DISTRIBUTED ANOMALY INTRUSION DETECTION SYSTEM BASED ON MULTI-AGENTS

Arokia Renjit J.[1], **Shunmuganathan K.L**[2]

[1]Researh Scholar, Sathyabama University, TamilNadu, India – 600119.
[2]Professor & Head, Department of CSE, RMK Engineering College, TamilNadu, India – 601 206.
Email: [1]arokiarenjith@gmail.com

Abstract

Networks have the problem of security attacks like denial of service attacks and others. The firewalls and encrypted software's does not provide a complete security solution for those attacks. Network intrusion detection aims at distinguishing the attacks on the Internet from normal use of the Internet. It is an indispensable part of the information security system. Due to the variety of network behaviors and the rapid development of attack fashions, it is necessary to develop fast machine-learning-based intrusion detection algorithms with high detection rates and low false-alarm rates. In this paper, we have proposed an effective Intrusion Detection System in which local agent collects data from its own system and it classifies anomaly behaviors using SVM classifier. Each local agent is capable of removing the host system from the network on successful detection of attacks. The mobile agent gathers information from the local agent before it allows the system to send data. Our system identifies successful attacks from the anomaly behaviors. Experimental results show that the proposed system has high detection rate and low false alarm rate which encourages the proposed system.

Keywords: Mobile agents, Intrusion detection System, Network Security

I. INTRODUCTION

Intrusion detection is the process of monitoring the events occurring in a system or network and analyzing them for possible attacks or incidents which are violations of computer security policies. Intrusion prevention is the process of performing intrusion detection and attempting to stop detected possible incidents. Intrusion detection and prevention systems (IDPS) are primarily focused on identifying possible incidents, logging information about them, attempting to stop them and reporting them to security administrators. Also, organizations use IDS for identifying problems with security policies, documenting existing threats and deterring individuals from violating security policies. Intrusion detection system is a new network security technology in recent years [1]. Intrusion detection system can analyze and monitor system activity, identify and reflect the activity patterns that can be compared to discover attacks. Host based Intrusion detection and network based approaches are the two major variants of intrusion detection systems. Host based systems collect local data from sources internal to a computer, usually at the operating system level. This gives the advantage of collecting high quality data directly at the source. Unfortunately most of the attacks cannot be detected from a single location. Network based intrusion detection system monitors packets on the wire by setting the network interface to promiscuous mode and analyzing the network traffic. But network based intrusion systems suffer from scalability problems in case of high network traffic and have problems when encrypted communication is used. So recent approach is the development of distributed architectures, where sensors (host and network based) collect data, preprocess it and send it to a centralized analyzing station. But whenever one or more agents fails to follow the common set or rules due to spontaneous failure, tampering or even to malicious behaviors [2], the system's safety is under risk. Intrusion detection compares the set of baselines of the system with the current behavior of the system. It is assumed that normal and abnormal behaviors of the system can always be categorized. Anomaly detection and misuse or signature detection are the two techniques used for Intrusion detection system. Anomaly detection describes the abnormal patterns of behavior whereas misuse detection relies on the use of specifically known patterns of unauthorized behavior. These techniques rely on sniffing packets and using the sniffed packets for analysis. Anomaly detection compares the defined profiles against the actual usage patterns to detect abnormal activity patterns. These patterns will be considered as intrusions. However, Julish pointed out that data mining based intrusion detection usually relies on unrealistic assumptions on the availability and quality of training data [3], which causes detection models built on such training data to gradually lose efficiency in detecting intrusions as the real time environment undergoes continuous change. Instead of using static components in a IDS, mobile agent based systems has the advantages of overcoming network latency, reducing network load, autonomous execution, platform independence, dynamic adaptation and scalability issues.

II. RELATED WORKS

Denning [4] proposes a statistical method for intrusion detection. According to audit data, a profile is constructed to describe a given user or a given task. Several metrics are defined for the profiles. The Gaussian models of the metrics are constructed to detect intrusions. Li et al. [5] utilize statistical characteristics of n-grams to detect intrusions in the host system. Vigna and Kemmerer [6] use data that are sourced from network nodes, rather than the audit data, to construct profiles, enlightening the research on network-based intrusion detection. Caberera et al. [7] assume that the first derivative of the number of observed events in a time segment obeys the Poisson distribution, from which the Kolmogorov statistical values are extracted to measure the dissimilarity between observation network and normal behavior signals. Ye et al. [8] represent a sequence of events in time order as a Markov stochastic process. The joint probability for a particular sequence of events is used to distinguish between normal network behaviors and intrusions. In recent years, the hidden Markov model has been used in intrusion detection based on host audit data [9].

