INTERNATIONAL JOURNAL FOR DEVELOPMENT IN COMPUTER SCIENCE & TECHNOLOGY ISSN-2320-7884 (ONLINE)
VOLUME-1, ISSUE-II IS NOW AVAILABLE AT: WWW.IJDCST.COM ISSN-2321-0257 (PRINT)

# Radial Basis Function Driven Intrusion Detection Systems

**Mounika N.K #1, CH. Venkata Phani krishna#2,**

**#1, Student, K L University, Vaddeswaram, Guntur (dt),**

**#2, Associate professor, K L University, Vaddeswaram, Guntur (dt),**

**ABSTRACT: In this paper, we use 'Radial Basis Functions' for the task of Intrusion Detection. In this paper, we propose a system that combining both misuse and anomaly attacks. Intrusion detection systems have become a key component in ensuring the safety of systems and networks. As networks grow in size and speed continues to increase, it is crucial that efficient scalable techniques should be developed for Intrusion Detection systems. Hybrid intrusion detection is a novel kind of model combining the advantages of anomaly detection and misuse detection. The signature-based and the anomaly-based systems and is known as the Hybrid intrusion System We design a new hybrid intrusion system based on Radial Basis Functions. The proposed model is promising in terms of detection accuracy and computational efficiency.**

## I INTRODUCTION

Intrusion detection is defined as "the problem of identifying individuals who are using a computer system without authorization and those who have legitimate access to the system but are abusing their privileges. Intrusion detection techniques are traditionally categorized into two methodologies: anomaly detection and signature or misuse detection. Anomaly detection is based on the normal behavior of a subject (e.g., a user or a system); any action that significantly deviates from the normal behavior is considered intrusive. Misuse detection catches intrusions in terms of the characteristics of known attacks or system vulnerabilities; any action that conforms to the pattern of a known attack or vulnerability is considered intrusive.

In this paper we propose and evaluate the use of the Radial Basis Functions, a novel technique for the task of Intrusion Detection, and experimentally show that they have higher detection accuracy when compared to any other technique for the same task. Our system is a Hybrid system as it builds a model based on both the normal and the anomalous data and hence has the advantages of both the Signature based and the anomaly based systems. The Signature based systems build a model based on the available knowledge of the attacks and hence extract out signatures which are used to build a classifier to detect same or similar patterns when deployed online. This is also known as Misuse Detection. Complementary to these are the behavior based systems which build a model based on the available knowledge of the normal use of the system. Once deployed, they classify any significant deviation from the learnt behavior as attack. This is also called as Anomaly Detection.

Once we have data then we train our system using Radial Basis Function. Our Hybrid intrusion detection is a novel kind of model combining the advantages of anomaly based intrusion detection and signature based intrusion detection. Intrusion and anomalies are two different kinds of abnormal traffic events in an open network environment. The newly proposed HIDS is designed to solve problems with much enhanced performance.

## II RELATED WORK

The field of Intrusion Detection and Network Security is not new and a number of methods have been proposed and a number of systems have been built to detect intrusions. We now briefly discuss some of the techniques with regards to the task of Intrusion Detection.

Decision Trees have also been used for Intrusion Detection. The Decision Trees method selects the best attribute for each decision node during the construction of the tree based on some well defined criteria. One such criterion is to use the gain ratio as used in C4.5. Decision Trees can be easily used for building the Misuse Detection Systems, but, it is very difficult to construct Anomaly Detection System. Though, the Neural Networks can work effectively with noisy data, they require large amount of data during training and it is often hard to select the best possible Neural Network architecture. Support Vector Machines which maps real valued input feature vector to higher dimensional feature space through nonlinear mapping have been used for detecting intrusions. The Support Vector Machines provide real time detection capability and can deal with large dimensionality of data. However, they are used effectively for binary-class classification only. Along with these, other techniques for detecting intrusion includes the use of Genetic Algorithms, Autonomous Agents for Intrusion Detection and Probabilistic Agent based Intrusion Detection.

Association rules and frequent episodes are used to learn the record patterns that describe user behavior. These methods can deal with symbolic data, and the features can be defined in the form of packet and connection details. However, mining of features is limited to entry level of the packet and requires the number of records to be large and sparsely populated; otherwise, this approach, tend to produce a large number of rules that increase the complexity of the system.

One of the main drawbacks of the clustering technique is that it is based on calculating numeric distance between the observations, and hence, the observations must be numeric. Observations with symbolic features cannot be easily used for clustering; resulting in inaccuracy .In addition, the clustering methods consider the features independently and are unable to capture the relationship between different features of a single record, which further degrades attack detection accuracy. Naive Bayes classifiers have also been used for intrusion detection. However, this system, make strict independence assumption between the features in an observation resulting in lower attack detection accuracy when the features are correlated, which is often the case for intrusion detection.

