UTM UNIVERSITI TEKNOLOGI MALAYSIA
INTERNATIONAL JOURNAL OF INNOVATIVE COMPUTING
ISSN 2180-4370
Journal Homepage : https://ijic.utm.my/

A Review on Network Intrusion Detection System Using Machine Learning

Bello Nazifi Kagara & Maheyzah Md Siraj
Faculty of Computing
Universiti Teknologi Malaysia
81310 UTM Johor Bahru, Johor, Malaysia
nazefballo@gmail.com; maheyzah@utm.my

Submitted: 19/01/2020. Revised edition: 30/04/2020. Accepted: 1/04/2020. Published online: 20/05/2020
DOI: https://doi.org/10.11113/ijic.v10n1.252

***Abstract*—The quality or state of being secure is the crucial concern of our daily life usage of any network. However, with the rapid breakthrough in network technology, attacks are becoming more trailblazing than defenses. It is a daunting task to design an effective and reliable intrusion detection system (IDS), while maintaining minimal complexity. The concept of machine learning is considered an important method used in intrusion detection systems to detect irregular network traffic activities. The use of machine learning is the current trend in developing IDS in order to mitigate false positives (FP) and False Negatives (FN) in the anomalous IDS. This paper targets to present a holistic approach to intrusion detection system and the popular machine learning techniques applied on IDS systems, bearing In mind the need to help research scholars in this continuous burgeoning field of Intrusion detection (ID).**

***Keywords*—Intrusion detection system (IDS), Machine Learning Feature selection, Data mining**

I. INTRODUCTION

Network attacks are increasing daily. Intrusion is considered the most widely reported attack on network traffic. The intrusion detection system has been used to detect intrusion and preserve information security goals. With the enormous increase in the usage of network, network traffic security is becoming a significant issue in the network system with the spectacular development of the internet [1]. Conventional detection systems for intrusion are limited and do not provide a full solution to this issue [2]. They search the network traffic for possible malicious activities, they often succeed in finding correct security breaches and anomalies. In many cases though, they fail to detect malicious (false negative) actions or fire alarms when there is nothing wrong in the network (false positive).

The intrusion of a network is an unauthorized network operation that steals any sensitive or classified data. It is also, sometimes the reason network services are unavailable. The unexpected anomaly occurs frequently, and the Internet cyber world is suffering a significant loss in terms of data security, and the security of potential information, etc. The security system must, therefore, be stable, efficient and well-configured. There are mainly two forms of detecting network intrusion. One is based on signatures and another is a detection system based on anomalies. Signature-based intrusion detection systems include monitoring network traffic for a set of bytes or packet sequences considered to be an anomaly. A significant advantage of this identification scheme is that, when one understands what network activity needs to be detected, signatures are relatively easier to develop and understand. Detection based on signatures also has some disadvantages. For each attack, a signature must be generated, and they can detect only those attacks. However, if signatures are unknown to the detection scheme, they cannot detect any other novel attacks.

Intrusion detection enhances these systems of defense to boost system security. However, although the proactive security mechanisms will effectively secure information systems, it is still important to know what intrusions have occurred or are happening, so that we can recognize the dangers and vulnerabilities to protect and thus, be better prepared for future attacks. IDS's are not alternatives for proactive security mechanisms, like access control and encryption, despite such value. Nevertheless, Intrusion Detection Systems cannot offer proper protection to information systems itself. As an illustrative example, tracking

the attackers cannot lower the harm at all if an intruder erases all the records in an information system. Therefore, as part of a broader protection system, intrusion detection systems must be implemented along with the other preventive security mechanisms.

This manuscript is organized in an easy to follow format and sections, it includes an introduction to IDS, and the challenges they face, followed by the related scientific researches in IDS, also summarized in a table for better comprehension. In the fourth section, types of IDS were discussed followed by a taxonomy of IDS for a holistic view. Machine learning algorithms applied on IDS and data mining were discussed succinctly. And lastly conclusion and an acknowledgement.

