Intrusion Detection Systems: Conceptual Study and Review

Charvi Vij
Computer Science and Engineering
Jaypee University of Information Technology
Himachal Pradesh, India
charvi2617@gmail.com

Hemraj Saini
Computer Science and Engineering
Jaypee University of Information Technology
Himachal Pradesh, India
hemraj1977@yahoo.co.in

2021 6th International Conference on Signal Processing, Computing and Control (ISPCC) | 978-1-6654-2554-4/21/$31.00 ©2021 IEEE | DOI: 10.1109/ISPCC53510.2021.9609481

Abstract— **Recently, the increase in the usage of the internet and digital devices has influenced a lot of people and it has resulted in the transmission of data on a very large scale. Data transmissions have lots of issues and problems such as data security, data privacy, and data confidentiality. A lot of security mechanisms are implemented to tackle these issues and to improve computer system security. However, these issues are the biggest problem in the digital world as we cannot say that any device is 100% secure at this moment. An antivirus software is traditionally used over the years to provide security and safety to the digital devices. Nowadays, an Intrusion Detection System that works like an antivirus has been used and helped in detecting and preventing intrusions and attacks. In this survey paper, we presented a survey on the Intrusion Detection System (IDS) with its introduction, its types, and related work.**

Keywords— Intrusion Detection System, Anomaly based Intrusion Detection System, Signature based Intrusion Detection System, Host IDS, Network IDS

I. INTRODUCTION

The inclusion of the internet in our daily life made us an entity of the network that can access every resource of the network. A network is a set or group of digital devices connected to share every resource over a wired or wireless connection. An attack in the network can be initiated at any time by any intruder. An inside attack is done by someone who is known and who tries to access the resources for which he/she is not authorized to access. An outside attack is done by someone who is not known and who tries to access the system using the wrong ways and credentials.

There are also hardware attacks and software attacks. Hardware attacks are easy to detect whereas software attacks are difficult to detect and can damage the data or information present inside the system.

The ability of monitoring, preventing, and reacting to a computer or network misuse can be defined as Intrusion Detection. An Intrusion Detection System (IDS) is a system (digital device or software) which includes the monitoring of the traffic flowing in the network or a system for any policy breach, issues, and malicious activity. If any kind of harmful activity is discovered, it is reported to the administration [18].

The attempted attacks result in security issues that breach the Confidentiality, Integrity, and Availability (CIA) triad of Computer Security. The National Institute of Standards and Technology (NIST) defines Intrusion as an attempt to compromise the CIA triad or to breach the security mechanisms of a computer network [21].

Regardless of the mechanism it is based on, an Ideal Intrusion Detection System should tackle the following issues [19] [20]:

- The system must run and work continuously without human supervision. It must be reliable enough to allow it to run in the background.
- It should not be a black box. Its internal work should be visible from the outside.
- It must have fault tolerance ability so that it does not have to rebuild its database whenever there is a system crash.
- It must resist subversion. The system should be able to monitor itself to ensure that it has not been overthrown.
- It must impose minimum traffic and overhead on the system. It should not slow down a system.
- It must be able to observe harmful behaviour from normal behaviour.
- It must be easily adaptable to a system as each system has different functionalities and patterns.
- It must be able to deal with the changing behaviour of the system whenever a new application or software is being installed in the system.
- It must be difficult to deceive.

Since a basic IDS generates a lot of traffic and congestion in its database, the solution is to generate alerts on interesting events only. An effective and efficient IDS has a low rate of false positives and false negatives. The idea of Ideal IDS maximizes the percentage of chances of true positives and true negatives [17] [20].

A. Performance Metrices

The performance of an ID will be defined by the confusion matrix shown in Table I. It provides the classification between the original values and predicted values computed by the defined model. There are four parameters of a confusion matrix True Positive, False Positive, True Negative and False Negative. We have defined the confusion matrix according to the information predicted by the IDS [25]:

- True Positive (TP) is a parameter which shows that the model accurately predicts the positive class (alerts when there exists malicious traffic).
- False Positive (FP) is a parameter which shows that the model inaccurately predicts the positive class (alerts when there exists benign traffic).

978-1-6654-2554-4/21/$31.00 ©2021 IEEE

- True Negative (TN) is a parameter which shows that the model accurately predicts the negative class (does not alerts when there exists benign traffic).
- False Negative (FN) is a parameter which shows that the model inaccurately predicts the negative class (does not alerts when there exists malicious traffic).

