Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 9, November 2021: 8231-8254

Review of Intelligent Techniques for Intrusion Detection Systems in Network Environments

Parul[1]
Research scholar
Baba Mastnath University, Rohtak
dparul79@gmail.com
Deepak Kumar[2]
Assistant Professor
Department of Computer Science,
Government college, Panchkula
dks.karnal@gmail.com
Dr . Pooja Vyas[3]
Assistant Professor
Department of management
Indira Gandhi University, Meerpur, Rewari, Haryana
pooja.vyas9@gmail.com

Abstract – Networks performs a key role in current existence; network security has become a vital research place. An intrusion detection device (IDS)which is vibrant cyber security technique, monitors the state of software and hardware running in the network. Regardless of a long-time development, existing IDSs still face challenges in enhancing the detection accuracy, lowering the false alarm rate, and detecting unknown assaults. The Network Intrusion Detection System (NIDS) plays a crucial role in preserving information protection and especially classifying various attacks on contemporary networks. In the current situation, the option of an appropriate combination of anomaly detection features is more important in the NIDS. Swarm Intelligence (SI) techniques commonly used to select the features in the high dimensional dataset to improve the accuracy. Machine learning (ML) techniques exhibited high ability to develop the intrusion detection algorithm in the network field. Deep Learning techniques has been widely used to improve the performance on a NIDS to detect various network attacks. The implementation of an intelligent algorithm to resolve a wide range of NIDS issues is investigating namely the Swarm intelligence algorithm, machine learning algorithm, deep learning algorithm based on exploratory analysis to identify the benefit of using intrusion detection enhancement techniques.
Keywords*:* Intrusion Detection System (IDS), Network Intrusion Detection System (NIDS), Swarm Intelligence (SI), Machine Learning (ML), Deep Learning (DL).

1. Introduction

Over the last decade, information technology has developed rapidly, and security has become one of the main concerns of almost every sector. Cyber-security focuses on key areas such as application security, catastrophe security, information protection, and network security. Recently, numerous IDSs have been suggesting focusing mainly on rule-based systems, because their performance depends on the rules identified by the safety experts [1]. However,

the volume of network traffic is large and therefore, the process of encoding rules is both insufficient and slow.

Today, one big obstacle for intrusion detection is the collection of features from network traffic data. The Network Intrusion Detection System (NIDS) is a software-based program or hardware tool used to detect malicious activity on the network. Based on the detection techniques, intrusion detection is classifying as anomaly-based and signature-based [20000]. The NIDS developers use a wide range of a technique to detect intrusion. Information and Communication Technology (ICT) and networks accommodate a variety of complex user data that are vulnerable to multiple attacks by internal and external intruders. These attacks can be physical or caused by a computer. Malicious cyber-attacks pose significant security issues requiring a versatile and more robust intrusion detection system(IDS).

The IDS is a helpful intrusion detection system tool used to identify and recognize intrusion attempts or violations of security polices automatically at network-level and host-level organization. Based on intrusive behavior, intrusion detection is classifying as network-based intrusion detection(NIDS) and host-based intrusion detection (HIDS). The network-based IDS detects intrusion as traffic tracking by network equipment such as routers, switches, and network taps. Develop an effective NIDS model is one of the major research challenges. In Figure [1] shows the different types of IDS types.

NIDS may be implementing using two types of detection techniques. One of the types is Signature-based detection and the other one is Anomaly-based detection. Signature-based detection is a knowledge-based intrusion detection system that matches patterns in the intruder's network traffic to detect all known threats. Anomaly-based detection is a behavioral-based intrusion detection system that detects variations in the system's regular patterns by the structure of the system that observing. Effectively identify known threats with minimum number of alarm [3]. Frequently database upgradation needed. Novel or undefined (zero-day) assaults cannot be identifying by misuse of techniques. It's getting high false-positive rates. The HIDS examine the host system for activities. Host-based IDS might also observe the OS, system calls, error message, and audit logs on the host system.

Using intelligent algorithms identified the threats and their pattern in the computer system. Developing hybrid cyber-security methods and building computational systems that integrate with intelligent algorithms to analyze big data, mitigate threats, and protect against new invaders.

The Swarm intelligence algorithm is an emerging field of optimization. Optimization techniques applying from a variety of perspectives, such as parameter tuning, maximizing or minimizing an objective function, weight value optimization, feature selection, meeting multiple criteria, search strategy and finding a trade-off solution. The Study of the NIDS swarm intelligence algorithm, which inspires the optimization method in the development of the individual, has received much attention from swarm intelligence research.

Machine learning (ML) is constantly gaining strength in a wide range of applications, such as medical imaging, pattern recognition, signal processing problems, intrusion detection, etc.

Artificial Intelligence involving the creation of a self-learning algorithm to obtain information from data in order to make predictions and data-driven decisions [4].

In the present situation, machine learning methods are most commonly used in intrusion detection system [5]. Generally, ML techniques are built to classify the threats in NIDS, develop to find the threats in NIDS, thus helping managers to avoid inappropriate measures in network interventions.

