networks (NN), navie Bayes and fuzzy set-based approach. The signature-based IDS techniques consider a runtime features that match a specific pattern of intrusions. The signature-based IDS results in a low false positive rate. Compared to others, the signature-based IDS techniques are more effective for detecting the unknown intrusions.

1.4.1 Decision Trees and Neural Networks

A decision tree is a decision support tool which is built in the model of the tree-like graph. A decision tree includes a set of nonterminal nodes and branches [16]. The terminal nodes represent a test on a attribute, whereas the branch represents the node relationship and its test results. Each leaf node represents a class label, and the classification rules are represented as the paths from the root to a leaf node. To classify the particular data item, the decision tree starts from the root node and follow the assertions until it reaches a leaf node. When a terminal node approaches, the decision is taken for the data item. A decision tree is a special form of a rule set, and the hierarchical organization of rules characterizes the decision tree. The decision tree algorithms work from top to down in multiple steps. To split a set of items accurately, the decision tree selects a variable at each step. For intrusion detection, the decision is deployed for differentiating the malicious nodes from the legitimate nodes.

The neural networks are a computing system which is inspired by the biological neural networks. Such neural networks learn the malicious packet features and classify the tasks without task-specific programming. The neural network is a set of connected input or output units, and each connection is assigned to a weight. By adjusting the connection weight, the neural networks learn in the learning phase and predicts the correct class label. An artificial neural network includes a connected set of processing units and weighted connections represent the unit interrelationship. The input and output nodes represent the subset of units, whereas the remaining nodes represent the hidden layer. Each activated input node is propagated through the hidden layers towards the output node. Most of the decision tree and neural network algorithms exploit traffic pattern matching technique for detecting the intruders. In a WLAN, each packet has nearly 155 features. Providing all the features as a reference to the decision does not improve the classification accuracy. Thus, the optimal number of WLAN features is selected and deployed as the decision tree rules or output units. Using decision tree rules or functional

units, the intrusion detection systems classify the network traffic according to the identified malicious packet features.

1.4.2 Naive Bayesian and Fuzzy Sets

The Naive Bayesian classifiers utilize the Bayes theorem to classify the new instances of a data sample [16]. Each instance has a set of features, and according to the feature values, the instance is assigned to a class. According to the Bayes theorem, the instance belongs to the class, when it returns maximum posterior probability for instance. For categorical data, the posterior probability is estimated as the frequency of the value over a total number of instances. For continuous-valued attributes, it is assumed that the instances are distributed without loss of generality according to the Gaussian algorithm. Moreover, the naive based Bayesian algorithm assumes that the attributes are conditionally independent with each other. Most of the spam mail and message classification techniques have exploited Naive Bayesian classifier, since they are least prone to the uncertainty. The limitation of the naive Bayes classifier is the requirement of the prior probabilities. To represent the uncertain information, the fuzzy sets are used. The fuzzy sets not only deal the incomplete, noisy, imprecise data, but it also assists in developing an uncertain model of the data. The uncertainty model for classifier provides an intelligent and smoother performance than conventional classification systems.

1.5 Scope and Motivation

The WLAN has become pervasive in recent years. Due to the ease of deployment and provision of high capacity, the WLAN provides convenient network access to the users. For instance, the users enjoy high-speed Internet access in airports, hotels, and coffee shops worldwide using WLAN. A Recent survey concludes that 549 million households around the world had installed a WLAN at 2016. Recently, Google has signed to provide free Wi-Fi at Indian railway stations to deliver the Internet service to hundreds of thousands of people. Besides, the importance of WLAN is growing every year; it also sets new demands for security. The WLAN is exposed to security threats. The security is

an essential issue with WLAN servers because they often share the sensitive data. For instance, a WLAN device connected to a medical data server transmits the confidential patient information across the network. Therefore, strong security is essential to prevent unauthorized users from monitoring this information. To prevent hackers from spoofing the server information to gain access to the network, the security measures must be considered in the security schemes. Recent research focuses on the classification-based intrusion detection systems against the WLAN intrusions.

1.6 Problem Statement

The IDS models commonly exploit the classifiers as detectors for identifying the 802.11 specific intrusions. Three main problems need to be solved for designing an efficient IDS are as follows. Firstly, all the MAC features are not advantageous for the IDS and the extraneous features contain false correlations that reduces the efficiency of intrusion detection process. In other words, the irrelevant and redundant features introduce noise in the learning process, and it increases the runtime and computational complexity of the IDS. The second main issue is the selection of classifier for identifying the known and unknown intrusions effectively. It is proved that the decision tree classifier is the most prevalent due to the characteristics of rule Interoperability. However, the randomly selected tie-breaking threshold reduces the accuracy of intrusion classification. The third problem with the IDS models is the bias information. The conventional IDS techniques measure the information gain using entropy and Bias compensation factor. The bias compensation factor is measured using an optimal feature set with many distinct values. Since some of the features have distinct values in nature, it tends to reduce the classification accuracy with biased information. Providing equal importance to the frequency of distinct values and difference of feature values diminishes the accuracy of intrusion detection systems. Both types of features have a different impact on the classification accuracy. It is essential to differentiate the high frequency of attributes with small distinct values and less frequency of attributes with large distinct values. The proposed methodology attempts to solve the previous issues and improve the intrusion detection accuracy.

1.7 Aims and Objectives

The prime objectives of the research work are as follows:

- To select a set of optimal features which are most relevant to intrusions with the knowledge of accurate bias compensation factor.
- To design an effective IDS based on an improved decision tree algorithm and to accurately detect the MAC 802.11 specific intrusions.
- To detect both the known and unknown intrusions accurately, by modifying the tie-breaking threshold in decision tree algorithm based on the bias compensation factor.

1.8 Contributions

The main contributions of the proposed work are as follows.

- The main contribution of the proposed Normalized information gain and Tie Breaking threshold-based Decision Tree (N-TBTDT) is to improve the detection accuracy of MAC 802.11 intrusions using Optimal Feature Selection and improved decision tree classifier.
- Optimal Feature Selection considers the Normalized Information Gain (NIG) and Normalized Bias Compensation (NBC) and facilitates the system to select highly appeared features with small distinct values, which is most relevant to the intrusions.
- Chaotic Particle Swarm Optimization (PSO) based feature selection properly controls the feature velocity and avoids the earlier standstill of the features before reaching the global optima. This assists the N-TBTDT to correctly clusters the intrusions under a set of optimal features.
- Dynamic tie threshold measurement based on bias compensation factor increments the decision tree when the new data packets arrive and successfully classifies the unknown intrusions.

- Finally, the performance of N-TBTDT system attains excellent detection accuracy with the reduced number of features compared to the Novel anomaly-based IDS for MAC Intrusions (NMI) work.

1.9 Organization of Thesis

The rest of the thesis is organized as follows:

Chapter 2 explains the different 802.11 standards in WLAN. This chapter discusses the effectiveness of IDS in WLANs security and the requirements of IDS in WLANs. It also discusses various IDS techniques along with its different processes. Besides, this chapter explains the data analysis and the data mining techniques used in the IDS models.

Chapter 3 surveys the MAC intrusions in WLANs. This chapter discusses the common WLAN protection mechanisms, and its drawbacks. It provides the literature survey of different IDS Techniques in WLAN security. This chapter also discusses the feature reduction techniques for and data mining approaches for IDSs.

Chapter 4 suggests an specification-based IDS technique, named as Normalized information gain and Tie Breaking threshold based Decision Tree (N-TBTDT). This chapter explains the process of optimal feature selection, which is based on the normalized information gain and the chaotic PSO algorithm. Moreover, this chapter explains intrusion classification using an improved decision tree classifier.

Chapter 5 evaluates the performance of the N-TBTDT system by comparing it with the NMI technique. It describes the experimental setup and different metrics that are used to evaluate the performance of N-TBTDT.

Chapter 6 concludes the thesis, and it also describes the future directions.

CHAPTER 2 BACKGROUND

This chapter provides details of wireless local area networks and discusses the effectiveness of IDS in WLAN protocols. It provides details about the requirements of intrusion detection systems in WLANs. Moreover, this chapter discusses the various types of IDS techniques. This chapter elucidates a detailed discussion of different data collection and analysis techniques. It discusses the importance of data mining techniques and explains the necessity of feature reduction techniques for WLAN. It describes the various types of clustering and classification techniques that are employed for detecting intrusions in existing IDS techniques.

2.1 Wireless Local Area Network (WLAN)

The IEEE 802.11 families of standards define the wireless local area networks. The WLAN technology enables multiple users to share resources concurrently and simplifies the networking process [1]. The examples of resources are a broadband Internet connection, network printers, data files, and audio/video streaming [17]. Compared to wired 10BASE-T network, the WLAN has the same capacity and provides the same speed without the complications that are associated with laying wire. Without having a connector cable throughout an office building or home, the devices can be connected using WLAN. The WLAN permits the devices to move freely anywhere in the office or home without Ethernet cables. The WLAN includes the following four fundamental architectural components.

Wireless Medium (WM): The wireless medium is used for transmitting the 802.11 WLAN frames between two wireless devices.

Distribution System: The distribution system permits the network to expand using multiple APs by enabling the wireless interconnection in an IEEE 802.11 WLAN.

Wireless Station: The wireless station is a device which utilizes the 802.11 protocol. The examples of the wireless station are smartphones, laptops, personal computers, personal digital assistant, and APs.

Access Point: The access point is a specialized wireless device, which connects multiple access points and connects the distribution system.

Basic Service Set: The Basic Service Set is a primary building block of WLAN, and it has a group of wireless stations. They connect with each other using the wireless medium. The capacity of signal propagation of the wireless medium decides the coverage area.

2.1.1 Different 802.11 WLAN standards

The WiFi family consists of a set of half-duplex over-the-air modulation techniques. The 802.11 is a first wireless networking standard in the family. The 802.11b is an is an amendment to the basic protocol. Moreover, 802.11b is the first widely accepted standard for the wireless communication, followed by 802.11a, 802.11g, 802.11n, 802.11ac, and 802.11ad. Different 802.11 standards of WLAN are listed as follows. Table 2.1 shows various amendments of 802.11 standards and their specifications. The 802.11b standard exploits the coding technique based on a direct sequence spread spectrum. In contrast, the 802.11a has operated on the more recently allocated band of 5GHz Unlicensed National Information Infrastructure (UNII). Using the same band used in 802.11b, the third modulation standard for wireless LANs, 802.11g extends the throughput up to 54 Mbit/s. Although the mobile device is compatible with the 802.11b, it reduces the speed of a network. To provide high data rate, the 802.11g chooses the primary modulation method such as Orthogonal Frequency Division Multiplex (OFDM). The 802.11n improves the 802.11g in the amount of bandwidth using the MIMO technology.

Table 2.1: Different WLAN Standards

Standard	Data Rate	Frequency	Channel Bandwidth	Range	
				indoor	Outdoor
802.11a	54 Mbps	5 GHz	20 MHz	100 ft	400 ft
802.11b	11 Mbps	2.4 GHz	20 MHz	100 ft	450 ft
802.11g	54 Mbps	2.4 GHz	20 MHz	125 ft	450 ft
802.11n	65 to 600 Mbps	2.4 GHz & 5 GHz	40 and 60 MHz	225 ft	825 ft
802.11ac	78 Mbps to 3.2 Gbps	5 GHz	20,40,80, and 160 MHz	90 ft	1000 ft
802.11ad	6.76 Gbps	60 GHz	2.16 GHz	30 ft	1000 ft

This standard provides a speed up to 300 Mbps. The main advantages of the 802.11n are the fastest maximum speed and the best signal range. However, this standard is not yet finalized, and moreover, multiple signal usage with MIMO concept greatly interfere with nearby 802.11b/g based networks. The newest generation of WLAN standard 802.11ac exploits the dual-band wireless technology and supports simultaneous connections on both the 2.4 GHz and 5 GHz Wi-Fi bands. The 802.11ac offers backward compatibility to 802.11b/g/n and provides bandwidth up to 1300 Mbps on the 5 GHz band. When longer ranges are essential for network devices, the 802.11ac standard is preferred. The primary goal of IEEE 802.11ad standard is to provide high speeds up to 7 Gbps. To achieve these speeds, the 802.11ad exploits the 60 GHz ISM band, and it ensures the fewer interference levels. A new standard called as 802.11 ax is expected to replace current standards by the end of 2019. The 802.11 ax is known High Efficiency WLAN (HEW) for both 2.4G and 5G and is also expected to be backward compatible with other standards.

2.1.2 MAC 802.11 specific Intrusions on the WLAN

All 802.11 standards are vulnerable to replay intruders, and the replay intrusion includes the messages of the probe, associate, authenticate, disassociate, and deauthenticate users

from the WLAN [18][19][20]. The intrusions on protected WLAN are mainly divided into three broad categories such as injection, impersonation, and flooding. The injection intruder simply monitors the specific packets and then proceed with the key cracking process offline. Moreover, the impersonation intrusion enables the malicious devices to crack the secret key and, exploits the key stream to forge and inject packets into the network. In flooding attack, the intruder is likely to make service unavailable through fake packet flooding in the network. Table 2.2 shows the different types of MAC 802.11 intrusions.

Table 2.2: Types of WLAN Intrusion

Intrusion	Type	Impact on	Effect	By target	Attack Severity	Remark
Injection	ARP-Injection	WEP	Key Cracking	Network	Moderate	Requires resources
	Chop-Chop		Packet Decryption		Moderate	Requires upto 64 bits
	Fragmentation		Packet Decryption			
Impersonation	Caffe-Latte	WEP	Secret Key Cracking		High	Not possible in all operating systems
	Hirte		Secret Key Cracking		High	Fast key cracking
	Honeypot		Loss of Privacy	Client	Moderate	Targets WEB protected devices only
	EvilTwin		Loss of Privacy		Moderate	Requires Secret key knowledge

Flooding	Deauth	Upto 802.11n	Connectivity Loss		High	Affects all standards
Flooding	Disassociation	Upto 802.11n	Connectivity Loss	Client	High	
	Block Ack	802.11n	Annoyance		High	Requires accuracy
	Fake PS	Upto 802.11n	Annoyance		High	
	Authentication Request	Upto 802.11n	Inability to join the network	Client	Low	Ineffective in most secured devices
	CTS and RTS flooding	Upto 802.11n	Annoyance		Low	Target Client resources
	Probe Request dAnd Response	Upto 802.11n	Annoyance		Low	Requires Secret key knowledge
	Rogue AP	None	Privacy Loss		Moderate	Requires wired n/w support

2.1.3 Effectiveness of IDS in WLANs Security

WPA is the second generation of the security mechanism, and it provides a powerful encryption mechanism like temporal key integrity protocol. Thus, it can provide a reliable security against eavesdropping intrusions. Subsequently, the other version of 802.11i security protocol, called WPA2 is introduced with the aim of providing security to the users. Both the WPA and WPA2 provide authentication and integrity to the users.

However, they permit the WLAN devices to access the firmware of AP directly, and it is vulnerable to password cracking and brute force intrusions. Even though, the cryptographic techniques improve the data confidentiality during communication, it incurs an additional computational power and imposes high latency. An intelligent intruder utilizes the communication delay to deny or modify the secure keys and to inject the malicious behavior into the network using the keys. It necessitates the implementation of external protection schemes like intrusion detection systems in WLAN [23-31]. An IDS is a tool employed to detect the unauthorized access to a system. Due to the lack of physical boundaries in the network, an intrusion is likely to be perpetrated from anywhere, and several intrusions exploit this vulnerability to undermine the integrity and security of the network. The conventional wired IDS system does not ensure absolute security. With an issue of wireless security, implementing a wireless IDS system is the optimal solution to detect the WLAN attacks. The comparison of WPA, and WPA2 security protocols are listed in Table 2.3.

