International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 9 (2017) pp. 1844-1847

Enhanced Intrusion Detection System Using Data Reduction: An Ant Colony Optimization Approach

K. Kanaka Vardhini and Dr. T. Sitamahalakshmi
Associate Professor, Department of CSE, Audisankara College of Engineering and Technology
Professor, Department of CSE, GITAM University, Visakhapatnam, India.

[1]Author ID:57188758025 & ORCID:00000-0002-6860-3583

Abstract
Intrusion Detection Systems (IDS) are playing a critical role in present day data security systems, wherein there are so many intrusion detection techniques contributing to the purpose partially or fully. In particular, Ant Colony Optimization (ACO) based intrusion detection methods are in action for effective intrusion detection. In this paper, an enhanced heuristic algorithm has been proposed with sensitivity and specificity factors for increasing the success rate in the intrusion detection. Data Patterns have been extracted from the network data effectively using the proposed heuristic algorithm. The Knowledge discovery from databases (KDD) dataset is used as input and the proposed approach increased the percentage of intrusion detection as compared with the conventional methods.

Keywords: Classifier, Ant Colony Optimization, KDD, Intrusion Detection System (IDS), Sensitivity, Specificity, Optimization.

INTRODUCTION

In the present day context, maintenance and management of the data in the secured environment and maximizing the intrusion detection has become the prime target for the researchers. The possibilities of unauthorized attempts to access and manipulate the information are limitless. Such types of attempts find the vulnerabilities in the operating system as well as in application programs. The intrusion detection systems [1] are expected to detect unauthorized attempts without compromising confidentiality, integrity and authentication. Because of vulnerability and penetration, it is not always possible to prevent the intruder hence intrusion detection with alarms is required for avoiding the damages.
The analysis of audit data, log files provide a better way to detect intruders which is shown in the Figure1

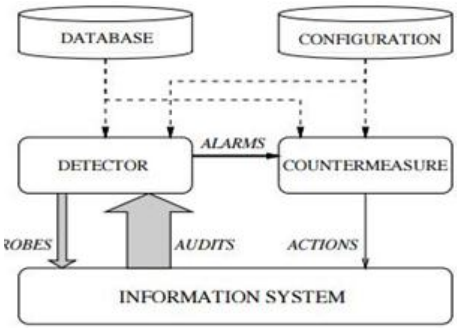

Figure1: Intruder detection form database analysis

The intruder detection system analyzes the information or audits data like source and destination number, number of login failures, application version number to detect an attack. The outcome of the analysis will be resulting in the generation of the classifier which is able to provide the countermeasure to stop the attacking.
In this paper KDD data [2] is used as the input dataset. The KDD dataset consists of 41 attributes along with target class label. The KDD contains symptomatic features that describes 4 types of attacks

- **Denial of Service (DOS) :** Making Server busy unnecessarily by sending sync packets
- **Probe:** It is an attempt trying to collect network information for analysis
- **User to Root attack(U2R):** Person trying to exploit vulnerability to gain root access
- **Remote to Login (R2L):** Unauthorized access from a remote machine.

LITERATURE REVIEW

There are some attempts made towards intrusion detection systems using mobile agent mechanism. N.Jaisankar et al., [3] implemented a new anomaly detection system which is based on both the user activity and program operation. They also proved that proposed method provides a security using mobile agent mechanisms. They found that the layered approach can quickly detect the intruder when compared to existing ones.
Ant colony and particle swarm optimizations are used by M.Sailaja et al.,[4] for Intrusion detection. They found that the hybrid approach reduces the false alarm rate with high accuracy and the effectiveness of IDS depends on the data collection. They also proved that the quality of the intrusion detection system depends on the data provided to the intrusion detection engine. The hybrid swarm intelligence algorithm (PSO/ACO) was used for intrusion detection and the results were compared with the support vector machine algorithm.
M. Poongodi and S.Bose [5] designed intrusion detection system using trust evaluation metrics. Firefly -based security procedures with Dynamic Growing Self-Organizing Tree Algorithm is used to detect Distributed Denial of Service attacks. This approach involves trust-based evaluation wherein the intrusion detection is done using secured trust evaluation policies. The proposed system combines the existing Firecol based security procedures with Dynamic Growing Self-Organizing Tree Algorithm in the trust evaluation-based environment. Mohammed Alweshah et al., [6] hybridized the stimulated annealing with firefly algorithm to perform classification. They used 11 different datasets to