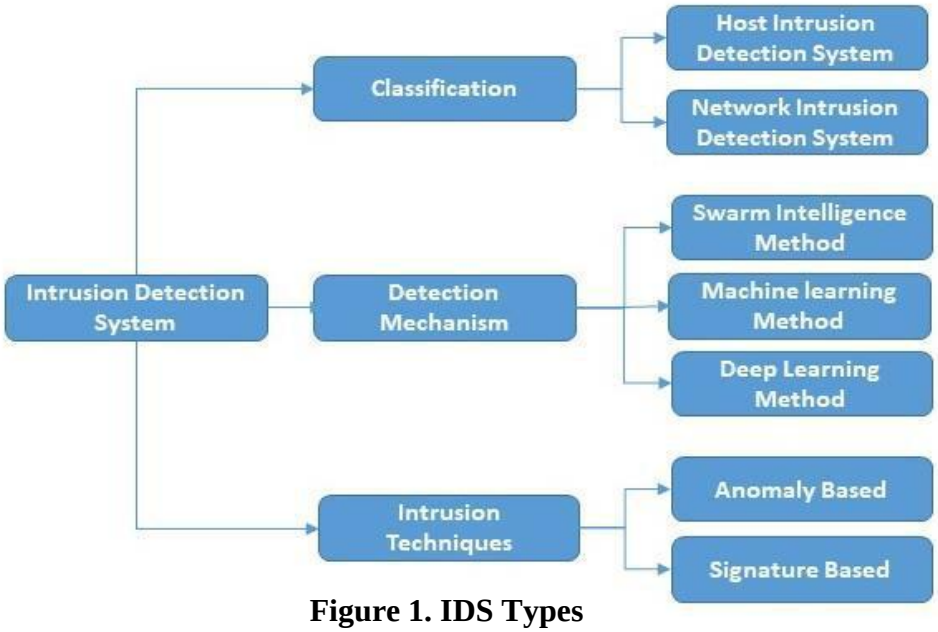

Figure 1. IDS Types

2. Network Intrusion Detection System (NIDS)

Intrusions are a set of associated malicious acts performed by an internal or external intruder that aims to compromise the targeted device. Intrusion detection includes control of computer systems and network traffic and evaluating behaviors to identified potential intrusions affecting the device. For this function, a collection of tools and mechanisms known as the Intrusion Detection System (IDS) [6].

The IDS starts with the collection of data from the events observed. Comprehensive logging of the event-related Data and events correlating from multiple sources. The discovery engine is the cornerstone of the IDS, which uses a wide range of methodologies and related techniques, depending on the situation.

2.1 Intrusion Detection Methodologies

Signature-based detection, anomaly-based detection, and specification-based detection are the most common methods of intrusion detection. They used together, either integrated or separately,to increase the accuracy of detection [7]. In the Table [1] describe the different types of Intrusion Detection.

2.2 Signature-based detection

A signature is a pre-configured type that suits an intrusion that is well-known. Signature-based detection defining in [3] as "the process of comparing signatures to observed events," identifies potential incidents. Signature-based detection referring to as misused detection or Information-based detection due to the use of information obtained from previous intrusions and vulnerability. However, this method is not sufficient to detect unknown intrusions and variants of known intrusions, as their patterns are unfamiliar. Moreover, keeping knowledge up-to-date is another problem, as it is a time-consuming and difficult process. Signature-based identification has a very low positive rate and high classification accuracy.

2.3 Anomaly-based Detection

An anomaly is any deviation from normal behavior. Detection based on the anomaly, also called Behavior-based detection describing as "the process of comparing irregular activities with normal activities events observed to identify significant deviations". Anomaly-based detection consists of general modules: 1) Parameterization: representing the observed behavior in a profile that consists of different attributes and characteristics of what needs to investigating, such as network connections, hosts, and applications. 2) Training After parameterization, normal and abnormal behavior can be differentiating by the creation of a classification model.3) Detection: using a built classification model to detect new traffic anomalies [8].

2.4 Specification-based Detection

Specification-based intrusion detection to detecting threats physically stated program depends on the combination of both misuse detection and anomaly detection. Specification-based detection identifies a breach of predefined rules and it is undesirable to rely on expert knowledge during the definition of the rule.

Table 1. Types of Intrusion Detection

	Signature – based	**Anomaly-based**	**Specification-based**
Procedure	Pattern matching a known attacks	Pattern identifies a unusual activity	Identifies a breach of predefined rules.
Detection Rate (DR)	High	Low	High
False Alarm Rate (FAR)	Low	High	Low
Detection of unknown attacks	Impotent	Capable	Impotent
Limit	The updating of signatures is burdensome	Computing any machine learning process is heavy	Relying on expert knowledge during the definition of rules is undesirable

Different types of attacks are there in the networking system they are 1. DoS (Denial of service) attack in which the hackers make a computing resources too demanding or too occupied to attend reasonable networking requests. 2. R2L (Remote to Local) attack is an assault in which a user sends packets over the internet to a device that he/she does not have access to in order to reveal the vulnerabilities of the device and to manipulate the privileges that the local user will have on the machine.3.U2R (User to Root) attack attainment with the

user account and try to access super user privilege. 4.In order to detect threats, Probe attack hackers monitors a system or devices, which later exploit the system [9]. In the Table [2] describe the various types of attacks.

Table 2: Various types of attacks

Types of Attacks	**Properties**	**Examples**
DoS (Denial of service)	- Contain malicious events that massively transmit unnecessary requests to disrupt computer resources	Back, Land, Neptune, Pod, Smurf, Teardrop, Mail Bomb, Process table, Udpstorm,
R2L(Remote to Local)	- send packetsto a remote device over a network without having an account on that system and gain access to it to damage the operation of the system.	Spy,Phf,Imap,Guess-pwd,Multihop,Httptunnel,snmp-guess, Xlock, Xsnoop, Warezclient, Warezmaseter. ,
U2R(User to Root)	- Can break vulnerabilities to obtain device super user privileges while beginning as a legitimate user.	Load module, Perl, Rootkit, Buffer-overflow, Sqlattack, Xterm, Ps
Probe	- Check the device and network infrastructure for vulnerabilities. - Provide an attacker with vulnerability lists, such as SMBv1 and open ports, to hack victims	Ipsweep, Nmap, Portsweep, Satan, Mscan, Saint

3. KDD-Cup'99/NSL-KDD Dataset

DARPA Intrusion Detection Data Sets [21], under the direction of DARPA and AFRL / SNHS, are collected and released by the Cyber Systems and Development Division (formerly the DARPA Intrusion Detection Evaluation Division) of the MIT Lincoln Laboratory for the assessment of computer network intrusion detection systems. In Table [3], different datasets are compared. In Table [4], dataset features are described [10].