Table 2.3: Comparison of WLAN security protocols

Security Protocols	**Encryption Keys**	**Key Rotation**	**Authentication**	**Key Update Function**	**Key Distribution**	**Attacks and Vulnerabilities**
WPA	Temporal Key Integrity Protocol (TKIP)	Dynamic Session keys	802.1x and EAP	Available	Automatic distribution	Chopchop, DoS attacks
WPA2	Counter Mode with Cipher block Chaining Message Authentication Code Protocol (CCMP)	Dynamic Session keys	802.1x and EAP	Available	Automatic distribution	DoS attacks due to unencrypted management and control frames, MAC address spoofing due to Deauthentication, Offline dictionary attacks in WPA2-Personal

2.2 Types of Intrusion Detection Techniques

To protect the WLAN from intrusions, external intrusion detection techniques are used [32]. The conventional IDS models distinguish the abnormal activities of devices from the normal activities and discover intrusions successfully. There are three main techniques have been used for classifying the intrusive activities. Those are Misuse, anomaly, and specification-based intrusion detection systems.

Misuse Based IDS: The misuse-based intrusion detection is also named as signature-based detection scheme. This kind of IDS techniques generates a signature for the previously known intrusion. Instead of analyzing the legitimate behaviors of devices, the device misbehavior is measured through the similarity of activities to the signature. To identify the intruders, the misuse-based IDS exploits the signature as a reference. The device activities against the rules of networking are defined as anomalies. Mostly, the misuse-based IDS exploits the information of Interval, Retransmission, Integrity, and Delay as rules. This IDS model can accurately detect and isolate the well-known intrusions. However, it fails to cope up with the new intrusions in WLAN. Thus, the misuse-based IDS model is lightweight intrusion detection, and it is suitable for resource-constrained mobile devices in MANET. The detection of novel attacks is not possible. Another difficulty is that signatures must update all the possible variations of the pertinent intrusion. Otherwise, the intrusion detection accuracy is reduced drastically.

Anomaly-Based IDS: The anomaly-based IDS techniques solve the issues in misuse detection techniques [33][34]. Instead of malicious behavior, this kind of IDS techniques focuses on the normal behaviors. Initially, the activities of the legitimate device are described in the intrusion detection system and the IDS flags as an intruder when the device activities are varied from the defined legitimate behavior in a statistically significant amount. It paves the way for detecting the novel intrusions also. There are two problems associated with the anomaly-based IDS in WLAN. Firstly, a device performs legitimate functions, but previously unseen behavior. This tends the IDS to flag the

legitimate device as an intruder and it increases the substantial false alarm rate. Secondly, if an intrusion does not exhibit an anomalous behavior, it is not identified as the intruder.

Specification-Based IDS: The specification-based IDS model utilizes the advantages of both the misuse and anomaly-based IDS models. It defines the normal behavior of legitimate devices and identifies the intrusions based on the deviation from normal behavior. The malicious activities which are previously unseen behaviors also not incur a high false alarm rate. Since it is decided according to the deviations from legitimate behaviors, it can still detect previously unknown intrusions. The IDS verifies the activities of every device in the network. The main difference of specification-based IDS from the anomaly model is that the first model identifies the malicious activities and the latter identifies the already known malicious activities. Due to the time-consuming process of specification and constraint selection, the cost of specification-based IDS model is high.

2.3 Conventional Data Collection Processes

The effectiveness of the data collection mechanism is essential because the detection accuracy of the system against intrusions is based on the collected data. The data collection process degrades the accuracy of intrusion detection when the data is incomplete. If the data is acquired with incorrect information, the false sense of IDS occurs. These problems have been identified with conventional data collection processes. The conventional data collection mechanisms exploit two different methods such as behavior-based data collection and traffic based data collection [35][36][37].

2.3.1 Host and Network-Based Data Collection

Most of the intrusion detection systems detect the intrusions based on the actions performed by the intruders in a host. The intruders launch malicious activities through command execution, service access, and improper data provision. Most of the intrusions are performed on the end host, except the DDoS intrusion since the data flooding

intrusion is performed only on the client devices to prevent the legitimate packets from flowing. Most of the existing intrusion detection systems claim that they detect the intrusions at the end hosts. For example, an IDS identifies a ping flooding intruder at the ICMP layer by observing the frequency of echo request packets. For flooding intrusions, the network-based data collection is suitable. However, for other intrusions, the host-based data collection methods are efficient, compared to the host-based data collection model.

The main advantages of using a host-based data collection model for intrusion detection system are as follows.

• Host-based data collection allows the collection of data with respect to what is happening on the host, instead of observing the data packets flowing into the network for intrusion detection.

• In high traffic networks, an efficient data collection is impossible due to high packet loss caused by the high traffic, whereas properly implemented host monitoring application report every single packet transferred to each host.

• Network-based data collection mechanisms are subject to intrusion insertion and evasion. These problems do not occur on host-based data collection.

2.3.2 Direct and Indirect Monitoring for Data Collection

The direct network monitoring collects the data from the device which generates the packet, or which receives the packet [38-40]. In contrast, the indirect data collection obtains the data from a source device, since the source exactly reflects the behavior of the object that is being monitored. Indirect monitoring of the host observes the log file to train the classifier for intrusion detection. The indirect monitoring is performed by monitoring the packets destined to the appropriate ports in the host. The advantages of direct monitoring compared to the indirect monitoring for data collection are listed as follows.

• An intruder can alternate the data gathered using an indirect data collection model before the IDS software or hardware exploits the data for classifier training and intrusion detection.

• Some data packets may not be captured using an indirect data source. For example, not every action performed by a device gets recorded in a log file. Furthermore, indirect data monitoring applications are incapable of accessing the specific internal information on the object being monitored. For example, TCP-Wrappers is incapable of examining the internal operations of the device and the monitoring application can only obtain the data passed through the external interface of the device.

• With indirect monitoring, the data is generated by mechanisms that do not know the needs of the intrusion detection system that will be using the data. For this reason, indirect data sources usually carry a high volume of data. For example, every packet transmission builds 50K-500K records in a device.

• Compared to the direct data monitoring model, the indirect data sources collect more irrelevant data to the intrusions and tends the IDS to spend more resources on filtering and reducing the data, before utilizing the log file knowledge for detection purposes.

2.4 Data Analysis Techniques

Any unapproved activities of mobile devices to degrade the performance of WLAN or disclose the data confidentiality is named as an intrusion. A collection of tools and algorithms are used to design an intrusion detection system. Several IDS techniques are proposed in WLAN. According to the IDS models, the conventional IDS can detect the intrusions using pattern matching, or data mining, or hybrid algorithms. The IDS techniques involve mainly three steps for detecting the intrusions.

• **Defining Features and Extraction:** To categorize the data, analyzing and identifying the features that are relevant to the intrusions is essential. Initially, the dataset is

partitioned based on the selected features and the rules are generated from features [41] [42].

• **Rules Extraction**: By applying the appropriate method based on either pattern matching, data mining, or hybrid method, the partitioned data are analyzed to find the common rules in the data set.

• **Applying Rules for Intrusion Detection**: After defining the rules in the dataset, the device activity which is going out of the predefined rules is classified as intrusions.

2.4.1 Pattern Matching Analysis Techniques

The pattern matching algorithms exploit the signature of intrusions and identify the malicious activities performed in the network. The pattern matching considers the previously identified intrusions as a reference and differentiates the intruders from legitimate devices. The significant advantage of IDS is high detection accuracy with a low false positive rate, due to the predefined intrusion signature. However, it lacks in identifying the unknown patterns, and those patterns are misclassified as normal patterns. The limitation of pattern matching is that it can only detect the already known intrusions, and it cannot identify the new malicious activities. To detect the new intrusions, it is essential to update the signature records constantly. The following rules are generated by the signature-based IDS to monitor and identify the network anomalies.

• **Interval rule:** Difference between the arrival time of two consecutive data packets at a time.

• **Retransmission rule:** The selected devices are involved in the data forwarding.

• **Integrity rule:** The receiver ensures that the received data packet does not deviate from the original packet generated by the sender.

• **Delay rule:** The routing nodes forward the received packets immediately when the transmission medium is idle.

• **Radio transmission range:** The mobile device must receive the messages only from the neighboring devices.

• **Jamming Rule:** Collision must be maintained in a certain limit.

2.4.2 Data Mining Techniques

Data mining techniques aim at extracting the hidden knowledge from the collected data. The hidden knowledge is represented as patterns, relationships, groups or classes. Instead of processing the bulk data periodically, the data mining methods, process the bulk amount of data at once and identify the hidden knowledge. The retrieved knowledge is used as a reference to identify the intrusions in the future. Different data mining techniques are used for extracting the known intrusion behavior from the database of collected records. The data mining-based IDS models apply clustering and classification algorithms to select an optimal feature set and classify the intrusions. Both the clustering and classification techniques in data mining are widely applied to a variety of real-time applications and can deal with a significant amount of data.

• **Clustering:** Clustering defines the process of grouping the similar data into a class using the intrinsic similarity characteristics of data. Several clustering techniques are used in the data mining applications, and each clustering technique has different classification rules. Based on the desired output, the clustering technique is selected. There are two types of clustering techniques such as hierarchical and non-hierarchical. The hierarchical method builds a large cluster, in which a set of nested clusters is nested. In the non-hierarchical method, a large data are divided into multiple clusters, with or without overlapping.

• **Classification:** Classification takes each instance of a dataset and assigns it to a class. It extracts models to define a data class. Such models are called classifiers. The classification model classifies the device activities into anomalous behavior and normal behavior. The commonly used data mining classification techniques are a naive Bayesian classifier, Support Vector Machine (SVM), and decision tree classifier. The classification techniques build a learning algorithm using a collected data as the training set. The conventional data mining consists of different classification methods, and each method

has its advantages and disadvantages. Most of the classification algorithms involve following steps:

- Data sample collection and training
- Features retrieval and class identification
- Optimal feature selection for classification
- Learning model based on training
- Classifying the known and unknown intrusions

The learning model efficiency depends on the feature selection. During this process, the set of attributes or features are extracted. There are two models in the literature for feature selected such as the filtering and the wrapper module. In the filtering model, information, gain, correlation coefficient, and normalized gain is used for selecting the set of optimal features. The wrapper model selects an optimal feature set based on the predictive accuracy of the learning model. Table 2.4 shows the comparison between various data analysis techniques.

Table 2.4: Comparison of Different Data Analysis Techniques

Method	Concept	Advantages	Disadvantages
Pattern Matching	IDS matches the existing patterns with the incoming traffic patterns to detect the known intrusions	Simple to implement and less resource consumption	New intrusions cannot be detected using existing patterns
Classification	Applying the classification models to detect the intrusions when the incoming traffic is high. It extracts	This automated process detects both misuse and anomalous intrusions.	High resource consumption and needs human intervention

	different type of knowledge and classify the intrusions using the extracted knowledge		
Clustering	The training phase is used for learning and testing phase is used for detecting the intrusions in the incoming traffic.	Used for detecting both misuse and anomalous intrusions.	High time consumption

2.5 Data Classification Methods in IDS Design

There are three different classification techniques such as supervised, semi-supervised, and unsupervised learning models, which are applied in intrusion detection systems [42]. In the supervised learning algorithm, the training data is analyzed to produce an inferred function and to map with the incoming traffic. The semi-supervised learning is a subclass of supervised learning, and it uses a small amount of labeled data to a large amount of unlabeled incoming traffic.

2.5.1 Supervised Learning

Supervised classification builds a model and differentiates two classes according to the selected optimal numerical features with minimal errors. To build that model, the classifier exploits the dataset with labeling features, and moreover, the dataset must include both normal and intrusions samples. Supervised learning model provides classifier with more information compared to the semi-supervised and unsupervised techniques. This improves the detection accuracy significantly. However, supervised learning faces some problems. (i) The given datasets in the training time uncover all

legitimate aspects. (ii) It cannot guarantee the accurate labels, particularly when the data set contains noise and incorrect information.

2.5.2 Semi-Supervised and Unsupervised Learning

Semi-supervised learning models work in-between supervised and unsupervised methods. In the application of real-time anomaly intrusion detection, the semi-supervised method is more practical. The labeled data of the normal class is the only requirement. However, such method is not widely used. Since the labels for possible anomalies in the training time is mostly available. Unsupervised learning does not require labeled data, and it partitions the data into normal and anomalies using statistical models, and moreover, it does not require any prior knowledge. However, it assumes the following (i) It presumes that the normal data size is high, and the intrusions represent a very small amount data. (ii) The normal and intrusion data are different from each other statistically. Table 2.5 shows the different machine learning methods with its advantages and limitations.

Table 2.5: Comparison of Different Classification Technique

Machine Learning Technique	Learning Method	Features	Advantages	Limitations
Decision Tree	Supervised and can be adopted Semi-supervised	1. Interpretation is easy for data models 2. Support both discrete and continuous values 3. It can deal with noise data	1. Decision Tree works well with heterogeneous data 2. Better detection accuracy	1. Small variation leads to increased number of leafs 2. Small training data is inefficient

SVM	Supervised and can be adopted Semi-supervised	1. Handling classification issues, even if the collected data are not linearly separated by the features	1. Insensitive to the dimensions of incoming data 2. Better learning ability	1. Training phase consumes more time 2. It does not provide additional information about the detected intrusions
Naive Bayesian	Supervised and can be adopted Semi-supervised	1. Solves classification and prediction issues by considering the probabilistic relationships between among the features	1. Incorporating both prior and current data improves the detection accuracy	1. It does not handle continuous features 2. Accuracy of prior knowledge decides the detection accuracy
Neural Networks	Supervised and can be adopted Semi-supervised	1. Few features can be adjusted without requiring the reprogramming	1. It can handle noise and incomplete data	1. Slow learning 2. High processing time during testing phase
K-Nearest neighbor	Unsupervised	1. Support multi-model classes 2. Handling classification	1. It can handle noise data	1. High time consuming when the training data

		issues, even if the features do not linearly separate the collected data		size is high 2. Accuracy depends on the number of dimensions in data

SUMMARY

Chapter 2 has provided a background information for WLAN and security protocols. It has also discussed different types of 802.11 standards. The common security mechanisms for 802.11 security protocols such as WEB, WPA, and WPA2 are addressed with its unsolved issues. This chapter explained the need for an IDS protection over WLAN and discussed various IDS techniques. The types of data collection and analysis techniques in IDS design has also been discussed. Besides, this chapter has provided a detailed survey of different types of classification models used in intrusion detection systems.

CHAPTER 3 LITERATURE REVIEW

This chapter surveys the MAC intrusion in wireless local area networks and discusses the basic security schemes in WLAN. It provides a detailed review of different types of IDS techniques, such as anomaly, specification, and signature-based IDSs. This chapter elucidates a detailed discussion of feature reduction techniques. This section discusses the filtering, wrapping, and hybrid models in feature reduction. It describes the various types of data mining techniques that are employed for detecting intrusions in existing IDS techniques.

3.1 Survey on MAC intrusions in WLAN

In WLANs, the intruders perform malicious activities to damage the data integrity and performance of the network. Due to the rapid enhancement of WLAN, there is an ever-growing risk of security and privacy. Conventionally, three basic security schemes are developed for WLANs. The conventional security schemes are not sufficient to maintain the network security in a WLAN, due to the arrival of intelligent intrusions. The IDS is a defense mechanism, and it decides whether the data traffic is normal or malicious using data mining techniques. There are different IDS techniques which are being used to identify the various types of intrusions in WLANs. The survey in [21] presents an elaborate taxonomy of intrusions to the network, and it provides a security architecture for a wireless network. This taxonomy provides a systematic approach and analyzes all the security intrusions in the wireless network. The most common MAC intrusions in WLAN are DDoS, Eavesdropping, and so on [25]. The work in [43] provides an overview of WPA/WPA2 and reviews various types of intrusions against WPA/WPA2. The survey on security vulnerabilities of IEEE 802.11 WLANs is provided in [2]. The work in [2] proposes two techniques to enhance safety and to overcome some known vulnerabilities. The work in [7] discusses the major intrusion types concerning IEEE 802.11 family of networks, particularly the latest 802.11i security standard. It elaborates on 802.11i specific intrusions, and it experimentally investigates the mitigation of 802.11 specific intrusions by properly designing an IDS. Intrusions in a WLAN is classified into

two types such as active and passive intrusions. The main aim of active intrusion is to disturb the functions of a network, whereas the passive intrusions overhear the wireless medium without disrupting the normal network activities.