TABLE I. CONFUSION MATRIX

	Positive	**Negative**
True	Alerts when there is malicious network traffic	Do not alert when there is benign network traffic
False	Alerts when there is benign network traffic	Do not alert when there is malicious network traffic

Although, the performance measurement is also done by different other metrices than confusion matrix such as accuracy, precision, recall, F-score, Receiver Operating Characteristics Curve (ROC), Area Under Curve (AUC) and many more, but most of the researchers have not used all the metrices required to show the certainty of the Intrusion Detection System.

II. Classification of Intrusion Detection System

An Intrusion Detection System (IDS) monitors all the activity and data flowing in the network. It tries to detect the patterns of malicious activities and information and tried to prevent it from doing any harm to the system.

Fig. 1. shows different types of IDS that are implemented and categorized as per different criteria. There are broadly two different types of IDS which are based on Location and Data. Location based IDS protects the whole site on which it is deployed, whether it is a network or a host, from attacks or intrusions. Data driven IDS protects the network or system from the known as well as unknown malicious data being circulated. It analyzes the data through the behavior of the network/system or signatures of the attacks.

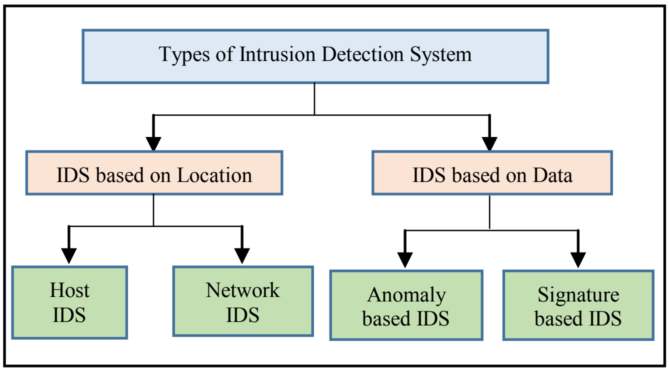

Fig. 1. Types of Intrusion Detection Systems

A. Host Intrusion Detection System

A Host-based IDS is an IDS that monitors and analyses the internal working of a system as well as the network traffic on its network interfaces. It is installed on a single system, and it keeps track of any malicious packet in the operating system files which can cause damage to the system. To monitor a particular host or its components which are not accessible by other systems, the HIDS controls the authorized access of that host. To detect an attack using HIDS, it should be installed on a host. Without installing HIDS, it cannot detect any attack because HIDS has restricted control of the whole network. For taking precautionary and preventive measures in mainframe systems where the outside interaction is rare or less, HIDS were introduced and implemented for the first time [23].

B. Network Intrusion Detection System

A Network-based IDS is an IDS that detects malicious and harmful packets in a network. To analyze and monitor all the traffic (unicast and multicast), NIDS requires all the network traffic access including the traffic which was not intended for it. It should easily tackle and withstand a large amount of network traffic. When any malicious threat is detected, either the administrators are notified, or the IP address should be blocked from accessing the internet. It can be placed at 3 points: it can be connected inline; it can be connected using a network tap or it can be connected directly to a switch spanning port [23].

C. Anomaly-based Intrusion Detection

An Anomaly-based IDS is an IDS that monitors both the computer and network activities for any malicious attacks and classifies that attack as either malicious or benign. The classification of attacks is done by finding patterns that match the behavior of expected malicious actions. ABIDS is also capable of detecting "zero-day" attacks (an attack that happens on the day when a loophole is discovered). The customization of ABIDS can be done according to the systems, applications, or networks. The customization makes it difficult for the attacker to carry out malicious activities without being detected by the ABIDS. The major disadvantage of ABIDS is that it produces false positives at a higher rate [16].

D. Signature-based Intrusion Detection System

A Signature-based IDS is an IDS that detects the attacks by analyzing specific patterns that match with expected and/or known malicious activities used by attackers in a network. SBIDS used a database that contains attack signatures and if any pattern matches is found while analyzing in the database, then that attack is identified as malicious. The SBIDS is originated from antivirus software that works on the same technique of signature detection. The drawback of SBIDS is that the database is not updated regularly with new attack signatures due to which any new and unknown attack can take place. The advantage of SBIDS is that the database processing is so simple and effective, and it produces false positives at a lower rate [16].

