International Journal of Emerging Research in Management &Technology ISSN: 2278-9359 (Volume-6, Issue-9)

Research Article | September 2017

Optimizing Error Rate in Intrusion Detection System Using Artificial Neural Network Algorithm

S. Vijaya Rani
Assistant Professor, MCA Department, Brindavan College, Dwaraka Nagar, Bangalore,
Karnataka, India

Dr. G. N. K. Suresh Babu
Associate Professor, MCA Department, Acharya Institute of Technology, Bangalore,
Karnataka, India

Abstract:

The illegal hackers penetrate the servers and networks of corporate and financial institutions to gain money and extract vital information. The hacking varies from one computing system to many system. They gain access by sending malicious packets in the network through virus, worms, Trojan horses etc. The hackers scan a network through various tools and collect information of network and host. Hence it is very much essential to detect the attacks as they enter into a network. The methods available for intrusion detection are Naive Bayes, Decision tree, Support Vector Machine, K-Nearest Neighbor, Artificial Neural Networks. A neural network consists of processing units in complex manner and able to store information and make it functional for use. It acts like human brain and takes knowledge from the environment through training and learning process. Many algorithms are available for learning process This work carry out research on analysis of malicious packets and predicting the error rate in detection of injured packets through artificial neural network algorithms.

Key words: ABC algorithm, PSO algorithm, KDD cup 1999 data set, Optimization, Neural network, Intrusion detection system, Malicious packets.

I. INTRODUCTION

The ABC algorithm requires fresh fitness tests on the fresh parameters to enhance performance and also demerits in use of secondary information about the problem[5].

The procedure involves, gathering data from KDD Cup1999 data set (KDD CUP 99 training set and test set contain 4898431 and 311027 intrusion and normal records)[1], Configure the proposed network, Initialize the weights and instant biases, Train the network through the Artificial bee colony algorithm and Particle Swarm Optimization algorithm. The efficiency of detecting attacked dataset may be calculated.

II. ARTIFICIAL BEE COLONY ALGORITHM

ABC is designed by Karaboga in 2005 and is a kind of swarm optimization algorithm. It resembles the intelligent behavior of honey bees. A group of honey bees is called swarm and a swarm can complete assigned work through better coordination and social cooperation. Here, the bees involved are employed bees, onlooker bees, and scout bees. The food would be searched by employee bee. After finding the food source, the information will be transfered to onlooker bees.

ABC algorithm is good in exploration but not good at exploitation. ABC algorithm is tested on three different data sets of KDD cup 1999 data set containing all types of attack signatures for optimization problems.

The number of employed bees or the onlooker bees is equal to the number of solutions in the swarm. The ABC automates a randomly distributed initial population of SN solutions (i.e food sources), which denotes the swarm size[4].

Consider $X_i= \{ x_1,1,x_2,2,…..,x_n,n\}$ as the i^{th} solution in the swarm, n is the dimension size. Each employed bee X_i creates a new candidate solution V_i adjacent to the present position as represented by

$$V_{ik} =X_{ik}+\phi_{ik} x(X_{ik} -X_{jk})$$

[6]Where X_j is a randomly selected candidate solution (i not equal to j), K is a random dimension index selected from the set ,{1,2…n } and ϕ_{ik} is a random number within [1,-1] . Once the new candidate solution V_i is generated, a greedy selection is used. If the fitness value of V_i is better than that of its parent X_i, then update X_i with V_i otherwise X_i should not be changed. The employed bees finish the search process and share the information of their food sources with the onlooker bees through dances. An onlooker bee evaluates the nectar information taken from all employed bees and chooses a food source with a probability related to its nectar amount. This probabilistic selection is really a roulette wheel selection mechanism which is described as under[3]:

$$P_i = fit_i / \Sigma_j\, fit_j$$

Where fit_i is the fitness value of the i^{th} solution in the swarm[6]. Better the solution i, the higher the probability of the i^{th} food source selected. If a position cannot be improved over a predefined number of cycles, then the food source is abandoned. Assume that the abandoned source is X_i, and then the scout bee discovers a new food source to be replaced with i^{th} as[6] :

$$X_{ik} = lb_j + \text{rand}\,(0,1) * (ub_j - lb_j)$$

III. METHODOLOGY

Input	→	Malicious packets/KDD 99 data set
Activation function	→	ABC algorithm and PSO algorithm
Output	→	Detection of malicious packets . Further error rate of detection of malicious packets can also be calculated.

[2]The mechanism of attack recognition is proposed to be achieved through several stages such as monitoring stage, detection stage and decision stage. Neural network may receive malicious packets from the data set and analyzes their undesirable features and behaviors for misuse intrusion. The undesirable behaviors of the malicious packets are identified as attack signatures and used to create the new patterns. Some agents may be trained on the new patterns in order to prevent packets which are matched with new patterns. On the other hand, desirable behaviors are extracted from the normal records in the data set KDD99. Neural network also receives the normal packets from the data set and analyzes their desirable behavior for normal packets.

