Volume 7, Issue 4, April 2017 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

A Novel Method to Enhance the Security of Computer System using AIS in Intrusion Detection System

[1]Sharad Gangele[*], [2]Dr. Anshuman Sharma
[1] M.Tech (CSE), Research Scholar,
[1, 2] Department of Computer Science & Engineering, School of Research and Technology, People's University, Bhopal, Madhya Pradesh, India

Abstract— Now a days, with the growing use of the computers, a consistent case of the vulnerabilities is also increased frequently. To improve the security of computer system various approaches have been proposed from different areas. One such approach is make a use of AIS in Intrusion Detection System. The threats and intrusions in computer network have similarities with human diseases, therefore IDS basically can be compared to biologic immune system. In this case human body deals with an effective approach namely HIS (Human Immune System) which affords a high level security to human body from the invasion of pathogens and can be utilized for the identification and protection of unnoticed intruders. To enhance the efficacy of IDS, AIS (Artificial Immune System) can play a vital role, which provides influential features in the design of IDS. This paper gives a new way to the anomaly detection problem. To detect the anomaly problems and unnoticed intruders, negative selection algorithm method is found more suitable than K-Mean Clustering method.

Keywords— Clustering, AIS, Negative Selection Algorithm, Security, IDS.

I. INTRODUCTION

In today's scenario, computer security becomes a critical issue, which provides protection against harm or damage to the computer system and the services they afford. The arenas are of growing importance due to the increasing dependence on computer systems. Therefore it becomes necessity to reduce vulnerability or attack of intruders in computer network with the help of computer security [1],[2].This computer security will not only help to protect against intruders but also provide network security by intrusion recommendation for overall organization. The foremost aim of IDS is to find unauthorized use, misuse and abuse of computer systems. Network based IDS for network intrusion detection helps in monitoring the number of hosts and network traffic [3]. It is basically involved with two main components first is anomaly detector and other is misuse detector. Anomaly based IDS have been explored as protective technique to detect unnoticed intruders, which is unknown to the system. The anomaly detector, detects the intrusions by comparing the deviation of profile of normal behaviour of users, systems, network traffics and services. This detector firstly defines the normal system behaviour and defines all other as abnormal. Misuse detection is another approach to detecting computer attack, which is signature based detection [4],[5]. In this method, abnormal behaviour is defined first and other behaviour is defined as normal. This method is basically concerned with detection intruders who are trying to break into a system through known vulnerabilities. This approach investigates whether these misuse signatures are present in the auditing trails or not. At present various techniques have been established for host and network based IDS. However, there is a more scope for some unsolved problems through development of an effective network based IDS.[6-8] In this work, we have suggested a comprehensive comparative study of clustering algorithm model which is centroid based algorithm and AIS for examining large intrusions detection datasets. Moreover, our work examined the aptness of centroid based clustering algorithms and empirically compared with negative selection algorithm to detect the unnoticed intruders in terms of run time efficiency. The novelty of this work is to build a network based IDS using AIS, to find unnoticed or unknown attack, and to generate neural network technique, which allowed high level information in the pattern of clustering.

II. LITERATURE SURVEY

A lot of research works have been carried out in the literature for intrusion detection system (IDS) and some of them have motivated us to take up this research. Han and Cho (2006) presented a novel intrusion-detection technique based on evolutionary neural networks. It takes less time to obtain superior neural networks than when using conventional approaches. Hong, Xin and Li (2002) proposed a third party IDS with candidate signatures and also suggest an algorithm called signature apriority for it. Chebrolu, Abraham and Thomas (2005) suggested an important input features in building IDS that is computationally efficient and effective. In which performance of two feature selection algorithms involving Bayesian networks (BN), classification and regression trees (CART) were suggested. Hong , Zhang and Wu (2004) discussed an intrusion detection method that combines rough sets and SVM algorithm. They also

