Wireless Personal Communications (2022) 125:1385–1403
https://doi.org/10.1007/s11277-022-09609-x

Intrusion Detection System Using Deep Belief Network & Particle Swarm Optimization

P. J. Sajith[1] · **G. Nagarajan**[1]

Accepted: 7 February 2022 / Published online: 1 March 2022

Abstract
Securing the services of security such as data integrity, confidentiality and availability is one of the great challenges. Failure to secure above will potentially lead many cyber-attacks. One of the greatest hits for detecting intrusion is an intrusion detection system (IDS) and there are so many advances put forward by many researchers. Even though there exists a large number of Intrusion Detection Systems intruders are still continuing with their job. Another evolving and yet revolutionized strategies is Deep Learning. So, integrating these two systems to create an effective model that could potentially find normal or malicious attacks. In this paper, we classify intrusion using Deep Belief Network and Particle Swarm Optimization into categories like Normal, Probe, DoS, U2R, R2L. The dataset used for applying this model is DARPA 1999 and they are evaluated under various measures. Also, the proposed system is compared with other system like ANFIS, HHO, Fuzzy GNP in which our system outperforms better with greater accuracy of 96.5%.

Keywords Adaptive neuro fuzzy inference system · Deep learning · Deep belief network · Harris Hawks optimization · Particle swarm optimization

Abbreviations

ML	Machine learning
DL	Deep learning
SVM	Support vector machine
KNN	K nearest neighbor
ANFIS	Adaptive neuro fuzzy inference system
DBN	Deep belief network
CNN	Convolution neural network
RNN	Recurrent neural network
ANN	Artificial neural network
HHO	Harris Hawkins optimization

✉ P. J. Sajith
pjsajith@gmail.com

G. Nagarajan
nagarajanme@yahoo.co.in

[1] Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai 600127, India

F-GNP	Fuzzy graph neural processes
IDS	Intrusion detection system
NIDS	Network intrusion detection system
RBM	Restricted Boltzmann machine
DAE	Denoising auto encoder
U2R	User to root
R2L	Remote to local
DoS	Denial of service
SNN	Spiking neural network
MF	Member function

1 Introduction

Intrusion detection systems (IDS) observe gadgets that have been added to the safety mass to prevent harmful action on a framework. This work centers around network intrusion detection systems (NIDSs) fundamentally on the grounds that they can recognize the broadest scope of assaults contrasted with different sorts of IDSs. Organization IDSs dissect traffic to identify on-going and approaching assaults on an organization. These days, business IDSs fundamentally utilize an information base of rules, called marks, to attempt to identify assaults on an organization or on a host system. Interruption discovery frameworks are checking gadgets which are utilized to identify interruptions on a PC or an organization. Interruptions are unapproved and atypical exercises which were characterized by Li et al. [1] as an arrangement of associated activities carried out by a noxious foe that results in the tradeoff of an objective framework. An interruption recognition framework is a fundamental apparatus for network directors in light of the fact that without such a gadget, it is difficult to investigate the immense measure of parcels navigating current organizations consistently. After over 30 years of concentrated exploration on interruption location frameworks, the field is as yet open to additional examinations particularly with respect to the exactness of the discovery. Also, variations of referred to assaults too as new assaults can frequently go through the framework without being recognized. Basic IDS architecture is depicted in Fig. 1.

In NIDS, the recognition framework is reviewing the approaching and active organization traffic from all hosts continuously and dependent on specific standards, it can recognize and distinguish the assault, at that point, take the reasonable safety efforts to stop or hinder it, which essentially decreases the danger of harm to the organization [2, 3]. In any case, because of the fast expansion in the intricacy of the network safety assaults, the current techniques utilized in NIDS are neglecting to adequately address this issue. Area of classification engine is the essential section of IDS structure that selects whether or not the change from above includes the vector and fits the interference definition [4]. Mostly this phase may be both executed as arranged signature where the classifications are finished with from now on defined imprints or peculiarity located in which by configuring the data flows of previous, the system's conventional flow beam is learned. Albeit the initial one has a time of decent choice and dependability; consequently, generate a low recognition rate because it can't identify new sorts of assaults. Simultaneously, the last one has adequate flexibility, robustness, and adaptability. Along these lines, for executing a dynamic IDS framework, principally the abnormality location techniques are favored particularly with a DL instrument [5, 6].

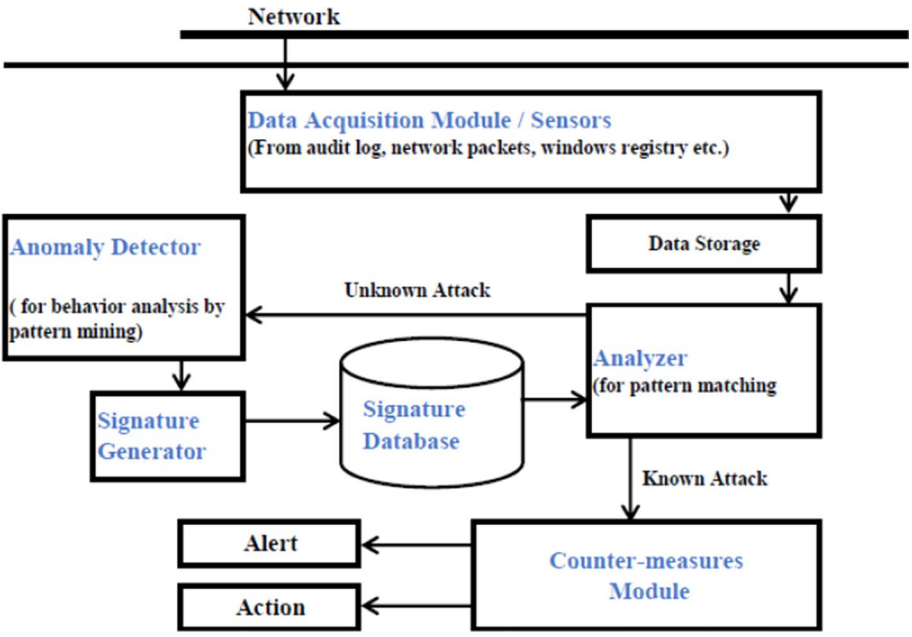


Fig. 1 Traditional IDS

To resolve the above problems, experts have started to focus on creating IDSs using AI techniques. AI is a kind of computerized reasoning strategy that can naturally find valuable data from gigantic datasets [7, 8]. AI-based IDS can reach attractive area levels when there is sufficient planning data and AI models are sufficiently generalizable to perceive varieties of attacks and new attacks. Furthermore, AI put together IDSs don't depend vivaciously with respect to space data; henceforth, they are not hard to design and create. Significant learning is an element of AI that can reach remarkable spectacles. Differentiated and regular ML methodology, significant learning methodologies are better at overseeing immense data [9]. Moreover, significant learning methods can normally take in incorporate depictions from unrefined data and thereafter yield results; they work from top to bottom and are utilitarian. A noteworthy quality of deep learning is deep construction, which contains many secret layers [10]. Interestingly, the usual ML models, such as SVM and KNN, contain no or just a secret layer. As a result, these usual AI models are also referred to as shallow models.