II. CHALLENGES FACING THE CONVENTIONAL IDS'S

Distinguishing between intrusion and regular network traffic activities is quite challenging and takes much effort into the process. With the inherent use of networks day in day out, it is without doubt that malicious activities will be unveiling every day in the society. Networks have become essential for almost every aspect of our lives, it is without doubt that same energy equivalence has to be put in place in order to safeguard these networks and enforce a uniform orderly manner of usage. An analyst will examine all of the information in the process of figuring the intrusion chain on the network connection which is extensive. A fairly optimized approach is desired to build a system, to model the networks intrusion into the existing network traffic. Thus, it is crucial to have a mechanism to identify and classify all attacks in detail and also to mitigate the false alarm frequency. The most prominent challenges facing traditional IDS are their inability to identify unique or unfamiliar attacks as they are based on signature [3]

Networks are playing a vital role in the whole world today and the consequences of having a downfall or an event where the networks become crippled is unacceptable. New techniques and methods have to be devised in order to meet up with the prevailing attacks by perpetrators in order to gain illicit access to networks and carry out their dubious acts. The main issue currently facing the Intrusion Detection Systems (IDS) is high false positives (FP) and false negatives (FN). Each day, the intrusion detection system produces 15,000 alerts, and 1,000 are FP. These errors created by the IDS's, subsequently makes the network administrators and end-users to lose confidence in security warnings recorded by these devices [4]

The moment an irregular behavior is detected in a network, the IDS triggers an alert and constantly pass it to the network administrator, and in the case of IPS, a proactive measure is taken to block the suspicious traffic. Nevertheless, there is no perfect detection strategy, which can differentiate flawlessly between malicious and ordinary traffic. Traditionally, these detection mechanisms are capable of misinterpreting normal traffic as malicious, thereby triggering a false positive (FP), or malicious traffic as an ordinary, triggering a false negative (FN). FPs and FNs pose a variety of problems. With the constant misinterpretation of the network traffic, too many FPs could mask real attacks and thus overwhelm the security operator. In the case of actual attacks, that is true positives, will be completely hidden within the false FPs, making it feasible for the security operator to ignore them [5].

III. LITERATURE REVIEW

Researchers have lately proposed several intrusion detection systems that use the technique of selecting features and conventional methods to improve classification accuracy.

Abdullahi *et al.* (2018), implemented an intrusion detection method with a range of attributes which utilizes the NSLKDD dataset, on the basis of splitting the input data set into various subsets depending on each threat. The optimal set of features were generated by merging the outline of subsets extracted using the info gain filter.

Khammassi and Krichen (2017), selected the best sub-set of features utilized as search strategy and as a learning algorithm, logistic regression was employed for the development of a network intrusion detection system using wrapper techniques based on a genetic algorithm. Their methodology supported the KDD99 dataset with effective detection rate with 18 features and 20 features for the UNSWNB15 dataset.

Paulaukas and Aukalnis (2017), also proposed an ensemble model and employed the use of four distinct classifiers, J48, C5.0, Naïve Bayes and PART, with focus on putting together poorer classifiers to compose richer ones. Their result of the ensemble model achieved more accurate results in the Intrusion detection system.

Farid *et al.* (2016), the authors addressed the complexity of the intrusion detection datasets, as many of them are robust and have several attributes. Some of these features may be redundant or may not significantly contribute to intrusion detection. The purpose of this work was to define useful attributes from the training dataset to build a classifier using data mining algorithms. Experimental results on the KDD'99 data set for intrusion detection show that the proposed method achieves high classification rate and decreases false positives in such an environment with limited computational resources.

Malik *et al.* (2015), suggested a model using Particle Swamp Optimization (PSO) and Random Forest (RF) The model proposed achieves higher accuracy compared to other classification algorithms in line with a low false-positive rate.