analyzed the ability rough sets have to decrease the amount of data and get rid of redundancy, the technique can reduce the amount of training data and overcome the SVM defect of slow running speed when processing large datasets. Uppal, Jabad and Arshad (2014) presented a fundamental aspect of intrusion detection system and also studied different techniques that are commonly used and helpful for the new researchers who want to know the basic knowledge of intrusion detection systems. Carver *et al.* (2001) examined a technique for limiting uncertainty in adaptive intrusion response systems and specifically in the adaptive, agent-based intrusion response system. Cohen has also explored the inadequacy of manual intrusion response and the need for automatic intrusion response. Caberera, Ravichandran and Mehera (2000) advocate statistical traffic model in particular case of intrusion detection and develop some result on it. Freund and Schapire (1997) examined the multiplicative weight update Littlestone Warmuth rule that can be adapted in model, yielding bonds that are slightly weaker in some cases, but applicable to a considerably more general class of learning problems. Depren *et al.*(2004) focused on normal traffic data and a mathematical model was used for describing normal traffic and a test was conducted based on the deviations from the mathematical model. In study a self-organizing map (SOM) structure was used for constructing the mathematical model describing normal traffic and anomaly detection. Debar, Becker and Siboni (1992) attempt for a designing a possible application of neural networks as a component of an intrusion detection system. The approach of user behavior modeling that takes advantage of the properties of neural algorithms and display results obtained on preliminary testing of data. Khannous, Rghioui, Elouaai and Bouhorma conducted a study on integrates and combines the basic concepts of intrusion detection system which is based on the role of T cells described by the negative selection algorithm. Lee and Stolfo (1998) have given a thought on general and systematic methods for intrusion detection. They used data mining techniques to discover consistent and useful patterns of system features that describe program and user behavior with the help of set of relevant system. Tsong, Tsung and Yuh (2007) have reviewed the a three-tier architecture of intrusion detection system which consists of a blacklist, a white list and a multi-class support vector machine classifier. The design also provides the flexibility for the practical usage. Sridevi, Jagajothi and Chattemvelli (2012) have innovatively presented artificial immune systems which were based on the human immune system (HIS). Human immune system can detect and defend against harmful and previously unknown invaders. Rassam , Maarof and Zainal (2010) have put on a new look on the detection rate by reducing the network traffic features and to investigate the feasibility of bio-inspired Immune network approach for clustering different kinds of attacks and some novel attacks. Lei and Ghorbani (2004) have given description on detecting network intrusions based on a competitive learning neural network. The approach was compared to that of the self-organizing map (SOM), which was a popular unsupervised training algorithm used in intrusion detection. Rassam and Maarof (2012) investigated an application of bio-inspired clustering approach, named artificial immune network, for clustering attacks for intrusion detection systems. Kim and Bentley (2001) performed a study on how to choose appropriate detector and antigen sample sizes. These ideal sizes allow the AIS to achieve a good non-self antigen detection rate with a very low rate of self antigen detection. One more thing concluded that the embedded negative selection operator plays an important role in the AIS by helping it to maintain a low false positive detection rate. George (2012) explored the decrease in execution time for the classification as they reduce the dimension of the input data and also the precision and recall parameter values of the classification algorithm shown the SVM with PCA method was more accurate as the number of misclassification decreases. Lee, Stolfo and Mok (1999) elaborated a data mining framework for adaptively building intrusion detection (ID) models. The central idea was to utilize auditing programs to extract an extensive set of features that describe each network connection or host session. Levin, Ziyyon and Israe proposed a novel approach on data-mining tool based on building the optimal decision forest. The tool won second place in the KDD99 classifier learning contest. One more contribution was given by Fan, Lee, Stolfo and Miller (2000) for the problem of building cost-sensitive intrusion detection models to be used for real time detection and major cost factors in IDS, including consequential and operational costs were discussed. Chung (2012) has a useful contribution on agent-based artificial immune system (ABAIS) was adapted to intrusion detection system (IDS). An agent-based IDS (ABIDS) inspired by the danger theory of human immune system was proposed. Khalkhali *et al.*(2011) examined a host-based web anomaly detection system and analyzed the POST and GET requests. Web access log file was introduced which eliminates the shortcomings of common log files for defining legitimate users sessions boundaries. Yanbin (2015) suggested a network intrusion detection system based on artificial immune principle of the new model. Through the optimization algorithm, the model could improve the ability of the immune response of the system.

III. PERCEPTION OF IDS AND AIS

In this segment, we have shown a diverse study of these two terms *i.e* IDS and AIS pertaining to our work.