1.1 Key Highlights

The focus of this paper to bring an efficient DL-based detection model for classifying certain intrusion with best accuracy in which some of key notes is listed below;

a. Proposing an efficient intrusion detection system using DBN and PSO
b. Model is applied over DARPA 1999 dataset
c. Proposed system is compared with ANFIS, HHO, Fuzzy GNP
d. Accuracy of the proposed system outperforms with other cutting-edge systems

1.2 Organization of the paper

As we came across the introductory part in Sect. 1, the rest as follows; Sect. 2 depict the related works which is proposed for recent years about IDS, Sect. 3 depicts the methodology of detecting potential intrusions, Sect. 4 presents the implementation and results and concludes in Sect. 5.

2 Related Works

In this paper [11], a profound learning approach for interruption discovery utilizing Recurrent Neural Network (RNN-IDS) is proposed. RNN helps in improving the precision of classifier to accomplish compelling interruption location. It can recollect past data and can apply it to the current yield which makes it damn compelling than past DL approaches like Feed Forward Neural Networks. The display of suggested method is surveyed utilizing the dataset of NSL-KDD and is focused on multi and dual class queries and differentiated and other methodologies of ML-based such as SVM, ANN, J48, and so on. Impacts of number of neurons and different learning rates are further analyzed. The outcomes show that the use of RNN for representation expands the precision appropriately and that its presentation is higher than standard AI application procedures in dual and multi-class collection.

In this work [12], a significant philosophy based on learning in the network interference area using Denoising Auto-Encoder (DAE) is completed. Reduction of weight work is consolidated which helps to select the predetermined amount of huge features to decrease the dimensionality of the components. The picked data is then portrayed using multi-facet perceptron (MLP) as classifier. Examinations are performed using the UNSW-NB data set. The results show that the assurance of the components gives an attractive revelation execution with low memory and power requirements.

In [13], the importance of assessing the significance of characteristics was investigated for IDS with the most generally used dataset, KDDCup 99. For every part, they had the opportunity to share the relevance of the component to the information obtained. What's more, they introduced the most applicable highlights for each class name. In [14] examined arbitrary woodland strategies in abuse recognition by learning examples of interruptions, inconsistency identification with exception discovery instrument, and hybrid location by consolidating both the abuse and irregularity discovery. They indicated that the maltreatment method works better than the KDDCup 99 test outcome's winning segments, and moreover eccentricity ID worked better stood out from other dispersed independent peculiarity revelation methods. By and large, it was reasoned that the half breed framework upgrades the presentation with the upside of joining both the abuse and abnormality discovery approaches [15–17]. With a computationally quicker calculation, greater identification rate and a lower counterfeit alert rate their structure contrasted higher when compared to different distributed outcomes. Nonetheless, the method of steady learning is cannot accept by it, this is its disadvantage. In [18], the SNN-based model's introduction was considered and discovered as the excellent calculation with an excessive recognizable proof rate. With declined data set, they had the alternative to the reason that the SNN carried out nicely for the scheme of 'U2R' attack with K-infers. In any case, their work does not show overall results of the test dataset.

This paper [19] explores the capacity of utilizing significant learning designs dynamically for the area of network interference. For perceiving the class of assault with a pattern of multinomial, to expect whether is there any interference, a model structure of cloud facilitated was developed and that joins an order model of significant binomial learning. The model framework coordinates profound learning models constructed utilizing the H_2O system, also an informing administration to caution the organization manager. A study of comparison was once finished utilizing the notable dataset of benchmarked NSL-KDD to differentiate the models gathered with Naive Bayes, Random Forest, DeepLearning4J, Lib-SVM and Logistics Regression, and the significant models of learning H_2O. The outcomes confirmed that commonly the various models were beaten by the model of H_2O deep learning, for each multinomial and binomial characterization accomplishing more than 83% precision on the test dataset and over 99.5% exactness utilizing cross-approval on the preparation dataset.

In this work [20] not all ascribes are required for recognizing assaults diminished number of highlights can diminish the discovery rate or increment the recognition. Henceforth, we join channel and covering based way to deal with select proper element for IDS. The all-inclusive work is in progress utilizing GPU offices to diminish the time taken for calculation and improved outcomes. Utilizing this methodology, we consolidated channel and covering based way to deal with select proper highlights for distinguishing Network Intrusion. The inspiration of the work is in lessening the quantity of highlights with improved execution for a positive identification rate. The proposed work centers around NIDS. Despite the fact that different procedures exist in the writing for NIDS regarding determination of highlights, classifiers, the proposed work focuses on the Meta heuristic methodology called firefly strategy for include choice and C4.5 classifier and contrasted and Bayesian organization classifier.

3 Methodology

Figure 2 represents the overall methodology of DL-based IDS in which initially sufficient data are collected from DARPA 1999 dataset and these are pre-processed to avoid certain anomalies and analysis of network is performed. Then these are passed over to extract or select sufficient features using PSO and are given to DBN classifier for effective classification. Once training is completed, testing is also applied and thereby gaining 5 categories of intrusion. The main advantage of using the DBN is that it requires only a very small dataset and only requires a very short training time. The accuracy is also high for DBN.

3.1 Dataset Description

For this model for execution, we utilized DARPA 1999 dataset created in MIT Lincoln Laboratories. The Dataset is made by presenting physically produced network-based assaults [21]. Different attacks that can be potentially find in an organization is characterized in a brief form [22] concerning DARPA interruption discovery evaluation dataset [23].

a. The test information of the DARPA1999 included 190 examples of the 57 attacks which included 8 Probes, 17 DoS, 17 R2L and 15 U2R with subtleties of attack types given in Table 1 [24].