III. Data Set

It shows a collection of raw security data which consists of different features and attributes that can help to build effective IDS to achieve our goals. Mostly researchers use Knowledge Discovery in Databases (KDD), Network Security Laboratory- Knowledge Discovery in Databases (NSL-KDD) data sets that contains features related to intrusions or attacks. Selection of a data set generally represent that how our objectives will be defined and achieved. Below, we have listed all the popular datasets used for Intrusion Detection in the order of precedence of usage over the years [22] [24]:

- KDD Cup 1999 dataset includes suitable data for the IDS evaluation which has been used thoroughly over the years. It has approximately five million unstructured and raw data, in which 80% is accounted for attack data.
- NSL-KDD dataset is the improved and advanced version of KDD dataset. It includes the raw data of KDD Cup 1999 but without the redundant data which helps in improving the machine learning algorithms.
- UNSW-NB15 dataset includes the genuine and authentic versions of the different attacks that appear usually nowadays.
- CICIDS2017 dataset includes a simulation of real-world data. The data was captured in the period of 5 days starting from Monday (3rd July 2017 – 9 A.M.) to Friday (7th July 2017 – 5 P.M.).

IV. LITERATURE REVIEW

We have surveyed about types of Intrusion Detection System that works in various fields using different datasets, techniques and metrices. Table II shows the year and number of papers studied and surveyed for the review.

TABLE II. YEAR AND NUMBER OF PAPERS REVIEWED ON IDSs

Type of IDS	Year (No. of paper)
Host IDS	2000(1), 2001 (1), 2005(1) 2008 (1), 2010 (1), 2012(2), 2013 (1), 2014 (1), 2015 (2), 2016(2), 2017 (1), 2018 (1), 2019 (1), 2020(3)
Network IDS	2000(1), 2001 (1), 2005(1) 2008 (2), 2009 (1), 2010 (1), 2012(2), 2013 (1), 2014 (2), 2015 (2), 2016(2), 2017 (1), 2019 (1), 2020(3)
Anomaly based IDS	2000(1), 2001 (1), 2005(1) 2008 (1), 2009 (1), 2010 (1), 2011(1), 2012(2), 2013 (1), 2015 (3), 2016(2), 2017 (1), 2018 (2), 2019 (1), 2020(3), 2021(2)
Signature Based IDS	2000(1), 2001 (1), 2005(1) 2008 (1), 2010 (1), 2011(1), 2012(2), 2013 (1), 2015 (3), 2016(2), 2017 (1), 2019 (1), 2020(3)

Aggarwal, Palvi, and Varun Dutt [1]., objectives were to find out that how attackers and defenders make decisions under the influence of available IDS and interdependence information about the opponent's action in a cybersecurity game. By using behavioral cybersecurity, they have developed research questions that show the available IDS and interdependence information using lab experiments. They used computational cognitive modeling and Instance-based learning Theory (IBLT) for the reason of decision taken by attackers and defenders. When IDS were absent, the attacking actions got reduced in Info Condition, but not in No Info Condition. The communication between IDS availability and Information availability was not enough to defend actions.

Alaba, Adebola, Stephen Maitanmi, and Oluwabukola Ajayi [2], aimed to attain higher detection rates of any attacks, to decrease the false positives, and to maintain computing efficiency while targeting the above objectives. They have used the Stacking Ensemble technique in which the stacks contain C4.5, Random Forest, Naïve Bayes classification models, and Support Vector Machines. The models used for base classification were: C4.5, Random Forest, and Naïve Bayes Classification. The model used for meta classification was Support vector Machines. They have found that the proposed ensemble with the proposed algorithm classified 99.5% instances correctly and 0.6% incorrectly. The proposed ensemble's performance metrices were calculated as: Accuracy: 99.5%, Detection rate: 99.5%, False Positive rate: 0.6% and Precision: 99.5%.

Aljawarneh, Shadi, Monther Aldwairi, and Muneer Bani Yassein [3]., purpose was to calculate the intrusion scope threshold degree based on data's important training features, to minimize the time complexity while determining the feature association impact scale while using a new hybrid classification approach, to improve the accuracy, and false positive and false negative rate. They have used a hybrid approach of classification for Intrusion Detection model by using Decision trees, Neural networks, and Nearest neighbor. For feature selection, they have used the NSL-KDD dataset and WEKA automated data mining tool with 20% testing and 80% training data. The 80% training data were classified using J48, Random Tree, Naïve Bayes, and proposed hybrid model in which the proposed model gives out 99.81% by classifying the instances. The proposed model also has an improved false positive rate at 0.003% and a true positive rate at 0.997%.

