International Journal of Modern Engineering Research (IJMER)
www.ijmer.com Vol. 3, Issue. 5, Sep - Oct. 2013 pp-2770-2772 ISSN: 2249-6645

Layered Approach & HMM for Network Based Intrusion Dection

Archana patil, A. T. Bhole

ABSTRACT: *In this, we are using two techniques together as signature based and anomaly based called as Hybrid technique. Anomaly detection, where the strategy is to suspect of what is considered an unusual activity for the subject (users, processes, etc.) and carry on further investigation. This approach is particularly effective against novel (i.e. previously unknown) attacks. Signature based detection systems detects previously known attack in a timely and efficient way. The main issue of this approach is that in order to detect an intrusion this must to be previously detected. This Hybrid technique gives better result than signature based and anomaly based technique. Also we are using here layered approach to get result faster ,because in layered approach we have different four layers as prob,U2R,R2L,DOS and we assigned different features to different layer so that if any layer find attack at that layer that attack will fix ,that attack should not go further .Main aim of this paper is to increase accuracy and efficiency .*

Index Terms: *Intrusion detection, Layered Approach, Hidden Markov Model, network security, decision trees, naive Bayes.*

I. INTRODUCTION

Intrusion detection is defined as ``the problem of identifying individuals who are using a computer system without authorization (i.e., `crackers') and those who have legitimate access to the system but are abusing their privileges (i.e., the `insider threat')".Also we can say that the identification of attempts to use a computer system without authorization or to abuse existing privileges. According to Heady et al. where an intrusion is defined as ``any set of actions that attempt to compromise the integrity, confidentiality, or availability of a resource", disregarding the success or failure of those actions.[12] The definition of an intrusion detection system does not include preventing the intrusion from occurring, only detecting it and reporting it to an operator.[12]

There are two types of intrusion detection depending on way their components are distributed

1. A centralized intrusion detection system is one where the analysis of the data is performed in a fixed number of locations, independent of how many hosts are being monitored. We do not consider the location of the data collection components, only the location of the analysis components.

Eg: IDES, IDIOT .

2. A distributed intrusion detection system is one where the analysis of the data is performed in a number of locations proportional to the number of hosts that are being monitored. Again, we only consider the locations and number of the data analysis components, not the data collection components .

Eg: DIDS, GrIDS.

Also Intrusion detection is divided into :

1. Anomaly detection, where the strategy is to suspect of what is considered an unusual activity for the subject (users, processes, etc.) and carry on further investigation. This approach is particularly effective against novel (i.e. previously unknown) attacks. Its main drawback is the high rate of false positives, because any legitimate but new activity can rise an alert.
2. Signature detection, where the strategy is to look for some special activity (signature) of previously known attacks. Signature based detection systems detects previously known attack in a timely and efficient way. The main issue of this approach is that in order to detect an intrusion this must to be previously detected.

Previously there is only one technique is used at a time but In this we are using both as signature based and anomaly based combine called as hybrid based technique .That is we are developing hybrid system using HMM based layered approach for NIDS. We also integrate the Layered Approach with the HMMs to gain the benefits of computational efficiency and high accuracy of detection in a single system. By using this we get fast result because we are using layered approach .Layered approach means we have different four layers as PROBE , DOS , U2R ,R2L and for every layer different different features are assigned and whenever we got some malicious attack that attack must be detected at that moment ,that attack should not go further. Due to this technique speed of our operation increase.

A hidden Markov model(HMM) is a statistical generative model in which the system being modelled is assumed to be a Markov process with unobserved state. An HMM can be considered as the simplest dynamic Bayesian network. An HMM is like a finite state machine in which not only transitions are probabilistic but also output. An HMM is a doubly stochastic process with an underlying stochastic process that is not observable, and can only be observed through another set of stochastic processes that produce the sequence of observed symbols . HMM is a useful tool to model sequence information. This model can be thought of as a graph with N nodes called ‘state’ and edges representing transitions between those states. Each state node contains initial state distribution and observation probabilities at which a given symbol is to be observed. An edge maintains a transition probability with which a state transition from one state to another state is made.

II. RESEARCH ELABORATION

Layered Approach

The goal of using a layered model is to reduce computationand the overall time required to detect anomalous events. The time required to detect an intrusive event is significant and can be reduced by eliminating the communication overhead among different layers. This can be achieved by making thelayers autonomous and self-sufficient to block an attack without the need of a central decision-maker. Every layer in the LIDS framework is trained separately and then deployed sequentially. We define four layers that correspond to the four attack groups mentioned in the data set. They are Probe layer, DoS layer, R2L layer, and U2R layer. Each layer is then separately trained with a small set of relevant features. Feature selection is significant for Layered Approach and discussed in the next section. In order to make the layers independent, some features may be present in more than one layer. The layers essentially act as filters that block any anomalous connection.We have four different attacks probe attack,dos attack, u2r attack,r2l attack corresponding to four different layers.As Probe layer, R2L layer,U2R layer,DOS layer.

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

DoS Layer : The DoS attacks are meant to force the target to stop the service(s) that is (are) provided by flooding it with illegitimate requests.