## III RADIAL BASIS FUNCTION

There are two categories of training algorithms: supervised and unsupervised. RBF networks are used mainly in supervised applications. In a supervised application, we are provided with a set of data samples called training set for which the corresponding network outputs are known. In this case the network parameters are found such that they minimize a cost function:

$$\min \Sigma^{Q}_{i=1} (Y_K (X_i) - F_K (X_i))^T ((Y_K (X_i) - F_K (X_i))$$

where Q is the total number of vectors from the training set, Yk (X i ) denotes the RBF output vector and Fk (X i ) represents the output vector associated with the a data sample X i from the training set. In unsupervised training the output assignment is not available for the given set. A large variety of training algorithms has been tested for training RBF networks.

In order to use a Radial Basis Function Network we need to specify the hidden unit activation function, the number of processing units, a criterion for modeling a given task and a training algorithm for finding the parameters of the network. Finding the RBF weights is called network training. If we have at hand a set of input-output pairs, called training set, we optimize the network parameters in order to fit the network outputs to the given inputs. The fit is evaluated by means of a cost function, usually assumed to be the mean square error. On-line training algorithms adapt the network parameters to the changing data statistics. RBF networks have been successfully applied to a large diversity of applications including interpolation, chaotic time-series modeling and soon

## IV HYBRID INTRUSION DETECTION SYSTEM

. The Hybrid Intrusion Detection System integrates the flexibility of Anomalous Detection System with the accuracy of a signature-based Intrusion Detection System. Anomalous Detection System is designed by acquiring the volatile data when there is no any anomalous activity. Here we train our system using Radial Basis Function for the anomalous activity. This new approach automatically enables HIDS to detect similar anomalous attacks in the future. In this hybrid approach we design both the anomaly based detection system and signature based detection system. Here the system is running in parallel. In this system we have two main blocks one is signature based detection and other is anomaly based detection. For the signature based intrusion detection we use standard dataset and in anomaly based intrusion detection we collect data from the system when there is no anomalous activity. In this we select layers which are corresponds to the particular attacks and for these attacks again we select features from the dataset. Then during the system training we use conditional random fields for labeling these features as normal or attack.

For this we collect the volatile data from the system when there is no any anomalous activity. For this data collection we investigate user patterns, such as profiling the programs executed daily or the privileged processes executed with access to resources that are inaccessible to ordinary user. Once we have data then we train our system using Radial Basis Function. Hybrid intrusion detection is a novel kind of model combining the advantages of anomaly based intrusion detection and signature based intrusion detection.

### *A) SIGNATURE BASED INTRUSION DETECTION*

The collected signatures are used to match with incoming traffic to detect intrusions. These are conventional systems that detect known attacks with low false alarms. A signature-based intrusion detection system employs a priori knowledge of attack signatures. The signatures are manually constructed by security experts by analyzing previous attacks. However, the signature-based IDS cannot detect unknown attacks without any pre collected signatures or lack of attack classifiers. In this system we consider four different attacks:

Probe Layer: The probe attacks are aimed at acquiring information about the target network from a source that is often external to the network. e.g. port scanning.

DoS: Denial of service attacks are meant to force the target to stop the service(s) that is (are) provided by flooding it with illegitimate requests. e.g. SYN flood, land attack;

R2L Layer: The Root to Local attacks are one of the most difficult to detect as this system, involve the network level and the host level features. e.g. guessing password;

U2R Layer: The U2R attacks involve the semantic details that are very difficult to capture at an early stage. Such attacks are often content based and target an application e.g. various buffer overflow attacks.

### *B) ANOMALY DETECTION*

Intrusion detection is the process of monitoring computers or networks for unauthorized entrance, activity, or file modification. Anomaly-based IDS establish a baseline of normal usage patterns, and anything that widely deviates from it gets flagged as a possible intrusion. Anomaly detection method can investigate user patterns, such as profiling the programs executed daily or the privileged processes executed with access to resources that are inaccessible to ordinary user. The major advantage of anomaly-based IDS is their ability to detect attempts to exploit new and unforeseen vulnerabilities. Their complexity, due to the inherently dynamic nature of computer networks, is their major disadvantage, as well as their high false alarm rate because the entire scope of the behavior of an information system may not be covered during the learning phase. To overcome this drawback we train our system using Radial Basis Function as they are more accurate and due to this false alarm rate goes on decreasing.