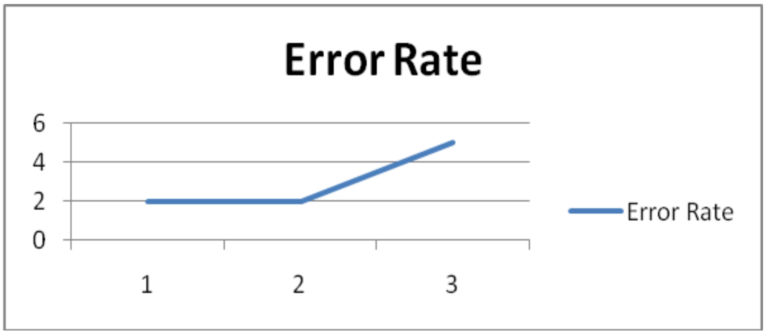
IV. STUDY OF THE PERFORMANCE OF ABC ALGORITHM

The study of ABC algorithm for detection of malicious packets among 3 trials of KDD 99 data set is as under:

Table1. Result given by ABC algorithm

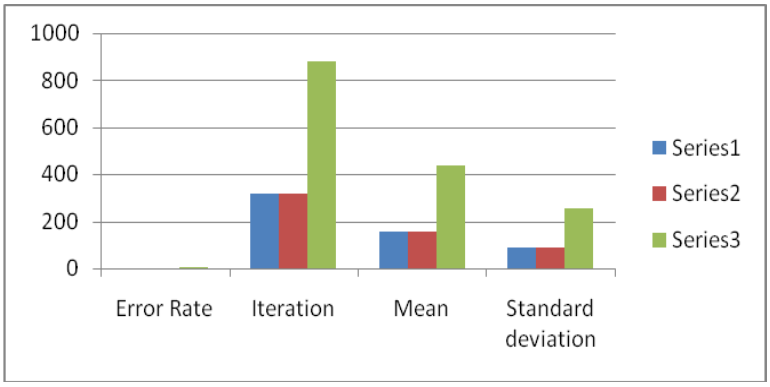
Error Rate	Iteration	Mean	Standard deviation
2%	318	159	92.0869
2%	320	160	92.6606
5%	878	439	253.74

The graphical view of result obtained by ABC algorithm is represented as below:



Graph1. Representation of Error Rate

It is clear from the above graph that the ABC algorithm gives the optimum error rate for various inputs given.



Graph2.Representation of output parameters

V. PSO ALGORITHM

It was introduced by James Kennedy and Russel Ebhart in 1995. It creates a swam of particles which move in space around the problem space to search the best suitable fitness function. It is capable of optimizing nonlinear and multi dimensional problems. The ideas behind the optimizing properties are

a) A single particle can find its good position by sharing its fittest position and the knowledge obtained from the other particles.

b) A stochastic factor in each particle's velocity makes it move through unknown problem space.

The above two properties enable an extensive exploration of the problem space and gives a best solution for the problem.

VI. WORKING PROCEDURE

The particles of a swam moves in a problem space and evaluate their position through a fitness function. Once the problem space is defined , a set of particles focus in it and the corresponding position and velocities are updated by iterations according to the algorithm.

[7] In this algorithm the position of the particle at any instance is given by

$$x_{i,d}\,(it+1) = x_{i,d}\,(it) + v_{i,d}\,(it+1)$$

And the velocity for corresponding particle is

$$\begin{aligned} v_{i,d}\,(it+1) = v_{i,d}\,(it) &+ C_1 * Rnd(0,1) * [pb_{i,d}(it) - x_{i,d}(it)] \\ &+ C_2 * Rnd(0,1) * [gb_d(it) - x_{i,d}(it)] \end{aligned}$$

Where,

i→ Index of the Particle

d→ Dimension

it→ Iteration

$x_{i,d}$→Position of the particle i in dimension d

$v_{i,d}$→Velocity of the particle i in dimension d

C_1→Acceleration constant for the cognitive component

C_2→Acceleration constant for the social component

Rnd→Random value from 0 to 1

$pb_{i,d}$→ the location in dimension d with the best fitness of all the visited locations in that dimension of particle i.