3.1 Concept of Intrusion Detection Systems-

Because of the increasing reliance in computer internet network, the problem of intrusion by unauthorized users is rising. An intrusion is unauthorized attack or attempted attack into or unauthorized activity in a computer system. An intrusion detection system is a system which monitors the frequency and features of Internet attacks on a computer system and also provides the filters for attack alerts [9-12]. The system detects unauthorized users which makes an effort to enter into a computer system by comparing a user profile, detects events that gives an idea of unauthorized entry into the system, reports a regulate function about the unauthorized users[13]. The events that shows an unauthorized access into the computer system and has a control function which automatically takes action in response to the event. The IDS observe attacks on other hosts and analyse whether these attacks are general attacks or attacks from specific computer network which produces a corresponding signal [14-16]. The IDS also examine a computer networks vulnerability to attacks, which

detected on the other monitored hosts. The IDS automatically builds user signature based profile data for each user which can be used to find normal actions for each user to reduce the existence of false alarms and to improve detection of intruders [17,18].

3.2 Classification of Intrusion Detection Systems-

Classifications of Intrusions are as follows **[13]**:

1) Attempted break-ins- It is an Anomaly based detection system that is detected through distinctive profile behavior or violations of constriction security.
2) Security control system penetration- Detection by monitoring of activity for definite patterns.
3) Denial of service - Detection of intrusion takes place by a typical usage of system resources.
4) Leakage-Detection of intrusions by different usage of system resources.
5) Masquerade attacks- Detection on the basis of distinctive profile behavior or violations of security constraints.

3.3 Intrusion Detection using clustering-

Intrusion detection system comprises, recognizing a set of unpleasant activities that negotiated with the elementary security necessities like confidentiality, availability and integrity of information resources. The colossal increase of attacks on network and is the key reason for the data mining based intrusion detection techniques which is enormously beneficial in detecting the attacks[19],[20]. This paper describes a system which gives idea about network data mining that is the usage of data mining method which helps in capturing flow data and data packet in a network together with comparative study of other methods. Moreover, In IDS, preceding work has been done with respect to K-mean clustering algorithm. Training data having out of order flow records which are separated into normal traffic and irregular one [21]. The analogues of cluster centroids having utility for distance-based detection of anomalies in data. Previous methods of network intrusion detection are associated with saved patterns of recognized attacks, detection of intruder take place through the comparison of network connection characteristics with the pattern of attacks by specialists. The foremost disadvantage of traditional method is to be unable to detect unknown attack. Furthermore, if novel attacks were identified, this new pattern would have to be updated in system manually. The neural networks are extensively known as a proficient method to adaptively classify patterns, but for the intrusion detection problem their long training cycles and high intensity delayed the application [22].

IV. PROBLEM TO BE ADDRESSED

4.1 IDS using clustering-

Clustering is an unsubstantiated learning technique which assists the network security proficient with labelling network traffic records as normal or intrusive. Its main objective is to determine structure in a collection of data which is not labelled. A significant benefit of using clustering or unsupervised learning to detect network attacks is having the ability to find new attacks not seen before. This infers that attack types with unknown intrusion can be detected [23].

4.2 K-means algorithm-

The K-means clustering is a traditional clustering type of algorithm which solves the well-known clustering problem. The process takes place in a simple way where given data is classified through a certain number of clusters (k clusters).The foremost idea is to define k centers first for individual cluster, then take each point fitting to a given data set and associate it to the closest center. A loop has been generated and as an effect of this loop we may observe that the k centers change their position step by step until centers do not move any more [23-25].

Algorithm:

1. Pick Initial value of k.
2. Randomly assign points to k group/Clusters.
3. Repeat steps 3 to 5 until convergence.
4. Determine the cluster to which source data belongs Use Euclidean/City block distance formula. Add element to cluster with min (Distance (x_i, y_j)).
4. Calculate the means of the clusters.
5. Change cluster centroids to means obtained Using Step 3.

The Main Disadvantage of K-Mean algorithm is that algorithm may take a large number of iterations through dense data sets before it can converge to produce the optimal set of centroids. This can be incompetent on large data sets due to its unbounded convergence of cluster centroid. Also, the count of clusters should match the data. If incorrect value of k is picked, then it will invalidate the whole process. An empirical way to find the best number of clusters is to try K-means clustering with different number of clusters and measure the resulting sum of squares.