Güneş *et al.* (2015), performed a feature importance study on the dataset KDD CUP 99 to identify the impact of features on machine attack. Based on Adaptive Resonance Theory (ART) and Principal Component Analysis (PCA).

Chabathula *et al.* (2015), proposed a main approach to component analysis in feature reduction and feature selection using machine learning. Proposed a principal component

Analysis approach using machine learning for feature reduction and selection. The authors adopt a different approach to network analysis in order to reduce data features. NSL-KDD dataset was used in the research.

Hota and Shrivas (2014), suggested a model using multiple feature selection methods to eliminate unimportant data set attributes and create a robust and effective classifier. The results revealed better results for the combination of C4.5 and inforGain and acquired a precision of 99.68 percent with 17 attributes.

Horng *et al.* (2011), proposed an SVM-based intrusion detection model that includes a hierarchical clustering algorithm, simple selection process for features, and SVM technique. There have been fewer abstract and higher qualified training instances of the hierarchical clustering algorithm, extracted from the KDD-Cup 1999 training set. It could considerably reduce the practice time, and also enhance the success of the resulting SVM. The straightforward feature selection technique was administered to exclude irrelevant features in the training set in order to enable the obtained SVM model to interpret the network traffic information more precisely. The data set used to test the proposed system is the KDD-Cup 1999. The model showed better performance in detecting DoS and Probe attacks and the best overall accuracy performance compared to other intrusion detection systems centered on the same dataset.

Table 1.

TABLE 1 A synopsis of researches in IDS design

Author/Year	Aim of the study	Data set	Methodology	Results
Y Zhou *et al.*/2019	To develop an efficient IDS with high accuracy and low false alarms.	NSL-KDD, KDDCUP 99, CIC-IDS2017	Combination of correlation-base feature selection (CFS), Bat Algorithm (BA), proposed CFS-BA. As feature selection techniques to improve classification efficiency.	Accuracy and precision have significantly increase with the proposed CFS-BA.
Abdullahi *et al*./2018	To build a framework of intrusion detection with minimum number of features in the dataset.	NSL-KDD	Used multiple features selection techniques to get the optimal data subsets.	Highest accuracy obtained using Random Forest and PART classifiers under combination methods.
Paulaukas and Aukalnis /2017	In order to create a stronger learner.	NSL-KDD	An ensemble of four different classifiers: J48, C5.0, Naïve Bayes and PART. Which depends on the idea of combining multiple weak learners	Their results prove that their ensemble model produces more accurate results for an IDS.
Kamassi and Krichen/2017	To reduce the dimensionality of the subsets	KDD99-2 & UNSW-NB1 5	Applied wrapper method based on Genetic Algorithm as the search strategy and Logistic Regression as the learning algorithm	Their method provides high detection rate with a subset of only 18 features for the KDD99 dataset & 20 features for the UNSW-NB15 dataset.
Hota and Shrivas/ 2014	In order to obtain more robust and effective classifier	NSL-KDD	Four feature selection techniques: symmetrical uncertainty, Relief, correlation, & Infor Gain, combined with C4.5 decision tree technique	C4.5 with Infor. Gain had better results and achieved and accuracy of 99.68% with only 17 features. Symmetrical Uncertainty with C4.5 is also promising with 99.64% with 11 features.
Gupta and Shrivastava /2015	To achieve high quality performance of Intrusion Detection System.	KDD99 data set	SVM was used to classify normal attacks and BC to enhance performance improvements in IDS.	Increased accuracy.
Horng *et al*./2011	Proposed SVM-based IDS which combines a hierarchical clustering algorithm a simple feature selection procedure.	KDD- Cup 1999	An SVM-based intrusion detection system with BIRCH algorithm is proposed.	greatly shorten training time, improved the performance of resultant SVM and detects DoS and Probe attacks better than previous methods.

