Implementation Of Intrusion Detection System Through Data Mining

Rakesh Yadav
Research Scholar
Department of computer scienc & Engg.
JIT Borawan ,India

Mahesh Malaviya Research Scholar
Prof. & Head
Department of Comp. Sci. & Engg.
JIT Borawan ,India

Abstract

Security is major issue now in these days in different application level as well as in the network level applications and utilities. This paper is based on a new approach based on process mining. In daily use we use various computer based application and interacted through different processes. Some of the process is well known and they provide support for smart works. But some processes are malicious and interrupting different kinds of applications, in this project we are going to introduce the malicious processes classification for using it over IDS development. For that purpose we make efforts for analysing different processes collected from the server to client's machines.

1. Introduction

An intrusion detection system (IDS) inspects all inbound and outbound network activity and identifies suspicious patterns that may indicate a network or system attack from someone attempting to break into or compromise a system. In this era of technology their various decisions are taken from the machine, this decision making capability of any machine is developed through machine learning concept. Using the machine learning concept we make such algorithms by which machine make intelligences decisions from the previously provided examples and learning sets.

Computer network is the source of transmission between the machines, important information travels across the network and on the global network (Internet) but that information is accessible [1,2]. Network has become very risky and unsecure, precarious security is provided but every internet user wants to secure concern network up to optimal and acceptable extend, any machine that is connected to the global network (Internet) directly or under another domain, it has security threats [1,3]. The security issues of preceding environment computer

Network and internet have become obligatory; machines connected to the network are easily assessable [2, 3]. Firewalls and routers are used to detect massive kind of virus and worms but it is possible when the virus or worm is already defined in signatures and it can also be detected when the virus is already spread [4]. In order to spot these suspicious actions we use Intrusion Detection Systems (IDSs) and Intrusion Detection System is usually categorized as misuse based system or anomaly based system [6, 7, 8, 11]. Further it can be classified as hybrid based system [10]. Intrusion Detection System (IDS) is normally practiced for identifying malicious activities and their resources. Misuse based system maintains records for description of attacks and signatures that are used to detect the attacks where anomaly based system has feature of detecting previously unknown attacks [1]. Hybrid based system perform intrusion detection by using expertise of both misuse based and anomaly based system [5, 10]. Some of appropriate techniques based on Intrusion Detection System (IDS) are analyzed in detail, in this paper we discussed advantages and shortcomings of analyzed techniques that are still in use for intrusion detection.

In the rest of this paper, section II describes Literature review; Section III describes proposed work and section IV describes conclusion and future work.

2. Literature Review

To start working to design a new IDS system we study some research papers that are included in this section.-

Intrusion Detection System (IDS) one of the appropriate anomaly based intrusion detection technique is Bayesian Event Classifier [6]. According to author, Intrusion Detection Systems is meant to identify the predefined intrusion attacks where unlike Intrusion Detection System (IDS) in anomaly based approach there is a chance of detecting unknown attacks [6]. In same context author has anticipated an event classification technique that is based on Bayesian networks. The technique is based on two

main problems; the first problem is stated as the positive false alarm that how the decision of a particular event should be treated, whether it is a normal activity or it should be detected as anomalous activity, it was detected in a simpleminded way. The second problem was, the system was not able to differentiate the anomalous behavior caused by unusual authentic action. According to author such information can be accessed by system health monitoring [6]. Further [6] has tried to rescind these two problems by classifying the previous scenarios with the change of Bayesian network by identifying Denial of Service (DOS), Masquerader attack and unknown attacks. Author has recommended this new technique as proposed solution for anomaly based intrusion detection. Some weaknesses of this proposed solution is that, this mechanism is very complex other than that we can also risk the system performance in the proposed technique. [7] Introduces the conditional random fields technique, the technique is used in a toolkit (CRF++) [7] as a model. Author has anticipated the technique as best among the previous techniques, and defined the Conditional Random Fields (CRT) as a unique technique for task of intrusion detection [7]. In the experiment among the other techniques it was recorded a very high rate of accurate results for intrusion detection. It is also one of the best feature in [7] among other techniques that proposed technique can be used without client server environment, where number of other techniques are proposed for the client server environment (research labs etc.). The Projected technique is a directionless graphical model, it used for the task of sequence classification and labeling [7], unlike the other models which prefer joint distribution the Conditional Random Fields (CRT) model favor conditional distribution. Proposed technique also avoids the observation and label bias problems. The technique of Conditional Random Fields (CRT) was proposed by identifying twenty four altered types of enormous network attacks which were further categorized in four groups of Denial of Service (DOS), Probing, R2L (unauthorized access from a remote machine) and U2R (unauthorized access to root). The experiment was matched with two other known best techniques Decision Tree [13] and Naïve Bayes [13] and the results were very efficient and effective. Weakness of proposed solution is that newly born unknown attacks cannot be detected by suggested technique.

3. Proposed work

Our proposed work includes the implementation of data mining based processes analysis and the intrusion detection system development. By conducting this study we gain information related to the different IDS system exist and their working process.

Additionally here we provide the process mining based IDS architecture that are used to analysis the different process of running on client machine. by implementation of the proposed system user is able to detect the malicious process running on the client machines in the network systems. In the proposed system we work with the security of internal network and analyse the processes running on the different network machines. for that purpose we propose an agent based IDS development scheme, where all the network computers read their running processes and send them to the server end, server machine contains a multithreaded program that respond all the connected machines. And collect all the system processes and save them over a data base where all the incoming processes are stored.

