INTRUSION DETECTION SYSTEM USING PARTICLE SWARM OPTIMIZATION AND BACK PROPAGATION NEURAL NETWORK

[1] Amit Kumar, [2] Basant Kumar , [3] Devendra Kumar Singh
[1] Student , [2] Student , [3]Assistant Professor,
Dept. of Computer Science & Engineering , SoS in Engg & Tech,
Guru Ghasidas Vishwavidyalaya ,Bilaspur(C.G) ,India

***Abstract* :** Intrusion detection system research field has grown tremendously in the past decade. Many people have proposed various models to improve performance of attack classification. Intrusion detection system play a vital role in network security, its check the system for integrity, confidentiality and availability in an individual manner like anomaly intrusion detection and misuse intrusion detection. The main objective of this paper is to develop a hybrid model for intrusion detection system so that achieve the best detection accuracy performance. We have given the idea of classification based on optimization and feature selection using Particle Swarm Optimization (PSO) for discriminating features selection and Backpropagation neural network (BPNN) for classification of attack on Network. This BPNN-PSO Model has experimented on KDD-CUP'99 dataset to achieve maximum detection rate with minimum false alarm rate.

Keywords: Intrusion Detection System, KDDCUP'99 Dataset, PSO, BPNN , Detection rate

I. INTRODUCTION

The Network is vulnerable to attacks and threat so need a security mechanism that can protect our computer system from such kind of problems. Intrusion Detection System is system software which is monitor computer network to detect misuse or anomalous behavior. It does not act or take any action when an intrusion detected only report the system administrator about the intrusion and administrator takes an action on the intrusions, Generally IDS[1] check the system for integrity, confidentiality and availability in an individual manner like Anomaly detection looks for abnormal or unauthorized activities in network and misuse detection tries to match the attack data with previous record data or known attack pattern but sometimes it may failed to detect a new kind of attack form so mostly having to focused on anomaly detection approaches. Anomaly-based approaches have based on Neural network, Machine learning and others techniques[1].In this paper we have developed hybrid model by using KDDCUP'99[6] dataset to categorized network attack into five different types Normal, Denial of service (DoS), Probing attacks (Probe), Remote to local attacks (R2L), User to root attacks (U2R). Further Various approaches have been used for intrusion detection but we have given the idea based on Artificial neural network technique like Backpropagation neural network (BPNN) and Particle swarm optimization (PSO) following some advantage like ability to detect known and unknown attack basically BPNN uses a supervised learning approaches for training due this it has maximum detection rate with minimum false alarm rate as compared to other approaches [1]. So the main goal of this paper to develop a model using BPNN-PSO algorithms for detection and classification of the attack.

II. INTRUSION DETECTION SYSTEM

An intrusion detection system is basically hardware or software that control and prevent the network from malicious activity [2]. The range of IDS varies from a single host machine to a combination of the computer which forms a network [2]. The IDS is considered to be the firewall security of the network [29].

2.1 CLASSIFICATIONS OF IDS

2.1.1 NETWORK INTRUSION DETECTION SYSTEM

NIDS is usually used to detect abnormal behavior of traffic within the computer system or subnet of the network [3].It detects the malicious activity by comparing abnormal behaviors with in network from the database which stored known attack. If any matches found then it will alert to the administrator and it tries to detect malicious activity by monitoring the network traffic as well as by reading the inbound packet of network and capable of monitoring the entire network [4].NIDS are deployed on a fixed point in the boundary between the network, wireless sensor network, remote access server [5]. The selection of fixed point is based on traffic of the network [5].

2.1.2 HOST BASED INTRUSION DETECTION SYSTEM

HBIDS runs on an individual host or individual device by checking the inbound and outbound packet it provides information to the administrator if any malicious event occurs [3].It takes the snapshot or data of the existing file system and matches with previous file system if any Modification found in critical file system then it will inform to administrator [7].There is no need of extra hardware to implement HIDS [7]. But this technique such as long analysis and offline integrity checking takes an undesirable delay where NIDS does not take that much amount of delay so that NIDS is usually detecting malicious activity in real time. HBIDS was the first IDS which was designed in which there is no interaction between the outside of the system [30].