IV. CATEGORIES OF IDS

1) Based on Known or Unknown Attack Patterns

i- Anomaly Based Intrusion Detection System

Detection based on anomaly defends against threats that have not been identified. An "anomaly" is something unusual. If any abnormal traffic is obtained from the model, a suspected intrusion alert will be triggered by the IDS. Having a lot of Telnet in less than twenty-four hours, Hypertext Transfer Protocol (HTTP) on nonstandard port, and bulky Simple Network Management Protocol (SNMP) traffic are a few examples of abnormal behavior. IDS initially establishes a baseline profile reflecting normal traffic behavior [6].

A model profile is developed that the IDS uses for some time to learn about traffic to study activity and behaviors during peak times, non-peak times, late times, and early business times, according to the nature of the organizational network. The traffic acquired in a designated amount of time is subsequently learnt statistically and a baseline profile is developed. When the ID system is switched either detection or prevention mode from learning mode, it begins to compare the usual traffic with the initially created profile, and any detected anomalous activity deviating from the base line profile will trigger an alarm cautioning the administrator to possible intrusion or otherwise prevent it if set to prevention mode. With specific traffic behaviors, user defined profiles can also be generated, for example the amount of e-mails published by a user and user access attempts [7] Examples are Snort, and BroIDS are detection system based on anomaly for ID systems [6].

ii- IDS based on Signatures

A knowledge-based ID system also known as signature-based intrusion detection system, refers to a device-known list of earlier attack signatures and security bugs. When we speak of IDS, the meaning of the word signature is recorded as evidence of an intrusion or invasion. Every intrusion tends to leave a footprint behind (e.g., data packet nature, failed application execution attempt, failed logins, access to files, and access to folders). Such footprints are called signatures and may be used in the future to identify and prevent the same attacks. Knowledge-based IDS detects intrusion attempts on the basis of these signatures [8] The IDS analyzes the information gathered and compares it with massive attack signature databases. Mostly, the IDS is searching for a specific attack already documented. Suspicious data are recognized as a virus detection system by comparing the new captured instances with the known malicious activities already stored. An alarm will be triggered once there is a match [9].

2) Classification Based on Location

i- Network Intrusion Detection System

The Network Intrusion Detection System (NIDS) watch-over network traffic and examines the intrusion transitory traffic. Once a packet is marked as an intrusion or viewed as an abnormal behavior, an alarm may warn the administrator. The NIDS can detect four major types of attacks: probe, user to root, and remote client, DoS. Snort IDS and Cisco Safe IDS are just a few examples of NIDS technology. (Mohammad and others, 2014).

Advantages:

a. Adaptable to the environment of the cross-platform.
b. The NIDS are handled centrally.

Disadvantages:

a. Needs additional training.
b. Use up the bandwidth of the LAN
c. The rate of failure is high.

ii- Host-Based Intrusion Detection System

Host-Based Network Intrusion Detection (HIDS) focuses on safeguarding a specific computer system, as it is incorporated within that particular system, to monitor internal or external intrusions over the system. As far as the internal attack is concerned, it watches which resources are accessed from whom and from whom and also to confirm whether there is any breach of security. A word processor, for example, suddenly begins to access and modify the system password database. HIDS inspects come and go from a computer system on its interfaces in the second part, which is the external attacks. HIDS acts by logging the activity and making it known to the designated authority. In HIDS, in order to monitor security, anti-virus applications are installed on the system, examples include spyware, antivirus, while HIDS instances includes open source tripwire, GFI LAN guard S.E.L.M [10].

iii- Real-Time Intrusion Detection System

This type of detection system works online, that is, it captures live network packets to detect abnormal activity. The overall output of the real-time IDS relies primarily on the number of selected attributes because it has to compare those attributes at a high count with the attributes of the incoming packets. The number of attributes also affects the real-time system's resource use [11]. These systems have the advantages of detecting abnormal behavior while it occurs, which is desired from a detection system for intrusion. Disadvantages are systems in real time require more resources and can turn into the bottleneck.