The flow chart for the process as under

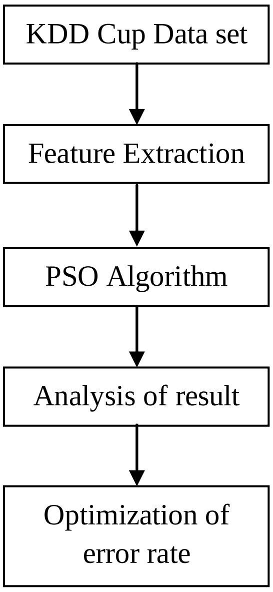


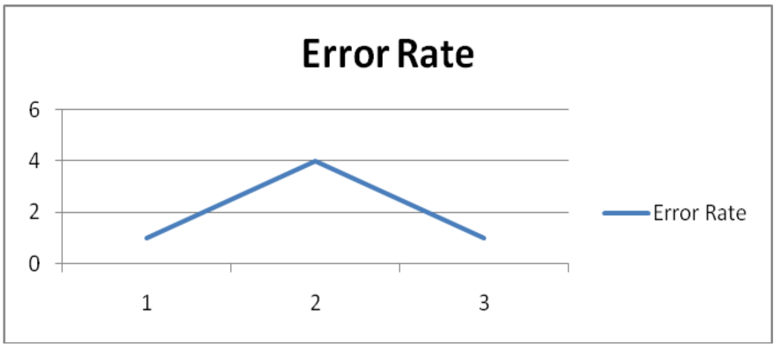
Figure 1 : Process step

In this algorithm the instance value of particles are taken as real numbers. The typical range is 20-40. The dimension and range of the particle is required to be optimized. The learning factors C_1 and C_2 were usually taken as 2.Based on the algorithm the results are obtained as under,

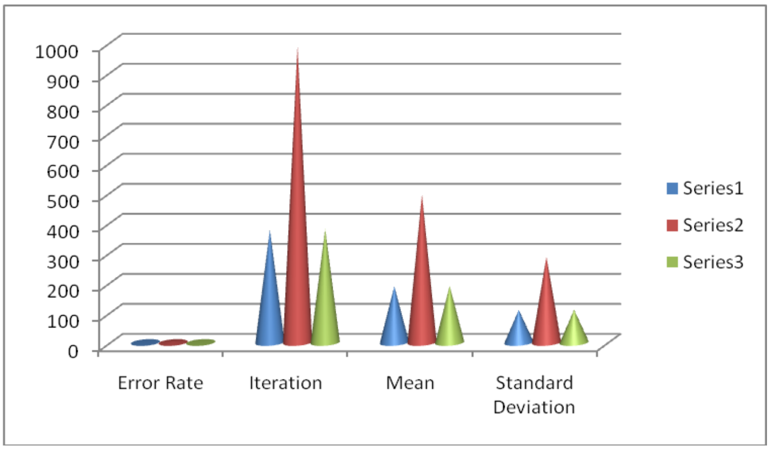
Table 2: Result obtained from PSO Algorithm

Error Rate	Iteration	Mean	Standard Deviation
1 %	376	188	108.83
4 %	989	494	285.78
1 %	379	189	109.69

The graphical representation of the above study with PSO algorithm is as below,



Graph3. Representation of Error Rate for PSO Algorithm



Graph4. Representation of output parameters for PSO Algorithm

VII. CONCLUSION

The study of ABC algorithm gives the error rate of 2%, 2% and 5% in three instances, whereas the PSO algorithm reduced the error rate to 1%,1% and 3% respectively. The research could be further widened by applying optimization to other parameters for various other algorithms such as Fuzzy C-Means, Fuzzy K-Means, SVM etc, and the error rate can be compared and finally one or more algorithms be modified for getting minimum or zero error rate in detection of malicious packets.

REFERENCES

[1] "Intrusion Detection using Artificial Neural Network" By Poojitha G, 978-1-4244-6589-7/10/$26.00@2010 IEEE.

[2] "Classifying Attacks in a Network Intrusion Detection System based on Artificial Neural Networks" by Mohammad Reza Norouzian, ISBN 978-89-5519-155-4, Feb. 13-16,2011 ICACT2011.

[3] "Intrusion Detection based on neural networks and Artificial Bee Colony algorithm" by Quan Qian, 978-1-4799-4860-4/14/$31.00 copyright 2014 IEEE ICIS2014, June 4-6, 2014, Taiyuan, China.

[4] "A Real time IDS Based on Artificial Bee Colony- Support Vector Machine algorithm" by Jun Wang, 978-1-4244-6337-4/10/$26.00@2010 IEEE.

[5] "A New Network Intrusion Detection Identification Model Research" by WenJie Tian, 978-1-4244-5194-4/10/$21.00@2010 IEEE.

[6] https://en.wikipedia.org/wiki/Artificial_bee_colony_algorithm.

[7] http://web.ist.utl.pt/gdgp/VA/pso.htm, Page 2/19.

ABOUT AUTHOR

1. S.Vijaya Rani, MCA,M.Phil,(PhD). Assistant Professor, MCA Dept,Brindavan College of MCA/MBA, Bangalore-560063.
She is a Research Scholar in Computer Science of Bharathiar University, Tamilnadu and having 10 years of teaching experience. The research area is Intrusion detection system using Artificial Neural Network. She has published 05 articles in various Journals and conference proceedings. She also attended various international ,national conferences and workshops.

2. Dr.G.N.K.Suresh Babu, M.E.,M.C.A.,M.Phil., Ph.D, Associate Professor,MCA Dept, Acharya Institute of Technology, Bangalore.
He is having 25 years of teaching and guiding experience. The research area is data mining and Artificial Neural Networks.He has published many articles in Journals and attended International and national conferences.