2.2 ATTACK TYPES

2.2.1 Probe Attack: An external source try to damage the target network. Some basic network connection level features like the "source bytes" "duration of connection" are crucial part for detecting probes while some features like "number of files accessed" doesn't share the information for detecting probe attacks[9].

2.2.2 Denial of Services (DoS) Attack: DoS is cyber-attack which is caused due to flooding of targeted machine and resource with superfluous request in an attempt to overload system and prevent some request for fulfilled [10].There are mainly two types of denial of service attack (a) flooding (b) flaw exploitation [31].

2.2.3 Remote to Local (R2L) Attack: R2L consists of two level first one is network level and second one is host level [11]. It became complex to detect attack due to presence of two level. Network level features consist of "duration of connection " and "service requested" whereas host level feature consist of "number of failed login attempt" [10].

2.2.4 User to Root (U2R) Attack: In this type of attack a hacker get access of system with normal user account and start doing some malicious work on the system in order to gain super user privilege [11].

2.3 DETECTION METHOD

There are mainly two type of detection approach [3].

2.3.1 Signature Based: In this approach the attack is recognized by comparing the specific pattern of attack and compare with known attack. It is similar like antiviruses which is used in various system. It takes less time to identify the attack and more accurate as compare to others [8].

2.3.2 Anomaly Based: It is different than signature based detection approach because it is used to detect unknown attack found in network with the help of basic approach which is used in machine learning for creating trustworthy activity [3].The main advantage of this approach to detect unknown attack with low maintenance cost [8] and the longer the system run the accuracy level will be more.

III. PREVIOUS WORK

Here, Some Previous paper works and result described related BPNN approaches [1].

In 2009, Ning-Qing Sun et.al [15]. Emphasized upon developing a Detection System by using Back Propagation neural network. First, apply feature optimization operation upon the data set and then applying the BPN algorithm on the optimized data set to yield accurate and precise attack classification.

In 2011, Changjun Han et.al [13] Authors have trained data using Back propagation neural network model for attack types classification. Used 1325 connections for training and 1245 for testing and they obtained results like the detection rate approx. 80.5%, with false alarm rate7.4%.

In 2011, Sufyan T. Faraj et.al [14]. They have trained dataset by using BPNN algorithm to classify and detect abnormal activities in network and in case any abnormal events are detected then classified into attack categories. The model having attack detection rate about 90% for test dataset and approximately 60-85% attack type classification.

In 2011, Mukhopadhyay I et.al [17]. Authors used Neural Network approaches like BPNN algorithm for detection and classification of attacks type (DoS, U2R, Prob, R2L). Intrusion detection system model gets success rate 73.9% for test dataset.

In 2012,Vladimir Bukhtoyarovf et.al [16]. Apply neural network approach and mainly their work was focused on classifying probe, DoS, R2L, U2R attacks using joint usage of the trained neural network. Found that 99.87% detection rate for probe attacks.

In 2013,V. Jaiganesh et.al [12]. Analyzed and classified the attacks into four classes DoS, U2R, R2L,Probe by using the BPNN techniques. Basically they have worked on intrusion detection rates for the attack. DoS attack the detection rate is accuracy approx.78.15%.

Many researchers have proposed model by using BPNN algorithm to classified events into attack type and got better detection rate as result also found that Neural network is efficient technique to build IDS.

Table 1: Previous work based on IDS

S. No	Title of paper	Year	Publication	Result (%)	Dataset	Methodology
1	A new intrusion detection approach using PSO based multiple criteria Linear programming [34]	2015	Procedia computer science	Detection rate 97.4 False alarm rate 3.63	KDD CUP 99	MCLP(multi criteria learning programming)
2	Network intrusion detection using PSO based on Adaptive and genetic Algorithm [35]	2014	IJSER	Detection rate 89.94	KDD CUP 99	Genetic Algorithm
3	A Novel anomaly intrusion detection based on SMO [36]	2017	JMM	Detection rate 95.17	KDD CUP 99	K-means ,SVM, ELM
4	An implementation of intrusion detection system using genetic algorithm [37]	2012	IJNSA	Detection rate of Probe 71.1 Dos 99.4 U2R 18.9	KDD CUP 99	Genetic Algorithm
5	A review on intelligent data mining and soft computing technique [38]	2017	IJSRCSEIT	Detection rate 99.16	NSL-KDD	SVM
6	Machine learning approach for intrusion detection on cloud virtual machine [39]	2013	IJAEIM	Detection rate 98 False rate 10	KDD CUP 99	SVM