test the efficiency of algorithm. Levy flight within firefly algorithm is used to explore the search space.

P. Amudha et al.,[7] proposed hybrid algorithm to detect intruders in the network. They combined artificial bee colony with enhanced particle swarm optimization to find the abnormal patterns of the network traffic. Friedman test and ANOVA test were applied to test the classifier accuracy. Ibrahim Aljarah and Simone A. Ludwig [8] used parallel particle swarm optimization with map reduce methodology for intrusion detection. They performed clustering on data to find the intruder in the network. Experimental results proved that proposed algorithm scales very well with large data sets. Mahmod S. Mahmod et al.,[9] combined artificial bee colony and multi layer perceptron to build IDS. Multilayer perceptron was used for classifier implementation which was trained by artificial bee colony. The result showed that hybrid approach is good at finding intruder in the network. Seyed Mojtaba Hosseini Bamakan et al.,[10] used multiple criteria linear programming for implementation of classifier to detect intruders. Particle Swarm Optimization technique has been used to improve the performance of the classifier.

METHODOLOGY

The KDD dataset is used as input dataset. The size of the dataset is reduced by considering unique values from each and every column. The functional architecture of proposed system is shown in Figure 2.

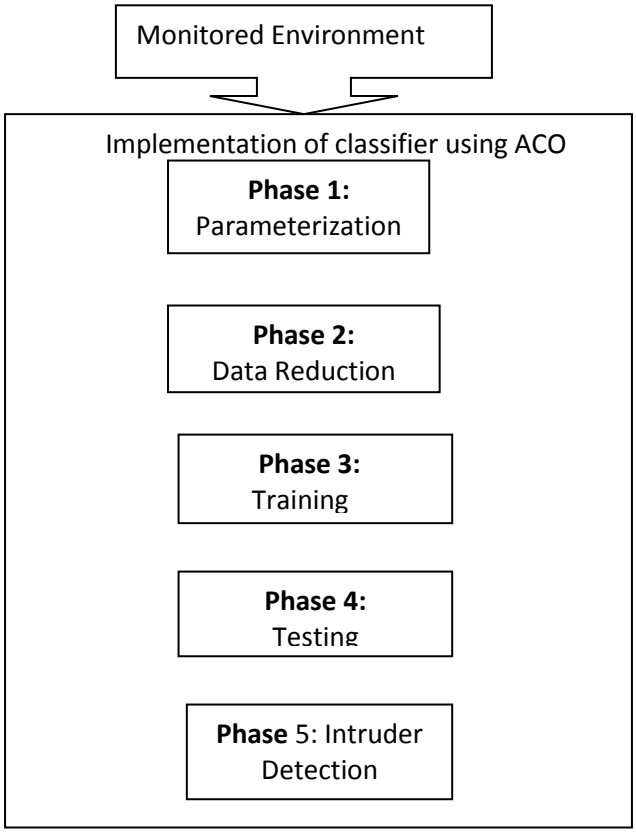

Figure 2: Functional architecture of proposed method

The whole method is divided into five phases, namely parameterization, Data reduction, training, testing and detection.

The categorical values have been changed to continuous values in the parameterization stage. In reduction stage, data is selected based on uniqueness in the dataset which as shown in Table 1. This resulted in the reduction of the data size from 41X 25192 to 41X 3922.

Table 1: Number of unique values in each attributes.