Alqahtani, Hamed [4]., tried to analyze and study various machine learning algorithms used for intrusion detection systems, to test the effectiveness by doing several experiments on the dataset which have categories of attacks, and to calculate the effectiveness by various performance matrices like recall, F1-score, precision, and accuracy for all the machine learning algorithms. They have used data-driven IDS modeling in which data exploration, data processing, and machine learning-based security modeling have been processed. They have used KDD'99 cup data which includes 48,98,431 instances. They have used 10-fold cross-validation on the data set for testing and finding out the average accuracy. They have found that for detecting intrusions Random Forest classifier-based IDS model was the best choice that performed better than the others. Out of all the classification techniques, Random Forest perform better in terms of Precision > 98%, Recall > 92%, and F1Score > 96%.

Cemerlic, Alma, Li Yang, and Joseph M. Kizza [5]., objectives were to implement a Bayesian approach-based IDS which can detect novel attacks, to calculate the effectiveness and feasibility of the IDS using DARPA dataset and real-world traffic, and to improve the detection of the IDS by using the feedback method. They have proposed an IDS using the Bayesian approach which has 6 modules for detection: Data gathering and parsing, Bayesian Network inference, Knowledge base, System Configuration, Response, Bayesian Network Learning. They implemented Bayesian model, used the K2 learning algorithm and training data as their base of knowledge through which it can build a Directed Acyclic Graphs (DAG) in which nodes are connected to increase the probability of the network. They found that the real-world network traffic can detect a never-before-seen attack by using a tcpdump format to group the packets into connections and then selecting a connection-specific feature which was when fed to the BN, it can easily show which connection is normal or malicious. Results according to DARPA dataset: true negative: 93.89%, true positive: 97.88%, error rate: 2.881%.

Singh, Amrit Pal, and Manik Deep Singh [6]., intended to analyze and compare Host-based IDS and Network-based IDS and to calculate the efficiency of HIDS and NIDS using various tools. They have used OSSEC (HIDS) which provides monitoring of logs and generates alerts based on log alerts and other functionality like integrity checks and many more. They have used SNORT (NIDS) which provides sniffing and storage of log packets that entered our network and generate an alert for an attack. OSSEC was installed on 1 server which is based on Linux/Unix environment. Also, OSSEC clients were the computers that are connected through LAN. SNORT was installed between the router and a firewall through which all the traffic monitoring is possible. Also, it does not require to be installed on each host or client. They have found that OSSEC is using system logs to send alerts to more than one server, storing the output in a database, and sending the output to a hybrid IDS. They have found that SNORT can be placed anywhere in the network through which all data can that has been transferred between the computes can be monitored.

Singh, Dinesh [7]., targeted to propose a Novel Collaborative IDS (CIDS) for the cloud, to detect the known attacks using SNORT, to detect the unknown attacks using an Anomaly Detection System (ADS), and to reduce the impact of the attack and gain higher accuracy and performance. For detecting network attacks in the cloud, they have used the NIDS module. Correlation Unit (CU) is included in each cluster of cloud to change the low-level attacks into high-level confirmed attacks. NIDS is used to detect attacks inside a cluster and the CU helps with the correlation between the two. The placement of CU is done by bully election algorithm. They have used SNORT and an ADS with a Decision tree and SVM classifiers through which they will detect unknown attacks and whether the attacks were malicious or not. They have evaluated their model on different datasets of KDD99, NSL-KDD, ITOC and captured the details. They have evaluated the model on different metrices with different datasets: Learning time: 46.616 seconds, accuracy: 98.92%, intrusion detected: 99.4% intrusion missed: 0.60%, true alarms: 98.31% and false alarms: 1.69%.

Somwanshi, Akshay A., and S. A. Joshi [8]., proposed to protect the server data by implementing or including fake servers in the network like the Honeypot tool, to capture the information about the attacks through honeypot, and to block the attackers and take actions against them. They have used Honeypot for server security as honeypot provides fake services same as an actual server service, so whenever an attack happens and it accesses the honeypot services, the attacker will think that it has access to a real server and its services. They have made a system based on the implementation of honeypot in a network to find an illegal attack. The proposed system is working on extracting the honeypot through which an attack can be detected by gaining the source and target address for which an attack was initialized. After having the addresses, it will hold up the load and implement and change the honeypot according to the attack. It has been shown that a honeypot can be easily deployed within a server. The results showed that the honeypot efficiently collects the information and behavior about the attack and the attacker so that proper defenses can be taken. Also, honeypot uses a load balancer which reduces the load on the actual server and can process multiple requests in a short period while increasing the speed simultaneously.