IV. METHODOLOGY

In methodology we have discussed about the procedure and method which is used in proposed model .We have used KDD CUP'99 Dataset [6] ,PSO ,BPNN for intrusion detection and attacks classification.

4.1 KDD CUP'99 DATASET

In 1998, KDD CUP'99 Dataset was introduced at Lincoln laboratory in M.I.T[17]. It's a compromise of Training Dataset and Test Dataset and widely used for anomaly detection in IDS. Approximately KDDCUP'99 consists of seven weeks of network traffic in the form of binary data having size of compressed 4 Gigabytes. Which can be further processed into about 5 million network connection records and the size of each record is 100 bytes[17]. The KDD CUP'99 training dataset consists of approx. 4,900,000 connection vector and each one consists 41 features and every connection classified into normal or attacks using classification technique so that four types of attack can be identified .While the KDD CUP'99 testing dataset comprises of 2 weeks network traffic approximately 2 million connection records [18]. The dataset consists of 41 features are classified into 3 groups: Basic features, Traffic features, Content features and below table show the feature description [6].

Table 2: Features description of dataset

S.no	Features name	Description
1	duration	Connection time
2	Protocol type	Protocol class
3	Service	Destination Network service
4	flag	Connection status either normal or error
5	Src_bytes	No.of data bytes from (src) to (dest)
6	dst_bytes	No.of data bytes from (src) to (dest)
7	land	0 if conn. not from same source(src) otherwise 1
8	Wrong fragmnet	counting of wrong fragments
9	urgent	counting of urgent packets
10	Hot	No. of “hot” indicator
11	Num_failed_login	counting of failed login
12	Logged_ in	Login successfully then 1 otherwise 0
13	Num_compromised	No.of compromised condition
14	Root_shell	0 in case of root shell is not obtained otherwise 1.
15	Su_attempted	0 in case “su root” command didn’t attepted otherwise 1.
16	Num_root	root access count
17	Num_file_creation	No. of file creation operations
18	Num_shell	shell prompts count
19	Num_access_files	No. of operation on access control files
20	Num_outbounds_cmds	counting of outbound commands in ftp conn.
21	Is_hot_login	0 in case login doesn’t belong to“hot” otherwise 1
22	Is_guest_login	0 in case login doesn’t belong to“guest” else 1
23	Count	same hst connection num
24	Srv_count	Counting of conn. to the same service
25	Serror_rate	“SYN “error % of conn.
26	Srv_serror_rate	“SYN” error % of conn.
27	Rerror_rate	“REJ” error % of conn.
28	srv_rerror_rate	``REJ" error % of conn.
29	same_srv_rate	% of conn.to the same srv
30	diff_srv_rate	% of conn. to different srv
31	srv_diff_hst_rate	% of conn. to different hst
32	dst_hst_count	count for dst hst
33	dst_hst_srv_count	srv_count for dst hst
34	dst_hst_same_srv_rate	same_srv_rate for dst hst
35	dst_hst_diff_srv_rate	diff_srv_rate for dst hst
36	dst_hst_same_src_port_rate	same_src_port_rate for dst hst
37	dst_hst_srv_diff_hst_rate	same_src_port_rate for dst hst
38	dst_hst_serror_rate	serror_rate for dst hst
39	dst_hst_srv_serror_rate	srv_serror_rate for dst hst
40	dst_hst_rerror_rate	rerror_rate for dst hst
41	dst_hst_srv_rerror_rate	srv_serror_rate for dst hst