Active Intrusions: An intruder actively participates in the data forwarding to disturb the normal operation of the WLAN. An intruder modifies the contents of the data packets or introduces false information into the network. The WLAN is highly vulnerable to the MAC Layer Denial of Service, DoS intrusions [27]. Moreover, there is no complete solution to prevent MAC Layer DoS intrusions, as most of the conventional algorithms provide a partial settlement of the problem.

Passive Intrusions: The passive intrusion attempts to hack the secret information, transmitted or received on the air medium. These types of intrusions are usually tough to detect, as the intruder does not modify or disturb the data flow in the network. Table 3.1 illustrates the active and passive intrusions in WLAN along with the defensive scheme.

Table 3.1: Different Types of WLAN Intrusions

Attack	Type	Impact	Defense System
Unauthorized access	Active intrusion	Without obtaining access rights from the data owner, viewing the secret data of others	Data encryption
Rogue Access Point		Without the authorization from administrator, the rogue AP is installed to collect the information from devices	Intrusion alerts and containment mechanism

Man in the middle attack	Active intrusion	Data capturing during communication, when there is a necessity to use the third party	Data encryption
Denial Of service attack		Due to unnecessary message flooding, authorized users are prevented from accessing the network services	Authentication
Replay attacks		Data transmitted by the authorized users is maliciously or fraudulently repeated or delayed	Ensuring the data freshness using time and sequence number
Session Hijacking		Attack on a particular user session over a protected network	Comparing each session to a single IP address
Traffic Analysis	Passive Intrusion	Inferring data from multiple transmissions	Authentication
Eavesdropping		Overhearing the data transmission to obtain a confidential data	Direct sequence spread spectrum

3.2 Common WLAN protection mechanisms, and its drawbacks

Due to the security vulnerabilities and intrusions, the provision of WLAN protection is difficult. Conventionally, several security solutions have been developed with the aim of avoiding the intrusions against MAC 802.11 networks. The evolution of security standards for WLAN begins with 802.11. The basic encryption standards such as WEP, WPA, and WPA2 has some security weaknesses, because both the WEP and WPA standards ignore the authentication [28]. To identify various types of DDoS intrusions against the control frames of MAC layer in WLANs, new protocols are proposed to improve the WEP and WPA [44]. The Rivest Cipher 4, RC4 algorithm is used for encrypting the data packets and the secret keys are exchanged using initialization vector (IV). The mutual authentication mechanism is implemented to provide an authentication between AP and wireless devices. In WEP and WPA, an implementation of the RC4 encryption technique ensures strong data privacy to them. However, the WEP fails to protect the network from eavesdropping. With WEP protected networks, it is possible for an intrusion to identify the key of the knowledge of the keystream alone. There is a possibility to use the keystream to forge and inject the malicious packets, and this tends to more severe intrusions in WLAN. Due to the proliferation of readily available hacking tools, the WEP protected WLAN is insufficient for securing the enterprise-wide distributed processing environments. To overcome the limitations of WEP, the IEEE 802.11i and WPA2 are introduced. The 802.11i/WPA2 includes the new features of AES, message integrity, and fast-roaming support. Moreover, the vendor interoperability and forward and backward compatibility are consistent with the IEEE and Wi-Fi Alliance. The WEP, WPA, and WPA2 security protocols are compared in Table 3.2.

Table 3.2: Features of WLAN Basic Security Schemes

Security Property	**Scheme**		
	WEP	**WPA**	**WPA2**
Algorithm	RC4	TKIP	CCMP and AES

Size of security key	40/104 bits	1. 128 bits (encryption) 2. 64 bits (authentication)	128 bits
Key Mixing	Concatenation with base key	TKIP mixing function	TKIP mixing function
Packet Key	Concatenated	Mixing function	Mixing function
Key management	None	EAP Based	EAP Based
Header integrity	None	MIC	CCM
Data integrity	CRC-32		
Replay intrusion	No security scheme	Sequence-based replay intrusion detection	

3.3 Review of IDS Techniques in WLAN

Intrusion detection is the second level of defense mechanism for protecting the WLAN from intrusions. An IDS performs the intrusion detection process using three different methods. Based on different detection methods, the existing IDS techniques are classified into three types such as anomaly based, misuse based, and specification-based IDS [32]. Several IDS techniques have been proposed in each method to detect intrusions in WLAN.

3.3.1 Based on Anomaly Model

The anomaly-based IDS model depends on the heuristics or the rules, rather than using patterns or signatures, and it identifies any misuse that falls out of legitimate node activities. In [34], a node-based anomaly IDS detects the anomalies with the support of cross-feature analysis technique. The packet features used in MAC layer are used for

deviating the malicious behavior of mobile devices from normal behavior. An anomaly-based IDS technique detects the unknown intrusions. The anomaly detection technique detects packet dropping intrusions using the MAC features that are selected as an optimal set [45]. A novel anomaly-based intrusion detection approach in [46] extracts interpretable fuzzy IF-THEN rules from network data traffic and applies it for classifying the anomalies from legitimate traffic. The fuzzy rule-based system is used for wrapping the MAC features and searches for an optimal feature subset for reducing the dimensionality of the input data. By applying diverse baseline filters, the fuzzy rule-based system retains only the relevant features from the entire features and removes the irrelevant features. Thus, this system improves the classification accuracy. An online k-means algorithm in [47] clusters the network traffic to learn the classifiers and to detect the intruders. Specifically, the k-means clustering considers the distance from the largest cluster, and the main advantage of the k-means based IDS is the selection of nominal features, instead of numerical features. However, the marginal detection rate and strong assumptions of most legitimate network activities are the major drawbacks. Table 3.3 illustrates different data mining techniques with its advantages and limitations.

Table 3.3: Compilation of IDS Techniques

Technique/Dataset	Data Mining Methods used	Process	Advantages	Limitations
Chitrakar, Roshan, and Chuanhe Huang [48] (Kyoto 2006+)	1. SVM Classification 2. K Medoids clustering	1. K Medoids clustering clusters the similar data 2. SVM is used for classifying the incoming traffic	1. Better accuracy in the presence of noise data	The time complexity is more when the dataset is very large

Chitrakar, Roshan, and Chuanhe Huang [49] (Kyoto 2006+)	1. K Medoids clustering 2. Naive Bayesian	1. K Medoids clustering clusters the similar data 2. Naive Bayesian is used for classifying the incoming traffic	Mean time of false alarm rate is reduced	Hard to predict when a Naive Bayes classifier is used for high dimensional data
Fu, Song, Jianguo Liu, and Husanbir Pannu [51] (Captured own Data set)	1. One class and two class SVM	1. One class SVM measures abnormality score 2. Two class SVM classifies the new instances	There is no need to provide a prior failure history as well as it modifies the learning model based on the observed failure events	High false positive rate under large scale networks
Farid, Dewan Md, et al. [50] (KDD99)	1. Naive Bayesian 2. Decision tree	1. Both provide balance intrusion defections	Maintain false positives at acceptable level	Performs poor under high dimensional data
Yasami, Yasser, et al. [53] (Captured own data	1. K means 2. ID3 Decision	1. K means used in training phase 2. Decision Tree is	Outperforms the individual k-	This approach is limited to

set)	Tree Learning	used in testing phase	Means	specific intrusions.
Peddabachigari, Sandhya, et al. [54] (KDD cup 99)	1. Decision Tree algorithm 2. SVM	1. Decision tree groups the similar data 2. SVM is used for classifying the incoming traffic	Handles the noise and incorrect data	The results are equals to the individual SVM classifier
Peddabachigari, Sandhya, et al. [54] (KDD cup 99)	1. Ensemble approach	1. Results from different classifiers are combined to take the final decision	Outperforms the individual classifier	Manual intervention is necessary for base classifier selection.

The algorithms in [48][49] utilize the k-Means algorithm. It differs in the representation of the various clusters. Each cluster is represented by the most centric object in the cluster, instead of using the implicit mean value. Moreover, the k-medoids method provides better results than the k-means algorithm in the presence of noise and outliers. Compared to the mean value, a medoid is less influenced by outliers. This method detects network anomalies and produces much better results than k-Means. The machine learning classification tree models in [50][51][52][53] are also called as a prediction model or decision tree. The classification tree models follow the tree model graph structure, where the internal nodes, branch, and the nodes represent the test property, test result, and the class to which any data belongs respectively. Conventionally used classification tree models are ID3 and C4.5. There are two methods for tree construction such as top-down tree construction and bottom-up pruning. The previous decision tree classification models belong to top-down tree construction. When compared to the naive Bayes classification, the result obtained from decision trees is more accurate.

3.3.2 Based on Specification Model

The specification-based detection approaches predefine the legitimate network activities, and it labels the packets that do not match the specifications as intrusions. A

specification-based IDS engine is built according to the functionality and limitations of the 802.11 MAC protocol [55]. The IDS engine deployed at each node performs the intrusion detection process using a set of specifications that describe the proper operation of the MAC protocol. Also, the IDS can effectively identify both known and unknown intrusions in real time in considerable overhead. The Behavioral Monitoring Specification Language (BMSL) is proposed in [56] to specify both legitimate and malicious behaviors for an IDS using traffic and behavior collection. The BMSL models the information and sequence of events happened in the network. The BMSL programs are converted into intrusion detection engines. The combination of BMSL approach and specification-based IDS provides the detection rate of a signature-based IDS. However, the requirement of extensive attack dictionary eliminates the advantages of specification over signature-based designs. The work in [57] presents Extended Finite State Automata (EFSAs) and establishes a specification for a traffic-based IDS. The performance of EFSAs is evaluated using DARPA/Lincoln Labs data sets. The EFSA approach is combined with the anomaly-based intrusion detection. It exploits an unsupervised machine learning to classify the legitimate device activities from the anomalies. As IDS needs to execute a processor and core memory intensive machine learning module, the anomaly-based designs incur high computational complexity.

3.3.3 Based on Signature Model

Unlike anomaly-based IDS, the misuse based IDSs depend on the use of specifically known patterns of malicious behavior, and it can detect only the known intrusions. The work in [58] considers the signature-based IDS technique to facilitate intrusion detection over wireless networks. In [59], the sneeze algorithm is used in identifying the intruders on a WLAN and isolate them from the network. It follows the biomimetic approach to detect and expel an intruder at the network edge. The basic idea is to enable the APs to monitor one another by searching for unrecognized network activity. If an AP running Sneeze determines an unrecognized activity, it applies changes to the key and notifies the system administrator. The response system includes the rekeying and the human searching the area surrounding the reporting AP to remove the rogue AP. Sneeze exploits the direct data collection process and simple pattern matching. The sneeze focuses on

intrusions involving a rogue AP/man in the middle. Table 3.4 discusses different IDS techniques along with its collection approaches.

Table 3.4: Comparison of various Types of IDSs

IDS techniques	Type	Collection Approach	Analysis
clustering-based IDS [60]	Anomaly	Network	Data Mining
Signal print-based IDS [61]	Anomaly	Network	Pattern matching
Multi agent-based IDS [62]	Anomaly + Signature	Direct + Network	Pattern matching
Lightweight agent-based IDS [63]	Anomaly + Signature	Network	Pattern matching
Four-layer IDS model [64]	Anomaly + Signature	Network	Data Mining
Biomimetic approach-based IDS [65]	Signature	Network	Pattern matching
Deterministic DCA [66]	Anomaly	Network	Data Mining
Adaptive Bayesian Based IDS [67]	Anomaly	Network	Pattern matching
Temporal Signatures [68]	Anomaly	Direct	Pattern matching
Signatures Apriori [69]	Signature	Network	Combined
Host based IDS [70]	Hybrid	Direct	Pattern matching
Peer based IDS	Signature	Direct	Combined
Specification based IDS [71]	Specification	Hybrid	Pattern matching
Specification based Anomaly detection [72]	Specification	Direct	Pattern matching
Adaptive intrusion response [73]	Hybrid	Hybrid	Pattern matching

A combined anomaly and signature-based IDS are proposed in [63]. The hybrid scheme exploits the direct data collection scheme. The two-tier analysis function executes the signature and the anomaly-based detection modules parallelly and builds the first stage. If the individual IDS models cannot classify the incoming traffic as an intrusion or legitimate, the hybrid scheme goes to the combined results. The advantage of that scheme is the use of realistic dataset for testing. However, lack of numerical results is a primary reason behind the false positive rate. The Multi-Agent Based Distributed WLAN IDS (MABDIDS) in [62] address the lack of IDS interoperability in the design of an efficient IDS. The two-tier analysis function exploits the Data Analysis Agent to perform coarse detection. The response agents are used for efficient and effective coordination of data collection on possible intruders. The MABDIDS follows a distributed design. However, the numerical results are not used, resulting in false positive rate. Table 3.1 shows the comparison between various IDS techniques.

3.4 Feature Reduction Techniques for IDS

The feature reduction is a process of selecting optimal features, and it is executed, before applying a learning algorithm. The reduction of the feature space reduces the complexity of training model and improves the classification accuracy of the classifier. There are two common methods for feature reduction such as filtering and wrapper method.

3.4.1 Filtering Techniques

Two feature selection algorithms are proposed in [74], and it provides a comparison of those algorithms with a mutual information-based feature selection technique. The feature goodness measurement such as correlation coefficient and mutual information for selecting feature is mainly used as the filtering measures. Moreover, it runs an IDS based on the machine learning algorithm, named as Least Squares SVM. The work in [75] exploits the Genetic Algorithm (GA) for selecting a set of optimal features from such large spaces and enhances the detection accuracy with reduced false positive rate [75]. Due to the combination of induction algorithm and IDS, and moreover the repeated

execution of the algorithm, the genetic algorithm consumes more time. In [76], a vital set of features is selected using the sequential backward technique. In each iteration, one irrelevant or redundant feature is removed, and such iterations tend to produce the optimal features and improve the detection accuracy. Despite the assurance of prominent performance of IDS using feature selection, there are some inherent complexities in the detection and classification of intrusions such as diversity among selected features. To tackle the problem of feature selection, diverse techniques have been proposed. The Gradually Feature Removal (GFR) method in [77] is used for selecting the number of optimal features among total features of the KDD Cup dataset. The GFR applies filtering method to remove the irrelevant features gradually. Because the irrelevant features do not contribute to the accuracy of the SVM classifier. A feature Vitality Based Reduction Method in [78] detects the crucial features using three different subset selection techniques such as correlation measurement, Information Gain, and Gain Ratio. The use of optimal features in Naive Bayes classifier improves the accuracy of intrusion detection. However, the filter approach has some drawbacks. The filter methods fail to consider the accuracy of the classifier. In other words, it identifies the features using the metrics such as metrics are gain, information gain, correlation coefficient and normalized gain. Those metrics represent the impact of a feature on attack classification. In this method, each feature is measured separately and thus does not consider the feature redundancy. The features with the same impact on classification accuracy are known as redundant features. Lack of considering the feature redundancy makes a negative impact on the performance.