Table 3: Comparison of the standard datasets in the IDS

Dataset & Year	**Features**	**Advantages**	**Limitations**
DARPA & 1998	-	- The first standard to evaluate the IDS. - Consists of a wide range of attacks.	- The models used to generate traffic have been too basic. - Synthesized data is not a representation of background traffic in actual networks.
KDD CUP '99 & 1999	41 Attributes [32 Numerical & 9 Categorical]	- Identify threats - Training Dataset and Testing Dataset are different.	- Identical records are available. - Enhancement needed to adapt new environment
NSL-KDD & 2009	41 Attributes [32 Numerical & 9 Categorical]	- Identical records not available - The Picked records are less than from the total records.	-Not ideal for representing the actual network that exist.
Kyoto & 2009	24 features [14 statistical features & 10 additional features)	-Identical records not available. - In real network it work well.	- Does not having any details on different forms of attacks.
UNSW-NB15 & 2015	49 Features	-It work good in modern network traffic and attacks.	

Table 4: Dataset Features Description

NO	Features_Name	Type	Description	Categories
1.	Duration		The time interval between connections.	Continuous
2.	Protocol_type		Categories of protocols	Symbolic
3.	Service		Terminal connection services	Symbolic

4.	Flag	Basic Features	Fault or Running status	Symbolic
5.	src_bytes		Volume of data transfer from sources to destination	Continuous
6.	dst_byte		Volume of data transfer from destination to source.	Continuous
7.	Land		1-data transfer from same source 0-Otherwise	Symbolic
8.	Wrong fragment		Incorrect packet no	Continuous
9.	urgent		Urgent fragment no	Continuous
10.	hot	Content Features	Current pointers no	Continuous
11.	Num_failed_logins		Unsuccessful login no	Continuous
12.	Logged_in		1-Successful login 0-Unsucessful login	Symbolic
13.	Num_compromised		Compromised condition no	Continuous
14.	Root_shell		1-terminal shell is obtained 0-terminal shell not attain	Symbolic
15.	Su_attempted		1-Sucide attempt is obtained 0-Sucide attempt is not Obtained	Symbolic
34.	dst_host_same_srv_rate	features	Host destination for same service count	Continuous
35.	dst_host_diff_srv		Host destination for different service	Continuou

	_rate		rate	s
36.	dst_host_same_src_port_rate		Host destination for same service	Continuous
37.	dst_host_srv_diff_host_rate		Host destination for different service	Continuous
38.	dst_host_serror_rate		Host destination for host eror	Continuous
39.	dst_host_srv_serror_rate		Host destination for same service	Continuous
40.	dst_host_rerror_rate		Host destination for same service error rate	Continuous
41.	Dst_host_srv_rerror_rate		RST error connection rate	Continuous

4. Review on intelligent Algorithm in NIDS

One of the nature-inspired algorithms is a swarm intelligence algorithm built based on the concept of collective actions of insects such as ants, bees and termites living in colonies. Swarm intelligence is a group behavior and not having any centralized processor [11]. In Figure [2] shows the different categories of Intelligent Algorithm.

4.1 Taxonomy on Swarm Intelligence Algorithm

Swam can be narrated as a bunch of individuals that are present enormous in number and reveals the same behavioral characteristics. Swarm Intelligence algorithms raised on the behavioral models of living organisms like ants, birds, fish and bees. These algorithms were constructed based on the adaptability flexibility and coordinating characteristics of the species. The members of a swarm exhibit a distinct pattern to be followed by each member of the population without any centralized commands for controlling the individuals present in the swarm [12]. The communication between the members can be direct or indirect. These algorithms can be used to address problems that are NP-hard such as traveling salesman, vehicular routing, scheduling optimization and intrusion detection. In the Table [5], describe the Swarm Intelligence algorithm for solving NIDS problem. Feature selection is one of the plays a major role to construct an efficient intrusion detection model.

Ant Colony Optimization (ACO) - is a heuristic search algorithm based on the communication of ants that interact with each other to find an optimal food source.ACO algorithm can be

16	Num_root		Terminal access count	Continuous
17	Num_file_creations		File creation process no	Continuous
18	Num_shells		Shell Prompts number	Continuous
19	Num_access_files		Access Control files number	Continuous
20	Num_outbound_cmds		In FTP session Outbound command number	Continuous
21	Is_hot_login		1 if the login belongs to the hot list; 0 otherwise	Symbolic

22	Is_guest_login		1 if the login is a guest login; 0 otherwise	Continuous
23	Count	Time based Traffic feature s	Number of connection to the same host	Continuous
24	Srv_count		Within connection Number	Continuous
25	Serror_rate		SYN error Connection %	Continuous
26	Srv_serror_rate		Within SYN error Connection %	Continuous
27	Rerror_rate		REJ error Connection %	Continuous
28	Srv_rerror_rate		Within REJ error Connection %	Continuous

29	Same_srv_rate		Same service Connection %	Continuous
30	Diff_srv_rate		Different service connection %	Continuous
31	Srv_diff_host_rate		Different host connection %	Continuous
32	dst_host_count	Host based traffic	Host destination count	Continuous
33	dst_host_srv_count		Host destination for service count	Continuous

hybridized with other heuristic algorithms to enhance the performance of the developed model
[12]. A group of ants work in coordination with each other synchronously or asynchronously to achieve an optimal solution.ACO exhibits robustness as it can be modified for addressing different optimization problems. But convergence rate of ACO to an optimal solution is slow compared to other heuristic-based algorithms and the time requirement is also uncertain moreover ACO does not perform well with the problems having a large search place.