Bonifacio et al. [10] propose an NN for distinguishing between intrusions and normal behaviors. They unify the coding of categorical fields and the coding of character string fields in order to map the network data to an NN. Rapaka et al. [11] use execution numbers of system calls in a host machine as the features of network behaviors to train the NN. Zhang et al. [12] propose an approach for intrusion detection using hierarchical NNs. Han and Cho [13] use evolutionary NNs to detect intrusions.

Mukkamala et al. [14] use SVMs to distinguish between normal network behaviors and intrusions and further identify important features for intrusion detection. Mill and Inoue [15] propose the TreeSVM and ArraySVM for solving the problem of inefficiency of the sequential minimal optimization algorithm for the large set of training data in intrusion detection. Zhang and Shen [16] propose an approach for online training of SVMs for real-time intrusion detection based on an improved text categorization model. Han et al. [17] analyze the content for network data packages and use the data-mining techniques to acquire attack signatures. Qin and Hwang [18] propose an approach, which dynamically omits some non functionary frequent episode rules, as a supplement to the data-mining-based approaches. Otey et al. [19] propose a general-purpose outlier detection algorithm that works on mixed attribute data in distributed settings. Furthermore, they extend their algorithm to handle dynamic and streaming data sets.

Guan et al. [20] propose a K-means-based clustering algorithm, which is named Y-means, for intrusion detection. Xian et al. [21] combine the fuzzy K-means method and a clonal selection algorithm to detect intrusions. Jiang et al. [22] use the incremental clustering algorithm that is an extension of the K-means algorithm to detect intrusions.

Hoglund et al. [23] extract features that describe network behaviors from audit data, and they use the SOM to detect intrusions. Kayacik et al. [24] propose a hierarchical SOM approach for intrusion detection. Specific attention is given to the hierarchical development of abstractions, which is sufficient to permit direct labeling of SOM nodes with connection type. Sarasamma et al. [25] propose a hierarchical SOM for intrusion detection. They use the classification capability of the SOM on selected dimensions of the data set to detect anomalies.

III. PROPOSED SYSTEM

Our aim is to design and develop an intelligent Intrusion detection system based on anomaly detection method that would be accurate and low in false alarms. The proposed system has the local agent deployed in all the systems connected to the network. Local agent is responsible for detecting the local anomalies. Apart from detecting anomalies, the local agent shares this information with all other systems in the network through mobile agent. Mobile agent gathers information from local agent and decides on allowing the system to communicate with other systems in the network and thereby provides a global security solution.

A. Mobile Agent:

In order to use mobile agents, all the hosts in the networks must have an agent platform installed, where the agents are going to be executed. The mobile agent based systems has the advantages of overcoming network latency, reducing network load, autonomous execution, platform independence, dynamic adoption, static adaptation, scalability.

B. Local Agent:

Local agent is implemented in every system in the network which gathers information's about its system. The three main functions of local agent are

1. It monitors its own system and its environment dynamically. It uses SVM classifier to find out the local anomaly.
2. Whenever a node wants to transfer information to another node, it broadcasts the message to its neighboring nodes. It gathers neighboring nodes information using mobile agents. It then calls the SVM classifier to find out the attacks with the help of trained test data.
3. It provides same type of security solution throughout the network.

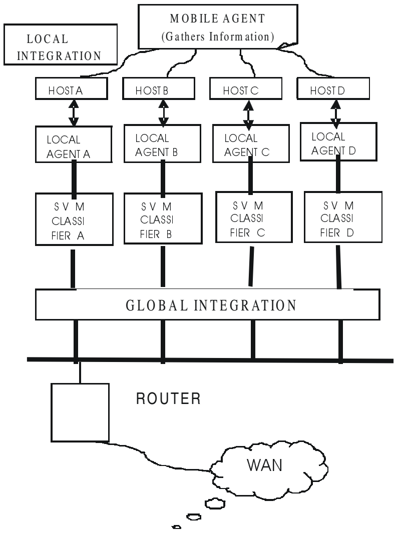

Fig 1. Proposed System Architecture

1. Local Monitoring in Current Node: Local agent is present in this system and it continuously monitors its own system. If an attacker packet arrives at this system to gather information, it calls SVM classifier to find out attacks. If an attack has been made, local agent will filter the respective system from the global networks.