Muthukumar, Balasundaram, and Praveen Kumar Rajendran [9]., aimed to make an Intelligent IDS to detect any private cloud attack based on the previous records of intrusions and by constantly updating an intrusion detection database, to predict the private cloud attacks by providing training to the system, and to implement the Intelligent IDS on a private cloud using hardware and software. Their research is divided into 3 phases: Training, Testing, and implementing, and updating the IDS. While implementing the IDS, they have proposed some algorithms under the name "Muthu-Praveen Algorithm of Intelligent Intrusion". All the algorithms have different functionality based upon the phase implementation. In the training phase, they have trained the hardware and the application on a sample database of attacks and if any matching is found, the attack is detected based on the given training. In the testing phase, they tested out that their training was successful or not. In the implementing and updating phase, they used real-time data in the system and if the system easily traces the intrusion, then it will alert the cloud admin and users. Five pages were developed, and each page has been tested out 25 times: Index page, Admin Login page, user login page, user home page, and Admin homepage. The average response time for all the pages was 6671 ms and the error percentage was 0%. The 0 error percentage shows that the application has cleared all the test cases. The average response time was less than 7 seconds (6.67s) which shows that at a particular time 25 users can use the application under 7 seconds which makes the performance better. Response time of 4 pages was less than 2.6 seconds which shows that 90% of the performance was good.

Seeber, Sebastian, and Gabi Dreo Rodosek [10]., objectives were to improve the monitoring of network and traffic steering using Software Defined Networking (SDN), to detect network attacks by taking data from network components with OpenFlow, to enable redirection of traffic to various IDSs including cloud-based IDS, to eliminate static function deployment of security in a network, and to investigate some Research Question (RQ) based on OpenFlow implementation. By using OpenFlow inside an SDN environment, they have proposed an architecture that can independently implement dynamic security features. Their approach focuses on amalgamating the dynamic function of security for a quick and immediate response towards an attack or detection. In the architecture, the SDN-

C was deciding independently about the decisions without considering the previous decisions, but after a connection establishment, the detection capabilities will be enhanced, and SDN-Cs can be designed as a distributed controller as well. Their approach was leveraging the SDN and distributed controller to implement a service chaining function that establishes a verified chain of security features. Using their approach, narrow-focused deployment and enhancement can be seen in various use cases such as SAAS, PAAS, and CIA. Their proposed solution shows the detection of traffic using multiple IDSs with cloud based IDSs and can be used on DDoS attacks as well.

Sharma, Aravendra Kumar, Sushil Kumar Saroj, and Prashant Kumar [11]., tried to study the Distributed Intrusion Detection System (DIDS) for Wireless Sensor Networks (WSN), to research about the security issues, need for IDS and Distributed IDS in a WSN, and to implement an architecture of DIDS which mostly eliminates all the security issues of WSN. Their proposed architecture for DIDS was based on an agent-based, intelligent, and distributed system as DIDS can be set inside the intelligent agents so that they can be positioned on a network. Their architecture includes 7 components: Local Packet Monitoring Module, NbPrimeter Module, Key Management Module, Local Detection Engine, Alert Region Module, Voting Module, and Local Response Module. They have found out that WSNs are implemented for physical interaction with the outside world, making it vulnerable to different types of attacks, but an IDS inside a WSN will add additional defensive measures to be taken at the time of any attack with already deployed defensive techniques. They have seen that the need of an IDS inside a WSN leads to a more secure network and some features of IDS in WSN should include auditing data, minimizing the use of resources, not trusting any node in WSN, should be present in the entire network (distributed), and resist the entry of any attack in the nodes.

Sedjelmaci, Hichem, and Mohamed Feham [12]., worked to propose a new hybrid Intrusion Detection System (IDS) by using anomaly detection with Support Vector Machines (SVM) and Misuse Detection for Clustered Wireless Sensor Network (WSN), to be able to detect routing attacks with low false positive rates, and to reduce workload and increase the energy consumption with the proposed IDS for the prolongation of the network. A distributed learning algorithm is used by anomaly detection to train SVM so that it will distinguish between 2 types of attack- normal and anomalous. Since they are using clustered WSN, they have used hierarchical topology to distribute the network into clusters so that each cluster should have a cluster head. They have used minimum number of IDS nodes to save the energy of the network. In the training phase, IDS nodes get the data from neighboring nodes through the communication links which are secured so that the trained data can be easily embed into the Hybrid IDS Module (HIDM) which turns out to be a lightweight and precise detection system. The IDS framework for ISN comprises of different modules to analyze the packets. It is found out that by using SVM we should have a distributed binary classifier for anomaly detection which will analyze the packets. For that Hyperplane with maximum margin was calculated. Also, to find the optimal hyperplane, they have solved the convex optimization problem by using Langrange multiplier and for calculating support vectors, they have used Kuhn-Tucker (KKT) conditions. They have found out that SVM classifier with anomaly detection was giving decent results with taking less time to train compared to neural networks and it is precise for new signatures attacks. SVM generates low generalization errors.