4.2 PARTICLE SWARM OPTIMIZATION

Particle swarm optimization is a heuristic global optimization technique or method. In 1995, put forward originally by Dr Russell Eberhart and James Kennedy [19]. It was an optimization method which is based on social behavior as a stylized representation of the movement of an organism in a bird flock or fish schooling [19]. The main goal of the algorithm to simulates the behavior of bird flock flying together in space(multidimensional) for search a optimum place [20], and adjusting their movements and distance for the better search (Pbest) means the particles are initialized randomly in start and then search for optimal by updating the velocity and position. So it’s basically a part of swarm intelligence [23].Overall PSO is a combination of local search method with global search method and local search means searching through self-experience and global search means searching through neighbors experience [24]. PSO is generally used in data mining [25], image processing [26].

The PSO algorithm mimics a particle flying in the search space (multidimensional) and moving towards the global optimum (Gbest) place [20]. So following equations are used to evaluate or update the velocity and position of particle .particle swarm optimization is a metaheuristic approach to find the solution of the problem[32].In another word, PSO is used to solve the optimal power flow problem [33].

$$Vi[t+1] = Vi[t] + c1*r1\,[Pbest(t) - p(t)] + c2*r2\,[Gbest(t) - g(t)] \qquad (1)$$

$$Pos(i)[t+1] = Pos(i)[t] + Vi[t+1] \qquad (2)$$

Where, Vi is the velocity of particle, Pbest is the personal best ,Gbest is the global best, Pos(i0 is position of particle, particle index(i)=1,2,…D and acceleration coff.($0 < c1,c2 < 2$) ,random value ($0 < r1,r2 < 1$).

The working pseudo code of the particle swarm optimization algorithm is as follows [20].

```
For each particle
Initialize the particle randomly
END
While (particle size)
{
Loop ()
Calculate the fitness value of each particle
If the fitness value is better from Pbest in history then update the new value with the Pbest
END Loop
Select the best fitness value particle from the others as Gbest
While
Min. error or Max. Iterations are not attained
{For each particle
Calculate particle velocity by from equation (1)
Update the particle position by from equation (2)
Next
}
}
```

4.3 BACK PROPAGATION NEURAL NETWORK

Back-propagation neural network algorithm is a supervised learning technique for neural networks, in 1974 which is first proposed by Paul Werbos in his PhD thesis [21]. However, it was rediscovered experimentally in 1986 by Rumelhart, Hinton and Williams [21] and after that BPNN became extensively used. The main objective of the BPNN algorithm computes the error in connection network by using the error evaluation function as well as modify or update weights of the connection network to minimize the error up to the difference between actual network output and the desired output [21]. Where, the working principle of the BPNN algorithm would be broken down into four major steps After randomly chosen the weights of the connection network, so the four steps as follow [22] a) Feed-forward computation b) Backpropagation to the output layer c) Backpropagation to the hidden layer d) Weight updates and stopped when the value of the error in connection network has become desired minimum [22]. Where, in figure represents the internal architecture of connection network having three layers as follow input layer, hidden layer and output layer [21].

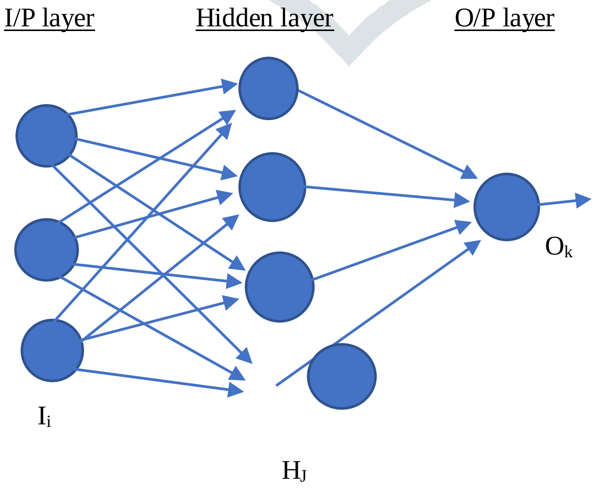

Figure 1: Architecture of BPNN layers

Eq.(3),it shows the expression of the error at the ith layer. However, for normalization purposes, the error evaluation function expression in Eq. (4) is used in the derivations of algorithm [21].

$$e_i = (t_i - o_i) \qquad (3)$$

$$E = 0.5 * \sum e_i^2 = 0.5 * \sum (t_i - o_i)^2 \qquad (4)$$

Where, i denotes the layer index, ti is the desired output, and oj is the actual network output. It is special case of automatic differentiation [27].The minimum of the error function along with weight is consider to be solution of learning process [28].The activation function for the back propagation network is sigmoid function which can be given as s(x)=1/(1+e^-cx) where c is constant [28].