V. PROPOSED APPROACH

In this work, we have studied the negative selection algorithm for Intrusion Detection System. The Negative selection algorithm is motivated by the chief mechanism in the thymus which provides a set of mature T-cells proficient of tying only non-self-antigens. The first negative selection algorithm have efficacy to detect data manipulation due to virus in a network. In Negative selection algorithm firstly, a set of self-strings "s" is generated, which shows the normal

state of the system followed by generation of a set of detectors "D", which only recognize the counterpart of S String. In order to categorize them, these detectors can then be utilize for new data which affords the fact that data has been manipulated. The Negative Selection Algorithm generates the set of detectors through the process given below.

input S_{seen}= set of seen known self elements

output: D = set of generated detectors in AIS

Begin

Repeat

- Randomly generate potential detectors and place them in a set *P*. Let P = {p_1, p_2….}
- Determine the resemblance of each member of *P* with each member of the self set S_{seen}
- If at least one element in *S* recognized a detector in *P* element according to a recognition threshold, then the detector is rejected, otherwise it is added to the set of available detectors *D*
- Until Stopping criterion has been met

End

The Negative Selection Algorithm was aimed for improve the detection rate, novelty detection, intrusion detection and similar pattern recognition and two-class categorization problem fields. The algorithm has been designed to have equilibrium between the number of detectors and quality of the String matches. The urge of reliance between detectors has not found which indicates that detector training and application is integrally similar and appropriate for a dispersed and parallel implementation.

VI. CONCLUSIONS

In this paper, the probability of unsubstantiated intrusion detection using centroid-based clustering algorithms has been studied. Since network traffic intrusions having the dynamic pattern, unsubstantiated intrusion detection has found more appropriate for anomaly detection than classify intrusion detection approaches. A study has been performed which is comprises of centroid based clustering and Artificial Immune System technique along with a case study of data acquired from KDDCUP98 data set. A comparison based investigation and assessment of the clustering algorithms generate rational intrusion detection rates. Our favorable clustering and detection outcomes inspired us to explore this work in future. Recognizing the specific attack class, related cluster and the insightful characteristics, which is distinctive to a specified cluster can be utilize for individual clusters comprehensive analysis. However, the arenas of clustering which is feature selection based will be explored, which will be very promising in detection of new attack categories. In Future, Self-labelling techniques can also be developed to enhance the activity of clustering-based intrusion detection.

REFERENCES

[1] Caberera, J.B.D., Ravichandran, B. and Mehera, R.K., "Statistical Traffic Modeling for Network Intrusion Detection", In Proceedings of International Symposium on Modeling, Analysis and Simulation of Computer and Tel. Sys., pp. 466-473, 2000

[2] Carver, C. A., Hill, J. M. D. and Pooch, U. W., "Limiting Uncertainty in Intrusion Response", IEEE Man Systems and Cybernetics Information Assurance Workshop, pp. 142-147, 2001

[3] Chebrolu, S.**,** Abraham A**.** and Thomas, J. P., "Feature Deduction and Ensemble Design of Intrusion Detection Systems," Computer & Security., Vol. 24, No. 4, pp. 295–307, 2005

[4] Chung, M.,O., "Host-based Intrusion Detection Systems adapted from Agent-based Artificial Immune Systems", Neuro Computing, Vol. 88, pp.78–86, 2012

[5] Debar, H., Becker, M. and Siboni, D., "A Neural Network Component for an Intrusion Detection System", Proc. 1992 IEEE Computer Society Symposium on Research in Computer Security and Privacy, pp. 240-250, 1992

[6] Depren , M.O., Topallar, M. , Anarim, E. and Ciliz, K., "Network-Based Anomaly Intrusion Detection System Using Soms," in Proc. IEEE 12th Signal Process. Communication Appl. Conf., pp. 76-79, Apr. 2004

[7] Fan W., Lee, W., Stolfo, S. , and Miller, M., "A Multiple Model Cost-Sensitive Approach for Intrusion Detection", Eleventh European Conference on Machine Learning ,2000

[8] Freund, Y. and Schapire, R. E.,"A Decision-Theoretic Generalization of online Learning and an Application to Boosting", Journal of Computer and System. Science, Vol. 55, No. 1, pp. 119–139, 1997