Babatope, Longe Olumide, Lawal Babatunde, and Ibitola Ayobami [13]., intended to study and examine about the placement of sensors in Network based Intrusion Detection System (NIDS), to study about the deployment of sensors in NIDS, to propose an extent model to deploy sensors in NIDS, and to catch the malicious packets that have been through firewall. For considering sensor deployments, they have considered some constraints: Analyze network topology, Critical components(servers), Infrastructure components (switches and routers), Security components (Firewalls and IDS components), Deployment within functional boundaries. The installation configuration is done to choose the location where an IDS sensor can be installed. The common ones are: Standalone sensor, Device management, Firewall sandwich, Remote sensor, and Standalone Sensor Configuration. After considering all the options, they have studied about Sensor placement in NIDS to detect intrusions which includes Network based IDPS, Appliance based sensor, Software only sensors, Network architectures and sensor locations, Spanning port, Network tap, and IDS load balancer. While deploying or placing a sensor, problem is encountered that can be defined as some of the nodes are not engaging with the network specification. A usual problem has been seen when the node dies due to more energy consumption like discharge of batteries, shirt circuits and many more.

Ehret, Christoph, and Ulrich Ultes-Nitsche [14]., worked to propose a hybrid IDS which amalgamated host and network-based components while focusing on HIDS with Artificial Immune System, to study about Human Immune System (HIS) and Artificial Immune System (AIS), and to introduce new and untouched AIS concepts that can help to upgrade the effective working of an IDS. Artificial Immune System (AIS) is a flexible system derived from observation of theory-based immunology and its functions that can be used in solving a problem. The need to implement an AIS was to interact with the world with different types of nodes and to communicate between them through the communication links or paths. The innate system used in HIS is used to keep viruses, bacteria and many more out of our body which can be compared with misuse detection of the IDS. Both HIS and IDS use database of signature to match the patterns and detect the virus or malicious activity respectively. AIS is based on the functionality of HIS. As it was seen, AIS can be applied for security of computers for detecting viruses, can be used as antivirus or as an ID component. Biological functions and rules can be seen benefitting the IDS. The immune system base inspired IDS considers many features while detecting anything anomalous but main features can be communication between misuse and anomaly detection, distribution, bypassing the failures happening at single points, positioning, affecting the single processes only.

Chen, Hao [15]., objectives were to propose an experiment with Genetic Algorithm (GA) and Multi-Objective Optimization (MOO) for sensor placements, to show that the experiment is an evidence of the proposed concept and will show the validation of the approach, and to show the first use of heuristic-based optimization techniques for IDS sensor placement in the network optimally. Using GA, which is a heuristics-based optimization technique, through which their approach can easily place IDS sensor optimally with the help of natural selection. GA can only be used with MOO which in turn provides more than one goal through which a problem can be seen as a set of goals, and it returns set of results in a single execution and each result shows a balance between multiple goals. Their approach is the first one to use heuristics-based optimization techniques to successfully place an IDS sensor. In their first experiment, they have checked the connection between numbers of sensors and detection parameter which contained detection rate and false alarm rate. It was seen in their implementation that with placement of 7 and 6 sensors in the nodes have 94.15% detection rate and 42.75% and 44.49% false alarm rate, respectively. In their second experiment, they have tried to find the minimum optimal cost or budget to detect attack. They found out that detection rate will increase if budget increases. In their third experiment, they have tried to use MOO for budget optimization. They program their approach in such a way that the cost comes out to be optimal (between 16 to 22 means -20% to +10% of original budget).

Gyanchandani, Manasi, J. L. Rana, and R. N. Yadav [16]., tried to study various techniques of Anomaly-based Intrusion Detection System (ABIDS), to define properties with advantages and disadvantage of each technique, and to study different example which have been implemented in the past as well was in the current projects. Signature based IDS is used to detect the signatures of already known attacks by matching them with the signatures that are stored in a database and if matching is found, then it is shown as an intrusion attack. Anomaly based IDS is used to detect unknown attacks and zero-day attacks based on the behavior of the network and if any deviation from the normal behavior is found, then it is an intrusion attack. They have studied and explained each category and subcategory of Anomaly-based IDS(ABIDS). Each approach has their advantages and disadvantages. Also, they have studied and mentioned the advantages and disadvantages of each category and subcategory.