## V PERFORMANCE

. If we consider all the 41 features given in the data set, we find that the time required to train and test the model is high. To address this, our integrated system is implementing a four-layer

system. The four layers correspond to Probe, DoS, R2L, and U2R. For each layer, we then selected a set of features that is sufficient to detect attacks at that particular layer. Feature selection for each layer enhances the performance of the entire system. The runtime (testing) performance of our model is comparable with other methods; however, the time required to train the model is slightly higher. We also observe that feature selection not only decreases the time required to test an instance, but it also increases the accuracy of attack detection. This is because using more features than required can generate superfluous rules often resulting in fitting irregularities in the data, which can misguide classification. the main strength of our method lies in detecting the R2L and the U2R attacks, which are not satisfactorily detected by other methods. Our method gives slight improvement for detecting Probe attacks and was similar in accuracy when compared with other methods for detecting the DoS attacks. The prime reason for better detection accuracy for the RBF is that they do not consider the observation features to be independent. RBF evaluate all the rules together, which are applicable for a given observation. This results in capturing the correlation among different features of the observation resulting in higher accuracy.

## VI CONCLUSION

The Radial Basis Function can be effectively used for the task of Intrusion Detection. RBFs are very effective in improving the attack detection rate and decreasing the false alarm rate. In this paper we proposed hybrid intrusion detection system which is accurate for the intrusion detection. It detects the known as well as unknown attack. The signature-based systems can be deployed at the periphery of a network to filter out attacks that are frequent and previously known, leaving the detection of new unknown attacks for anomaly systems. Our integrated system can effectively and efficiently detect such attacks like U2R and R2L. Our system can help in identifying an attack once it is detected at a particular layer, which expedites the intrusion response mechanism, thus minimizing the impact of an attack. our system has the advantage that the number of layers can be increased or decreased depending upon the environment in which the system is deployed, giving flexibility to the network administrators.

## VII REFERENCES

[1] Park, J., Sandberg, J. W. (1991), "Universal approximation using radial basis functions network," *Neural Computation,* vol. 3, pp. 246-257.

[2] K.K. Gupta, B. Nath, R. Kotagiri, and A. Kazi, ―Attacking Confidentiality: An Agent Based Approach,‖ Proc. IEEE Int'l Conf. Intelligence and Security Informatics (ISI '06), vol. 3975, pp. 285-296,2006 .

[3] T. Abraham, IDDM: Intrusion Detection Using Data Mining Techniques, http://www.dsto.defence./gov.au/publications/ 2345/ DSTO-GD-0286.pdf, 2008.

[4] H. Shah, J. Undercoffer, and A. Joshi, ―Fuzzy Clustering for Intrusion Detection,‖ Proc. 12th IEEE Int'l Conf. Fuzzy Systems (FUZZ-IEEE '03), vol. 2, pp. 1274-1278, 2003.

[5] N.B. Amor, S. Benferhat, and Z. Elouedi, ―Naive Bayes vs. Decision Trees in Intrusion Detection Systems,‖ Proc. ACM Symp. Applied Computing (SAC '04), pp. 420-424, 2004.

[6] Qingqing Zhang, Hongbian Yang, Kai Li ―Research on the Intrusion Detection Technology with Hybrid Model‖,2010 2nd Conference on Environmental Science and Information Application Technology,978-1-4244-7388-5/10 ,ESIAT,pp. 646-649,2010

[7] Bors, A.G., Pitas, I., (1996) "Median radial basis functions neural network," *IEEE Trans. on Neural Networks,* vol. 7, no. 6, pp. 1351-1364. Casdagli, M. (1989) "Nonlinear prediction of chaotic time series," *Physica D,* vol. 35, pp. 335-356.

[8] Leung K., C. Leckie, Unsupervised anomaly detection in network intrusion detection using clusters, In Proceedings of the Twenty-eighth Australasian conference on Computer Science - Volume 38, Newcastle, Australia, 2005, pp. 333 – 342.

[9] Kai Hwang, Fellow, IEEE, Min Cai, Member, IEEE, Ying Chen and Min Qin,‖ Hybrid Intrusion Detection with WeightedSignature Generation over Anomalous Internet Episodes‖, IEEE Transactions On Dependable and SecureComputing,Vol.4,No.1, pp.41-55, JAN-MAR 2007.

[10] Sekar R., M. Bendre, P. Dhurjati, D. Bullineni, A fast automaton-based method for detecting anomalous program behaviours, IEEE Symposium on Security and Privacy, 2001, S&P 2001, pp. 144 – 155.

[11] A. Lazarevic, L. Ertoz, V. Kumar, A. Ozgur, and J. Srivastava, ―A Comparative Study of Anomaly Detection Schemes in Network Intrusion Detection,‖ Proc. Third SIAM Conf. Data Mining, 2003, http://www.users.cs.umn.edu/~kumar/papers.