Particle Swarm Optimization (PSO) - is a global optimization algorithm based on the flocking behavior of birds or schooling behavior of fishes where the group of agents tries to converge to a common goal by observing the feedback from the other members of the swam. The PSO algorithm does not need to calculate any crossover mutation probabilities. Also, the speed of searching for an optimal particle is fast and the evaluations of PSO are simple

compared to the other algorithms. But the performance of the PSO algorithm depends on the selection of control parameters and hence, an optimal solution is not guaranteed.

Artificial Bee Colony (ABC) - algorithm is the optimization method inspired by the bees seeking their food, ABC algorithm is easy to implement, robust, and highly flexible. Addition or reduction in the number of bees does not require re-initialization of the algorithm this is because the performance of the algorithm is dependent on only two control parameters maximum cycle number and colony size.ABC algorithm can be hybridized with other algorithms to address various optimization problems. Also it requires very few parameters compared to other searching techniques. In order to improve the performance of the model, control parameters need to be modified that require new fitness tests.ABC requires a high amount of objective function evaluation to achieve an optimal solution.

Firefly Algorithm (FA) - inspired by the lightening of the firefly species to solve complex problems having high variance and complicated objective functions, compared to the other swarm algorithms firefly algorithm can be applied to various model functions very efficiently this algorithm is a population-based search approach, A is based on attractiveness between the fireflies and therefore, fireflies in the swarm can divide themselves into subgroups. The subgroups can move around each other or local optima to achieve the best global solution.

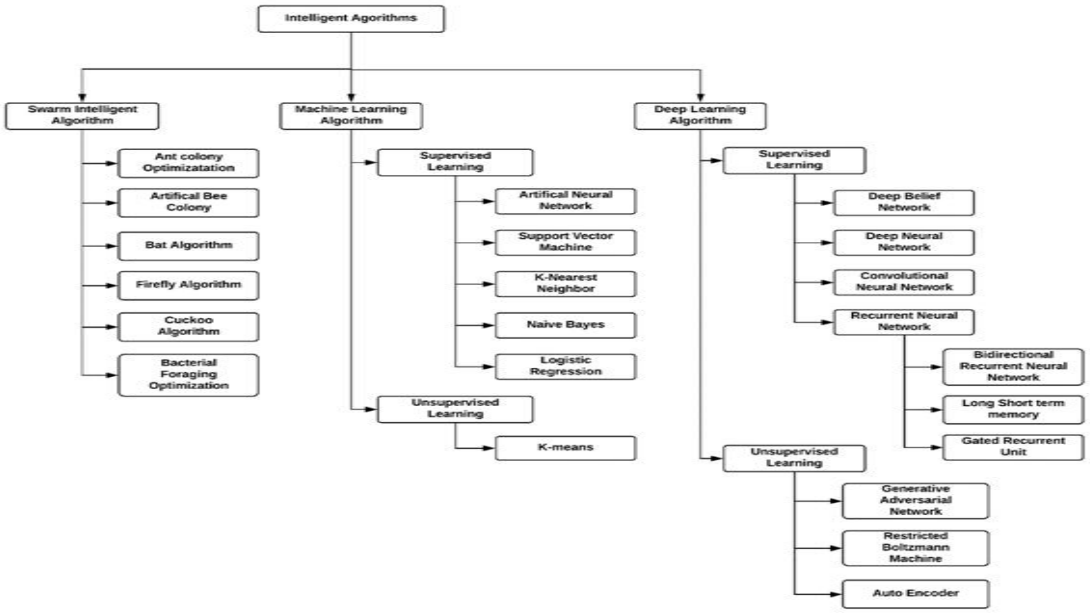

Figure 2. Categories of intelligent Algorithms

he division into subgroups enables the fireflies to find optimal solution simultaneously for the larger population size. But FA quite often gets trapped into local optima as they perform a local search to find the best solution and has a disadvantage of not memorizing any previous solution for each firefly present in the population. This results in fireflies moving in the search space regardless of the previous better solution and may result in completely missing the best solution.

Bat Algorithm

Bat Algorithm (BA) is inspired by the echolocation behavior of bats and is used to solve single objective and also multiple objective optimization problems.BA works well with

nonlinear problems and can be applied to a wide range of problems.BA has a quick transaction from exploration to exploitation. These characteristics of BA allows them to have a fast convergence rate.BA gives a good optimal solution but there is no mathematical analysis to establish the relation between the parameter and convergence rate and the accuracy of the BA is based on the system depending on the number of function evaluations.

4.2 Taxonomy on Machine Learning Algorithm

Machine learning (ML) techniques approaches are often used to learn from the system model and it is a subset of Artificial Intelligence. ML techniques are categories as supervised learning and unsupervised learning. In supervised learning techniques, the input data is mapped automatically to the output sample. For selecting the features and the classification process, machine learning techniques are briefly used for anomaly detection. In Table [6], describe the machine learning algorithm solving the NIDS problem. Machine learning methods can spontaneously discover the crucial difference between usual data and unusual data with peak accuracy. Two different types of machine learning models are supervised and unsupervised.

Supervised learning model -Supervised learning models depend on useful information in labeled data. The common machine learning algorithms used in IDS under supervised learning are Artificial Neural Network (ANN), Support Vector Machine (SVM), K- Nearest Neighbor(KNN), Naive Bayes, logistic regression(LR).

Artificial Neural Network (ANN)- The design scheme of an ANN is to imitate the way human brains work. An ANN carries an input layer, various hidden layers, and an output layer, the unit in adjacent layers are fully connected. It has a strong fitting ability, especially for nonlinear functions but due to the complex model structure, training ANN s is time-consuming. ANN models are Trained by the backpropagation algorithm that cannot be used to train deep networks.

Support Vector Machine (SVM)- The master plan of SVM s is to find a max-margin separation hyperplane in the n-dimension feature space. SVM s can attain gratifying results even with small-scale training sets because the separation hyperplane is determined only by a small number of support vectors. Learns useful information from small train set and have strong generation capability.SVM is sensitive to noise near the hyperplane and they are able to solve linear problems well.