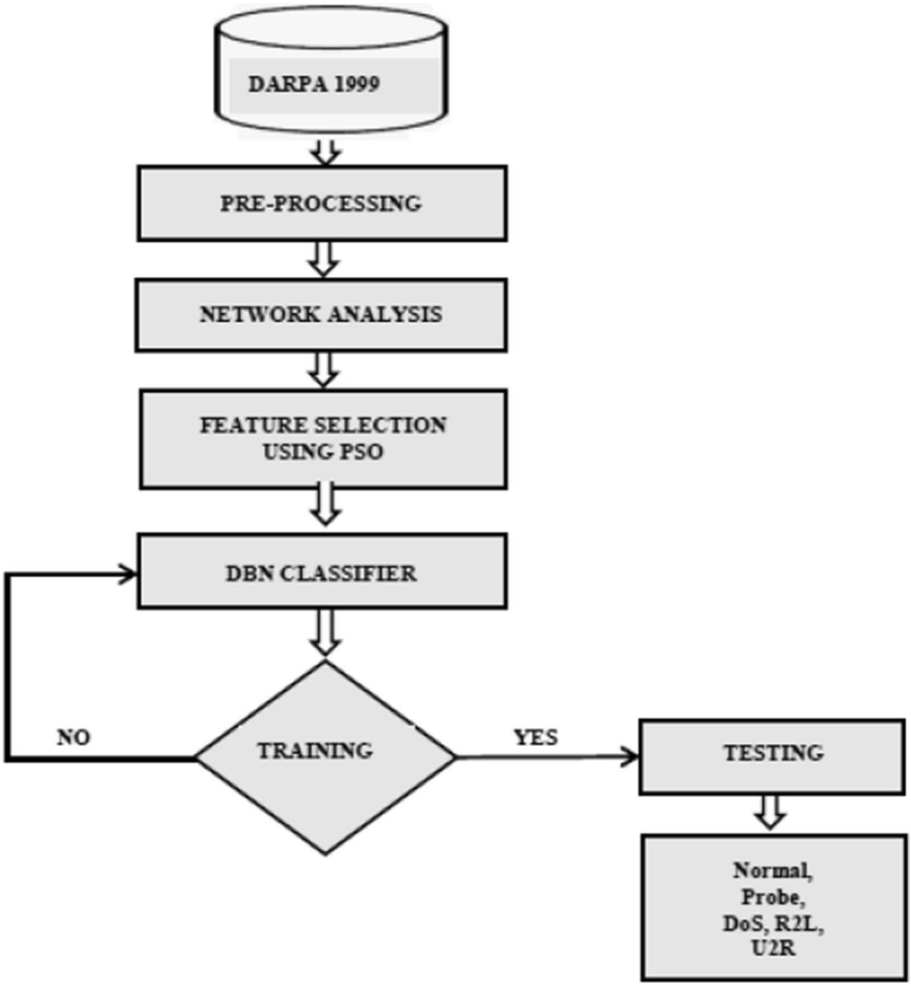


Fig. 2 Proposed system for IDS

Table 1 Dataset attack types and its classes

Attack class	Attack types
Probe	Portsweep, queso, msscan, lsdomain, illegal-snifer, ipsweep ntinfoscan, satan,
DoS	Selfping, dosnuke, back, tcpreset, syslogd, arppoison, mailbomb, teardrop, processtable, neptune, udpstorm, land, warezclient, apache2, crashiis, smurf, pod
R2L	Imap, xlock, sshtrojan, ppmacro, netbus, sendmail, snmpget, ncftp, httptunnel, xsnoop, named, dict, framespoof, netcat, guest, ftpwrite, phf
U2R	Sechole, ps, secret, perl, fdformat, casesen, ntfsdos, yaga, ppmacro, eject, loadmodule, nukepw, sqlattack, xterm, ffbconfig

b. The assaults ordered into four basic categories, Denial of Service assault (DoS), Probe assault, User to Remote assault (U2R) and Remote to local assault (R2L). The test assaults naturally examine a framework or organization in endeavor [24] to gather records of private frameworks in order to detect a IP address that is authenticated, have working framework sorts (mscan, queso) dynamic ports (mscan, portsweep), and perceived weaknesses (satan) [24].
c. The DoS assaults are plan to confound a host or organization administration toward off substantial clients from utilizing a help given by the framework [25]. These comprise of the Solaris working framework crash (selfping), effectively end all TCP associations for a host that is more specified which tcpreset, followed by arppoison for storing degenerated ARP, rest is crashiis for MW-NT and dosnuke for W-NT [24].
d. R2L assaults, [26] an aggressor who doesn't have a record or any entrance on a casualty machine and takes advantages of bugs or shortcoming in machine to gains nearby admittance to the machine (visitor, dict), eliminate documents from the machine (ppmacro) or changes information on the way to the machine(framespoof). New R2L assaults incorporate an internet browser assault called a man-in-center (framespoof), NT power point large scale assault (ppmacro), a Linux trojan SSH worker (sshtrojan), netbus for trojan apparatus of NT, ncfp for accessing document using FTP in linuxa [24].
e. Attacks by U2R Is nothing but client that acts as security over a framework and can only be accessible by W-NT. The main strategy is to eliminate documents which makes that particular area harm free. These incorporate mystery assaults, where a client who is approved to get right of passage to the unique documents eliminates them (ntfsdos, sqlattack) [24].

3.2 Pre-processing

Once the dataset is collected, these are now undergoing pre-processing stage where DARPA 1999 contain several attack type and classes which potentially be character type, some of them be discrete type in which all these variations will be converted to numeric forms. And afterward from the information, separate class features as label data and in the end by means of L2 standard, the data is standardized.

3.3 Network Analysis

This is the resulting advance in proposed system where affiliation analyzer dismantles the traffic information of the inaccessible sensor affiliation. A far away sensor network is a social connection of far-away sensor places. Transmission of information groups are the critical mean of correspondence inside inaccessible sensor community focuses. As per the IEEE convention guidelines, every information bundle has its fixed parcel design and every remote sensor hub follow that bundle information design. The point of organization examination is to dissect information parcels just as organization traffic.

3.4 Feature Selection

Here once network is analyzed and normalized, these are now passed over to feature selection process where you naturally or physically select those highlights which contribute most to your expectation variable or yield in which you are keen on. So, for that PSO is used here for feature extraction. PSO is an equal estimation, which has the advantages of straightforward execution, high exactness and quick assembly [23, 25, 26]. To track down the best arrangement, PSO instates some irregular arrangements in arrangement space, these arrangements are a few particles, where characterize the molecule speed v_i and the molecule position x_i. In the meantime, use the capacity of health to determine if the circumstances of the particles are ideal, use *gbest* and *pbest* to capture the individual's better circumstances and social opportunity independently. For every particle, note its well-being, if it is higher contrasted with *pbest* then it will also be *pbest*, and assume higher contrasted with *gbest* then it will be like *gbest*, the speed and location of the molecule is update by it. The speed of molecules and position update rules are according to the accompaniment;

$$v_i = wv_i + c_1 \times rand0 \times \left(\text{pbest}_{\,i} - x_i \right) + c_2 \times rand() \times \left(\text{gbest}_{\,i} - x_i \right) \tag{1}$$

$$x_i = x_i + v_i \tag{2}$$

where v_i is the speed of a molecule, inactivity weight is w, *rand()* is an irregular worth somewhere in the range of 0 and c_1 and c_2 is the current situation of the molecule c_1 and c_2 are speed increase factor. If the velocity or circumstance of the particles exceeds the degree of stroke, it will be defined as the most limiting velocity or the circumstance of the cutoff. At the point where the molecule has been reinvigorated, it will keep reheating till the better game plan is noticed. Regularly searching the better position or appearing at the most remarkable number of cycles will halt the demand. In DBN, the range of concealed core portion points influences the generation of the independent studying stage and coordinated studying stage's fine-tuning. Along these lines, the quantity of covered up layer hubs in the profound adapting should be enhanced by PSO calculation to improve the exhibition of the organization.

3.5 DBN Classifier

Now we have DARPA 1999 dataset, which potentially contains 190 instances in which 170 for training stages and 20 for testing stage is taken. As stated, early DBN consist of RBM layers in which initially RBN is needs to be defined. A two-layer association is RBM, all unit in the middle of the layer has a bi-directional affiliation, the units inside a comparative layer are not bound. An energy model is RBM, energy of the shared arrangement is characterized by it:

$$E(v, h) = -\sum_{i \in V} a_i v_i - \sum_{j=H} b_j h_j - \sum_{i,j} v_i h_j w_{ij} \tag{3}$$

Here v_i stands for the visible unit and h_j stands for the hidden units. The weight of edge between v_i and h_{j} is denoted by w_{ij}. Acknowledge that a value of 1 or 0 the center point of neuron have, where v_i and h_j independently address the mysterious layer unit j 's and

evident layer unit I 's combined state, the load among I and j is the w_{ij}, the tendency of the unit j is addressed by b_j and a_i tends to the inclination of the unit I.