3.4.2 Wrapper Techniques

The wrapper method employs the intended learning algorithm and identifies the best set of features. The filter method exploits either information gain, correlation measurement, and gain ratio to evaluate the optimal features according to the heuristics-based measurements. Unlike filtering models, the wrapper methods depend on the classification methods [79]. The works in [43][60] consider the standard MAC header features as input for learning the process of IDS, resulting in high computational time complexity and the

storage area of IDS. Consequently, the detection accuracy of IDS decreases drastically. Moreover, the selection of irrelevant and redundant features increases noise in learning the process of an IDS and degrade the attack detection accuracy of the IDS. The feature selection is the most crucial model to build the intrusion detection system. In feature selection, an optimal set of features that have a greater intrusion detection tendency is selected to construct the suitable detection algorithms. The wrapper approach in [80] [81] exploits genetic algorithms for intrusion detection. In [16] and [82], a naive Bayes and decision tree algorithm exploit a wrapper-based approach, and they reduce the features in n iterations. Moreover, the best choice has been used to generate rules and the tree that has high sensitivity and specificity are identified as the best trees. The main advantage of this genetic-based feature selection algorithm is that it selects only the necessary and contributing features for classification.

In [83], a decision tree is used for reducing the features, and in a tree, each nonterminal node represents a test or decision on the considered instance. Based on the outcome of the results, the tree branches are created. To classify a particular instance, the decision tree algorithms start from the root node to the terminal node. A decision is made when it reaches the terminal node. The decision trees can also be interpreted and characterized by the hierarchical organization of rules. The feed-forward neural network is involved in the training phase along with an augmented cross-entropy error function [84]. A new feature selection algorithm [85] follows the wrapper approach with the support of neural networks. During the feature selection process, the automatic determination of neural network architectures plays a vital role. The constructive approach considers the correlation information and select the optimal features to decide the neural network architectures. This reduces the redundant features successfully. In [16] and [82], the wrapper-based feature reduction approaches exploit the genetic algorithm. After obtaining results for n iterations, the decision tree is created with multiple branches. The main advantage of this genetic-based feature selection algorithm is the selection of contributing features for classification. Even though the wrapper method produces better feature subsets, it expands high running time to identify the best feature subset, when compared to the filter model.

3.4.3 Hybrid Techniques

An efficient hybrid ant colony optimization-based feature selection algorithm in [86] determines the optimal subset size. It overcomes the issues in both the filter and wrapper-based feature reduction models. The major advantage of the hybrid approach is the scalability and decentralization. A hybrid intrusion detection method based on hidden Markov model (HMM) and fuzzy logic in [87] classifies the anomaly profile from the normal network traffic. It provides the following advantages: it requires only less storage and reduced training time. When the processes of intrusions are similar to the normal behavior, the hybrid approach detects network-based intrusions only at high false positive rates. The following table provides some of the examples for the filter, wrapper, and hybrid approaches. Table 3.5 provides examples of feature reduction techniques with its advantages and limitations.

Table 3.5: Comparison of Various Feature Reduction Techniques

Feature Reduction Techniques	Type	Metrics/Classifier	Advantages	Limitations
G-LDA [88]	Filter	Mean and Model value	High classification accuracy with small sized dataset	Data pre-processing does not consider the instances from different classes
Lightweight IDS [83]	Wrapper	Neuro tree classifier	Low false alarm rate	Poor results under large data size
Genetic Algorithm [16] [82]	Wrapper	Clustering and classification	Selecting relevant features improves the detection accuracy	High running time to identify the best feature subset
Novel Hybrid Method [89]	Hybrid	Gain Ratio/ K-means classifier	Low false positives	Cluster size is inadequate under large dataset size
Hybrid Method [90]	Hybrid	Weighted Mutual information	Reduces the number of iterations in feature selection process	Fails to identify all feature combinations under high dimensional data

3.5 Data mining approaches for IDS

Due to large data size, the manual labeling is extremely challenging and expensive. Large size data increases the complexity of auditing and data analysis. The data mining techniques have a significant advantage in data extraction over a large and highly dimensional data. The usage of data mining techniques in intrusion detection system can enhance the detection rate by reducing the false alarm rate of an IDS. The work in [48] exploits the hybrid learning approach, in which k-Medoids and Naive Bayes classification technique are combined. The real fact is that k-Medoids clustering techniques handle the real-world scenario of data distribution, and the combined k-Medoids and Naive Bayes can group the whole data into corresponding clusters accurately, compared to the k-means algorithm.

The data mining techniques apply the clustering algorithms and provide labeling to the data automatically. A novel flow-based detection scheme in [91] exploits the K-mean clustering algorithm. The K-means clustering algorithm divides the dataset into different clusters such as legitimate and intrusions. The resulting cluster centroids assist the IDS to detect the anomalies in a new monitoring data based on simple distance calculations. The work in [92] exploits k-means clustering algorithm to detect intrusions and analyzes the dataset. An IDS based on a parallel PSO clustering algorithm utilizes the MapReduce methodology to avoid the sensitivity problem of initial cluster centroids and premature convergence [94]. The PSO based IDS can process large datasets on commodity hardware. It can detect the new intrusions in the network, but increases the false positive rate, when increasing the number of new intrusions. The work in [95] presents a new agent-based IDS, and it uses the rough set theory to classify intrusions. It can manage noise and uncertainty in data, but the rough set classifier is computationally expensive, as it needs the knowledge of entire features. The Naive Bayesian Classifier [96] employs a Bayesian approach to classify the instances in the network. The NBC assumes that the features have conditional independence among them and this assumption significantly increases the accuracy of IDS. However, it is not accurate always. Moreover, the storage

space and computational complexity of NBC based IDS are biased for the stupendous dataset.

The work in [96] combines principal component analysis (PCA) and PSO to improve the performance of support vector machine, SVM. It applies the improved SVM to the intrusion detection and enhances the detection accuracy. The improved SVM exploits the PCA and attains the dimensionality reduction. Secondly, it utilizes the PSO algorithm for selecting the punishment factor and kernel parameters in SVM optimally. In [97], an IDS model has been proposed to detect malicious behavior in wireless networks. It selects optimal features using Information Gain measurement. The SVM parameters decide the performance of classifier and the swarm intelligence algorithms. A hybrid model in [89] optimally decides the best set of features and classifies the normal profiles from the 802.11 specific intrusions. The hybrid model of feature selection employs the information gain ratio measurement to estimate the relevance of each feature. It selects the optimal set of MAC layer features using k-means classifier to improve the accuracy of IDS. Moreover, the hybrid model reduces the learning time of the classifier by selecting an optimal feature set as input.

A wrapper model for feature selection and parameter optimization in an SVM [98] selects optimal features using the binary PSO. The utilization of continuous PSO optimizes the parameters in the kernel function of SVM simultaneously, and it tends to better classification performance. However, only one dataset with a small number of features is considered in the experiments. This cannot demonstrate the exact performance of the proposed algorithm. The work in [99] proposes two distributed IDS approaches in a hierarchical and a completely distributed architecture respectively. The IDS exploited in both the architectures identifies the intrusions using the SVM classification algorithm. They use a set of parameters derived from the network layer and detects the intrusion. A cooperative and distributed intrusion detection system in [100] that uses data from the MAC, routing and application layers, coupled with a Bayesian classifier, detect malicious activities.

An effective intrusion detection method in [101] employs SVM classifier and detects higher layer intrusions. The distance measurement in feature reduction removes redundant and noisy points the space. The k-means neighboring algorithm successfully eliminates the noise data in the dataset. The remaining samples are taken as testing instances. If the training data set is unbalanced, the detection accuracy is reduced. The characteristics of the SVM classifier are used to categorize the intrusion pattern from the normal one using a predefined collection of historical information [102]. The use of one class SVM with the IDS significantly deviate the intrusions from the normal profiles and achieves a better detection accuracy even when the training samples are too low and high. The work in [102] and [103] distinguishes only the intrusions. However, it fails in classifying the erratic attack pattern under a different type of classes. To overcome this, a multi-class SVM [104] partitions the unbalanced attack underclasses, but the accuracy of IDS is minimum in multi-class SVM due to the knowledge extraction from entire training data.

SUMMARY

Chapter 3 has provided a survey of various active and passive intrusions in WLAN. It has also surveyed about different types of 802.11 standards. This chapter reviews different intrusion detection systems. Different IDS techniques are compared with its advantages and limitations. The feature reduction techniques such as wrapper, filtering, and hybrid models are discussed in detail. The data mining techniques in IDS design has also been discussed.

CHAPTER 4 PROPOSED APPROACH

This chapter proposes a specification-based IDS, named as N-TBTDT that exploits the data mining techniques such as Chaotic PSO and improved decision tree classifier in detecting the 802.11 MAC intrusions. This chapter explains the two phases such as training and testing in N-TBTDT and learns the critical MAC layer features to identify intrusions effectively. Identifying useful features to train the classifier is an important research area, especially in the IDS. It explains the techniques that are exploited by N-TBTDT system.

4.1 Introduction

Due to the proliferation of handheld devices such as smartphones, tablets, and laptops, the adoption of WLAN in the society has increased drastically. The International Telecommunications Union releases the statistics, i.e., the number of mobile users has reached to 6.8 billion worldwide, and almost 40% of the world's population is now using the Internet. Due to the communication held through a wireless medium, the wireless devices are being abused for unlawful cyber-criminal activities such as malicious intrusions, computer hacking, data forging, and financial information theft. The Norton Cybercrime reports that this kind of intrusions causes a direct loss of about 83 billion Euros with an estimated 556 million users worldwide who have been impacted by cyber-crime 2012. To fight against cyber-criminal activities, it is essential to enhance the security of wireless communications. Moreover, the widely reported security weakness of the 802.11 networks leads businesses to face tremendous risks associated with the Wi-Fi networks. It is hard to secure the WLAN, compared to the wired network security, since the transmission medium is air. For providing security to the WLAN, the encryption techniques are used. Due to the transmission of encrypted data traffic, the cost is increased, and the network performance is shrinking, yet lots of organizations deploy 802.11 standard based wireless infrastructures. For most WLAN users, there are three core issues such as denial of service inject packets, impersonation, and injection into the network and unauthorized access. It necessitates the IDS implementation in WLAN.

A common approach in intrusion detection models is to use classifiers as intrusion detectors. Selecting the best set of features plays a vital role in ensuring the performance of the classifier. Most intrusion detection techniques exploit the information gain measure to reduce the number of features and to train the classifiers. Conventionally, the information gain is measured using the entropy and Bias compensation factor. The bias compensation factor in intrusion detection is measured using an optimal feature set with a large number of distinct values. The bias compensation factor provides equal importance to the frequency of distinct values and difference of feature values. However, both have a different impact on the classification accuracy. Differentiating the high frequency of attributes with small distinct values and less frequency of attributes with large distinct values is essential. That means, highly appeared features with small distinct values is most relevant to the attackers, compared to the difference of feature values. As some of the features have distinct values in nature, it tends to reduce the classification accuracy with biased information. To handle the issues, the proposed concept includes two primary components. The first component is an optimal feature selection using novel normalized information gain and Chaotic based particle swarm optimization and the second component is intrusion detection using a decision tree classifier. The decision classifier can add the new possible scenarios in a tree quickly. This reduces the execution time without reducing the classification accuracy even under a new class.

4.1.1 Importance of Feature Reduction in IDS

An intrusion detection technique is a valuable security tool in identifying and isolating the intrusions from the WLAN. The IDS is used as an intrusion pattern recognition system. To do the pattern recognition, the feature reduction is a critical pre-processing step. The quality of the feature construction and feature selection algorithms decides the effectiveness of an IDS. The main aim of feature reduction is to identify the relevant features and trains the classifier, without having a negative impact on the classification accuracy [12]. Conventionally, the feature selection techniques are carried out manually with the domain knowledge. After that, the automatic feature construction methods such

as filter and wrapper are frequently applied. The wrapper method produces better feature subsets. However, it expands high running time to identify the set of best features when compared to the filter model. Most of the filtering models apply information gain measurement for feature reduction, however lack of removing the redundant features makes a negative impact on the performance. Thus, the proposed scheme utilizes the advantages of bias compensation factor and certainty factor in feature reduction.

4.1.2 Decision Tree Algorithm

The proposed work improves the Very Fast Decision Tree (VFDT) for intrusion detection. The VFDT is a lightweight data mining technique, and it can process a large amount of data by utilizing considerable memory space. The VFDT builds the tree starting from the scratch, and it is an efficient solution to identify the intrusions. Therefore, the VFDT is selected and applied for detecting 802.11 MAC intrusions in WLAN efficiently. Although the VFDT differentiates the intrusions from the legitimate network traffic, it has certain limitations. It cannot handle noisy data, and classification accuracy decreases with the increase in noise. It is essential to provide variations in the VFDT to handle the noise data. The random selection of tie-breaking threshold increases the processing time and decreases the overall accuracy of the decision tree. Moreover, it is inappropriate for the resource-constrained network. To solve the issues in VFDT, a fixed tie-breaking threshold is applied When the difference between two information gains is small, the fixed tie-breaking threshold takes the decision quickly and solves the issues. However, an excessive tie-breaking value reduces the performance of VFDT on noisy and complex data.

To solve this issue, an adaptive tie-breaking threshold using Hoeffding bound, instead of using fixed tie-breaking threshold. The value of Hoeffding bound fluctuates with the increase in the number of classes. However, the random selection of threshold is unfair for the intrusion classification. Thus, the proposed work measures the certainty factor to decide the tie-breaking threshold, instead of information gain and identifies the necessity

of adding new leaf node for the intruder quickly, compared to the existing decision tree algorithm.

4.2 Overview of the N-TBTDT

The first component estimates the information gain, in terms of entropy. To utilize the advantage of bias compensation factor measurement in optimal feature reduction, the certainty measure on information gain is applied. The certainty measurement differentiates the impact of high frequency of attributes with small distinct values and less frequency of attributes with large distinct values using a mathematical model.

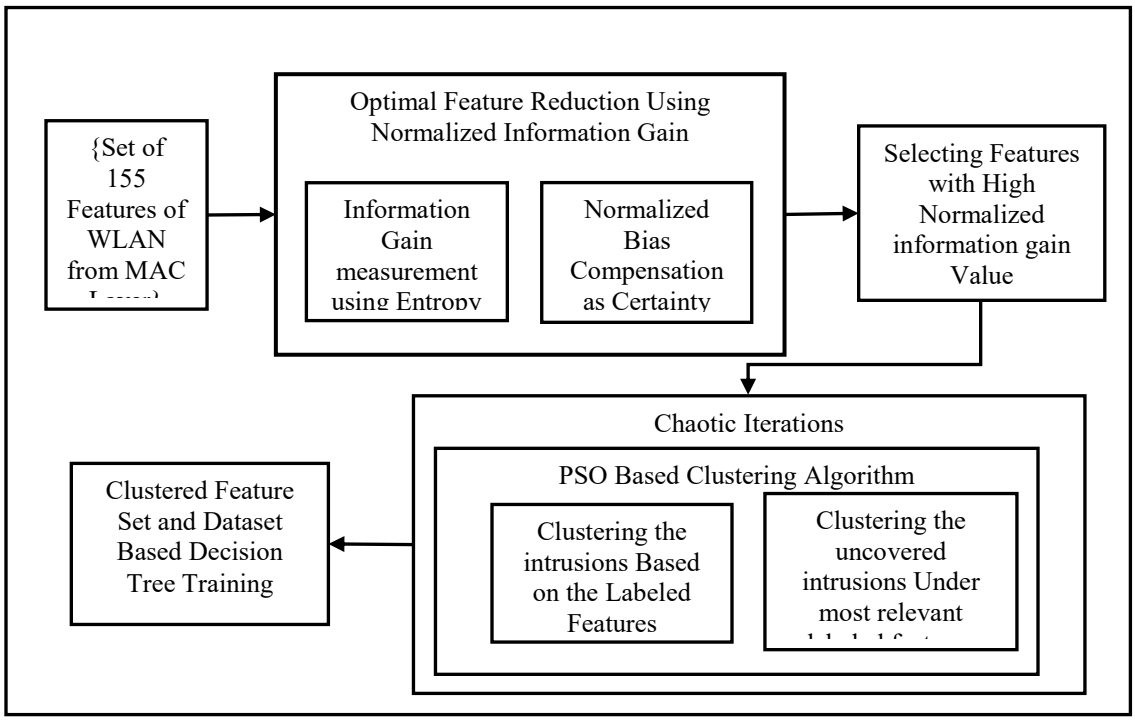

Figure 4.1: Optimal Feature Selection Using Normalized Information Gain and Chaotic PSO Optimization Algorithm

The normalized information gain is the weighted multiplication of information gain and certainty on information gain. This improves the efficiency of feature reduction. The first component ranks the features using normalized information gain value and selects the optimal set of features using Chaotic based particle swarm optimization. The PSO

algorithm only searches the best feature in local optima, and it results in an earlier standstill of the features before reaching the global optima. This problem is called as premature convergence. The Chaotic algorithm properly controls the feature velocity and determines the optimum solution accurately. Figure 4.1 shows the optimal feature reduction of proposed work. The chaotic based on particle swarm optimization uses labeled and unlabeled features simultaneously and determines a group of optimal features. By applying the optimal set of features, the decision tree classifier classifies the attacks under the appropriate classes. The major drawbacks of the decision tree classifier are poor accuracy with the dataset, where the information gain in decision trees is biased in favor of those attributes with more levels. Number of uncertainty values due to biased values increase the complexity of decision tree classification.