Yeo, Liu Hua, Xiangdong Che, and Shalini Lakkaraju [17]., objectives were to study about basics of IDSs and their classification, to study about various algorithms that are used to detect malicious activities, to compare the different methods for detection of intrusion activities and study the various measures and necessity of IDS in Security. Intrusion Detection can be defined as to detect the illegal access and use of a computer system. IDS carry out tasks that can help in identifying certain attacks in the network and to be defensive in nature against these attacks. They have defined some issues that should be solved by an Ideal IDS like it should work without any human intervention, must be able to differentiate between normal and malicious behaviors, should have low false positive and negatives rates and many more. They have also shown the classification of IDS based on the network in which they can be deployed: HIDS, NIDS, Signature based IDS and Anomaly based IDS. They have also stated about an Intrusion Prevention System (IPS) which usually works when there are distrust attacks which needs response and control, so basically it tries to prevent the attacks from attacking the network and is mostly made up of firewalls or routers and IDS. Current research in information security shows that at 45.6% of the users are targeted at least one time. It has been seen that prior personnel and outsiders can be the biggest threat to an organization. To detect the attacks in prior stage, IDs can collect and correlate data from the attackers and try to store possible evidence for an attack.

V. RECOMMENDATION BASED ON LITERATURE SURVEY

Security is very important and crucial for any organization or an individual. Research on Intrusion Detection Systems has been done over the past years to enhance their flaws and challenges. This paper gives an overview of the study done on IDS. Based on our study, we personally recommend measures that can be applied to various IDSs to, not completely but partially, overcome their research issues. HIDS works on host so it is needed that in a certain type of network, each host must have an HIDS installed to protect itself from the attacks. The NIDS should be able to withstand and analyze the large number of packets which requires more efficient resources and service than expected. Installation of Anomaly based IDS should be fast enough to setup in a new environment and the detection of different types of attacks should be done using a multi-class classifier using AI or ML techniques. To eliminate multiple signatures of same type of an attack, the Signature based IDS can be used with specific type of system, data set or multiple IDS engines. Table III summarizes the recommendation measures for each type of IDS reviewed.

TABLE III. PREVENTIVE MEASURES FOR IDSs

Type of IDS	Recommendations
HIDS	• Each host must have an HIDS installed with signatures of the attacks to detect the intrusions. • Each host should have enough and basic resources such as greater bandwidth, less latency, more battery power, and many more to work to prevent resource hogging.
NIDS	• It can work with HIDS to prevent resource fluctuations in the network. • The placement of NIDS can be considered on various points, such as at the perimeter of the network, on point where the network is connected to any extranets or intranets, or on the point where any remote systems are connected to the network. But the main point considered can be the connection to the internet as it is easy to stop the attacks at the main access point of the internet.
Anomaly based IDS	• Installation of Anomaly based IDS should be fast which

	can be achieved by using it with Signature based IDS. • To detect each type of novel attacks, a multi-class classifier or a multiple-one-vs-all binary classifier can be used with AI or ML techniques.
Signature based IDS	• To continually update the signatures of attacks without creating multiple signatures of a specific attack, a language based script, an AI based detection or multiple IDS engines with multi processors can be used which matches the signature with the predefined signatures. • It is feasible to use it with specific systems or domains like DNS, SMTP, and many more through which the database of the signatures will contain specific type of signatures of the attacks.

VI. Conclusion and future work

Intrusion Detection System is a software application that manages and monitors a network or system for any intrusion activity or violation attack. Each attack or malicious activity is reported to the administrator and necessary steps will be taken against it. The security of digital devices is at bar as not a single device is free from attacks or intrusions whether from inside or outside. We have explained the existing literature on the IDS and their contribution in securing the systems or networks from attacks which helped us to gain the research gaps. Some of the measures explained have not been implemented so far and yields objectives that can be carried out in future work. We would like to work on the problems explained in literature review. Future work includes working on Anomaly-based IDS to improve its performance and accuracy through which novel attacks can be detected.

References

[1] Aggarwal, P. and Dutt, V., 2020. The role of information about opponent's actions and intrusion-detection alerts on cyber decisions in cyber security games. Cyber Security: A Peer-Reviewed Journal, 3(4), pp.363-378.

[2] Alaba, A., Maitanmi, S. and Ajayi, O., 2019. An Ensemble of classification techniques for Intrusion Detection Systems. International Journal of Computer Science and Information Security (IJCSIS), 17(11).