V. CONCLUSION AND FUTURE WORK

In this paper, we have proposed the model (PSO-BPPN) for Intrusion detection system after done feasible studies and the related research task is to detect, classify abnormal activities or events into attack type by using Particle swarm optimization (PSO) and Backpropagation neural network (BPNN) algorithm and evaluate the detection rate for attacks classification. In recent time many researchers have proposed model by using Neural network approaches to classified events into their attack type and got better detection rate as result also found that Neural network is efficient technique to build IDS. Next, we are going to apply our proposed model on KDD CUP'99 dataset which shows that result with the highest detection rate and the lowest false alarm rate.

REFERENCES

[1] C.vaidya ,Dr.M. Raghuwanshi ,Roshani gaidhane "Intrusion detection and attack classification using back-propagation neural network "(IJERT) ISSN: 2278- 0181 vol .3 –no-2014

[2] Dr. S.Vijayarani and Ms. Maria Sylviaa.S " intrusion detection system– A study" International Journal of Security, Privacy and Trust Management (IJSPTM) Vol 4 -no- 2015

[3] https://en.wikipedia.org/wiki/Intrusion_detection_system

[4] www.techpedia.com/definition/network based intrusion detection system.

[5] Hussain Ahmad madni uppal,memoona javed and MJ.Arshad."An overview of intrusion detection system" (IJST) vol 5 -no-2014

[6] https://kdd.ics.uci.edu/databases/kddcup99/task.html

[7] Kopelo letou, dhruwajita devi and Y. jayant singh "host based intrusion detection system and prevention system". international journal of computer application (0975-8887) volume 69-no 2013

[8] www.IUP-org/Intrusion detection system.

[9] Sapna S Kaushik, Dr. prof P.R desmukh "Detection of attack in intrusion detection system" (IJCSIT)vol2(3),2011-982-986.ISSN-0975-9646

[10] https://en.wikipedia.org/wiki/Denial-of-service_attack

[11] Swati paliwal, Ravindra gupta "Denial-of service, Remote to user, Probing detection by using genetic algorithm" IJCA(0975-8887) vol 60-no- 2017

[12] Sumathi P. Mangayarkarasi S and Jaiganesh V, "An Analysis of Intrusion Detection System using Back Propagation Neural Network",IEEE publication DOI: 10.1109/ICICES.2013.6508202 2013

[13] Yi Lv, Yang D., Hao Y.and Han C., "An Intrusion Detection System Based on Neural Network", International Conference on Mechatronic Science, Electric Engineering and Computer IEEE Publication,-2011

[14] Al-Janabi,Saeed H and Faraj S,, "A Neural Network Based Anomaly Intrusion Detection System", Developments in E-systems Engineering, DOI 10.1109/DeSE. IEEE publication 2011

[15] Yang LI, FGIT and Ning-Qing Sun "Intrusion Detection Based on Back-propagation Neural Network and Feature Selection Mechanism" LCNS, vol 5899-no- 2009

[16] Semenkin E.and Bukhtoyarovf V. "Neural Networks Ensemble Approach for Detecting Attacks in Computer Networks," IEEE World Congress on Computational Intelligence, June, 10-15, 2012

[17] Mukhopadhyay I, , Chakrabarti S and Chakraborty M, Chatterjee T, " Back-Propagation Neural Network Approach to Intrusion Detection System", International Conference on Recent Trends in Information Systems, IEEE publication 2011