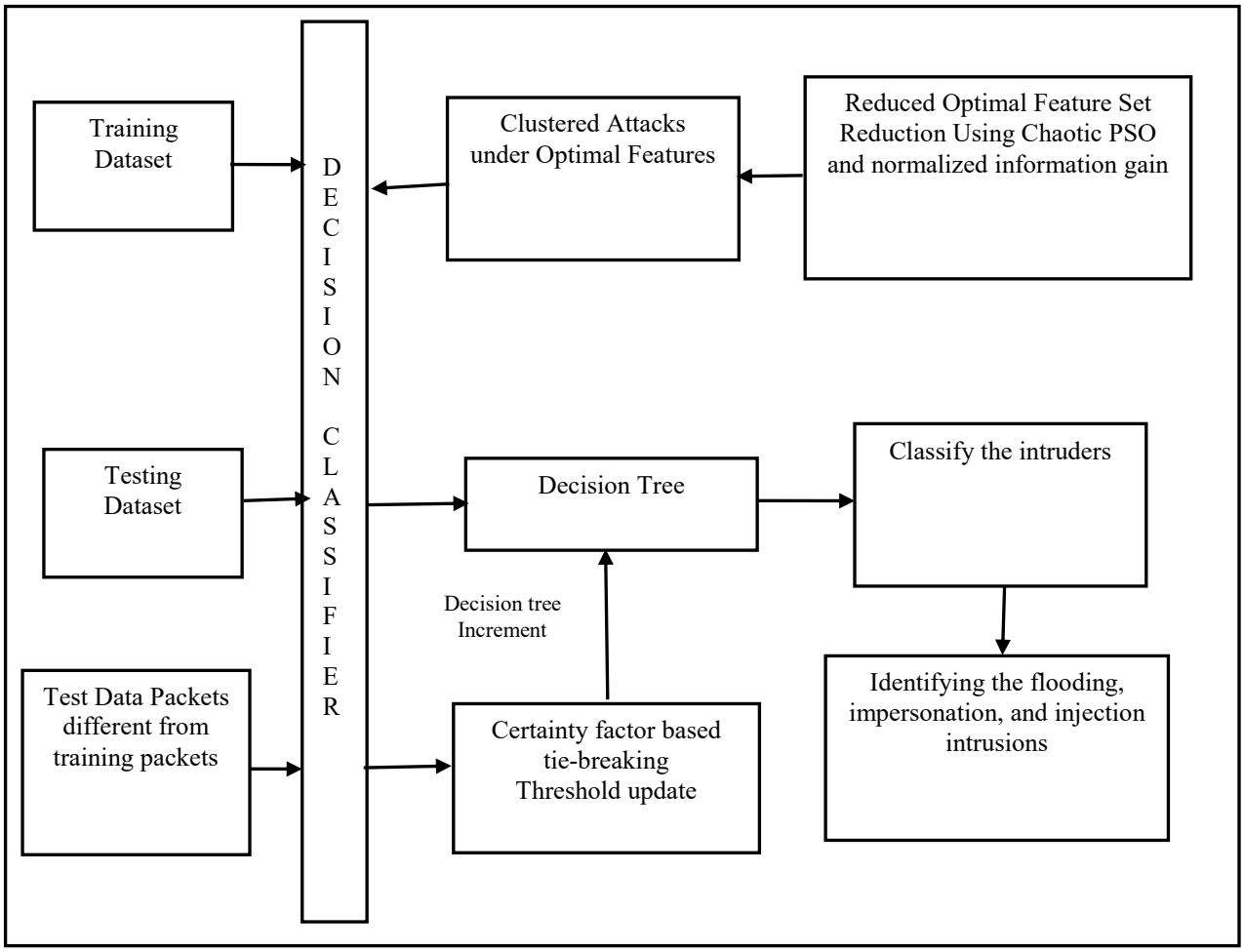

Figure 4.2: Decision Tree Classifier Based Intrusion Detection System

The measurement of normalized information gain using both entropy and certainty factor improves the accuracy of decision tree classification, as shown in figure 4.2. The main contribution of the classifier is to increment the decision tree when the new data packets arrive. Most of the attacks are categorized into flooding, impersonation, and injection. With the development of defense system against those attacks, the severity of intrusions is also increased. In such cases, the information gain measurement algorithm is frequently executed. When the information gain difference exceeds the tie-breaking threshold, a new leaf node is appended. However, the random selection of threshold and frequent measurement of information gain is unfair for the intrusion classification. Thus, the proposed work measures the certainty factor to decide the tie-breaking threshold, instead of information gain and identifies the necessity of adding new leaf node for the intruder quickly, compared to the existing decision tree algorithm. Thus, the proposed work classifies the intrusions quickly, without degrading the classifier accuracy.

4.3 Optimal Feature Selection

To identify the optimal MAC features, the N-TBTDT enables the IDS to measure the normalized information gain and to classify the intrusions. The selected most relevant features of the intrusions play an important role in training the IDS. Initially, the N-TBTDT takes into account all the features from the standard 802.11 MAC frame header and measures the normalized information gain for each feature to assign the rank value. The collected AWID dataset with 155 features includes normal and four classes. Among them, one class if normal and others are injection, impersonation, and flooding [39]. The proposed scheme partitions the dataset into training and testing set. In the training dataset, each feature has different values in packets generated by various intrusions. A set of values of a feature is split into distinct values for different classes. The N-TBTDT estimates the normalized information gain at the different breakpoints of classes of each intrusion, and it classifies the intrusions under features by considering the normalized information gain threshold. The proposed scheme rank the feature according to the count of intrusions clustered under the corresponding feature, and it selects the top-ranked features over entire feature space.

An intrusion can uncover under the top-ranked features. In case of leaving that how many attacks are uncovered by the top features, it degrades the accuracy of improved VFDT classifier. Moreover, the proposed scheme exploits the Chaotic PSO scheme for measuring the fitness of uncovered attack to cluster it under the most relevant labeled feature. The chaotic PSO examines the uncovered intrusion and relevance of each labeled feature along with the uncovered intrusions using normalized information gain value. After that, the chaotic PSO imposes the corresponding intrusion under the appropriately labeled feature. The N-TBTDT continues the process when all the uncovered intrusions are clustered under the labeled feature. Thus, the selection of optimal features from all the extracted features minimizes the learning time of IDS, while improving the detection accuracy of the classifier.

4.3.1 Normalized Information Gain Measurement for Feature Reduction

The Normalized information gain and Tie Breaking threshold-based Decision Tree, N-TBTDT aims at reducing the features while improving the accuracy of IDS. To reduce the features, the proposed methodology considers the Normalized Information Gain, NIG, Normalized Bias Compensation, NBC as certainty measure, Chaotic PSO, and improved decision tree algorithm. An Information Gain (IG) defines the impurity level of each feature. However, the information gain is biased if it has more distinct values in the class. The existing work attempts to reduce the error occurring in the bias of information gain using breakpoint measurement. The breakpoint information refers the inconsistent mean value of a feature (x) in between two adjacent subclasses (sc), which refers a subset of class value (v). Providing equal importance to the frequency of distinct values and difference of feature values impacts the classification accuracy.

$$NIG\ (x) = IG\ (x) / \log_2(NBC + 1) \quad \ldots\ldots\ldots\ldots (1)$$

$$IG\ (x) = Entropy\ (x) - [\sum|x,v|/x * Entropy\ (x,v)] \quad \ldots\ldots\ldots\ldots .. (2)$$

$$Entropy\ (x) = -\sum P_x \log_2 P_x \quad \text{Where P stands for probability}$$

$$NBC\ (x) = \{\alpha * (sc_i - sc_j)\} + \{(1-\alpha) * [(2*n_i*n_j)/n_i+n_j]\} \quad \ldots\ldots\ldots\ldots\ldots (3)$$

While calculating Entropy, when the feature that has no breakpoints, the denominator term in NIG value is determined as infinity and is not selected as the feature which has zero breakpoints have no significance in attack detection. Applying the equations (3) and (2) in (1), the value of NIG is returned. Where the sc_i and sc_j are the two sub-classes in a class v. $n_{i,}$ and $n_{j,}$ represent the frequency of subclasses i and j respectively. The weighting factors α and 1-α differentiate the high frequency of attributes with small distinct values and less frequency of attributes with large distinct values. If α is assigned as 0.1, the second term is assigned as 0.9. This gives high weight to the frequency compared to the difference of subclass values. α is computed for every feature x depending on the frequency of its appearance. Highly appeared features with small distinct values is most relevant to the attackers, compared to the difference of feature values. The accurate biased information improves the feature reduction efficiency and classification accuracy.

4.3.2 Chaotic PSO Based Attack Clustering

The PSO algorithm only searches the best feature in local optima, and it results in an earlier standstill of the features before reaching the global optima, due to the low-level diversification among particles. This problem is called as premature convergence. For instance, if the PSO is applied to recognize and cluster the most relevant features of each attack, some of the essential features are missed, due to the premature convergence. It tends to attack misclassification and negatively affects the accuracy of attack detection. The chaotic mapping function assists the particles to break away the local optima when it meets the premature condition in each iterative searching process and improves the forecasting accuracy. The Chaotic algorithm properly controls the feature velocity and determines the optimum solution accurately. To prevent the PSO algorithm from being trapped into a premature convergence, there is a necessity to apply control over the global exploration and local exploitation. The term exploration represents the heuristic search for new features, while the term exploitation means taking advantage of previous best solutions. Obviously, the PSO performance dramatically relies on its parameters, especially the previous velocity. The previous velocity provides the necessary momentum for attaining best-fit clustering solution over the search space.

Mainly, the chaotic algorithm exploits a dynamic inertia weight to control the impact of the previous velocity of the current one and effectively maintains the trade-off between exploration and exploitation in PSO. Larger inertia weight guides the particles to global search, while the smaller factor leads particles to local search using the previously estimated best solutions. Thus, proper control of the particle velocity plays a vital role in determining the optimum solution accurately and efficiently. The features that have low NIG value are selected as the optimal features. Each attack is classified under the optimum feature. Some of the attackers are unlabeled, due to which they do not have the same impact on the optimal features. To cluster the unlabeled attacks under optimal features, the Chaotic PSO is used. The PSO algorithm faces the issue of premature convergence. This results in an earlier standstill of the attacks before classifying under the most appropriate feature. Thus, the proposed methodology exploits the Chaotic PSO to measure the fitness of uncovered attack to cluster it under the most relevant labeled feature.

$$V_i^{t+1} = W * V_i^t + c_1 * rand() *(P_{Besti}^t - X_i^t) + c2 * rand() *(G_{Besti}^t - X_i^t) \text{ ..... (4)}$$

$$W = \begin{cases} W_{min} + (W\alpha) & \text{if NIG} \geq NIG_{avg} \\ W_{max} & \text{if NIG} < NIG_{avg} \end{cases} \text{ .....(5)}$$

Notably, the unlabeled attacks act as particles in Chaotic PSO. The inertia weight W decides the velocity of the particle. The Chaotic PSO generates k solutions for a particle i in the iteration of t, and among them, the best solution is considered as $G_{Best(i)}$. $P_{Best(i)}$ represents the previous best solution of the particle. The inertia weight tends to global search when the NIG value is less than the average NIG value. Otherwise, the W value is taken as a minimum, and it tends the particle i to local search. Thus, the Chaotic PSO effectively clusters the unlabeled attacks under optimal features.

4.4 Improving Decision Tree Algorithm

The proposed scheme improves the performance of VFDT classifier using dynamic tie-breaking threshold value. The proposed classification algorithm is an enhancement of original VFDT to make it efficient for the detection of new intrusions in wireless local area networks. The improved VFDT classification algorithm simultaneously trains and tests the decision tree based on learning traffic patterns and dynamically updated tie threshold value. After that, the testing phase classifies the intrusions accurately based on these learned patterns. The randomly selected or fixed tie-breaking threshold value degrade the classification performance of VFDT. To overcome this, enhanced VFDT in the proposed scheme takes the certainty factor values for tie-breaking threshold measurement. Moreover, it compares the Hoeffding bound and dynamically estimated tie-breaking threshold to decide about the insertion of the node to the decision tree. It results in accurate detection of MAC intrusions in WLAN.

4.4.1 Tie Breaking Threshold Based Decision Tree Algorithm

The improved decision tree algorithm aims at incrementing the decision tree when the new data packets arrive. Most of the attacks are categorized into flooding, impersonation, and injection. The decision tree algorithm randomly selects the tie-breaking threshold and frequently measure the information gain. When the information gain difference exceeds the tie-breaking threshold, a new leaf node is appended. However, the random selection of threshold is unfair for the intrusion classification. Thus, the proposed work measures the bias compensation factor-based tie-breaking threshold T and adds a new leaf node for the new intruder quickly.

$$T = 1/\text{ Difference } (f_i, f_j) \quad .....(6)$$

Where f_i and f_j represent the NBC value of two optimal features. For every consecutive and non-overlapping pair of optimal features, the T and hoeffeding bound are compared

to add the new leaf. If both the features have similar NBC value, the T value is increased. That means, both the features have a similar impact on the attack detection efficiency. It has less possibility to appear less than the hoeffeding bound. Thus, the proposed methodology decides to add a leaf for any one of the features. Otherwise, a new leaf is added for both the features, that have a different impact on the classification accuracy. Thus, the proposed work classifies the intrusions quickly, without degrading the classifier accuracy.

4.5 Intrusion Detection and Categorization

To precisely differentiate the malicious activities from the legitimate profiles, the N-TBTDT exploits the use of a set of optimal features and chaotic PSO algorithm. It utilizes the improved VFDT classifier for the classification task, as the VFDT is an efficient tool that they learn very effectively with new data, and it can update the decision tree arbitrarily, whenever there is a necessity to create the node for the new pattern in intrusion classification. Also, the VFDT classifier avoids a redundant pattern leads to create the new node in the decision tree construction process. The improved VFDT classifier necessarily depends on the relevance of feature dependency and optimally builds the decision tree for accurate intrusion classification.