K-Nearest Neighbor (KNN)-The fundamental scheme of KNN is based on the manifold hypothesis. If most of a sample's neighbors belong to the same class, the sample has a huge likelihood of belonging to the class. Thus the classification result is only related to the top-k nearest neighbors. The Constraint(K) purely depends on the model of KNN.In small k, the more complex the model is the higher the risk of overfitting. conversely in larger k, the simpler the model the weaker the fitting ability.

Naïve Bayes- The Naive Bayes algorithm is based on the conditional probability and the hypothesis of the attribute independence. This classifier computes the various classes of

conditional probabilities. They are robust to noise and is able to learn incrementally. But Naive Bayes does not perform well on attribute-related data.

Logistic Regression(LR)-The LR is a type of Logarithm linear model. The LR algorithm computes the probabilities of different classes through parametric logistic distribution.LR is simple, can be trained rapidly, and has automatic scale features but LR does not perform well on nonlinear data, it limits its application.

Unsupervised Learning Model - Unsupervised learning brings out valuable feature information from unlabeled data making it much effortless to obtain training data. The machine learning algorithm used in IDS under unsupervised learning is K-means.

K-Means -The shorter the distance between two objects, the more likely they are to be placed in the same cluster. The K-means algorithm adapts well to linear data, they are simple, can be trained rapidly, they have strong versatility and can fit big data. But they don't perform well on non-convex data, they are sensitive to initialization and to the parameter K.

4.3 Taxonomy on Deep-learning Algorithm

Deep learning methods have recourse to anomaly detection for both Dimensionality reductions and classification tasks. With the quick rise in transmitted traffic, manual feature engineering falls through to manage with multi-directional and comprehensive data, while on the contrary deep learning models spontaneously study knotty data. moreover, deep learning models manage the dynamic feature of network traffic and uninterrupted modification in offensive outlines. In consequence deep learning models are up-skilled with an abundance of documented facts to fabricate an anomaly detection model. The model categorizes the new traffic into one of two the normal or anomaly class. In a multi class categorization, the model can be furthermore classified from the infected traffic to distinct categories and subcategories of charge. In terms of complication, deep learning proceeds towards presume time absorbing and concentrated mathematical operations presented throughout numerous concealed layers and huge figure specification through the edifying stage. In Table [7], describe the deep-learning algorithm solving the NIDS problem.

The principle behind shallow and deep learning methods is the practice of advanced ANN architecture which is stimulated by the human mind and calculate absolutely through distinct technique than the conventional digitated methods. ANNs are machine learning algorithms that turn inputs to outputs using non-linear unexpressed processing of a set of feigned neurons this method is organized into shallow and deep learning. Lately, deep learning networks are broadly used for numerous design identification and network implementation, due to their potentiality towards studying some arithmetic operations in depth. In a NADS practice, shallow and deep networks need some instruction about the legalized data class methodically adjust the interconnections between neurons to study the mass of the network and acquire a model that can distinguish attacks from common actions. Deep learning networks are allocated in contrasting groups depending on their architectural blueprint that consists of hierarchical layers of non-direct processing levels. Based on Hodo et al, deep networks are grouped into generative and discriminatory architectures. The generative architecture determines joint likelihood dealing out from observed data with their classes, which comprises the following model.

Recurrent Neural Network (RNN)- is a supervised and/or unsupervised learning model. The thesis behind RNN is that information is connected in a course along with a layer-by-layer connection with a response loop. There is a direct rotation between its layers that expanding its, accuracy, with the ability to generate an internal memory for recording data of the preceding load.

Deep Auto Encoder (DAE)- used for studying methodical coding in an unconquered manner. The fundamental architecture of DAE requires an input layer, more than one buried layer, and an output layer that had the same amount of neurons in the input layer for remodeling.

Deep Boltzmann Machine(DAM)- is an incidental handpicked model that incorporates energy and theoretical units for the overall networks to manufacture binary results.

A Restricted Boltzmann Machine(RBM)- is appealed to lessen buried layers, which does not permit intra layer connections between buried units. Instructing a stack of DBM using untagged data as the input of the succeeding layer and pushing a layer for discrimination could open to establishing an architecture of DBM.

Deep Belief Network (DBM)-consists of many hidden layers, where a relationship is between layers not between units within each layer .it is an assemblage of unsupervised and supervised learning networks. The unsupervised model is studied by a greedy layer-by-layer connection at a time, whereas the supervised network is one or more layers connected for organizing tasks. The discriminative architecture calculates rear dispensations of classes constrained on the observed data that consists of RNN and Convolutional Neural Network(CNN).

RNN- use the discriminatory ability for categorization of the task, and this takes place when the output of the model is tagged in a series with the input.

Convolutional Neural Network(CNN)- is a space constant multi perceptron ANN, it has many buried layers, which generally comprise of convolution/complicated layers, pooling layers, fully connected layers and normalized layers. The complicated layers share many weights that have small frameworks, making the CNN uncomplicated in the instructing process compared to the other models with the same amount of buried /hidden layers.

Multiple research studies put in deep learning techniques to NADS. Using a DBN -based NADS for forming a greedy layer-by-layer learning algorithm to study each stack of RBM at a time for finding interruptions of events. A deep auto encoder technique was flourished to bring down measurements that were considered a pre-stage for allocating network observations. A shallow ANN algorithm was appealed as a classifier to evaluate the productiveness of an auto-encoder technique compared with the PCA and factor analysis algorithms. Proposed RNN-based NADS IDS for recognizing malicious network instances. The investigation was conducted on different hidden nodes and learning rate values.