S.no	Name of the Feature	No of records after selection	Name of the Feature	Number of records after selection
1	Duration	758	is_guest_login	2
2	protocol_type	3	Count	466
3	Service	66	srv_count	414
4	Flag	11	serror_rate	70
5	src_bytes	1665	srv_serror_rate	56
6	dst_bytes	3922	rerror_rate	72
7	Land	2	srv_rerror_rate	42
8	wrong_fragment	3	same_srv_rate	97
9	Urgent	2	diff_srv_rate	79
10	Hot	22	srv_diff_host_rate	57
11	num_failed_logins	5	dst_host_count	256
12	logged_in	2	dst_host_srv_count	256
13	num_compromised	28	dst_host_same_srv_rate	101
14	root_shell	2	dst_host_diff_srv_rate	101
15	su_attempted	3	dst_host_same_src_port_rate	101
16	num_root	28	dst_host_srv_diff_host_rate	63
17	num_file_creations	20	dst_host_serror_rate	100
18	num_shells	2	dst_host_srv_serror_rate	88
19	num_access_files	7	dst_host_rerror_rate	101
20	num_outbound_cm	1	dst_host_srv_rerror_rate	100
21	is_host_login	1		

The KDD dataset contains 4 class levels excluding normal label. Performance metric (sensitivity, specificity) based ant colony optimization[11,12] is used to find out the best classification rules for intrusion detection in the later stage of the methodology as shown below

Proposed ACO Algorithm

Read reduced data set D of every class into matrix D1 of size N*M

N: number of rows in D1 and M: number of columns

L: number of class labels, A: Ants, RL: rule list

Number of Iterations=50

1. Initialize population;
2. Initialize pheromone;
3. Repeat
   For each ant
   Add term based on heuristic to the current Rule R i
   Update the pheromone at visited nodes
4. Evaporation of pheromone is done at unvisited nodes;
5. i=i+1; Until (K>=NOI)
6. Prune each constructed rule (R_i) by all ants.
7. Choose the best rule R_{best} among all the rules ;
8. Add R_{best} to the RL;

Description of Algorithm

Pheromone Initialization :

Initially all the edges between nodes are initialized with same amount of pheromone value using equation (1)

$$\text{Pheromone Initialization } (PI_{i)} = \frac{1}{\sum_{j=1}^{N} Dj} \quad \text{----------(1)}$$

Term Selection:

Terms to the rule list are added incrementally based on their probability value with a constraint that the other term containing the same attribute can't be considered once a term has been included in the rule. The probability value for the term selection is calculated as per the equation (2)

$$P_t = \frac{TP}{TP+FN} * \frac{TN}{TN+FP} \text{ ----------(2)}$$

Rule Quality and Pheromone Updation

After completion of rule construction the quality can be calculated using equation (3)

$$Q = \frac{TP}{Covered} + \frac{TP}{N} \text{ ----------(3)}$$

The pheromone value is updated based on the accuracy of the rules.

Rule Termination:

The rule construction is terminated when ant reaches stopping condition or all the records coming under the rule show the same target label. The termination also happens if there are no attributes to add in the rule.

EXPERIMENTAL RESULTS

In this paper performance metric based ACO is implemented to generate the classification rules using KDD dataset. The proposed classifier method is able to enhance the accuracy of the intrusion detection to 98.5%. The comparative analysis with that of the other classification techniques are shown in the Table 2 and the corresponding confusion matrix is shown in the Tables 3 & 4.