SUMMARY

This chapter proposes the specification-based IDS techniques that exploit chaotic PSO and enhanced VFDT classifier to detect the 802.11 intrusions. This chapter has introduced the WLAN security and the need for feature reduction in IDSs. The measurement of normalized information gain and normalized bias compensation factor are explained. The process of chaotic PSO is discussed. Finally, the intrusion detection and categorization using enhanced VFDT are explained.

CHAPTER 5 PERFORMANCE EVALUATION OF N-TBTDT

This chapter illustrates the experimental setup of N-TBTDT and evaluates the performance using the metrics such as Detection Accuracy, False Positive Rate, Precision, F-Score, and Classification Accuracy. The experimental setup creates two different scenarios by varying the Training Data Size (TDS) and a number of attacks, as well as analyzes the performance of N-TBTDT. This chapter evaluates the experimental results with detailed descriptions.

5.1 Introduction to N-TBTDT

Due to the ease of deployment and provision of convenient network access to the users, the WLAN becomes ubiquitous over the last decade. The public WLAN connections are free to access or have shared password, and it is attractive to the intrusions. Even though the public WLAN is insecure; people share their confidential data via public WLAN connections. This necessitates the WLAN security. To provide an effective IDS to the WLAN, this chapter presents the N-TBTDT that employs a specification-based IDS and improved decision tree classifier. An IDS model consists of training and testing phase. In the training phase, the critical MAC features are learned, whereas, in the testing phase, the improved decision tree classifier classifies the intrusions using learned features. The use of all the MAC features of classifier learning tends to extend learning time and computational complexity. Increasing the number of features does not contribute to improving the accuracy of IDS. Thus, the N - TBTDT system exploits the normalized information gain and Chaotic PSO in extracting the most relevant features that have a maximum number of intrusions, and the improved decision tree classifier is used in learning the classifier and to categorize the MAC 802.11 specific intrusions accurately.

5.1.1 Extracted Feature Set

To ensure the accurate intrusion detection, an IDS applies an optimal feature identification. The data records include both the normal and malicious traffic. For each feature, the N-TBTDT measures normalized information gain using information gain and normalized bias compensation factors. The top-ranked features are considered as an optimal set. Table 5.1 shows the extracted features that are used for intrusion detection.

Table 5.1: Extracted Feature Set

S. NO	Features	Description
1	frame.offset_shift	Time shift for this packet
2	frame.time_delta_displayed	Time delta from previous displayed frame
3	frame.time_relative	Time since reference or first frame
4	frame.pkt_len	Frame length
5	radiotap.present.ext	Ext
6	radiotap.flags.shortgi	Short GI
7	wlan.fc.version	Version
8	wlan.fc.type	Type
9	wlan.fc.frag	More Fragments
10	wlan.fc.pwrmgt	PWR MGT

5.2 AWID Dataset

The proposed work considers the AWID dataset [39] to evaluate the performance of N-TBTDT. The AWID dataset is a collection of normal traffic as well as the traffic from recent wireless intrusions such as injection, impersonation, and flooding over protected 802.11 networks. The AWID dataset is published in the year of 2015. The AWID divides the collected data into AWID-CLS-R-Trn and AWID-CLS-R-Tst. The AWID-CLS-R-Trn is utilized for training the classifier and AWID-CLS-R-Tst is used for testing purpose. The size of AWID-CLS-R-Trn and AWID-CLS-R-Tst is 886MB and 280.7MB respectively. The AWID-CLS-R-Trn has 17,95,575 records, whereas the testing set includes 16,33,190 normal traffic, and the rest of the dataset records, i.e., 162,385 records comprise a different type of intrusions. To create the AWID dataset, the WLAN is monitored for 60 minutes, in which the intrusion-free traffic is spanning for 45 minutes and the traffic that contains intrusions lasting for 15 minutes. The N-TBTDT system divides the AWID-CLS-R-Trn into small TDS having records about 16 intrusions. For each small TDS, the N-TBTDT runs the optimal feature selection algorithm to compute NIG value of the 155 features of the MAC 802.11 that are retrieved from the dataset.

5.3 Experimental Evaluation

The overall performance of N-TBTDT system is evaluated in two phases. The first phase of N-TBTDT chooses the set of optimal features among 155 features using Java Machine Learning Library. Secondly, the intrusion detection system executes the improved decision tree classifier in the Waikato Environment for Knowledge Analysis (WEKA). By extending basic Hoeffding tree, a new classifier N-TBTDT is created in WEKA. In WEKA, methods trySplit(), and getSplitPointCandidates() in GaussianConditionalSufficientState class. The N-TBTDT system experiments in Windows 7 with 4GB RAM. The N-TBTDT constructs different scenarios by varying different parameters and analyzing the performance of the N-TBTDT system. For examining the performance of N-TBTDT, it varies the training dataset from 177.2MB to 886MB, varies the number of intrusions considered in the dataset from 6 to 16.

5.3.1 Performance Metrics

The performance of the proposed N-TBTDT is evaluated by varying the TDS and number of intrusions. The effectiveness of the N-TBTDT system is examined using the performance metrics such as Detection Accuracy, False Positive Rate, Precision, F-Score, and Classification Accuracy.

Detection Accuracy: The detection accuracy is the ratio of the number of intrusion packets correctly identified by IDS to the total number of intrusion packets.

False Positive Rate: It is defined as the percentage of intrusion packets that are identified incorrectly as normal packets.

Precision: The precision is defined as the proportion of packets predicted as correctly to the total number of packets.

F-Score: It is a measure of a testing accuracy. It is measured using the precision and detection accuracy values.

Classification Accuracy: It is the ratio of packets successfully classified under intrusion classes to the total number of intrusion packets.

5.4 Experimental Results

This section discusses the comparative performance of N-TBTDT and NMI [105] under different scenarios. The reason behind the performance difference between two IDS models is described in detail.

5.4.1 TDS Vs. Detection Accuracy

Table 5.2: TDS Vs. Detection Accuracy

TDS (MB)	Detection Accuracy (%)	
	N-TBTDT	NMI

177.2	99.851	98.9
354.4	99.89	99.24
531.6	99.93	99.56
708.8	99.93	99.57
886	99.94	99.58

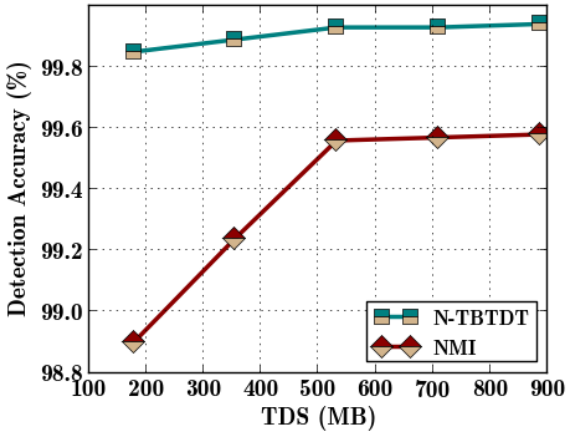

Figure 5.1: TDS Vs. Detection Accuracy

In Figure 5.1, the detection accuracy of N-TBTDT and NMI is evaluated by varying the TDS size. The numerical values of detection accuracy results are depicted in Table 5.1. While increasing the TDS value from 177.2 to 886 MB, the N-TBTDT shows no huge increment in the detection accuracy. Even with small TDS, N-TBTDT attains 99.851% of the detection accuracy. However, the NMI reaches only 98.9%. The reason is that the NMI fails to differentiate the frequency of distinct values and difference of feature values in breakpoint measurement. The N-TBTDT improves the detection accuracy from 99.851 to 99.94% when varying the TDS from 177.2 to 886 MB. As the large TDS contains an enormous number of records and it increases the detection accuracy significantly. The clustering of unlabeled features under labeled features using Chaotic PSO avoids the premature convergence and leads the N-TBTDT to classify the intrusions into the most related classes. However, the NMI approach suffers from the issues of premature convergence during clustering and breakpoint measurement. For example, at the point of

886MB, the detection accuracy of N-TBTDT is 99.94%, but the NMI attains only 99.54% of detection accuracy.

5.4.2 TDS Vs. False Positive Rate

Table 5.3: TDS Vs. False Positive Rate

TDS (MB)	False Positive Rate	
	N-TBTDT	NMI
177.2	0.15	1.1
354.4	0.062	0.76
531.6	0.060	0.43
708.8	0.059	0.39
886	0.057	0.37

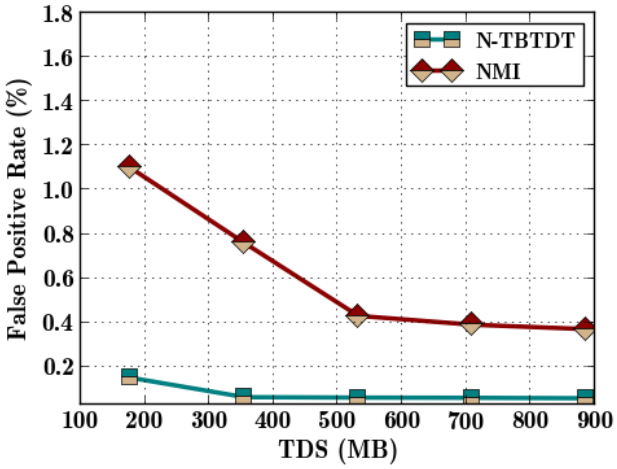

Figure 5.2: TDS Vs. False Positive Rate

Figure 5.2 depicts the false-positive rate results of N-TBTDT and NMI plotted based on the values in Table 5.2. While increasing the TDS from 177.2 to 886 MB, both the N-TBTDT and NMI is still are a considerable amount of false positives, due to the selection of the optimal feature set. However, the N-TBTDT outperforms NMI while increasing the size of training data. The N-TBTDT does not show huge decline in performance

when the TDS increases to 886MB from 177.2MB. Instead, the NMI increases the information gain value when the attributes with small distinct values appear in high frequency, and it led to the selection of an irrelevant feature in the network and increased false positive rate. Thus, the graph shows the apparent difference between the performance of N-TBTDT and NMI only with small TDS. Nevertheless, since the N-TBTDT considers the breakpoints with the knowledge of frequency and difference of distinct values, it can select optimal features that are more relevant to the intrusions, which results in a lower false positive rate in N-TBTDT, compared to the NMI. Also, the Chaotic PSO assists the N-TBTDT to efficient clustering and to maintain lower false positives from small to large sized TDS. For TDS of 177.2MB, the false positive rate of N-TBTDT decreases by 0.95%, when compared to NMI.

5.4.3 TDS Vs. Precision

Table 5.4: TDS Vs. Precision

TDS (MB)	Precision (%)	
	N-TBTDT	NMI
177.2	98.8	96.6
354.4	99.4	98.23
531.6	99.69	99.35
708.8	99.72	99.42
886	99.78	99.45

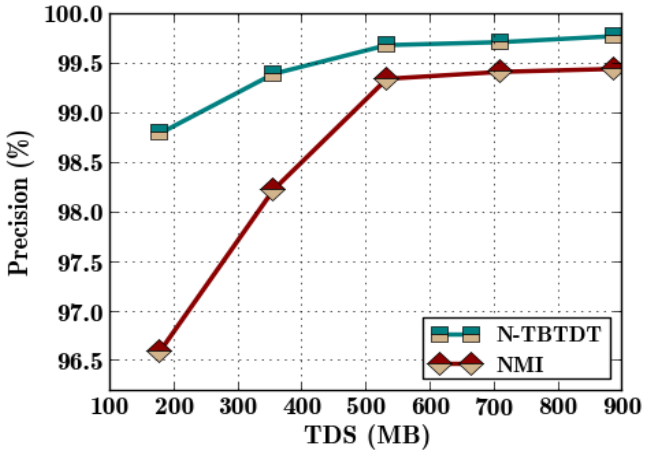

Figure 5.3: TDS Vs. Precision

Figure 5.3 portrays the precision results of the N-TBTDT and NMI. The Table 5.3 consists of the numerical values of precision results. Figure 5.3 shows that the precision of N-TBTDT increases slightly when the TDS is varied from 177.2MB to 886MB. This is because a large number of records with large sized TDS leads to measure the normalized information gain accurately and reinforce the intrusion behavior, resulting in a high true positive rate of the network. From Figure 5.3, the precision of N-TBTDT is high compared to NMI. At the same time, the NMI does not maintain its superior precision with low TDS compared to the scenario of high TDS, due to the selection of the optimal feature set relevant to the intrusions. For instance, Figure 5.3 depicts that the N-TBTDT and NMI attain 98.8% and 96.6% with small TDS.

5.4.4 TDS Vs. F-Score

Table 5.5: TDS Vs F-Score

TDS (MB)	**F-Score (%)**	
	N-TBTDT	**NMI**
177.2	99.3	97.7
354.4	99.6	98.73
531.6	99.81	99.46
708.8	99.82	99.49
886	99.860	99.51

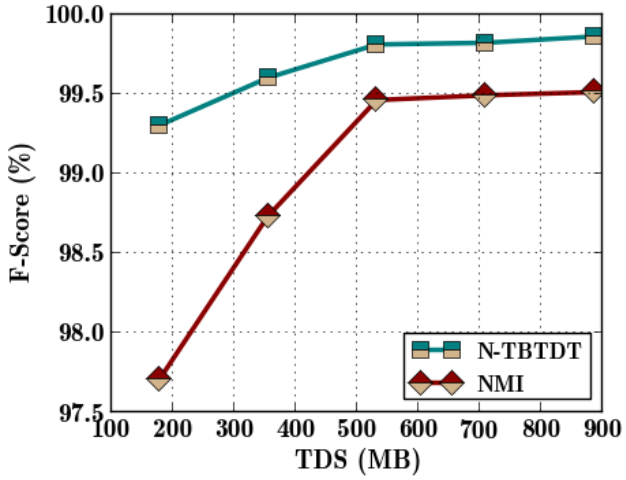

Figure 5.4: TDS Vs. F-Score

The high precision and detection accuracy value of NMI lead to attaining a high F-score. Figure 5.4 shows the experimental results of F-Score of N-TBTDTD and NMI by varying the TDS from 177.2 to 886 MB. Such experiments are carried out based on the numerical results of Table 5.4. When increasing the TDS, the number of malicious records also increases, resulting in improved precision, detection accuracy, and F-score in both the N-TBTDT and NMI. With sufficient training data, the NMI can identify the optimal feature set and improves the F-score. For instance, at the point of TDS is 886 MB, the N-TBTDT achieves the performance by 0.35% more than NMI. However, with small TDS of 177.2 MB, the difference between the N-TBTDT and NMI is 1.6%. It is apparent compared to the scenario of having large TDS.

5.4.5 Number of Attacks Vs. Detection Accuracy

Table 5.6: Number of Attacks Vs. Detection Accuracy

Number of Attacks	Detection Accuracy (%)	
	N-TBTDT	NMI
8	99.82	99.6
10	99.78	99.57

12	99.75	99.42
14	99.73	99.38
16	99.47	99.20

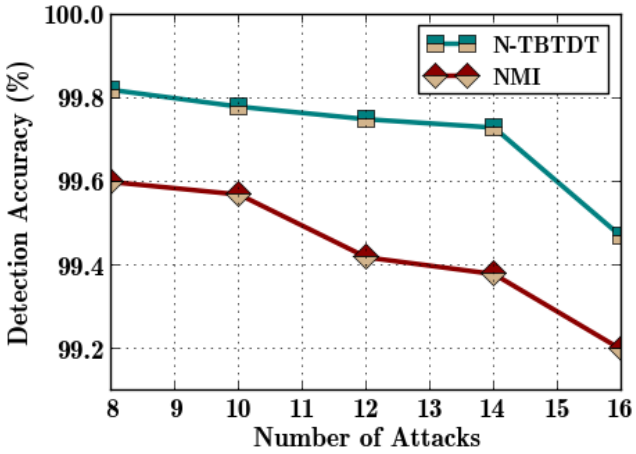

Figure 5.5: Number of Attacks Vs. Detection Accuracy

Figure 5.5 illustrates the detection accuracy of various intrusions with a fixed dataset of 177.2 MB. Table 5.5 shows the numerical values of the experiments carried out in Figure 5.5. The detection accuracy of both the N-TBTDT and NMI reduces when increasing the number of attacks. Even though, the N-TBTDT outperforms the NMI from a low number of attacks to the highly vulnerable scenario. It is due to the NMI fails to differentiate the high frequency of attributes with small distinct values and less frequency of attributes with large distinct values. Because, highly appeared features with small distinct values is most relevant to the intrusions, compared to the difference of feature values. As some of the features have distinct values in nature, it tends to reduce the detection accuracy with biased information in NMI. The result shows that the detection accuracy of N-TBTDT is prominently high compared to the NMI, due to the selection of optimal features and training of the classifier with the selected optimal features. In contrast, the NMI measuring the biased information with the high contribution of attribute frequency is not suitable for a large number of attacks, as it leads the NMI to decrease the intrusion detection accuracy. For instance, the N-TBTDT improves the detection accuracy by 0.27% compared to the NMI, when the number of attacks is equal to 16.