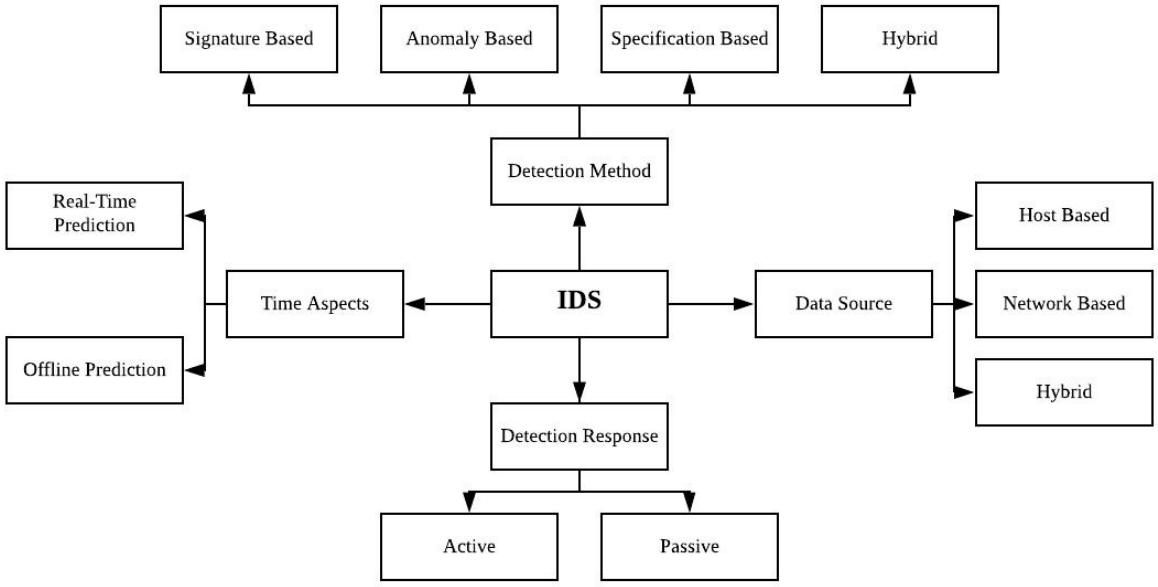

Fig.1. Taxonomy of Intrusion Detection System

V. DATA MINING METHODS APPLIED ON NIDS

Data mining (DM) techniques were increasingly being used to detect threats, anomalies or intrusions in a secure network environment [12]. As a result, data mining development has built a wide range of pattern recognition algorithms, machine learning, and database fields. It is possible to search for audit data using many types of algorithms. Lazar recognizes information algorithms from data heuristics and designs as a composition of data mining models [13].

Many data mining methods are used to detect intrusion. Below are some of them;

1. Classification:

Classification involves analyzing and assigning each dataset to either a normal or an abnormal class. As new instances, the existing structures are maintained. Classification can be used for both misuse and detection of anomalies but used for misuse more predominantly. By assigning data sets within predetermined sets, categorization in classification is achieved. Compared to clustering, it is more computational. In intrusion detection systems, various classification techniques are used, including Naive Bayes, k-nearest neighbor, vector support machine, and decision tree [14].

i- Decision Tree

The Decision Tree is a tree-like hierarchical grouping. each node on a branch denotes a binary field, one attribute represents the predicate's positive instances, and the other branch denotes the unwanted occurrences. The construction of a decision tree does not require prior expertise in a particular domain and has the advantage of handling high-dimensional information [14].

ii- K Nearest Neighbor.