Table 5:Swarm Intelligence Algorithm solving NIDS problem

Paper	**Algorithm**	**Dataset**	**Findings**	**Advantages**	**Limitation**
8	Linear correlation coefficient and cuttlefish algorithm	KDD-Cup'99	Accuracy - 95.03% False positive rate -1.65%	Understanding and reduction of data, limit storage space, reduction of processing cost.	Decrease the speed of detection rate
9	Feature Selection -Chi-Square Classifier -SVM MNB LPBoost	NSL-KDD	DOS and R2L attacks can be detected with 99% accuracy, probe with 98% and U2R with 100% accuracy	Interchangeable position in modular structure and algorithm changed anytime.	Contextual data cannot be identified .
10	Artificial bee colony	KDD-Cup'99	False alarm reduced	Resolving self-structuring multidimensional problems.	Limited search space
11	Chi-Square	KDD-Cup 99'	99% detection rate and 27% false positive rate	Compute easily and interpret	Need to improve non-linear data.
12.	Recursive feature Addition	ISCX 2012	Accuracy 92.9 And F-measures 92.9	RFA is an effective approach for identifying features.	Need to improve interpolation time
13	Ensemble Classifier	NSL-KDD	To detect 99.1%	Produce better results than single classifier.	Misconception between weak

			DDoS attacks		classifier combinations.
14	Ant Colony Optimization and SVM	KDD99	Accuracy 98.29%	To detect the intrusion in real time, the way of greedy approaches.	Slower convergence to detect intrusion
15	BAT and SVM	NSL-KDD	Accuracy 94.16% Detection rate 95.76% False Alarm rate 0.0408	Malleable and easy to implement in NIDS.	No centralized process.
16	Firefly Algorithm and C4.5 and Bayesian Networks	KDD-CUP ’99	Probe Accuracy 93.42%	Find good optimum solution in less number of detect.	No central control.
17	Cuckoo and ANN	KDD-Cup’99	Precision 0.98	Implementation is simpler.	Need high processing time.

Table 6: Machine Learning Algorithm solving NIDS problem

Paper	**Algorithm**	**Dataset**	**Findings**	**Advantages**	**Limitations**
18	Artificial Neural Network	KDDCUP’99	Detection accuracy Probe 98.79,U2R 96.51	Able to handle nonlinear data; Strong ability to fit;	Apt to over fit; prone to get stuck in a local optimum;
19	SVM,KNN,PSO	KDDCUP’99	Average elapsed time for ensemble experts is 0.459 s	Learn useful information from a small train set; Strong generation competency;	Sensitive to kernel parameters
20	Sparse	KDD-	Detection	Simply trained and	Need to improve

	Logistic Regression (SPLR)	Cup'99	rate 97.65%	automatically scaling features	non-linear data.
21	Genetic Algorithm and SVM	NSL-KDD	Accuracy 96.72%	Using graph, effectively evaluating network traffic threat.	Complex to fix the kernel limit value
22	K-Means and SVM	KDDCUP'99	Accuracy 95.75%	Simply trained and fit to big data.	Do not perform well on non-convex data; Improved method for initialization
23	Principal Component Analysis and Naïve Bayes	NSL-KDD	Accuracy 84%	Lesstraining and classify time.	Assuming the unknown parameter.
24	Random Forest and Average One-Dependence Estimator (AODE)	Kyoto	Accuracy 90.51% FAR 0.14	Flexible including missing data from previous node in the respective tree	Categorical variables difficult to find.
25	Random forest (RF) ,conditional informax feature extraction (CIFE),SVM,C5.0, Multilayer perceptro	KDD-CUP'99	Accuracy 99.2%, FAR 0.01, Error rate 0.007, Detection rate 99.7 Precision 99	Reduce variance and enhance accuracy.	Reliable towards outlier data.

	n neural network				
26	Ensemble Classifier	NSL-KDD	Accuracy 99.1%	Accurate than any one ofthe individual techniques	High cost

Table 7: Deep Learning Algorithm solving NIDS problem

Paper	**Algorithm**	**Dataset**	**Findings**	**Advantages**	**Cons**
27	Multi-layer perceptron (MLP) based DNN-IDS	NSL – KDD	In binary classification, accuracy 97%	Attributes are reduced inevitably and the outcome is optimum.	Fitting complex data models is high.
28	Deep feature extraction (Deep Belief Network) and multi-layer ensemble SVM	NSL-KDD, KDDCUP 99 UNSW-NB15, CICIDS2017	F-Measures NSL-KDD-94% UNSW-93% CICIDS-93% KDD99-97%	Rapid rate of convergence during the training stage.	Sensitive to the initial data.
29	Convolution Neural Network(CNN), CNN-RNN, CNN-LSTM, CNN-GRU	KDD-Cup’99	CNN 3 layer Accuracy RNN-96.9% GRU-97.7% LSTM-98.7%	Through Features selection, the data can be selected and turn out be more precise when its passed each level in the CNN.	Consume the usage of resources.
30	Long Short term memory, Recurrent neural network	KDD-CUP’99	LSTM-RNN DR-98.88%, FAR- 10.04%, Accuracy-96.93%	- Not any engineering attribute - Evaluating anomaly detection fast	Extremely moderate preparing time and not entirely interpretable