Table 2: Comparison of proposed Classifier accuracy with ABC, ACO and Naive Bayesian

Dataset	Proposed ACO	ABC	ACO	Naïve Bayesian
NSL KDD20%	98.5	97.23	98.0	97.4
10% KDD	99.00	98.4	98.5	98.5
KDD cup99	98.02	92.5	97.3	95.4

Table 3: Confusion matrix of Proposed ACO on NSL KDD20%

Type	DOS	Probe	U2R	R2L
DOS	82.86	0.41%	0.5%	2.04%
Probe	15.55	60.12%	0.19%	2.04%
U2R	1.2%	0.05%	0.00%	73.46%
R2L	0.08%	0.10%	89.06%	12.26%
NORMAL	0.31%	19.32%	10.25%	0.20%

Table4: Confusion matrix of Proposed ACO on KDD10%

Type	DOS	Probe	U2R	R2L
DOS	99.47%	4.78%	0.00%	5.88%
Probe	0.05%	70.30%	0.00%	0.01%
U2R	0.06%	3.54%	81.7%	0.45%
R2L	0.37%	0.5%	90.96	1.07%
NORMAL	1.07%	0.6%	0.23%	0.5%

CONCLUSION

Most of the traditional intrusion detection systems had many limitations due to heavy traffic in the network. In this paper, sensitivity and specificity are used as the part of heuristic function. The proposed ACO with new heuristic function enhanced the accuracy in finding the patterns which are useful for intruder detection. The usage of performance metrics as heuristic function increased the accuracy with less time complexity.

REFERENCES

[1] A. Kartit, A. Saidi, F. Bezzazi, M. El Marraki, A. Radi ,A New Approach to Intrusion Detection System, Journal of Theoretical and Applied Information Technology 29th February 2012. Vol. 36 No.2

[2] Dr. L.Dhanabal, S.P. Shantharajah, A Study on NSL-KDD Dataset for Intrusion Detection System Based on Classification Algorithms, International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 6, June 2015

[3] N. Jaisankar, R. Saravanan, K. Durai Swamy, Intelligent Intrusion Detection System Framework Using Mobile Agents, International Journal of Network Security & Its Applications (IJNSA), Vol 1, No 2, July 2009

[4] M. Sailaja,R. Kiran Kumar ,P.Sita Rama Murty, P.E.S.N .Krishana Prasad A novel approach for intrustion detection sytem using Swarm

Intelligence, proceeding of the InConINDIA 2012,AISC 132,pp469-479,springer

[5] M. Poongodi, S. Bose, A Novel Intrusion Detection System Based on Trust Evaluation to Defend Against DDoS Attack in MANET, Arabian Journal for Science and Engineering, December 2015, Volume 40, Issue 12, pp 3583–3594

[6] Mohammed Alweshaha, Salwani Abdullahb, Hybridizing firefly algorithms with a probabilistic neural network for solving classification problems , Elsevier, Applied Soft Computing 35 (2015) 513–524

[7] P. Amudha, S. Karthik and S. Sivakumari, A Hybrid Swarm Intelligence Algorithm for Intrusion Detection Using Significant Features, The Scientific World Journal Volume 2015 (2015), Article ID 574589, 15 pages

[8] Ibrahim Aljarah, Simone A. Ludwig.: MapReduce intrusion detection system based on a particle swarm optimization clustering algorithm, IEEE congress on Evolutionary Computation, 20-23 June 2013

[9] Mahmod S. Mahmod , Zakaria A. Hamed Alnaish, Ismail Ahmed A. Al-hadi.: Hybrid Intrusion Detection System Using Artificial Bee Colony Algorithm and Multi-Layer Perceptron, International Journal of Computer Science and Information Security, Vol. 13, No. 2, February 2015

[10] Seyed Mojtaba Hosseini Bamakan, Behnam Amiri, Mahboubeh Mirzabagheri , Yong Shi, A New Intrusion Detection Approach using PSO based Multiple Criteria Linear Programming, Information Technology and Quantitative Management, Computer Science 55 (2015) 231 – 237

[11] M. Dorigo. Gambardella, Ant colonies for the traveling salesman problem. Bio Systems, vol. 43, no. 2, pp. 73–81, (1997).

[12] M. Dorigo, M Brattain, T. Stutzle, Ant colony optimization: Artificial ants as computational intelligence. IEEE Computational Intelligence Magazine | November, (2006)