This is among the most basic methods of classification. It measures the vastness of the dissimilar data points of the input vector then allocate the unclassified data point to its closest neighboring component. K is a major factor. K's value determines where it will be assigned when k=1, it will now be assigned to its closest neighbor. If K's value is huge, then prediction takes a long time and influences the accuracy by reducing the noise effect. The function of implementation is simple and easy to execute in parallel.

iii- Naive Bayes Classifier

This is a probability classifier; it forecast the class by the likelihood of the relationship. It evaluates the relationship between independent and dependent parameters in order to obtain conditional statistical likelihood. The design of Naive Bayes is not laborious as far as no complex recurrent factors are concerned. It can be extended to a large number of data points but decreases in time complexity [11].

iv- Support Vector Machine Method

Support vector machine is a monitored learning technique that is used to predict and classify. The SVM

separates two groups relatively by finding the hyperplane to enable better prediction and classification, also in multidimensional space separating classes as it is a binary classification. The two classes are divided into standard and suspicious data denoting + 1 and -1 respectively. The help vector machine's fundamental objective is to find the optimal linear hyperplane to optimize the margin between the two classes [15].

2. Clustering

Given that the network information is wide, it is costly and time-consuming to mark each instance or data point in the category. Clustering involves grouping related data elements into different groups based on their similarities and between equivalence classes or partitioning the data into clusters within a single group. The technique of clustering is divided into two groups depending on the configuration of the cluster, namely hierarchical connection-oriented and non-hierarchical.

i- K-Means Clustering Algorithm

This is the most straightforward and broadly adopted clustering technique proposed by James Macqueen.
Several clusters specified by the user in k in this algorithm mean classifying cases into a predefined number of clusters. The initial stage clustering Kmeans selects the k instances as the cluster center. First, assign the nearest cluster to each instance of the dataset. For example, assignment, use Euclidean distance to measure the length between the centroid and each instance, and assign each data point to the cluster according to the lowest distance. K –Means algorithm has taken less time to execute when it applied to a specific dataset. The execution time is directly proportional to the time taken by the data, and the maximum increase in point will result in the total execution time. It's a quick iterative algorithm, but it's sensitive to noise and outlier [16].

ii- K-Medoids Clustering Algorithm

K-Medoids clustering, also known as Partitioning Around Medoid, can be identified as a cluster point with the minimum differences with all other cluster points. Through partitioning the algorithm as a K-means algorithm, K-Medoids is clustering. Instead of taking the mean value of the objects in K-Means clustering, the most centrally located instance in a cluster is known as centroid. The reference point and medoid is called this centrally located artifact. To minimize the squared error, this reduces the distance between centroid and data points. When the amount of data points rises to the limit, K-Medoids technique works better than the K-Means technique. It is stable in the presence of noise and external, as outliers have less impact on the medoid, but it is costly to process [16].

VI. EXTRACTING FEATURES

In existing papers, feature reduction and selection are usually used to detect intrusion. The methods are often used interchangeably to describe specific points. A reduction or extraction function is the process of getting new subspaces of lesser size than the initial feature space [17]. Another extension to define the selection of features is that the attributed given by the selection of features have to often be a subset of an existing set of attributes, thus feature reduction reduces the dimensions of the original set's linear combination with new synthetic features [17].

i- Feature Reduction

Reduction of feature is by finding a new subspace that has fewer dimensions than the novel space of the feature. Linear Discriminant Analysis (LDA), Principal Component Analysis (PCA), Uncorrelated Linear Discriminant Analysis (ULDA) and Independent Component Analysis (ICA) are the widely used methods of feature reduction. With regard to the pre-processing technique of selecting items, it seems that in response to the current uproar, most hybrid techniques are being tested to improve the efficiency and reliability problems found in current data sets [18] created a multi-class classification to assist with improved efficiency and reliability in the establishment of new ID systems. In order to produce better results, the aim was to combine different classifiers. The authors highlight the significant advantages of using multi-class IDS, supported by recent studies and suggestions for analysis. The writer further notes that, during implementation, the overall design performance decreased. This method is caused by attempts to combine patterns of classification where features do not fully fit, resulting in data duplication [18]. The author suggests a plan for increasing unnecessary features and enhancing model efficiency.