2. Communication between nodes: Whenever any system transfers information to some other system in the network, it broadcast through intermediate systems. Before transferring message, it sends the mobile agent to the neighboring node gathers information from that node and it return back to the system. It then calls the SVM classifier to find out the attacks. If there is no suspicious activity, then it will forward the message to the neighboring node.

3. Feature extraction: Data collection module is included for each intrusion detection subsystem to collect the values of features, and then normal profile is created using the normal scenario and attack profile is created during the attack scenario.

4. Data preprocess: Data preprocess is a technique to process the information with the test train data. The audit data is stored in a file and it is smoothed so that it can be used for anomaly detection.

C. Local Integration:

Local Integration module concentrates on self system to find out the local anomaly attacks. Every systems under that network follow the same methodology to provide secure global networks.

D. Global Integration:

Global Integration module is used to find the intrusion result for entire network. It is used to find the status of neighboring nodes before taking decisions towards forwarding messages.

E. SVM Classifiers:

The SVM classifier used to classify the anomaly patterns is given in figure 2.

Support Vector Machines (SVM) are a set of related supervised learning methods that analyze data and recognize patterns, used for classification and regression analysis. Since SVM is a classifier, then given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that predicts whether a new example falls into one category or the other. An SVM

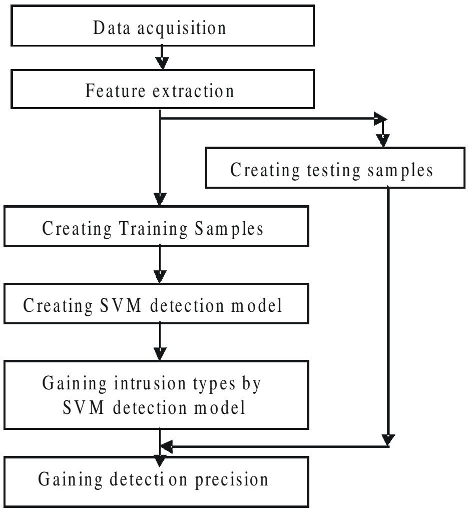
Data acquisition
Feature extraction
Creating testing samples
Creating Training Samples
Creating SVM detection model
Gaining intrusion types by SVM detection model
Gaining detecti on precision

Fig 2. Classifier Architecture

model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on.

Training mode:

Input: The file containing the features values logged during the learning phase

Output: files containing the mean, standard deviations and inverse matrices of feature set

Begin

```
for $i = 1$ to Num. of week days do
for $j = 1$ to Num. of hours in a day do
Read the feature values logged during learning phase;
for $k = 1$ to Num. of network features do
find sum of the values corresponding to the same hour and day of the week;
Compute Average values and standard deviation for each feature;
Compute
$\sum_{m=1}^{n} (X_1 - \mu)(x_m - \mu)^T$
where $n$ is the total number of features
Compute the Determinant of above covariance matrices
if Determinant $\leq 0$
```

Consider the neighbouring covariance matrix having positive Determinant

Compute inverse matrix corresponding to each covariance matrix

End

Detection mode:

Input : The file containing the network profile

Output : Sends alert in case a event is detected as intrusion

Begin

```
for $i = 1$ to Num .of week days do
for $j = 1$ to Num. of hours in a day do
for $k = 1$ to Num. of network features do
Read Average values and standard deviation for each feature;
Read the inverse matrices
Read the determinant matrix corresponding to each inverse matrix
Compute $(\mu + \sigma)$ for each parameter
If $(\mu - \sigma > x > \mu - \sigma)$ then
$x$ is intrusive
Compute $T^2 = (X - \mu)\, S^{-1}\, (X - \mu)^T$
If $T^2$ exceeds, the threshold flag alerts
Compute $g_t(X) = -1/2\ ln\,[S] - 1/2\,(X - \mu)^T S^{-1}\,(X - \mu) + ln\ p\,(l)$
If $g_1(X)$ exceeds the threshold flag alerts.
End
```

IV. EXPERIMENTATION RESULTS

This system not only blocks the security threats at the application level, but also stops some of the threats at the network level. Our results are compared with other recently published results in Table 1. which shows the proposed system is greatly competitive with others.