R2L Layer : R2L attacks are one of the most difficult to detect as they involve the network level and the host level features.

U2R Layer:The U2R attacks involve the semantic details that are very difficult to capture at an early stage.

Decision tree

Decision tree builds classification or regression models in the form of a tree structure.

Dataset is a collection of data, usually presented in a tabular form. It breaks down a dataset into smaller and smaller subsets while at the same time an associated decision tree is incrementally developed.A decision tree is composed of three basic elements:

1. A decision node specifying a test attributes.
2. An edge or a branch corresponding to the one of the possible attribute values which means one of the test attribute outcomes.
3. A leaf which is also named an answer node contains the class to which the object belongs.

Naive Bayesian

The Naive Bayesian classifier is based on Bayes' theorem with independence assumptions between predictors. A Naive Bayesian model is easy to build, with no complicated iterative parameter estimation which makes it particularly useful for very large datasets. Despite its simplicity, the Naive Bayesian classifier often does surprisingly well and is widely used because it often outperforms more sophisticated classification methods. [6]

A naive Bayes classifier assumes that the presence (or absence) of a particular feature of a class is unrelated to the presence (or absence) of any other feature, given the class variable.

III. RESULT

Table I : Comparision Between Difference Techniques

Technique	Detection Rate in percentage	Time Required
Layered Approach	98.71740059854639	Less than 1 sec.
Hidden Markov Process	95.05439161966156	12 sec
Decision Tree	95.0544	2 sec
Navie Bayes	93.5133	4sec

From this , layered approach gives very high detection rate and time required for detecting attack is also less.

IV. CONCLUSION

In this ,we can detect intrusion detection fast and accurately .Layered HMMs can be very effective in detecting the Probe, the U2R, and the R2L attacks as well as the DoS attacks. However, 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, we performed experiments with our integrated system by 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. By using layered approach we get high accuracy () and also time required for it is also less() than other two techniques HMM and WEKA (decision tree and navie bayes).

ACKNOWLEDGMENTS

The authors sincerely thank the anonymous reviewers whose comments have greatly helped clarify and improve this paper.

REFERENCES

[1] Kapil Kumar Gupta, Baikunth Nath, "Layered Approach Using ConditionalRandom Fields for Intrusion Detection" *IEEE Transaction On Dependable and Secure Computing*, Vol. 7, No. 1, January –March 2010.

[2] T. Abraham "IDDM: Intrusion Detection Using Data Mining Techniques" *International Journal of Network Security & Its Applications (IJNSA)*, Vol. 2, No. 2. 2008

[3] N.B. Amor, S. Benferhat, and Z. Elouedi, "Naive Bayes vs. Decision Trees in Intrusion Detection Systems", *Proc. ACM Symp.Applied Computing* , 2004.
[4] K.K. Gupta, B. Nath, and R. Kotagiri, "Conditional Random Fields for Intrusion Detection," *Proc. 21st Int'l Conf. Advanced Information Networking and Applications Workshops*, 2007.
[5] Yusufovna, S.F. "Integrating Intrusion Detection System and Data Mining" *IEEE Computer Society Washington* , Oct. 2008 , 978-0-7695-3427-5.
[6] Christopher Kruegel Darren Mutz William Robertson Fredrik Valeu " Bayesian Event Classification for Intrusion Detection" *Reliable Software Group University of California, Santa Barbara*
[7] "Understanding Intrusion Detection Systems *'' SANS Institute InfoSec Reading Room.*
[8] Ozalp Babaoglu "IDS:Intrusion Detection Systems", *Babaoglu* 2006.
[9] Y. Du, H. Wang, and Y. Pang "A Hidden Markov Models-Based Anomaly Intrusion Detection Method", *Intellegent Control and Automation* 2004.
[10] L. Portnoy, E. Eskin, and S. Stolfo, "Intrusion Detection with Unlabeled Data Using Clustering," *Proc. ACM Workshop Data Mining Applied to Security* (DMSA), 2001
[11] Y.-S. Wu, B. Foo, Y. Mei, and S. Bagchi, "Collaborative Intrusion Detection System (CIDS): A Framework for Accurate and Efficient IDS," *Proc. 19th Ann. Computer Security Applications Conf.* , 234-244, 2003.
[12] Autonomous Agents for Intrusion Detection, http://www.cerias.purdue.edu/research/aafid/, 2010.
[13] Probabilistic Agent Based Intrusion Detection, http://www.cse.sc.edu/research/isl/agentIDS.shtml, 2010.
[14] Shaik Akbar Dr.K.Nageswara Rao Dr.J.A.Chandulal "Intrusion Detection System Methodologies Based on Data Analysis" International Journal of Computer Applications , August 2010
[15] Wei Wang, Xiaohong Guan, Xiangliang Zhang, Liwei Yang "Profiling program behavior for anomaly intrusion detection based on the transition and frequency property of computer audit data" Center for Networked Systems and Information Security (CNSIS) and State Key Laboratory for Manufacturing Systems Engineering (SKLMSE), Xi'an Jiaotong University, Xi'an 710049, China,2006.
[16] Sam Drazin and Matt Montag "Decision Tree Analysis using Weka", *University of Miami.*