31	RNN,GRU,MLP and softmax module	NSL-KDD, KDD-CUP'99	Detection rate NSL-KDD-99.31%, KDD99-99.42%	Can process any length input.	RNN takes high processing time and it's a tedious process to access the data recurrently.
32	CNN with Gated Recurrent units(GRU)	ADFA	CNN with 600 GRU , Accuracy is 81%	The raw data is well-represented by a (deep) hierarchy of features, which can be modelled using a CNN. And the data we're working with has temporal properties which we want to model as well — hence the use of a RNN.	Gradient vanishing, exploding gradient. Large training data needed, don't encode the position and orientation of object.
33	Non-Symmetric Deep Auto Encoder(NADE)	KDD-CUP'99 NSL-KDD	Accuracy DBN-97.90% S-NDAE-97.85%	Unsupervised NDAE technique affords non-symmetric attributes dimensionality reduction minimize training time.	Extensive computation of resources and memory involved.
34	STL-IDS, SVM	NSL-KDD	Accuracy – 80.48%, Precision – 93.92, Recall -68.28, F-measures - 79%	Robust to noisy data	Carefully scrutiny requirement due to lack of guarantee for successful training due to overfitting problem
35	Wasserstein	KDD-	WGAN-CLIP -	- Not any	Extremely moderate

	generative adversarial networks (WGAN)	CUP'99	57.153 ± 0.5615%, WGAN-GP-47.627	engineering attribute - Evaluating anomaly detection fast	preparing time and not entirely interpretable
36	Restricted Boltzmann Machine (RBM)	ISCX	Accuracy for CD -88.6%, Accuracy for PCD-89%	Rapid rate of convergence during the training stage.	Sensitive to the initial data.

5. Conclusion and Future Direction

Deep learning had extracted the recognition and observation of researchers in unrelated fields. Deep models can handle complicated data and find connections among input characteristics without using human intervention. With the publication of new technologies and expeditious growth in transmitted traffic, researchers have been inquired into deep learning for intrusion detection [20]. This survey assessed and differentiated the key surveys contemplating deep learning for intrusion detection and building the contemporary survey upon the preceding ones. The study supplied a novel fine-grained taxonomy contemplating distinct blueprint aspects, including input data, detection, deployment, and evaluation strategies. Correspondingly, this survey furnished an accurate review of the interconnected experimental studies in deep learning - based IDS. Through the complete review, we had uncovered different discoveries and lessons. Deep learning is mostly used for feature learning in intrusion detection perspectives, even though some studies used deep learning models as classifiers. Although, we observed that most suggested approaches depend on the legacy benchmark datasets. Moreover, less attention has been given to announcing the productiveness of the proposed approaches. The present discovery reveals that additional endeavors are needed to ameliorate the current state-of-the-art, in the perspective of these discoveries. This survey also puts out several research challenges and future directions [36]. Since the benchmark datasets do not cross-examine the current advanced status of distinct types of networks, there is an acute necessity to use and generate more current datasets and real-time prototypes based on current hardware advances. Moreover, various domains should be reevaluated with deep learning approaches instead of shallow machine learning, also should conduct further relative studies and exploring hybrid and group architectures.

REFERENCES:

1. Moustafa, N., Hu, J., & Slay, J. (2019). A holistic review of network anomaly detection systems: A comprehensive survey. Journal of Network and Computer Applications, 128, 33-55.
2. Subashini, P., Krishnaveni, M., Dhivyaprabha, T. T., & Shanmugavalli, R. (2020). Review on Intelligent Algorithms for Cyber Security. In Handbook of Research on Machine and Deep Learning Applications for Cyber Security (pp. 1-22). IGI Global.
3. Aleroud, A., & Karabatis, G. (2017). Contextual information fusion for intrusion detection: a survey and taxonomy. Knowledge and Information Systems, 52(3), 563-619.

4. Ahmed, M., Naser Mahmood, A., & Hu, J. (2016). A survey of network anomaly detection techniques. Journal of Network and Computer Applications, 60, 19–31. doi: 10.1016/j.jnca.2015.11.016.
5. Aldweesh, A., Derhab, A., & Emam, A. Z. (2019). Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues. Knowledge-Based Systems, 105124. doi: 10.1016/j.knosys.2019.105124.
6. Kim, K., & Aminanto, M. E. (2017). Deep learning in intrusion detection perspective: Overview and furtherchallenges. 2017 International Workshop on Big Data and Information Security(IWBIS). doi:10.1109/iwbis.2017.8275095.
7. Thudumu, S., Branch, P., Jin, J., & Singh, J. J. (2020). A comprehensive survey of anomaly detection techniques for high dimensional big data. Journal of Big Data, 7(1), 1-30.
8. Mohammadi, S., Mirvaziri, H., Ghazizadeh-Ahsaee, M., & Karimipour, H. (2019). Cyber intrusion detection by combined feature selection algorithm. Journal of Information Security and Applications, 44, 80–88. doi: 10.1016/j.jisa.2018.11.007.
9. Thaseen, I. S., Kumar, C. A., & Ahmad, A. (2018). Integrated Intrusion Detection Model Using Chi-Square Feature Selection and Ensemble of Classifiers. Arabian Journal for Science and Engineering. doi:10.1007/s13369-018-3507-5.
10. Kanaka Vardhini, K., & Sitamahalakshmi, T. (2016). Implementation of Intrusion Detection System Using Artificial Bee Colony with Correlation-Based Feature Selection. Proceedings of the First International Conference on Computational Intelligence and Informatics, 107–115. doi:10.1007/978-981-10-2471-9_11.
11. Divyasree, T. H., & Sherly, K. K. (2018). A Network Intrusion Detection System Based On Ensemble CVM UsingEfficient Feature Selection Approach. Procedia Computer Science, 143, 442–449. doi: 10.1016/j.procs.2018.10.416.
12. Hamed, T., Dara, R., & Kremer, S. C. (2018). Network intrusion detection system based on recursive feature addition and bigram technique. Computers & Security, 73, 137–155. doi: 10.1016/j.cose.2017.10.011.
13. Das, S., Mahfouz, A. M., Venugopal, D., & Shiva, S. (2019). DDoS Intrusion Detection Through Machine Learning Ensemble. 2019 IEEE 19th International Conference on Software Quality, Reliability and Security Companion (QRS-C). doi:10.1109/qrs-c.2019.00090.
14. Mehmod, T., & Rais, H. B. M. (2016) "Ant Colony Optimization and Feature Selection for Intrusion Detection",Advances in Machine Learning and Signal Processing, 305–312. doi:10.1007/978-3-319-32213-1_27.
15. Enache, A.-C., & Sgarciu, V. (2014) "Anomaly intrusions detection based on support vector machines with bat algorithm" 2014 18th International Conference on System Theory, Control and Computing(ICSTCC). doi:10.1109/icstcc.2014.6982526.
16. B, S., & K, M. (2018) "Firefly algorithm based Feature Selection for Network Intrusion Detection" Computers &Security. doi: 10.1016/j.cose.2018.11.005.
17. Rithesh, K. (2019)" Anomaly-Based NIDS Using Artificial Neural Networks Optimised with Cuckoo Search Optimizer" Emerging Research in Electronics, Computer Science and Technology, 23–35. doi:10.1007/978-981-13-5802-9_3.
18. Akashdeep, Manzoor, I., & Kumar, N. (2017) "A feature reduced intrusion detection system using ANN classifier", Expert Systems with Applications, 88, 249–257. doi: 10.1016/j.eswa.2017.07.005.