[9] George, A., "Anomaly Detection based on Machine Learning: Dimensionality Reduction using PCA and Classification using SVM", International Journal of Computer Applications, Vol.47, No.21, pp. 5-8, 2012

[10] Han, S.J. and Cho, S.B.,"Evolutionary Neural Networks for anomaly Detection based on the Behavior of a Program," IEEE Trans. Syst., Man, Cybern. B, Cybern. , Vol. 36, No. 3, pp. 559-570, 2006

[11] Hong, H., Xin-L.L., Li,Y. R., "Using Data Mining to Discover Signatures in Network-Based Intrusion Detection," in Proc. Int. Conf. Mach. Learn. Cybern., Vol. 1, pp. 13-17, 2002

[12] Hong, P., Zhang, D. and Wu,T., "An Intrusion Detection Method Based on Rough Set and SVM Algorithm," in Proc. Int. Conf. Communication., Circuits Syst., Vol. 2, pp. 1127-1130, 2004

[13] Uppal , H. A., Jabad, M. and Arshad, M.J. ,"An Overview of Intrusion Detection System (IDS) along with its Commonly Used Techniques and Classifications", International Journal of Computer Science and Telecommunications, Vol. 5, Issue 2,pp.20-24, Feb. 2014

[14] Khalkhali, I., Azmi, R., Mozhgan A.K. and Khansari, M., "Host-based Web Anomaly Intrusion Detection System, an Artificial Immune System Approach" International Journal of Computer Science Issues, Vol. 8, Issue 5, No 2,pp. 14-24, Sep. 2011

[15] Khannous , A., Rghioui A., Elouaai F. and Bouhorma M.,"A New Approach to Artificial Immune System for Intrusion Detection of the Mobile Ad Hoc Networks" International Journal of Computer Applications Vol. 92 ,No.15,pp. 50-53, April 2014

[16] Kim, J. and Bentley, P.J., "The Artificial Immune System for Network Intrusion Detection: An Investigation of Clonal Selection with a Negative Selection Operator" Submitted to CEC2001, the Congress on Evolutionary Computation, Seoul, Korea, May 27-30, 2001

[17] Lee,W.K., Stolfo, S.J. and Mok, W., "A Data Mining Framework for Building Intrusion Detection Model", Proceedings of the IEEE Symposium on Security and Privacy ,Oakland, CA: IEEE Computer Society Press, pp.120-132, 1999

[18] Lee,W.K., Stolfo, S.J.,"Data Mining Approaches for Intrusion Detection", Proceedings of the Seventh USENIX Security Symposium , San Antonio, TX, Vol. 7,pp.6, 1998

[19] Lei,J.Z and Ghorbani,A.,"Network Intrusion Detection Using an Improved Competitive Learning Neural Network", in Proc. 2nd Annu. Conf. Communication Network Serv. Res., Vol. 4, pp. 190-197, 2004

[20] Levin , I., Ziyyon, R.L. and Israe, "Results of the Kdd99 Classifier Learning Contest LLSoft's results overview," SIGKDD Explor., Vol.1, No. 2, pp. 67-75,2000

[21] Rassam M. A., Maarof, M. A. and Zainal A., "Intrusion Detection System Using Unsupervised Immune Network Clustering with Reduced Features" Int. J. Advance. Soft Com. Appl., Vol. 2, No. 3, 244-263, Nov., 2010

[22] Rassam, M. A. and Maarof, M. A., "Artificial Immune Network Clustering approach for Anomaly Intrusion Detection", Journal of Advances in Information Technology, Vol. 3, No.3, pp.147-154, Aug. 2012

[23] Sridevi, R. , Jagajothi, G. and Chattemvelli, R., "A PCA-AIS Approach for Intrusion Detection", International Journal of Computer Science and Telecommunications ,Vol. 3, Issue 7,pp.104-108, July 2012

[24] Tsong, S. H. , Tsung J. L. and Yuh J. L.,"A Three Tier IDS via Data Mining Approach", MineNet ,pp.212-217,2007

[25] Yanbin, Z. "Network Intrusion Detection System Model Based On Artificial Immune", International Journal of Security and Its Applications, Vol.9, No.9, pp.359-370, 2015.