4 *v* and *h* joint probability's definition:

$$p(v,h) = \frac{1}{Z} e^{-E(v,h)} \tag{4}$$

The allocation function is called Z:

$$Z = \sum_{v,h} e^{-E(v,h)} \tag{5}$$

From the above 2 equation we can obtains its possibility by the Equation below;

$$p(v) = \frac{1}{Z} \sum_{h} e^{-E(v,h)} \tag{6}$$

Preparing a RBM is allowed the RBM to display to learn boundary) (ϕ) = w*a*b, that is to get familiar with the boundary when the probability likelihood arrives at the most extreme. The logarithmic probability work is utilized to determine the boundaries w:

$$\frac{\partial \log p(v)}{\partial w_{ij}} = \langle v_i h_j \rangle_{\text{data}} - < v_i h_j >_{\text{model}} \tag{7}$$

The gradient update criterion is:

$$\Delta w_{ij} = \varepsilon \left(< v_i h_j >_{\text{data}} - < v_i h_j >_{\text{model}} \right) \tag{8}$$

Epsilon is the learning rate. The point sections addendum shows the assumption for the sections under a specific likelihood conveyance. The assumption for the information conveyance is not difficult to acquire, conversely the assumptions for the model dispersion requirements to play out the long-lasting Gibbs Sampling. Hinton suggested a difference disparity calculation in 2002. The target work that should be advanced is changed from logarithmic probability capacity to differentiate uniqueness work, which speeds up the preparation speed, simultaneously the rules for the angle plunge is:

$$\Delta w_{ij} = \varepsilon \left(< v_i h_j >_{\text{data}} - \langle v_i h_j \rangle_{\text{recon}} \right) \tag{9}$$

These RBM networks are stacked by DBN in which the yield of the past RBM is the commitment of the last RBM. The information is a commitment to the organization through the basic RBM, the two fundamental layers are a bi-directional affiliation, and the unidirectional affiliation is the remainder of the layers, the association framework showed up in Fig. 3.

DBN learning is divided into two areas: preliminary preparation and fine-tuning [13]. Pre-planning is the model of solo get the shot of, beginning with the key RBM input information, by then the yield of the fundamental RBM as the responsibility of the 2nd RBM, so status layer by means of layer, till all layers are set up to wrap up. Afterward the presetting up, the restrictions of every layer, which includes the propensity and weight as the principal worth as far as possible, by then utilizing the BP calculation to change the constraints of the whole affiliation, which is a participation of oversight learning [11].

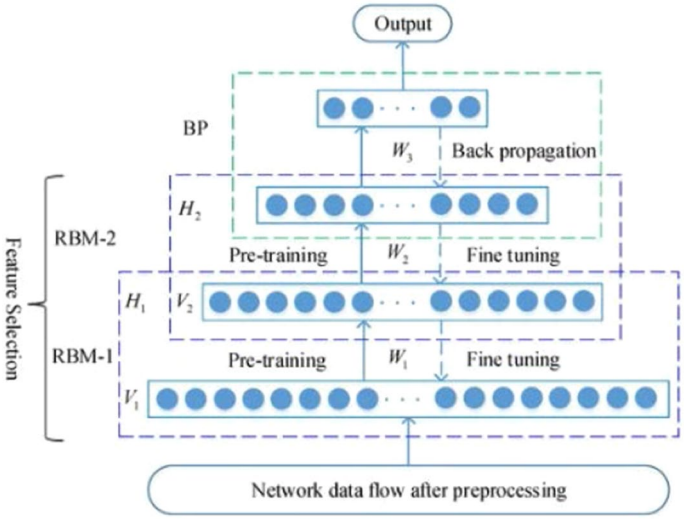


Fig. 3 DBN architecture by stacking RBM

Table 2 Comparative analysis

Models	False-Alarm rate	Accuracy (%)
HHO	0.05	76
F-GNP	1.9	85
ANFIS	3.4	92
DBN (Ours)	4.8	96.5

4.1 Dimensionality Reduction

The DBN is utilized to diminish the information. By unaided learning the crude information is planned to the area of low-dimensional in the pre-preparing stage. To excellent the significance of the affiliation design the managed learning is utilized in the tuning stage, the information is decreased to five assessments in this paper. Consisting of three secret layers, information layer and the yield layer, network of 5-layer engineering is planned. The portions of DARPA 1999 dataset is not especially excessive stood apart from the picture preparing, so the affiliation design that incorporate 3-secret layers fulfills the test prerequisites. Since the association plan of the DBN extraordinarily affects the learning execution, the amount of concealed layer centers really ought to be settled. In the event that the number is resolved physically, the ideal organization structure isn't really gotten, so we utilize the PSO calculation to track down the quantity of each covered up layer hub to decide the ideal organization structure.

5 Implementation and Results

Firstly, this model will be implemented over the Wamp server for performing several operations and also these are integrated with programming this neural network using an open-source library called TensorFlow. The proposed model is compared with other systems such as ANFIS, HHO, F-GNP. These comparing models also executed using Wamp server along with some Membership Function (MF) and some of results obtained is shown below. Also, we compared these systems using performance measures like False Alarm Rate (FAR) and Accuracy. Table 2 depicts the overall comparison of proposed model with other model in terms of performance measures. Figure 8 depict the comparative analysis of ANFIS, HHO, F-GNP, DBN.

5.1 ANFIS

The ANFIS structure maps contributions through input enrollment works and related boundaries, and afterward through yield participations and related boundaries to yields. During the learning interaction, the boundaries related with enrollment capacities changes. An incline vector empowers the figuring of these limits, giving an extent of how well the FIS models the data/yield data for a given game plan of limits. In the wake of procuring the tendency vector, any of the couple of smoothing out timetables could be applied to change the limits for reducing some mix-up measure. This learning method works correspondingly as that of neural associations. When appeared differently in relation to the generally FIS, ANFIS is really astounding. This makes the fuzzy system to acquire from the data they model (Fig. 4).