ii- Feature Selection

The selection of features is aimed at finding a subcategory of features within a given set that are sufficiently complete to reflect the information, and the elements in the subcategory are extremely important for prediction [19]. Function selection techniques can be broadly classified as filters, wrappers, and embedded approaches. The meaning of the features is evaluated from the dataset for the filter approach, and the attributes are chosen on the basis of statistics, the performance of the classification is used in the wrapper method because part of the evaluation and selection process of the function subcategories. Embedded approaches are arithmetically less detailed than wrappers in comparison to wrapper approaches as they incorporate communication with selection of features and learning procedures. Although embedded approaches integrate a more frequent risk

outcome to efficiently use variables classification attributes and predictor variables, changing the pattern of categorization to achieve more excellent performance is not easy. Current intrusion detection data sets typically include a wide range of redundant features can reduce the performance of data mining algorithms, resulting in undecidable redundant results [19]. The first step is therefore to decrease the dimensionality and choose the dataset's feature subset.

VII. BENCHMARK DATASET

The NSLKDD data was proposed by Tavallaee *et al*. extracted from the KDD'99 dataset in order to improve the dataset. The NSLKDD dataset is the refined version of the existing KDD99 dataset. The data sets include internet traffic records seen through a basic intrusion detection network The NSLKDD contains the same features as the KDD99, 41 attributes and one class attribute. Within the data set exists four different classes of attacks: Denial of Service (DoS), Probe, User to Root(U2R), and Remote to Local (R2L). This data set is comprised of four sub data sets: KDDTest+, KDDTest-21, KDDTrain+, KDDTrain+_20Percent, although KDDTest-21 and KDDTrain+_20Percent are subsets of the KDDTrain+ and KDDTest+. Several forms of study are performed by many NSL-KDD dataset experts using various tools and methods with a common goal of developing an efficient intrusion detection model [20].

The NSLKDD dataset has the following improvements which made it better than the KDD99 dataset:

i. The training set does not contain duplicate records, which will enable the classifiers not to be biased on recurring records.
ii. There are no redundant data in the proposed test sets; thus, learners' results are not biased.
iii. The levels of classification of distinct machine learning methods vary over a wider spectrum, making it more effective to provide reliable evaluations of specific learning techniques [20].

CONCLUSION

Network intrusion is the common goal of intruders. The increase in network usage is proportional to malicious attacks on these networks, hence, matching efforts have to be made in order to meet up with these high risks. This paper elucidated the commonly used data mining techniques applied on intrusion detection systems to eliminate unnecessary and redundant elements of the wide range of alerts received by administrators and focus on the significant attributes, alongside various machine learning machine learning techniques use to improve the detection rate in anomalous intrusion detection systems, which will aid classifiers predictions. The use of ensemble methods is a requisite to maintain the integrity of our networks as the involvement of human effort in the system is less fruitful.

ACKNOWLEDGMENT

I wish to extend my deep sense of sincere gratitude to my research supervisor, Dr. Maheyzah Md Siraj. The school at large, University Teknology Malaysia, for providing me with the necessary skills, knowledge and learning facilities to prepare this paper, and my father, Bello Kagara for his unending support and encouragement towards my academics.

REFERENCES

[1] S. A. Abhaya, K. Kumar, R. Jha. 2014. Data Mining Techniques for Intrusion Detection: A Review. *Int. J. Adv. Res. Comput. Commun. Eng.*, 3, 6938-6941.

[2] A. Youssef and A. Emam. (2011). Network Intrusion Detection Using Data Mining and Network Behaviour Analysis. *Int. J. Comput. Sci. Inf. Technol.*, 3(6), 87-98, Doi: 10.5121/ijcsit.2011.3607.

[3] S. S. Rajan and V. K. Cherukuri. (2010). An Overview of Intrusion Detection Systems. Retrieved May,. 12(3), 559-563, Doi: 10.1109/surv.2010.032210.00054.