[3] Aljawarneh, S., Aldwairi, M. and Yassein, M.B., 2018. Anomaly-based intrusion detection system through feature selection analysis and building hybrid efficient model. Journal of Computational Science, 25, pp.152-160.

[4] Alqahtani, H., Sarker, I.H., Kalim, A., Hossain, S.M.M., Ikhlaq, S. and Hossain, S., 2020, March. Cyber Intrusion Detection Using Machine Learning Classification Techniques. In International Conference on Computing Science, Communication and Security (pp. 121-131). Springer, Singapore.

[5] Cemerlic, A., Yang, L. and Kizza, J.M., 2008, July. Network Intrusion Detection Based on Bayesian Networks. In SEKE (pp. 791-794).

[6] Singh, A.P. and Singh, M.D., 2014. Analysis of Host-Based and Network-Based Intrusion Detection System. International Journal of Computer Network & Information Security, 6(8).

[7] Singh, D., Patel, D., Borisaniya, B. and Modi, C., 2016. Collaborative ids framework for cloud. International Journal of Network Security, 18(4), pp.699-709.

[8] Somwanshi, A.A. and Joshi, S.A., 2016. Implementation of honeypots for server security. International Research Journal of Engineering and Technology (IRJET), 3(03), pp.285-288.

[9] Muthukumar, B. and Rajendran, P.K., 2015, August. Intelligent intrusion detection system for private cloud environment. In International Symposium on Security in Computing and Communication (pp. 54-65). Springer, Cham.

[10] Seeber, S. and Rodosek, G.D., 2015, June. Towards an adaptive and effective IDS using OpenFlow. In IFIP International Conference on Autonomous Infrastructure, Management and Security (pp. 134-139). Springer, Cham.

[11] Sharma, A.K., Saroj, S.K. and Kumar, P., 2013. Distributed intrusion detection system for wireless sensor networks. IOSR Journal of Computer Engineering, 14(1), pp.61-70.

[12] Sedjelmaci, H. and Feham, M., 2011. Novel hybrid intrusion detection system for clustered wireless sensor network. arXiv preprint arXiv:1108.2656.

[13] Babatope, L.O., Babatunde, L. and Ayobami, I., 2014. Strategic sensor placement for intrusion detection in network-based IDS. International Journal of Intelligent Systems and Applications, 6(2), p.61.

[14] Ehret, C. and Ultes-Nitsche, U., 2008. Immune system based intrusion detection system. In Innovative Minds (Information Systems Security Association-ISSA 2008), Johannesburg, South Africa, July 2008.

[15] Chen, H., Clark, J.A., Shaikh, S.A., Chivers, H. and Nobles, P., 2010, February. Optimising IDS sensor placement. In 2010 International Conference on Availability, Reliability and Security (pp. 315-320). IEEE.

[16] Gyanchandani, M., Rana, J.L. and Yadav, R.N., 2012. Taxonomy of anomaly based intrusion detection system: a review. International Journal of Scientific and Research Publications, 2(12), pp.1-13.

[17] Yeo, L.H., Che, X. and Lakkaraju, S., 2017. Understanding Modern Intrusion Detection Systems: A Survey. arXiv preprint arXiv:1708.07174.

[18] Wanda, P., 2020. A survey of intrusion detection system. International Journal of Informatics and Computation, 1(1), pp.1-10.

[19] Vinchurkar, D.P. and Reshamwala, A., 2012. A review of intrusion detection system using neural network and machine learning. J. Eng. Sci. Innov. Technol, 1, pp.54-63.

[20] Bace, R.G. and Mell, P., 2001. Intrusion detection systems.

[21] Lazarevic, A., Kumar, V. and Srivastava, J., 2005. Intrusion detection: A survey. In Managing Cyber Threats (pp. 19-78). Springer, Boston, MA.

[22] Maseer, Z.K., Yusof, R., Bahaman, N., Mostafa, S.A. and Foozy, C.F.M., 2021. Benchmarking of machine learning for anomaly based intrusion detection systems in the CICIDS2017 dataset. IEEE Access, 9, pp.22351-22370.

[23] Gupta, M., 2015. Hybrid intrusion detection system: Technology and development. International Journal of Computer Applications, 115(9), pp.5-8.

[24] https://www.unb.ca/cic/datasets/ids-2017.html

[25] https://towardsdatascience.com/metrics-to-evaluate-your-machine-learning-algorithm-f10ba6e38234