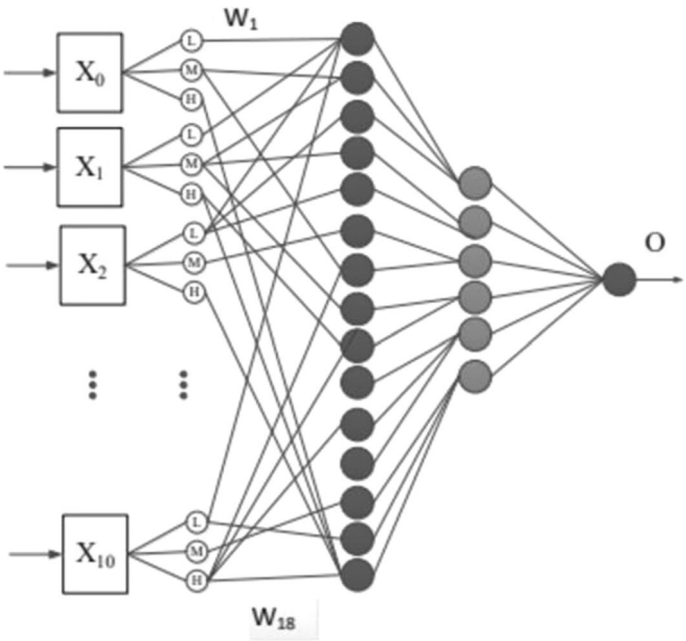


Fig. 4 ANFIS design

Because of the complex structure and the gradient learning behavior the computational cost for the ANFIS system is very high. This comes as a big issue when the application involved have large inputs. The major limitations of the ANFIS system include the problem of dimensionality, the location of a membership function and the number and type of membership functions.

5.2 HHO

HHO is a famous multitude based, angle free improvement calculation with a few dynamic and time-fluctuating periods of investigation. This calculation at first distributed by the lofty Journal of Future Generation Computer Systems (FGCS) in 2019, and right off the bat, it has acquired expanding consideration among scientists because of its adaptable design, elite, and excellent outcomes. The principal rationale of the HHO technique is planned dependent on the helpful conduct and pursuing styles of Harris' birds of prey in nature called "shock jump". Getting away from energy boundary has a unique randomized time-changing nature, which can additionally improve and fit the exploratory and exploitive examples of HHO. This factor likewise upholds HHO to lead a smooth progress between the investigation and misuse. Various LF-based examples with short-length bounces improve the manipulative practices of HHO while coordinating a nearby hunt (Fig. 5).

5.3 Fuzzy GNP

GNP has been supportive of acted like one of the transformative calculations. It was utilized to programmed program age for efficient specialist practices. GNP is addressed by chart structures which comprise of three kind hubs, i.e., start hub, judgment hub and preparing hub. These hubs are associated with one another as coordinated diagram

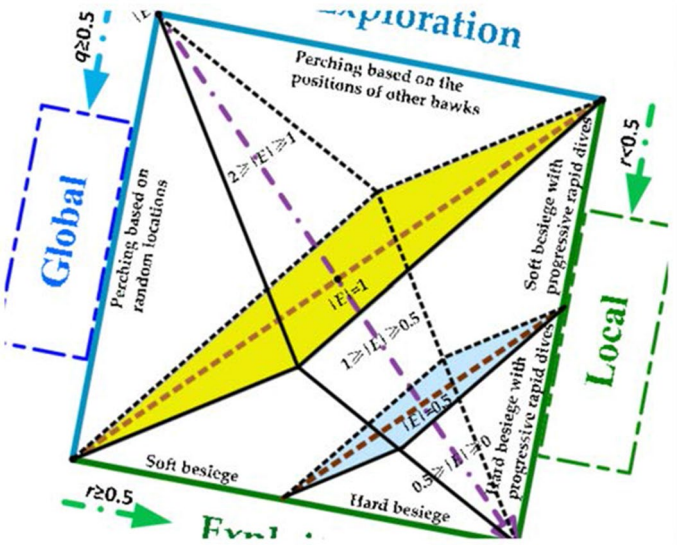


Fig. 5 HHO design

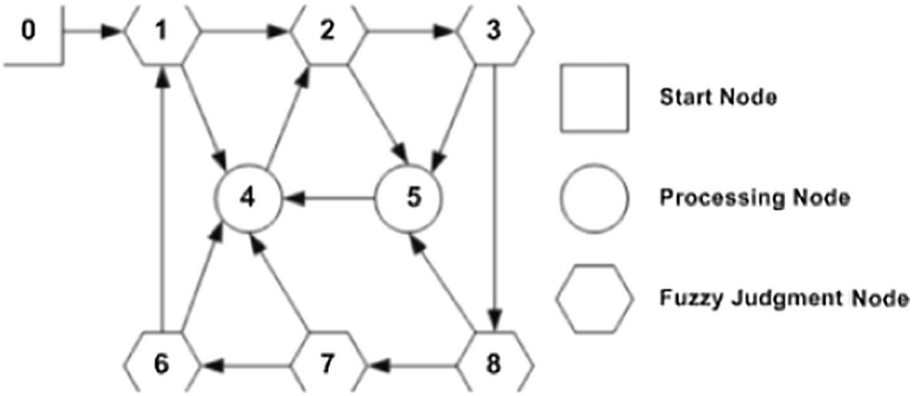


Fig. 6 Fuzzy GNP basic design

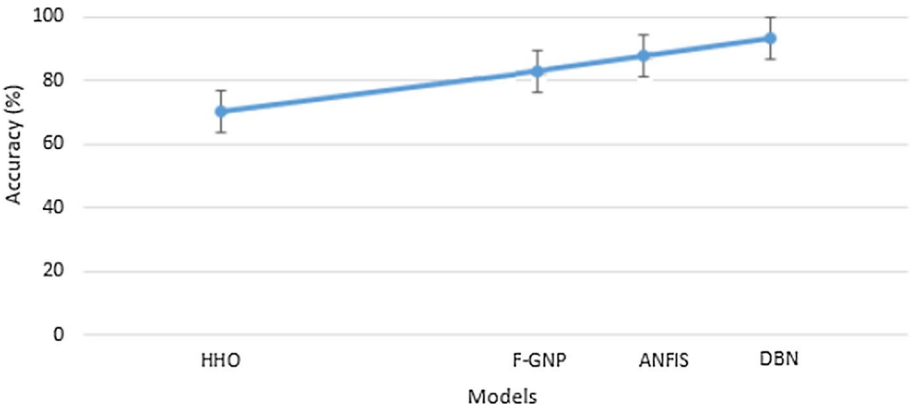


Fig. 7 Comparative analysis of HHO, F-GNP, ANFIS, DBN under accuracy

structures which give more benefits, i.e., reusability of hubs and flexibility to mostly discernible Markov choice issues. GNP has been effectively applied to the issues in unique conditions, for example, lift administrative control frameworks, stock exchanging markets and tile world (Fig. 6).

Table 2 shows the comparative analysis of HHO, F-GNP, ANFIS and our system under the performance measure FAR and detection rate. On examining the obtained values our model shows a false-alarm rate of 4.8 and it also shows a high detection rate of 96.5%. These values indicate that our model can be greatly recommended for detecting attacks. Figure 7 shows the comparative analysis.

Figure 8 depict the setting up of Wamp Server for the systems to be performed. Figure 9 depict the DARPA 1999 dataset importing. Figure 10 depict the operation of DBN, HHO, ANFIS, F-GNP over the server. Figure 11 depict the feature reduction and pre-processing stages of model.

Figure 12 depict the MF of ANFIS, HHO, F-GNP performance. Figure 13 depict the test data implementation of DBN.