[4] R. S. M. Tausif, J. Ferzund, S. Jabbar. 2017. Towards Designing Efficient Lightweight Ciphers for the Internet of Things, KSII Transactions on Internet and Information Systems, 11.

[5] C. Y. Ho, Y. C. Lai, I. W. Chen, F. Y. Wang, and W. H. Tai. (2012). Statistical Analysis of False Positives and False Negatives from Real Traffic with Intrusion Detection/Prevention Systems. *IEEE Commun. Mag.*,. 50(3), 146-154, Doi: 10.1109/MCOM.2012.6163595.

[6] V. Jyothsna, V. V. Rama Prasad, and K. Munivara Prasad. (2011). A Review of Anomaly based Intrusion Detection Systems. *Int. J. Comput. Appl.*, 28(7), 26-35, Doi: 10.5120/3399-4730.

[7] H. H. Soliman, N. A. Hikal, and N. A. Sakr. (2012). A Comparative Performance Evaluation of Intrusion Detection Techniques for Hierarchical Wireless Sensor Networks. *Egypt. Informatics J.*, 13(3), 225-238, Doi: 10.1016/j.eij.2012.10.003.

[8] L. K. and B. thuraisingham Masud, M., Masud, M., L. Khan and B. Thuraisingham. (2012). *Data Mining Tools for Malware Detection.* Boca Raton, FL: CRC Press,15-38.

[9] P. G. Scholar. (2015). Data Mining Techniques for Efficient Intrusion Detection System : A Survey. *Ijaceeonline.Com*, II(Xi).

[10] D. K. Li Y., Xia J., Zhang S., Yan J., Ai X. 2012. An Efficient Intrusion Detection System Based on Support Vector Machines and Gradually Feature Removal Method. *Expert Syst. with Appl.,* 39, 424-430.

[11] K. S. Lee S., Kim G. 2011. Self-adaptive and Dynamic Clustering for Online Anomaly Detection. *Expert Systems with Applications,* 38, 14891-14898.

[12] A. S. Shona, D. (2015). A Survey on Intrusion Detection using Data Mining Technique. *Int. J. Innov. Res. Comput. Commun.*, 3.

[13] S.-W. K. and C. F. T. Lin, C. Cann. 2015. An Intrusion Detection System Based on Combining Cluster Centers and Nearest Neighbours. *Knowledge-based Systems*, 78.

[14] I. R. Hind Tribak, Blanca Delgado, P. Rojas, Olga Valenzuela, Hector Pomares. 2012. Statistical Analysis of Different Artificial Intelligent Techniques Applied to Intrusion Detection System.

[15] D. A. Prasanna P., RaghavRamana A. V. T, Kumar R. K.

2012. Network Programming and Mining Classifier for Intrusion Detection Using Probability Classification. *Proceedings of the International Conference on Pattern Recognition, Informatics and Medical Engineering.*

[16] R. R. Chaudhari and S. P. Patil. (2017). Intrusion Detection System : Classification, Techniques and Datasets to Implement, 1860-1866.

[17] D. Zhang, F. and Wang. (2013). An Effective Feature Selection Approach for Network Intrusion Detection, Networking, Architecture and Storage, *IEEE Eighth Int. Conf.*, 307-311,

[18] E. E. and G. E. Wahba, Y. 2015. Improving the Performance of Multiclass Intrusion Detection Systems using Feature Reduction. *JCSI Int. J. Comput. Sci. Issues*, 12, 355-368.

[19] A. A.-B. Verónica Bolón-Canedo, Noelia Sánchez-Maroño. (2016). Feature Selection for High-dimensional Data. *Progress in Artifcial Intelligence,* 5(2), 65-75.

[20] L. M. Ibrahim, D. B. Taha, and M. S. Mahmod. (2013). A Comparison Study for Intrusion Database (KDD99, NSL-KDD) Based on Self Organization Map (SOM) Artificial Neural Network. *J. Eng. Sci. Technol.*, 8(1), 107-119.