[18] Mehbod Tavallaee, Ali.A.Ghorbani. Ebrahim.Bagheri, and Wei Lu "A detailed analysis of the KDDcup'99 dataset" 978-1-4244-3764-1/09/©2009 IEEE

[19] James kennedy and Russel Eberthart " Particle swarm optimization "0-7803-2768/IEEE 1995

[20] Muhammad Imran ,Rathiah Hashim and Noor Elaiza Abd Khalid "An overview of particle swarm Optimization Variants" Procedia Engineering 53 491 – 496 ,2013

[21] Alaeldin Suliman and Yun Zhang "A Review on Back-Propagation Neural Networks in the Application of Remote Sensing Image Classification" Journal of Earth Science and Engineering 5 (2015) 52-65 doi: 10.17265/2159-581X/2015. 01. 004

[22] Mirza.Cilimkovic"http://www.dataminingmasters.com/uploads/studentProjects/NeuralNetworks.pdf" Institute of Technology Blanchardstown

[23] J.kennedy,R.C,Eberthart,Y.shi,"Swarm intelligence,"morgan kaufmann publisher inc,CA San fransisco,CA,2001.

[24] Qing Zhu,Limin Qian,Yingchun Li,Shanjun Zhu,"An improved particle swarm optimization Algorithm for vechicle routing problem with time windows in:processing" IEEE congress on Evolutionary computation(CEC 2006),pp.1386-1390,2006.

[25] M.Ghannad-Rezai,H. soltanian-zadehand,M-R.Siadat, and K.V.Elisvich,"Medical data mining using swarm optimizationfor temporal lobe epilepsy" IEEE Congress on evolutionary computation CEC 2006,pages 761-768,2006.

[26] M .Omran, A.P. Engelbrecht and A.salman "Image classification using particle swarm optimization," proceeding of the fourth asia pacific conference on simulated Evolution and learning (SEAL,2002),pp. 370-374,2002.

[27] http://en.Wikipedia.org/wiki/Backpropagation

[28] Raul Rojas "The backpropagation algorithm of neural network"-A systematic introduction (ISBN 978-3540605058)

[29] Mohit Tiwari, Raj kumar ,Akash bharti, jai kishan "Intrusion detection system" international journal of technical research and application e ISSN:2320-8163, volume 5,issue2-2017),PP-38-44.

[30] Deber,Herve,dacier Marc Wespi,Andreas, "towards of taxonomy of IDS". IBM research division ,Zurich research laboratory saumarstress 1999

[31] Amrita Anand, Brajesh patel ,"An overview on IDS and types of Attack it can detect considering protocol". IJARCSSE volume 2,issue 8, 2012.

[32] https://en.wikipedia.org/wiki/Particle_swarm_optimization

[33] M.A Abido,"optimal power flow using particle swarm optimization using particle swarm optimization" electrical power and energy system,24 ,563-571,2002.

[34] Behnam amiri,seyed mojtaba Hosseini bamakan Mahboubeh mirzabaghari, yong shi, "A new intrusion detection approach using PSO based multiple criteria Linear programming".Procedia computer science (2015)

[35] .Bharat Rathi,dattaray v.jadhav, "Network intrusion detection using PSO based on Adaptive and genetic Algorithm" IJSER volume 5,issue 8,2014.

[36] Mehdi maukhafi ,Khalid yel hasini, seddik bri , " A Novel Anomaly intrusion detection based on SMO" .journal of mobile multimedia vol 13,no 3 &4,(2017)197-209

[37] Mohammad Sazzadul Hoque, Md.abdul Mukit,Md abu naser bikas. "An implementation of intrusion detection system using genetic algorithm" (IJNSA) Vol 4,No.2,2012.

[38] Suma S G,Dr Genapathy sannasi "A review on intelligent data mining and soft computing technique for effective intrusion detection." International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2017 IJSRCSEIT | Volume 2 | Issue 6 | ISSN : 2456-3307.

[39] Amjad Hussain Bhat , Sabyasachi Patra , Dr. Debasish Jena." Machine learning approach for intrusion detection on cloud virtual machine" international journal of application and innovation in engineering and management(IJAIEM).2013