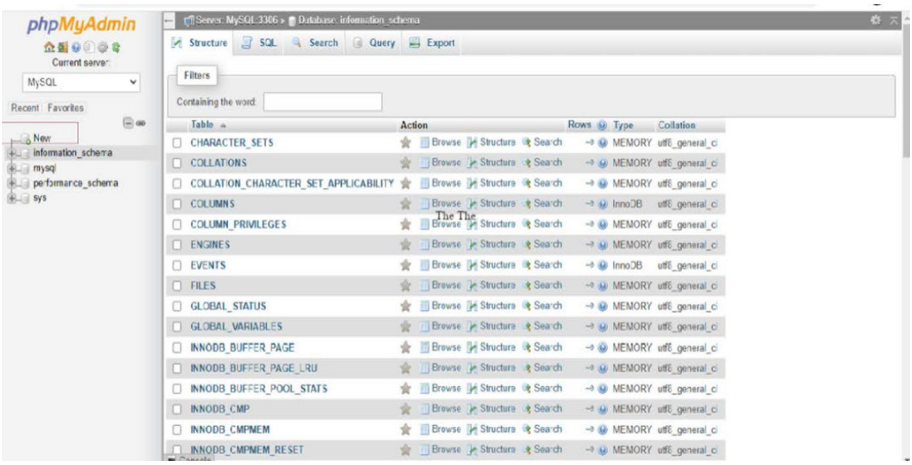


Fig. 8 Setting up of Wamp server

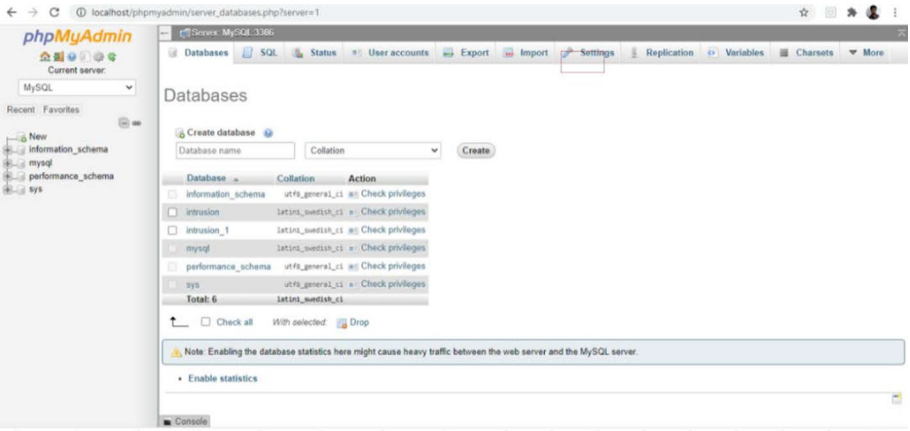
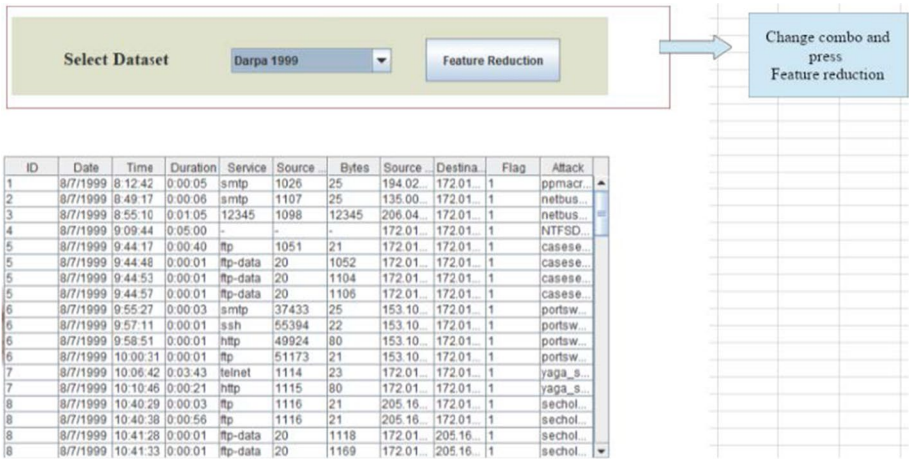
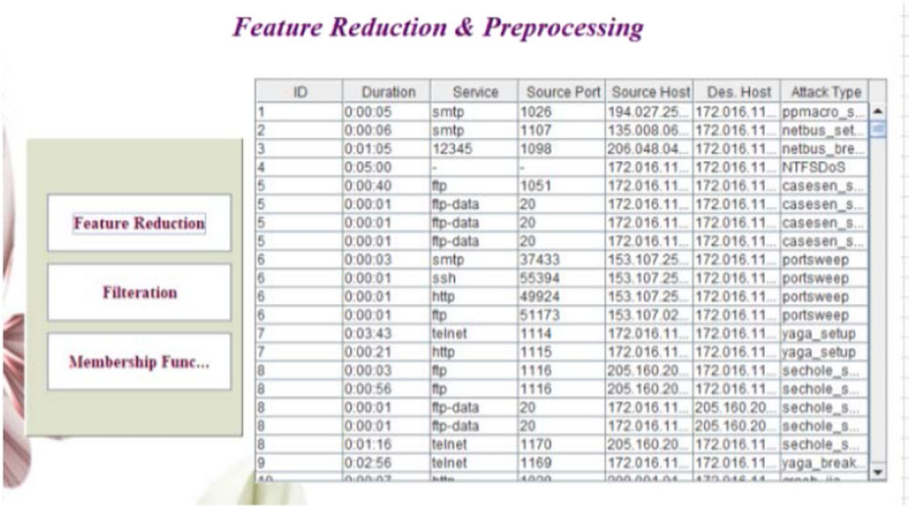


Fig. 9 Snapshot of setting up of database -DARPA 1999



ID	Date	Time	Duration	Service	Source ...	Bytes	Source ...	Destina...	Flag	Attack
1	8/7/1999	8:12:42	0:00:05	smtp	1026	25	194.02...	172.01...	1	ppmacr...
2	8/7/1999	8:49:17	0:00:06	smtp	1107	25	135.00...	172.01...	1	netbus...
3	8/7/1999	8:55:10	0:01:05	12345	1098	12345	206.04...	172.01...	1	netbus...
4	8/7/1999	9:09:44	0:05:00	-	-	-	172.01...	172.01...	1	NTFSD...
5	8/7/1999	9:44:17	0:00:40	ftp	1051	21	172.01...	172.01...	1	casese...
5	8/7/1999	9:44:48	0:00:01	ftp-data	20	1052	172.01...	172.01...	1	casese...
5	8/7/1999	9:44:53	0:00:01	ftp-data	20	1104	172.01...	172.01...	1	casese...
5	8/7/1999	9:44:57	0:00:01	ftp-data	20	1106	172.01...	172.01...	1	casese...
6	8/7/1999	9:55:27	0:00:03	smtp	37433	25	153.10...	172.01...	1	portsw...
6	8/7/1999	9:57:11	0:00:01	ssh	55394	22	153.10...	172.01...	1	portsw...
6	8/7/1999	9:58:51	0:00:01	http	49924	80	153.10...	172.01...	1	portsw...
6	8/7/1999	10:00:31	0:00:01	ftp	51173	21	153.10...	172.01...	1	portsw...
7	8/7/1999	10:06:42	0:03:43	telnet	1114	23	172.01...	172.01...	1	yaga_s...
7	8/7/1999	10:10:46	0:00:21	http	1115	80	172.01...	172.01...	1	yaga_s...
8	8/7/1999	10:40:29	0:00:03	ftp	1116	21	205.16...	172.01...	1	sechol...
8	8/7/1999	10:40:38	0:00:56	ftp	1116	21	205.16...	172.01...	1	sechol...
8	8/7/1999	10:41:28	0:00:01	ftp-data	20	1118	172.01...	205.16...	1	sechol...
8	8/7/1999	10:41:33	0:00:01	ftp-data	20	1169	172.01...	205.16...	1	sechol...