The server machine contains a decision mining algorithm that is trained using previously identified legitimate and malicious data patterns. this decision mining algorithm build a data model using the pre-classified data and using this data model upcoming data is analyzed.

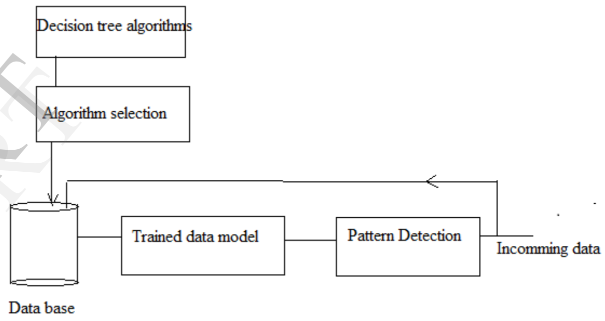

Fig. Basic server analysis flow

4. Conclusion and Future Work

In this manuscript we propose a data mining based IDS implementation for misuse analysis of the processes, here we propose an agent based scheme for malicious process detection and the implementation of the system is performed using visual studio IDE.

Inside this paper we keep focus to provide the processes mining and data mining based IDS implementation which is performed well but due to the small collection of malicious processes the performance of the IDS is varied over different systems and network parameters, in future we make a large collection of malicious processes and more data mining algorithms that provide much accurate results for classification of patterns.

5.References

[1] J. M. E Tapiador , J. E. Diaz Verdejo, “Detection of Web-Based Attacks through Markovian Protocol Parsing”,

Proceedings of the 10th IEEE Symposium on Computers and Communications, page 457-462, June 2005.

[2] F.Ullah and W. Tariq, "Operating System Based Analysis of Security Tools for Detecting Suspicious Events in Network Traffic", International Journal of Computer Science Issues(IJCSI), Vol. 8, Issue 6, No 2, page 418-422 November 2011.

[3] G. Wang, J. Hao, J. Ma and L.Huang, " A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering", Elsevier's journal of Expert Systems with Applications, Volume 37, Issue 9, page 6225-6232, September 2010.

[4] Mohammad M. Rasheed, O. Ghazali, N. MdNorwawi and Mohammed M. Kadhum," A Traffic Signature-based Algorithm for Detecting Scanning Internet Worms", International Journal of Communication Networks and nformation December 2009.

[5] T.Grandison and Evi mariaTerzi, "Intrusion Detection Technology", IBM Almad en Research enter, page 1 -7, September 7, 2007.

[6] C. Kr egel, D. Mutz, W. Robertso , and F. Vale ur, "Bayesian event classific tion for intru ion detection" In Proc. of the 19th Annual Computer Sec urity Applications Conference, Las Veges, NV , 2003.

[7] K.K. Gupta , B.Nath, and K.Ramamoha naro "Conditional Randomfields for intrusion detection", In 21st International Conference on advanced Information Networking and Applications Workshop ,IEEE pages203-208,2007.

[8]YihuaLi o , V. RaoVe muri, "Using T xt Categorization Techniques for Intrusion D etection", Proceedings of the 11 th USENIX Security Symposium, page 51-59, August 2002.

[9] C. Kruegel, F. Valeu , G. Vigna, a nd R. Kemmerer, "Stateful in trusion detecti on for high-sp eed networks" In Proceeding of the IEEE Sy mposium on Res earch on Security and Privacy . Oakland, CA: IEEE Press,Ma y 2002.

[10] D. Md. Farid, N.Huu Hoa, J.Darmont, N.Harbi, and M. ZahidurRahman, "Scaling up Detection R tes and Reducing False Positives in Intrusio n Detection us ing NBTree", In Proc. of th International Conference on Data Mining and Knowledge Engineering (ICDMKE 2010), page 18 6-190,April 2010.

[11] R. Se kar, A. Gupta, J. Frullo, T. Shanbhag, A. Tiwari, H. Yang, and S. Zhou, "Specification-based anoma ly detection: A new approach for detecting network intrusions ", In Proc. 9th ACM Conf. Computer an d Communication Security (C CS), page 265–2 74, 2002.

[12] M. Ku mar, Dr. M. H anumanthappa and Dr. T. V. S. Kumar, "Intrusion Detecti n System - False Positive Alert Reduction Technique", Pro ceeding. of Sec ond Internation al Conference on Advances i n Computer En gineering – A E 2011, page 1-4, Aug 2011.

[13] N. B. Amor, S. Be nferhat, and Z. Elouedi, "Naive bayesvs decision trees in intrusion detection systems", In Proceedings of the ACM S ymposium on A pplied Computin g, pages 420– 424, 2004.

[14] Yang, Y," Expert N etwork: Effective and efficie nt learning fr m human dec sions in text c ategorization and retrieval",Proceedings of the 17th Annual International AC M SIGIR Co nference on esearch and Development in Information Retrieval, 1994 .

[15] Improving the Performance Efficiency of an IDS by Exploiting Temporal Locality in Network Traffic , Govind Sreekar Shenoy , Jordi Tubella and Antonio Gonz ́alez , Department of Computer Architecture, Universitat Polit`e cnicade Catalunya, Barcelona, Spain

[16] A Semantic Approach to Host-based Intrusion Detection Systems Using Contiguous and Discontiguous System Call Patterns, Manuscript submitted on December 5, 2012, IEEE