5.4.6 Number of Attacks Vs. Classification Accuracy

Table 5.7: Number of Attacks Vs Classification Accuracy

Number of Attacks	Classification Accuracy (%)	
	N-TBTDT	NMI
8	99.81	99.51
10	99.79	99.45
12	99.76	99.42
14	99.73	99.38
16	99.7	99.23

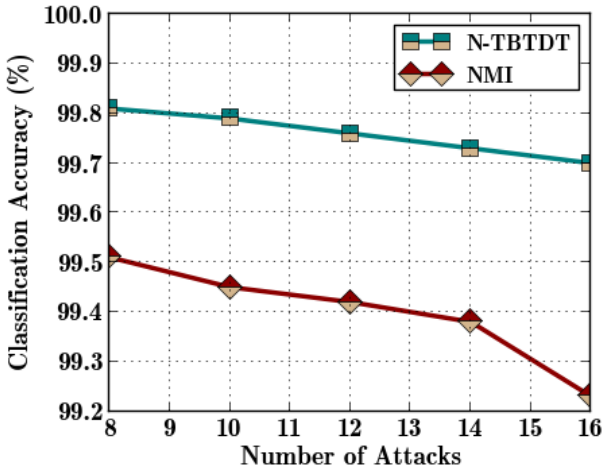

Figure 5.6: Number of Attacks Vs. Classification Accuracy

Figure 5.6 shows the result of classification accuracy of N-TBTDT and NMI, by varying the number of attacks in the dataset. The numerical results of classification accuracy are depicted in table 5.6. The performance difference between the N-TBTDT and NMI is apparent in the scenario of varying number of attacks from 8 to 16. From the figure 5.6, it is demonstrated that the performance of the N-TBTDT is higher than the performance of the NMI. The measurement of information gain without providing high importance to the difference between feature values in NMI may not reflect the impact of different types of

attacks on a selected feature. This tends the NMI to select non-optimal features and intrusion classification. For instance, at the point of 8 attacks, the N-TBTDT improves the classification performance by 0.3% more than the NMI. At the point of 16 attacks, the Figure 5.6 shows the classification accuracy of N-TBTDT is increased by 0.47%, more than that of NMI.

TEST RESULTS:

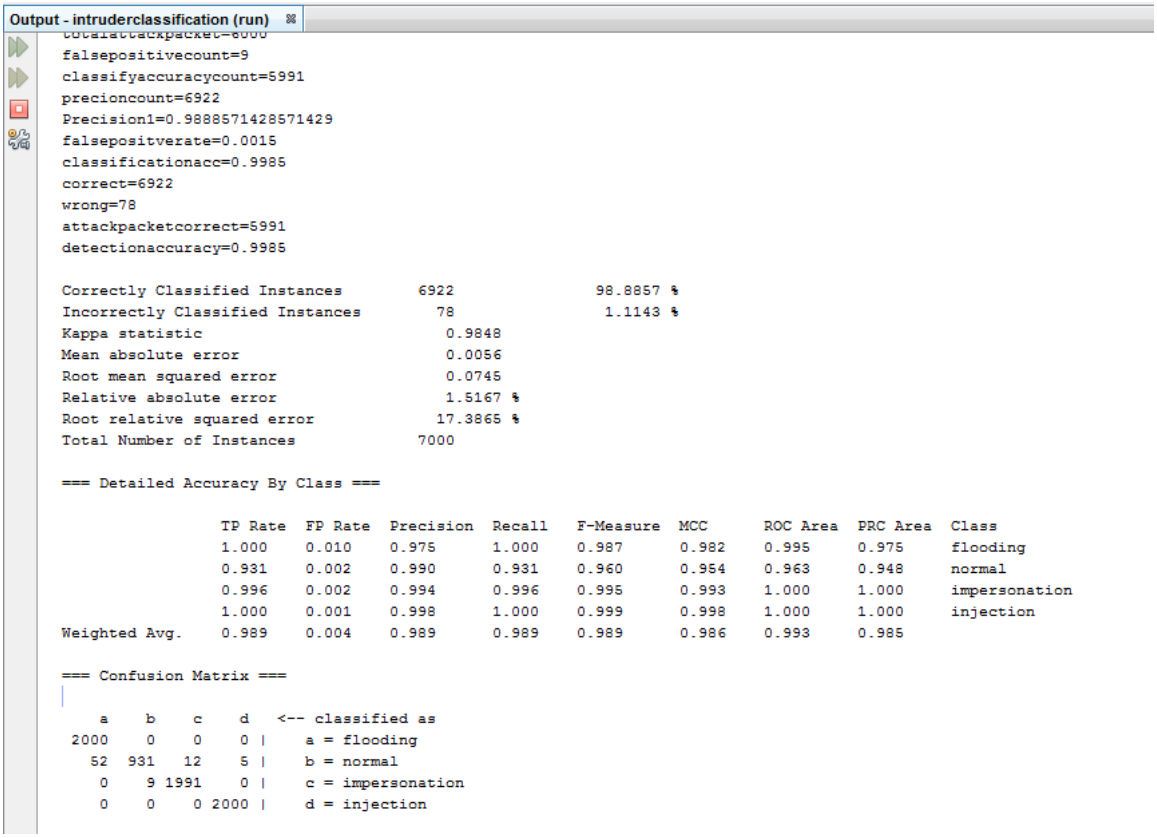

Figure 5.7: Test results

Considering 7000 test packets, 2000 packets are of flooding, 2000 packets of impersonation, 1000 packets of normal and 2000 packets of injection attacks. Test results proved efficient classification of flooding and injection attacks with marginal misclassification of impersonation and normal packets, achieving the classification accuracy of 98.8% which proves efficient results even with the small dataset.

SUMMARY

This chapter has discussed the experimental setup of the N - TBTDT system. The performance of the N - TBTDT system has been evaluated in terms of Detection Accuracy, False Positive Rate, Precision, F-Score, and Classification Accuracy. This chapter analyzed the performance of the N - TBTDT system under two different scenarios, and it has presented the experimental results as graphs with detailed descriptions. It has listed out the numerical values of experimental results in tables.

CHAPTER 6 CONCLUSION AND FUTURE DIRECTIONS

6.1 Conclusions

This work has a significant contribution in specification-based IDS technique that employs mining approaches to reduce the features and detect the intrusions effectively. The specification-based IDS technique, named as the N-TBTDT system has included a feature reduction and classifier training to detect and classify the MAC 802.11 specific intrusions more precisely. The N-TBTDT has measured the normalized information gain and normalized bias compensation factor and has selected the top-ranked features as optimal features. Further, it reduces the features using the Chaotic PSO technique that clusters the uncovered intrusions of unlabeled features with labeled features. The normalized information gain measurement with the knowledge of bias compensation factor is the main reason behind the improvement of performance of the classifier regarding learning speed, accuracy, and reliability. The bias compensation factor based tie-breaking threshold promises the efficient decision tree construction and provides a significant improvement in the clustering efficiency of N-TBTDT. The IDS learns the improved decision tree classifier using reduced feature set. The improved very fast decision tree, VFDT for intrusion classification eliminates the premature convergence issue during intrusion clustering under essential features, and so the clustering accuracy is increased.

The obtained experimental results show that the intrusion detection and classification accuracy of N-TBTDT are high when compared to the existing NMI technique. For example, at the point of 886MB, the detection accuracy of N-TBTDT is 99.94%, but the NMI attains only 99.54% of detection accuracy. The Chaotic PSO assists the N-TBTDT to efficient clustering and to maintain lower false positives from small to large sized TDS. For TDS of 177.2MB, the false positive rate of N-TBTDT decreases by 0.95%, when compared to NMI. The measurement of information gain by providing high

importance to the difference between feature values in N-TBTDT reflects the impact of different types of intrusions on a selected feature. This tends the N-TBTDT to select optimal features and intrusion classification. For instance, the N-TBTDT improves the detection accuracy by 0.27% compared to the NMI, when the number of attacks is equal to 16. Moreover, at the point of 8 attacks, the N-TBTDT improves the classification performance by 0.3% more than the NMI.

6.2 Future Directions

There are several possible directions for the proposed works to extend in the future, and those directions are summarized as follows.

- In future, the specification-based IDS techniques need to be evaluated under various data sets to show the exact performance.
- Detecting the set of optimal features for diverse intrusions separately improves the detection accuracy significantly.

A traffic-awareness detection system is an effective method to pursue IDS in the future that includes its state in deciding whether a suspect is normal or an intrusion.

REFERENCES

[1] Crow, Brian P., et al. "IEEE 802.11 wireless local area networks." *IEEE Communications magazine* 35.9 (1997): 116-126.

[2] Xiao, Yang, Chaitanya Bandela, and Yi Pan. "Vulnerabilities and security enhancements for the IEEE 802.11 WLANs." *Global Telecommunications Conference, 2005. GLOBECOM'05. IEEE*. Vol. 3. IEEE, 2005.

[3] Singh, Prashant, Mayank Mishra, and P. N. Barwal. "Analysis of security issues and their solutions in wireless LAN." *Information Communication and Embedded Systems (ICICES), 2014 International Conference on*. IEEE, 2014.

[4] Kyaw, Ar Kar, Pulin Agrawal, and Brian Cusack. "Wi-Pi: a study of WLAN security in Auckland CBD." *Proceedings of the Australasian Computer Science Week Multiconference*. ACM, 2016.

[5] Waliullah, Md, and Diane Gan. "Wireless LAN security threats & vulnerabilities." *International Journal of Advanced Computer Science and Applications* 5.1 (2014).

[6] Zou, Yulong, et al. "A survey on wireless security: Technical challenges, recent advances, and future trends." *Proceedings of the IEEE* 104.9 (2016): 1727-1765.

[7] Tsakountakis, Alexandros, Georgios Kambourakis, and Stefanos Gritzalis. "Towards effective wireless intrusion detection in IEEE 802.11 i." *Third International Workshop on Security, Privacy and Trust in Pervasive and Ubiquitous Computing.* IEEE, 2007.

[8] Kyaw, Ar Kar, Zhuang Tian, and Brian Cusack. "Wi-Pi: a study of WLAN security in Auckland City." *IJCSNS* 16.8 (2016): 68.

[9] Li, June, et al. "A detection method of WLAN security mechanisms based on MAC frame resolution." *Wuhan University Journal of Natural Sciences* 22.2 (2017): 93-102.

[10] Narayan, Shaneel, et al. "Impact of wireless IEEE802. 11n encryption methods on network performance of operating systems." *Emerging Trends in Engineering and Technology (ICETET), 2009 2nd International Conference on*. IEEE, 2009.

[11] Pandurang, Rathod Mahesh, and Deepak C. Karia. "Performance measurement of WEP and WPA2 on WLAN using OpenVPN." *Nascent Technologies in the Engineering Field (ICNTE), 2015 International Conference on*. IEEE, 2015.

[12] Chen, You, et al. "Survey and taxonomy of feature selection algorithms in intrusion detection system." *International Conference on Information Security and Cryptology*. Springer, Berlin, Heidelberg, 2006.

[13] Dash, Manoranjan, and Huan Liu. "Feature selection for classification." *Intelligent data analysis* 1.3 (1997): 131-156.

[14] Anjum, Farooq, Dhanant Subhadrabandhu, and Saswati Sarkar. "Signature based intrusion detection for wireless ad-hoc networks: A comparative study of various routing protocols." *Vehicular Technology Conference, 2003. VTC 2003-Fall. 2003 IEEE 58th*. Vol. 3. IEEE, 2003.

[15] Azer, Marianne A., Sherif Mohammed El-Kassas, and Magdy Saeed El-Soudani. "A survey on anomaly detection methods for ad hoc networks." *Ubiquitous Computing and Communication Journal* 2.3 (2005): 67-76.

[16] Benferhat, Salem, and Karim Tabia. "On the combination of naive bayes and decision trees for intrusion detection." *Computational Intelligence for Modelling, Control and Automation, 2005 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce,* Vol. 1. IEEE, 2005.

[17] Winfield, Alan FT, and O. E. Holland. "The application of wireless local area network technology to the control of mobile robots." *Microprocessors and Microsystems* 23.10 (2000): 597-607.

[18] Sheldon, Frederick T., et al. "The insecurity of wireless networks." *IEEE Security & Privacy* 10.4 (2012): 54-61.

[19] Potter, Bruce. "Wireless security's future." *IEEE Security & Privacy* 99.4 (2003): 68-72.

[20] Vihervuori, Olli. "Recent Developments in IEEE 802.11 Wireless Local Area Network Link-layer Security." *TKK T-110.5190 Seminar on Internetworking*. 2009.

[21] Welch, Donald, and Scott Lathrop. "Wireless security threat taxonomy." *Information Assurance Workshop, 2003. IEEE Systems, Man and Cybernetics Society*. IEEE, 2003.

[22] Mathews, Moffat, and Ray Hunt. "Evolution of wireless LAN security architecture to IEEE 802.11 i (WPA2)." *Proceedings of the fourth IASTED Asian conference on communication systems and networks*. 2007.

[23] Housley, Russ, and William Arbaugh. "Security problems in 802.11-based networks." *Communications of the ACM* 46.5 (2003): 31-34.

[24] Feng, Pan. "Wireless LAN security issues and solutions." *Robotics and Applications (ISRA), IEEE Symposium*, 2012

[25] Xiao, Yang, et al. "Security mechanisms, attacks and security enhancements for the IEEE 802.11 WLANs." *International journal of wireless and mobile computing* 1.3-4 (2006): 276-288.

[26] Liu, Yonglei, Zhigang Jin, and Ying Wang. "Survey on security scheme and attacking methods of WPA/WPA2." *6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM),* IEEE, 2010.

[27] Farooq, Taimur, David Llewellyn-Jones, and Madjid Merabti. "Mac layer DOS attacks in IEEE 802.11 networks." *The 11th Annual Conference on the Convergence of Telecommunications, Networking and Broadcasting (PGNet 2010), Liverpool, UK*. 2010.

[28] Pervaiz, Mohammad O., Mihaela Cardei, and Jie Wu, "Security in wireless local area networks", *Department of Computer Science & Engine*, 2007

[29] Bhattacharjee, Bala Srinivasanand Nandita. "Security analysis and improvements on WLANs." *Journal of Networks* 6.3 (2011): 470-281.

[30] ElGili, Mustafa, Samani A. Talab, and Awad H. Ali. "WEP and WPA Improvement." *Wireless Sensor Network* 2.03 (2010): 239-242.

[31] Stubblefield, Adam, John Ioannidis, and Aviel D. Rubin. "A key recovery attack on the 802.11 b wired equivalent privacy protocol (WEP)." *ACM transactions on information and system security (TISSEC)* 7.2 (2004): 319-332.