Fig. 10 Operation of DBN, ANFIS, HHO, F-GNP over DARPA 1999



ID	Duration	Service	Source Port	Source Host	Des. Host	Attack Type
1	0:00:05	smtp	1026	194.027.25...	172.016.11...	ppmacro_s...
2	0:00:06	smtp	1107	135.008.06...	172.016.11...	netbus_set...
3	0:01:05	12345	1098	206.048.04...	172.016.11...	netbus_bre...
4	0:05:00	-	-	172.016.11...	172.016.11...	NTFSDoS
5	0:00:40	ftp	1051	172.016.11...	172.016.11...	casesen_s...
5	0:00:01	ftp-data	20	172.016.11...	172.016.11...	casesen_s...
5	0:00:01	ftp-data	20	172.016.11...	172.016.11...	casesen_s...
5	0:00:01	ftp-data	20	172.016.11...	172.016.11...	casesen_s...
6	0:00:03	smtp	37433	153.107.25...	172.016.11...	portsweep
6	0:00:01	ssh	55394	153.107.25...	172.016.11...	portsweep
6	0:00:01	http	49924	153.107.25...	172.016.11...	portsweep
6	0:00:01	ftp	51173	153.107.02...	172.016.11...	portsweep
7	0:03:43	telnet	1114	172.016.11...	172.016.11...	yaga_setup
7	0:00:21	http	1115	172.016.11...	172.016.11...	yaga_setup
8	0:00:03	ftp	1116	205.160.20...	172.016.11...	sechole_s...
8	0:00:56	ftp	1116	205.160.20...	172.016.11...	sechole_s...
8	0:00:01	ftp-data	20	172.016.11...	205.160.20...	sechole_s...
8	0:00:01	ftp-data	20	172.016.11...	205.160.20...	sechole_s...
8	0:01:16	telnet	1170	205.160.20...	172.016.11...	sechole_s...
9	0:02:56	telnet	1169	172.016.11...	172.016.11...	yaga_break...

Fig. 11 Feature reduction & preprocessing stage of this model

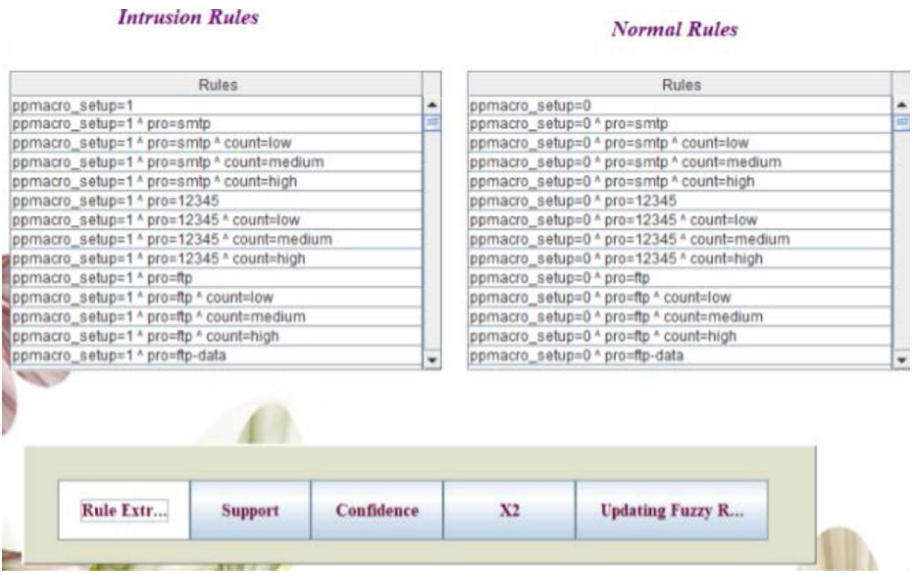


Fig. 12 MF performance of ANFIS, HHO, F-GNP

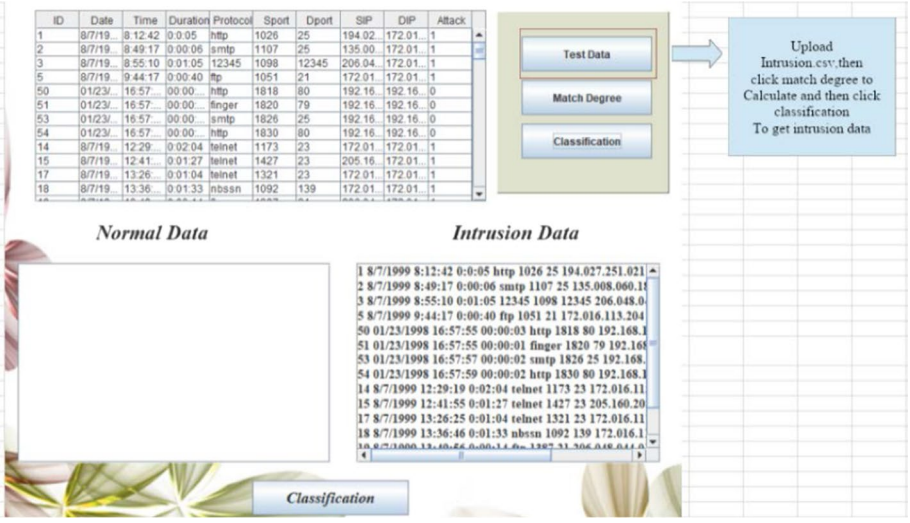


Fig. 13 Testing implementation of DBN model

6 Conclusion

With the revolutionized impact of DL, these get so much better over securing data integrity, confidentiality and availability from cyber-attacks. In this paper, we have seen an efficient DL-based IDS for efficient classify multi-class intrusion with greater accuracy. Here initially data are collected from DARPA 1999 with 190 instances and are normalized so that all types are converted over numeric version. Then these are passed over feature selection

where irrespective of other basic techniques, we used PSO for even more boosting to DBN. Then these are passed over for training and testing for effective classification into 5 categories of intrusion such as Normal, Probe, DoS, R2L, U2R. Also, these are evaluated with various other system such as ANFIS, HHO, F-GNP over Accuracy and FAR in which DBN out performs every other traditional system with 96.5%. In future many other advanced optimization techniques can be incorporated for obtaining more accurate results.

Funding The authors declared that no Funding was received for this work.

Data Availability The authors declare that no data or material was taken illegally. However, publically available benchmark datasets were taken for implementation.