19. Aburomman, A. A., & Ibne Reaz, M. B. (2016) "A novel SVM-kNN-PSO ensemble method for intrusion detection system", Applied Soft Computing, 38, 360–372 doi: 10.1016/j.asoc.2015.10.011.
20. Shah, R., Qian, Y., Kumar, D., Ali, M., & Alvi, M. (2017). Network Intrusion Detection through Discriminative Feature Selection by Using Sparse Logistic Regression. Future Internet, 9(4), 81. doi:10.3390/fi9040081.
21. Raman, M. G., Somu, N., Kirthivasan, K., Liscano, R., & Sriram, V. S. (2017). An efficient intrusion detection system based on Hypergraph-Genetic algorithm for parameter optimization and feature selection in support vector machine. Knowledge-Based Systems, 134, 1-12.
22. Al-Yaseen, W. L., Othman, Z. A., & Nazri, M. Z. A. (2017) "Multi-level hybrid support vector machine and extreme learning machine based on modified K-means for intrusion detection system" Expert Systems withApplications, 67, 296–303. doi: 10.1016/j.eswa.2016.09.041.
23. Sharmila, B. S., & Nagapadma, R. (2019) "Intrusion Detection System using Naive Bayes algorithm", 2019 IEEEInternational WIE Conference on Electrical and Computer Engineering (WIECON-ECE). doi:10.1109/wiecon-ece48653.2019.9019921.
24. Jabbar, M. A., Aluvalu, R., & Reddy S, S. S. (2017) "RFAODE: A Novel Ensemble Intrusion Detection System"Procedia Computer Science, 115, 226–234. doi: 10.1016/j.procs.2017.09.129.
25. Dominique, N., & Ma, Z. (2019) "Enhancing Network Intrusion Detection System Method (NIDS) Using MutualInformation (RF-CIFE)", NanoScience and Technology, 329–342. doi:10.1007/978-3-030-16946-6_26.
26. Das, S., Mahfouz, A. M., Venugopal, D., & Shiva, S. (2019) "DDoS Intrusion Detection Through Machine Learning Ensemble", 2019 IEEE 19th International Conference on Software Quality, Reliability and Security Companion (QRS-C). doi:10.1109/qrs-c.2019.00090.
27. Kasun Amarasinghe, Milos Manic (2018)" Improving User Trust on Deep Neural Networks based Intrusion Detection Systems" IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society.
28. Marir, N., Wang, H., Feng, G., Li, B., & Jia, M. (2018) "Distributed Abnormal Behavior Detection Approachbased on Deep Belief Network and Ensemble SVM using Spark" IEEE Access, 1– 1. doi:10.1109/access.2018.2875045.
29. Vinayakumar, R., Soman, K. P., & Poornachandran, P. (2017) "Applying convolutional neural network for network intrusion detection", 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI). doi:10.1109/icacci.2017.8126009.
30. Kim, J., Kim, J., Thu, H. L. T., & Kim, H. (2016) "Long Short Term Memory Recurrent Neural Network Classifier for Intrusion Detection", 2016 International Conference on Platform Technology and Service (PlatCon). doi:10.1109/platcon.2016.7456805.
31. Xu, C., Shen, J., Du, X., & Zhang, F. (2018)," An Intrusion Detection System Using a Deep Neural Network with Gated Recurrent Units", IEEE Access, 1–1. doi:10.1109/access.2018.2867564.

32. Chawla, A., Lee, B., Fallon, S., & Jacob, P. (2019),” Host Based Intrusion Detection System with Combined CNN/RNN Model”, Handbook of Experimental Pharmacology, 149–158. doi:10.1007/978-3-030-13453-2_12.
33. Shone, N., Ngoc, T. N., Phai, V. D., & Shi, Q. (2018) “A Deep Learning Approach to Network Intrusion Detection”, IEEE Transactions on Emerging Topics in Computational Intelligence, 2(1), 41–50. doi:10.1109/tetci.2017.2772792.
1. Al-Qatf, M., lasheng, Y., Alhabib, M., & Al-Sabahi, K. (2018), “Deep Learning Approach Combining Sparse Autoen-coder with SVM for NetworkIntrusionDetection”, IEEE Access, 1–doi:10.1109/access.2018.2869577.
35. Yan, Q., Wang, M., Huang, W., Luo, X., & Yu, F. R. (2019),” Automatically synthesizing DoS attack traces using generative adversarial networks”, International Journalof Machine Learning and Cybernetics. doi:10.1007/s13042-019-00925-6.
36. Aldwairi, T., Perera, D., & Novotny, M. A. (2018), “An evaluation of the performance of Restricted Boltzmann Machines as a model for anomaly network intrusion detection”,Computer Networks, 144, 111–119. doi: 10.1016/j.comnet.2018.07.025 .