TABLE 1. Results Comparison

Methods	FPR(%)	DR(%)
Genetic Clustering [15]	0.3	79
Hierarchal SOM [16]	2.19–3.99	90.94–93.46
Proposed System	5 - 9	89 - 98

The detection rate of anomaly in our proposed system is high and it encourages the system. The percentage of anomaly detection is calculated as follows.

Percentage of Anomaly detection

$$= \frac{\text{No. of Predicted abnormal class}}{\text{Total No. of traces}} \times 100$$

This system can act as Intrusion prevention system to detect and prevent the attacks. This system can be able to stop a number of attacks as well as the false positive rate of the proposed system is low. The proposed system is compared with existing system which uses Bayesian classifier program for training the classifier for anomaly detection. Our system uses SVM classifier for classification and train the data set. The proposed system is tested by introducing attacks in the network and finding out the detection rate over time period and it is compared with the anomaly detection rate of the existing system which uses Bayesian classifier. It is inferred from the results that our system has higher anomaly detection rate.

Our proposed system proved strong in detecting anomalies using agents in a distributed manner with the help of local agents and coordination among these agents are achieved using mobile agent. The information's required to classify anomalies are shared among the neighboring nodes before sending packets and thereby ensuring safety during communication between the network systems and therefore our intrusion detection system proves strong.

V. CONCLUSION

This paper provides a strong platform to detect anomalies. The proposed system is cooperative and distributive; it considers the anomaly detection result from the neighbor nodes and sends the current nodes result to its neighbor nodes. Our system also could differentiate congestive packet loss from malicious packet loss using a packet loss minimization algorithm implemented in routers by accessing the traffic rates and buffer sizes. Experimental results show that anomaly detection rate is higher when compared to existing mechanism. This mechanism proves strong in places where traditional security mechanisms like IDS and firewall have not been sufficient to provide security of networks.

ACKNOWLEDGEMENT

We take immense pleasure in thanking our Chairman Dr. Jeppiaar M.A, B.L, Ph.D, the Directors of Jeppiaar Engineering College Mr. Marie Wilson, B.Tech, MBA., (Ph.D) Mrs. Regeena Wilson, B.Tech, MBA., (Ph.D) and the Principal Dr. Sushil Lal Das M.Sc (Engg.), Ph.D for their continual support and guidance. We would like to extend our thanks to my guide, our friends and family members without whose inspiration and support our efforts would not have come to true. Above all, we would like to thank God for making all our efforts success.

REFERENCES

[1] Yixue Wang, A sort of Mult-Agent Cooperation Distributed Based Intrusion Detection System, Modern computer, 2008.

[2] Mukkamala. R.J. Gagnon and S. Jaiodia Integrating data mining techniques with intrusion detection methods. *Research Advances in Database and Information systems security*, 33-46, 2000.

[3] K. Julish, "*Data mining for intrusion detection: A critical review*", IBM, Feb 2002.

[4] D. Denning, "An intrusion detection model", *IEEE Trans. Softw. Eng.*, vol. SE-13, no. 2, pp. 222–232, Feb. 1987.

[5] Z.W. Li, A. Das and S. Nandi, "Utilizing statistical characteristics of *N*-grams for intrusion detection," in *Proc. Int. Conf. Cyberworlds*, Dec. 2003, pp. 486–493.

[6] G. Vigna and R.A. Kemmerer, "NetSTAT: A network-based intrusion detection approach," in *Proc. Comput. Secur. Appl. Conf.*, Dec. 1998, pp. 25–34.

[7] J.B.D. Caberera, B. Ravichandran and R.K. Mehra, "Statistical traffic modeling for network intrusion detection," in *Proc. Model., Anal. Simul. Comput. Telecommun. Syst.*, 2000, pp. 466–473.