Code Availability The authors declare that no exact code has been copied to carry out the research.

Declarations

Conflict of interest The authors declare that they have no conflicts of interest.

References

1. Wei, P., Li, Y., Zhang, Z., Hu, T., Li, Z., & Liu, D. (2019). An optimization method for intrusion detection classification model based on deep belief network. *IEEE Access, 7*, 87593–87605.
2. Prasad, R., & Rohokale, V. (2020). Artificial intelligence and machine learning in cyber security.". In R. Prasad & V. Rohokale (Eds.), *Cyber Security: The Lifeline of Information and Communication Technology* (pp. 231–247). Cham: Springer.
3. Gao, X., Shan, C., Hu, C., Niu, Z., & Liu, Z. (2019). An adaptive ensemble machine learning model for intrusion detection. *IEEE Access, 7*, 82512–82521.
4. Huang, X. (2021). Network intrusion detection based on an improved long-short-term memory model in combination with multiple spatiotemporal structures. *Wireless Communications and Mobile Computing, 2021*, 1.
5. Xiao, Y., Xing, C., Zhang, T., & Zhao, Z. (2019). An intrusion detection model based on feature reduction and convolutional neural networks. *IEEE Access, 7*, 42210–42219.
6. Vinayakumar, R., Alazab, M., Soman, K. P., Poornachandran, P., Al-Nemrat, A., & Venkatraman, S. (2019). Deep learning approach for intelligent intrusion detection system. *IEEE Access, 7*, 41525–41550.
7. Yang, Y., Zheng, K., Wu, B., Yang, Y., & Wang, X. (2020). Network intrusion detection based on supervised adversarial variational auto-encoder with regularization. *IEEE Access, 8*, 42169–42184.
8. Khan, F. A., Gumaei, A., Derhab, A., & Hussain, A. (2019). A novel two-stage deep learning model for efficient network intrusion detection. *IEEE Access, 7*, 30373–30385.
9. Naseer, S., Saleem, Y., Khalid, S., Bashir, M. K., Han, J., Iqbal, M. M., & Han, K. (2018). Enhanced network anomaly detection based on deep neural networks. *IEEE access, 6*, 48231–48246.
10. Papamartzivanos, D., Mármol, F. G., & Kambourakis, G. (2019). Introducing deep learning self-adaptive misuse network intrusion detection systems. *IEEE Access, 7*, 13546–13560.
11. Yin, C., Zhu, Y., Fei, J., & He, X. (2017). A deep learning approach for intrusion detection using recurrent neural networks. *IEEE Access, 5*, 21954–21961.
12. 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.
13. Xu, C., Shen, J., Du, X., & Zhang, F. (2018). An intrusion detection system using a deep neural network with gated recurrent units. *IEEE Access, 6*, 48697–48707.
14. Al-Qatf, M., Lasheng, Y., Al-Habib, M., & Al-Sabahi, K. (2018). Deep learning approach combining sparse autoencoder with SVM for network intrusion detection. *IEEE Access, 6*, 52843–52856.
15. Ali, M. H., Al Mohammed, B. A. D., Ismail, A., & Zolkipli, M. F. (2018). A new intrusion detection system based on fast learning network and particle swarm optimization. *IEEE Access, 6*, 20255–20261.

16. Yao, H., Fu, D., Zhang, P., Li, M., & Liu, Y. (2018). MSML: A novel multilevel semi-supervised machine learning framework for intrusion detection system. *IEEE Internet of Things Journal, 6*(2), 1949–1959.
17. Sahani, R., Rout, C., Badajena, J. C., Jena, A. K., & Das, H. (2018). Classification of intrusion detection using data mining techniques. In R. Sahani & C. Rout (Eds.), *Progress in computing, analytics and networking* (pp. 753–764). Springer.
18. Li, J., Qu, Y., Chao, F., Shum, H. P., Ho, E. S., & Yang, L. (2019). Machine learning algorithms for network intrusion detection. In L. F. Sikos (Ed.), *AI in cybersecurity* (pp. 151–179). Springer.
19. Parampottupadam, S., & Moldovann, A.-N. (2018) Cloud-based real-time network intrusion detection using deep learning. In *2018 International Conference on Cyber Security and Protection of Digital Services (Cyber Security)* (pp. 1–8). IEEE.
20. Ramprakash, P., Sakthivadivel, M., Krishnaraj, N., & Ramprasath, J. (2014). Host-based intrusion detection system using sequence of system calls. *International Journal of Engineering and Management Research, 4*(2), 241.
21. Kim, J., Kim, J., Kim, H., Shim, M., & Choi, E. (2020). CNN-based network intrusion detection against denial-of-service attacks. *Electronics, 9*(6), 916.
22. Karatas, G., Demir, O., & Sahingoz, O. K. (2020). Increasing the performance of machine learning-based IDSs on an imbalanced and up-to-date dataset. *IEEE Access, 8*, 32150–32162.
23. Tan, X., Su, S., Zuo, Z., Guo, X., & Sun, X. (2019). Intrusion detection of UAVs based on the deep belief network optimized by PSO. *Sensors, 19*(24), 5529.
24. Bhuyan, H. M., Bhattacharyya, D. K., & Kalita, J. K. (2015). Towards generating real-life datasets for network intrusion detection. *International Journal of Network Security, 17*(6), 683–701.
25. Kunhare, N., Tiwari, R., & Dhar, J. (2020). Particle swarm optimization and feature selection for intrusion detection system. *Sādhanā, 45*(1), 1–14.
26. Liu, J., Yang, D., Lian, M., & Li, M. (2021). Research on intrusion detection based on particle swarm optimization in IoT. *IEEE Access, 9*, 38254–38268.

P. J. Sajith received B. Tech Degree in Information Technology fromGovernment Engineering College Palakkad, Calicut University and M. E inComputer and Communication Engineering from PSNA College of Engineeringand Technology, Anna University, Tiruchirappalli in 2005 and 2011respectively. He is a faculty member of Department of Computer Science andEngineering, Toc H Institute of Science and Technology, Ernakulam, India. Hiscurrent research interests include Wireless Networks, Artificial Intelligence andMachine Learning.

G. Nagarajan received the B.E. degree in Electrical and Electronics Engineering from MS University and the M.E. degree in Applied Electronics from Anna University, in 2000 and 2005, respectively, and the M.E. degree in Computer Science and Engineering from Sathyabama University and the Ph.D. degree in Computer Science and Engineering from Sathyabama University, in 2007 and 2015. He is a Faculty Member of the Department of Computer Science and Engineering, School of Computing, Sathyabama Institute of Science and Technology, Chennai, India. His current research interests include Computer Vision, IoT, 5G, Edge/Fog Computing, Artificial Intelligence, Machine Learning, and Wireless Sensor Network. He has published more than 70 research papers in peer-reviewed journals such as IEEE, ACM, Springer-Verlag, Inderscience, and Elsevier. He also has contributed 15 book chapters thus far for various technology books. Finally, he has authored and edited 3 books thus far and is focusing on some of the emerging technologies such as the IoT, Edge/Fog Computing, Artificial Intelligence (AI), Data Science, Blockchain, Digital Twin, 5G, etc.