[32] Mishra, Amitabh, Ketan Nadkarni, and Animesh Patcha. "Intrusion detection in wireless ad hoc networks." *IEEE wireless communications* 11.1 (2004): 48-60.

[33] Garcia-Teodoro, Pedro, et al. "Anomaly-based network intrusion detection: Techniques, systems and challenges." *computers & security* 28.1-2 (2009): 18-28.

[34] Liu, Yu, Yang Li, and Hong Man. "MAC layer anomaly detection in ad hoc networks." *Information Assurance Workshop, 2005. IAW'05. Proceedings from the Sixth Annual IEEE SMC*. IEEE, 2005.

[35] Spafford, Eugene, and Diego Zamboni. "Data collection mechanisms for intrusion detection systems." *CERIAS Technical Report* 2000.8 (2000).

[36] Derrick, E. Joseph, Richard W. Tibbs, and Larry Lee Reynolds. "Investigating new approaches to data collection, management and analysis for network intrusion detection." *Proceedings of the 45th annual southeast regional conference*. ACM, 2007.

[37] Hu, Liang, Kuo Zhao, and Bo Li. "A data collection model for intrusion detection system based on simple random sampling." *Computational Methods*. Springer, Dordrecht, 2006. 1081-1085.

[38] Duffield, Nick G., and Matthias Grossglauser. "Trajectory sampling for direct traffic observation." *IEEE/ACM Transactions on Networking (ToN)* 9.3 (2001): 280-292.

[39] Kolias, Constantinos, et al. "Intrusion detection in 802.11 networks: empirical evaluation of threats and a public dataset." *IEEE Communications Surveys & Tutorials* 18.1 (2016): 184-208.

[40] Zhao, Kuo, et al. "Data collection for intrusion detection system based on stratified random sampling." *IEEE International Conference on Networking, Sensing and Control.* IEEE, 2007.

[41] Hsu, Hui-Huang, Cheng-Wei Hsieh, and Ming-Da Lu. "Hybrid feature selection by combining filters and wrappers." *Expert Systems with Applications* 38.7 (2011): 8144-8150.

[42] Ganapathy, Sannasi, et al. "Intelligent feature selection and classification techniques for intrusion detection in networks: a survey." *EURASIP Journal on Wireless Communications and Networking* 2013.1 (2013): 271.

[43] Y. H. Liu, D. X. Tian, and D. Wei, “A Wireless Intrusion Detection Method Based on Neural Network”, *Proc. Second IASTED International Conference Advances in Computer Science and Technology* (2006): 207-211

[44] Alliance, Wi-Fi. "Wi-Fi Protected Access: Strong, standards-based, interoperable security for today's Wi-Fi networks." *White paper, University of Cape Town* (2003): 492-495.

[45] Mitrokotsa, Aikaterini, Rosa Mavropodi, and Christos Douligeris. "Intrusion detection of packet dropping attacks in mobile ad hoc networks." *Proceedings of the International Conference on Intelligent Systems And Computing: Theory And Applications*. 2006.

[46] Tsang, Chi-Ho, Sam Kwong, and Hanli Wang. "Genetic-fuzzy rule mining approach and evaluation of feature selection techniques for anomaly intrusion detection." *Pattern Recognition* 40.9 (2007): 2373-2391.

[47] Wang, Yong, et al. "Computational intelligence algorithms analysis for smart grid cyber security." *International Conference in Swarm Intelligence*. Springer, Berlin, Heidelberg, 2010.

[48] Chitrakar, Roshan, and Chuanhe Huang. "Anomaly based intrusion detection using hybrid learning approach of combining k-medoids clustering and naive bayes classification."*8th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM).* IEEE, 2012.

[49] Chitrakar, Roshan, and Chuanhe Huang. "Anomaly detection using Support Vector Machine classification with k-Medoids clustering." *Third Asian Himalayas International Conference on Internet (AH-ICI)*. IEEE, 2012.

[50] Farid, Dewan Md, Nouria Harbi, and Mohammad Zahidur Rahman. "Combining naive bayes and decision tree for adaptive intrusion detection." *International Journal of Network Security & Its Applications (IJNSA)*, 2010.

[51] Fu, Song, Jianguo Liu, and Husanbir Pannu. "A hybrid anomaly detection framework in cloud computing using one-class and two-class support vector machines." *International Conference on Advanced Data Mining and Applications*. Springer, Berlin, Heidelberg, 2012.

[52] Tang D. H., Cao Z., "Machine Learning-based Intrusion Detection Algorithms." *Journal of Computational Information Systems*, 2009.

[53] Yasami, Yasser, and Saadat Pour Mozaffari. "A novel unsupervised classification approach for network anomaly detection by k-Means clustering and ID3 decision tree learning methods." *The Journal of Supercomputing* 53.1 (2010): 231-245.

[54] Peddabachigari, Sandhya, et al. "Modeling intrusion detection system using hybrid intelligent systems." *Journal of network and computer applications* 30.1 (2007): 114-132.

[55] Li, Hongjian, Ming Xu, and Yi Li. "The research of frame and key technologies for intrusion detection system in ieee 802.11-based wireless mesh networks." *International Conference on Complex, Intelligent and Software Intensive Systems, 2008*. IEEE, 2008.

[56] Uppuluri, Prem, and R. Sekar. "Experiences with specification-based intrusion detection." *International Workshop on Recent Advances in Intrusion Detection*. Springer, Berlin, Heidelberg, 2001.

[57] Sekar, R., et al. "Specification-based anomaly detection: a new approach for detecting network intrusions." *Proceedings of the 9th ACM conference on Computer and communications security*. ACM, 2002.

[58] Bao, Fenye, et al. "Trust-based intrusion detection in wireless sensor networks." *IEEE International Conference on Communications (ICC)*. IEEE, 2011.

[59] Sampangi, Raghav V., Saurabh Dey, and Vighnesh N. Viswanath. "The sneeze algorithm: A social network & biomimetic approach for intrusion detection in wireless networks." *Business Applications of Social Network Analysis (BASNA), 2010 IEEE International Workshop on*. IEEE, 2010.

[60] Zhong, Shi, Taghi M. Khoshgoftaar, and Shyam Varan Nath. "A clustering approach to wireless network intrusion detection." *17th IEEE International Conference on Tools with Artificial Intelligence, ICTAI 05.* IEEE, 2005.

[61] Mitchell, Rob, Ray Chen, and Mohamed Eltoweissy. "Signalprint-based intrusion detection in wireless networks." *International Workshop on Security in Emerging Wireless Communication and Networking Systems*. Springer, Berlin, Heidelberg, 2009.

[62] Hairui, Wang, and Wang Hua. "Research and design of multi agent-based intrusion detection system on wireless network." *International Symposium on Computational Intelligence and Design, 2008. ISCID'08,* Vol. 1. IEEE, 2008.

[63] Haddadi, Fariba, and Mehdi A. Sarram. "Wireless intrusion detection system using a lightweight agent." *Second International Conference on Computer and Network Technology (ICCNT)*. IEEE, 2010.

[64] Yuan, Song, Qi-juan Chen, and Peng Li. "Design of a four-layer ids model based on immune danger theory." *WiCom'09 5th International Conference on Wireless Communications, Networking and Mobile Computing, 2009.* IEEE, 2009.

[65] Sampangi, Raghav V., Saurabh Dey, and Vighnesh N. Viswanath. "The sneeze algorithm: A social network & biomimetic approach for intrusion detection in wireless networks." *IEEE International Workshop on Business Applications of Social Network Analysis (BASNA)*. IEEE, 2010.

[66] Stibor, Thomas, et al. "Geometrical insights into the dendritic cell algorithm." *Proceedings of the 11th Annual conference on Genetic and evolutionary computation*. ACM, 2009.

[67] Farid, Dewan Md, and Mohammad Zahidur Rahman. "Learning intrusion detection based on adaptive bayesian algorithm." *11th International Conference on Computer and Information Technology, 2008. ICCIT 2008.* IEEE, 2008.

[68] Jones, Anita, and Song Li. "Temporal signatures for intrusion detection." *Proceedings of 17th Annual Computer Security Applications Conference, ACSAC 2001.* IEEE, 2001.

[69] Han, Hong, Xin-Liang Lu, and Li-Yong Ren. "Using data mining to discover signatures in network-based intrusion detection." *Proceedings of International Conference on Machine Learning and Cybernetics, 2002.* Vol. 1. IEEE, 2002.

[70] Lin, Ying, Yan Zhang, and Yang-jia Ou. "The design and implementation of host-based intrusion detection system. *Third International Symposium on Intelligent Information Technology and Security Informatics (IITSI), 2010*. IEEE, 2010.

[71] Uppuluri, Prem, and R. Sekar. "Experiences with specification-based intrusion detection." *International Workshop on Recent Advances in Intrusion Detection*. Springer, Berlin, Heidelberg, 2001.

[72] Sekar, R., et al. "Specification-based anomaly detection: a new approach for detecting network intrusions." *Proceedings of the 9th ACM conference on Computer and communications security*. ACM, 2002.

[73] Foo, Bingrui, et al. "ADEPTS: adaptive intrusion response using attack graphs in an e-commerce environment." *Proceedings of International Conference on Dependable Systems and Networks, DSN 2005.* IEEE, 2005.

[74] Amiri, Fatemeh, et al. "Mutual information-based feature selection for intrusion detection systems." *Journal of Network and Computer Applications* 34.4 (2011): 1184-1199.

[75] Stein, Gary, et al. "Decision tree classifier for network intrusion detection with GA-based feature selection." *Proceedings of the 43rd annual Southeast regional conference-Volume 2*. ACM, 2005.

[76] Sung, Andrew H., and Srinivas Mukkamala. "The feature selection and intrusion detection problems." *Annual Asian Computing Science Conference*. Springer, Berlin, Heidelberg, 2004.

[77] Li, Yinhui, et al. "An efficient intrusion detection system based on support vector machines and gradually feature removal method." *Expert Systems with Applications* 39.1 (2012): 424-430.

[78] Mukherjee, Saurabh, and Neelam Sharma. "Intrusion detection using naive Bayes classifier with feature reduction." *Procedia Technology* 4 (2012): 119-128.

[79] Chandrashekar, Girish, and Ferat Sahin. "A survey on feature selection methods." *Computers & Electrical Engineering* 40.1 (2014): 16-28.

[80] Helmer, Guy, et al. "Feature selection using a genetic algorithm for intrusion detection." *Proceedings of the 1st Annual Conference on Genetic and Evolutionary Computation-Volume 2*. Morgan Kaufmann Publishers Inc., 1999.

[81] Helmer, Guy, et al. "Automated discovery of concise predictive rules for intrusion detection." *Journal of Systems and Software*60.3 (2002): 165-175.

[82] Liu, Huan, and Lei Yu. "Toward integrating feature selection algorithms for classification and clustering." *IEEE Transactions on knowledge and data engineering* 17.4 (2005): 491-502.

[83] Sindhu, Siva S. Sivatha, S. Geetha, and Arputharaj Kannan. "Decision tree based light weight intrusion detection using a wrapper approach." *Expert Systems with applications* 39.1 (2012): 129-141.

[84] Verikas, Antanas, and Marija Bacauskiene. "Feature selection with neural networks." *Pattern Recognition Letters* 23.11 (2002): 1323-1335.

[85] Kabir, Md Monirul, Md Monirul Islam, and Kazuyuki Murase. "A new wrapper feature selection approach using neural network." *Neurocomputing* 73.16-18 (2010): 3273-3283.

[86] Kabir, Md Monirul, Md Shahjahan, and Kazuyuki Murase. "A new hybrid ant colony optimization algorithm for feature selection." *Expert Systems with Applications* 39.3 (2012): 3747-3763.

[87] Li, Yongzhong, et al. "Intrusion detection method based on fuzzy hidden Markov model." *Sixth International Conference on Fuzzy Systems and Knowledge Discovery, 2009. FSKD'09.* Vol. 3. IEEE, 2009.

[88] Kasliwal, Bhavesh, et al. "A hybrid anomaly detection model using G-LDA." *Advance Computing Conference (IACC), 2014 IEEE International*. IEEE, 2014.

[89] Guennoun, Mouhcine, Aboubakr Lbekkouri, and Khalil El-Khatib. "Selecting the best set of features for efficient intrusion detection in 802.11 networks." *3rd International Conference on Information and Communication Technologies: From Theory to Applications, 2008. ICTTA 2008.* IEEE, 2008.

[90] Schaffernicht, Erik, and Horst-Michael Gross. "Weighted mutual information for feature selection." *International Conference on Artificial Neural Networks*. Springer, Berlin, Heidelberg, 2011.

[91] Münz, Gerhard, Sa Li, and Georg Carle. "Traffic anomaly detection using k-means clustering." *GI/ITG Workshop MMBnet*. 2007.

[92] Jianliang, Meng, Shang Haikun, and Bian Ling. "The application on intrusion detection based on k-means cluster algorithm." *International Forum on Information Technology and Applications, 2009. IFITA'09.* Vol. 1. IEEE, 2009.

[93] Aljarah, Ibrahim, and Simone A. Ludwig. "Mapreduce intrusion detection system based on a particle swarm optimization clustering algorithm." *IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2013.

[94] Bakar, Azuraliza Abu, et al. "An agent based rough classifier for data mining." *Eighth International Conference on Intelligent Systems Design and Applications, ISDA'08.* Vol. 1. IEEE, 2008.

[95] Chebrolu, Srilatha, Ajith Abraham, and Johnson P. Thomas. "Feature deduction and ensemble design of intrusion detection systems." *Computers & security* 24.4 (2005): 295-307.

[96] Wang, Hui, et al. "A novel intrusion detection method based on improved SVM by combining PCA and PSO." *Wuhan University Journal of Natural Sciences* 16.5 (2011): 409.

[97] Enache, Adriana-Cristina, and Victor Valeriu Patriciu. "Intrusions detection based on support vector machine optimized with swarm intelligence." *IEEE 9th International Symposium on Applied Computational Intelligence and Informatics (SACI)*. IEEE, 2014.

[98] Huang, Cheng-Lung, and Jian-Fan Dun. "A distributed PSO–SVM hybrid system with feature selection and parameter optimization." *Applied soft computing* 8.4 (2008): 1381-1391.

[99] Deng, Hongmei, Qing-An Zeng, and Dharma P. Agrawal. "SVM-based intrusion detection system for wireless ad hoc networks." *IEEE 58th Vehicular Technology Conference, 2003. VTC 2003-Fall*. Vol. 3. IEEE, 2003.

[100] Bose, S., S. Bharathimurugan, and A. Kannan. "Multi-layer integrated anomaly intrusion detection system for mobile adhoc networks." *International Conference on Signal Processing, Communications and Networking, 2007. ICSCN'07.* IEEE, 2007.

[101] Zhang, Wei, et al. "A cooperative network intrusion detection based on SVMs." *6th International Conference on Pervasive Computing and Applications (ICPCA)*. IEEE, 2011.

[102] Chen, Wun-Hwa, Sheng-Hsun Hsu, and Hwang-Pin Shen. "Application of SVM and ANN for intrusion detection." *Computers & Operations Research* 32.10 (2005): 2617-2634.

[103] Li, Kun-Lun, et al. "Improving one-class SVM for anomaly detection." *International Conference on Machine Learning and Cybernetics*. Vol. 5. IEEE, 2003.

[104] Ambwani, Tarun. "Multi class support vector machine implementation to intrusion detection." *Proceedings of the International Joint Conference on Neural Networks*. Vol. 3. IEEE, 2003.

[105] Usha, M., and P. Kavitha. "Anomaly based intrusion detection for 802.11 networks with optimal features using SVM classifier." *Wireless Networks* 23.8 (2017): 2431-2446.