[8] N. Ye, Y. Zhang and C. M. Borror, "Robustness of the Markov-chain model for cyber-attack detection," *IEEE Trans. Rel.*, vol. 53, no. 1, pp. 116–123, Mar. 2004.

[9] D. Yeung and Y. Ding, "Host-based intrusion detection using dynamic and static behavioral models," *Pattern Recognit.*, vol. 36, no. 1, pp. 229–243, Jan. 2003.

[10] J.M. Bonifacio, Jr., A.M. Cansian, A.C.P.L.F. De Carvalho and E. S. Moreira, "Neural networks applied in intrusion detection systems," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, 1998, vol. 1, pp. 205–210.

[11] A. Rapaka, A. Novokhodko and D. Wunsch, "Intrusion detection using radial basis function network on sequences of system calls," in *Proc. Int. Joint Conf. Neural Netw.*, 2003, vol. 3, pp. 1820–1825.

[12] C. Zhang, J. Jiang and M. Kamel, "Intrusion detection using hierarchical neural networks," *Pattern Recognit. Lett.*, vol. 26, no. 6, pp. 779–791, May 2005.

[13] S. J. Han and S. B. Cho, "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, Jun. 2006.

[14] S. Mukkamala, G. Janoski and A. H. Sung, "Intrusion detection using neural networks and support vector machines," in *Proc. Int. Joint Conf. Neural Netw.*, 2002, vol. 2, pp. 1702–1707.

[15] J. Mill and A. Inoue, “Support vector classifiers and network intrusion detection,” in *Proc. Int. Conf. Fuzzy Syst.*, 2004, vol. 1, pp. 407–410.

[16] Z. Zhang and H. Shen, “Online training of SVMs for real-time intrusion detection,” in *Proc. Int. Conf. Adv. Inf. Netw. Appl.*, 2004, vol. 1, pp. 568–573.

[17] H. Han, X. L. Lu and L. Y. Ren, “Using data mining to discover signatures in network-based intrusion detection,” in *Proc. Int. Conf. Mach. Learn. Cybern.*, 2002, vol. 1, pp. 13–17.

[18] M. Qin and K. Hwang, “Frequent episode rules for Internet anomaly detection,” in *Proc. IEEE Int. Symp. Netw. Comput. Appl.*, 2004, pp. 161–168.

[19] M. E. Otey, A. Ghoting and S. Parthasarathy, “Fast distributed outlier detection in mixed-attribute data sets,” *Data Min. Knowl. Discov.*, vol. 12, no. 2/3, pp. 203–228, May 2006.

[20] Y. Guan, A. A. Ghorbani and N. Belacel, “*Y* -means: A clustering method for intrusion detection,” in *Proc. IEEE Can. Conf. Electr. Comput. Eng.*, 2003, vol. 2, pp. 1083–1086.

[21] J. Xian, F. Lang and X. Tang, “A novel intrusion detection method based on clonal selection clustering algorithm,” in *Proc. Int. Conf. Mach. Learn. Cybern.*, 2005, vol. 6, pp. 3905–3910.

[22] S. Jiang, X. Song, H.Wang, J. Han and Q. Li, “A clustering-based method for unsupervised intrusion detections,” *Pattern Recognit. Lett.*, vol. 27, no. 7, pp. 802–810, May 2006.

[23] A.J. Hoglund, K. Hatonen and A. S. Sorvari, “A computer hostbased user anomaly detection system using the self-organizing map,” in *Proc. Int. Joint Conf. Neural Netw.*, 2000, vol. 5, pp. 411–416.

[24] H. G. Kayacik, A. N. Zincir-Heywood and M. I. Heywood, “On the capability of an SOM based intrusion detection system,” in *Proc. Int. Joint Conf. Neural Netw.*, Jul. 2003, vol. 3, pp. 1808–1813.

[25] S.T. Sarasamma, Q.A. Zhu and J. Huff, “Hierarchical Kohonenen net for anomaly detection in network security,” *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 35, no. 2, pp. 302–312, Apr. 2005.

J. Arokia Renjit B.E., M.E., (Ph.D) works as Associate Professor in Jeppiaar Engineering College and he has more than 9 years of teaching experience. His areas of specializations are Networks, Artificial Intelligence, Software Engineering.