ISSN (Online): 2456-0448
International Journal Of Innovative Research In Management, Engineering And Technology
Vol. 2, Issue 11, November 2017

NIDS: Survey Of Inrusion Detection Techniques Phase Wise Analyis Elucidation

[1] Gritto.D, Mohamed Suhail.M
M.Phil. Research Scholar, Department Of Computer Science, Vels University, Chennai
Assistant Professor, Department Of Bca & It,Vels University, Chennai

Abstract: ***The intrusion detection is a branch of cyber analytics that involves the recognition of network malicious activities and policy violations. The intrusions are deliberate actions spawned by the intruders against the security policies of the information system. The several attacks like, DoS, R2L, U2R, and Probing etc., are adverse that can divest the system information security. The IDS is used to detect any anomalous behavior, misuse or suspicious incident. The Network based intrusion detection (NIDS) is an IDS that monitors the network traffic to detect intrusion in real time. Presently enormous techniques and approaches were devised in the field of intrusion detection, even then the accuracy, rate of detection and the false alarm rate is under control. This paper is the survey of the contemporary data mining based NIDS detection techniques for ascertaining and categorizing the intrusion events. Deep emphasis is given for enhancing the attack detection rate and reducing the risks of false alarm rate. The rate of attack detection by data mining based algorithms like J48, Random tree and Random forest is analyzed using KDD Cup DARPA 99 data set.***

Keywords- ***NIDS; Packet Sniffing; Feature selection; Classification; False Alarm; J48, RF-Random Forest; RT-Random Tree***

I. INTRODUCTION

The internet and its applications like ICT, IoT are escalating gradually. As huge mass of sensitive fact traverse through the network, the risk of security attacks has also advanced exponentially. The prominent security attacks are DoS, Eavesdropping, Address Spoofing, Password-Based Attacks, Man-in-the-Middle Attack, Sniffer Attack, and Malicious code Attack etc. Securing the data cloud passing over the network becomes essential. Any event that distorts the CIA (Confidentiality-Integrity-Availability) of the information system is called the security attack. The organization have configured with the excellent technologies for detection and avoiding malicious attacks. Antivirus, Firewalls, IDS are the most familiar tools for strengthening the security infrastructure. Antivirus checks the programs, files or software that are already stored or installed in the system for any vulnerability. The firewall can also be called as IP filtering that analyzes packet header and prohibits the traffic from the intruders IP. The intrusion detection on the other hand analysis the whole packets. The header and the payload are detected for the availability of any attack patterns, if any positive signature is present then the alarm will be generated for alerting the security manager. The IDS provides dough security barrier among the other.

The IDS in general are classified based the method of detection, resource configured and reaction.**Signature based Anomaly Detection** isbased on the database of previous attack signatures and known system vulnerabilities. The signature database must be continually updated and maintained to identify attacks uniquely.

Misuse based Anomaly Detection is based on the references a baseline or learned pattern of normal system activity to identify active intrusion attempts. Higher false alarms are highly related.**Network based Detection** is based intrusion detection attempts to identify unauthorized, illicit, and anomalous behaviour based solely on network traffic. A network IDS, using either a network tap, span port, or hub collects packets that traverse a given network. Using the captured data, the IDS system processes and flags any suspicious traffic. Unlike an intrusion prevention system, an intrusion detection system does not actively block network traffic. The role of a network IDS is passive, only gathering, identifying, logging and alerting.
Host based Detection attempts to identify unauthorized,illicit, and anomalous behaviour on a specific device. Generally involves an agent installed on each system, monitoring and alerting on local OS and application activity. The installed agent uses a combination of signatures, rules, and heuristics to identify unauthorized activity. The role of a host IDS is passive, only gathering, identifying, logging, and alerting.
Hybrid Detection combines multiple techniques into a single hybrid system. It possesses the benefits of multiple approaches, while overcoming many of the drawbacks.

Active Reaction Based:
An active Intrusion Detection Systems (IDS) is also known as Intrusion Detection and Prevention System (IDPS). Intrusion Detection and Prevention System (IDPS) is configured to automatically block suspected attacks without any intervention required by an operator. Intrusion Detection and Prevention System (IDPS) has the advantage of providing real-time corrective action in response to an attack.
Passive Reaction Based:
These systems only monitor and analyze network traffic activity and alert an operator to potential vulnerabilities and attacks. A passive ID's is not capable of performing any protective or corrective functions on its own.

The deployment of HIDS will defense the local systems like server, router, gateway, DNS or any intersecting node where the NIDS protects the network secure and safe ideally. But the limitation with HIDS is that they are capable of monitoring the traffic that are directed a particular host in specific. They are inefficient for real time instantaneous attack detection since they rely on local system resource. They are also hard to integrate with gateway, DNS etc. The intruders can easily compromise is HIDS based system by cracking the host server through the control of C&C server. The NIDS in contrast, are tailored to detect serious intrusions like unauthorized access, DoS and bandwidth stealing etc. They are also suitable for real time based intrusion detection. In work the techniques related to network intrusion detection method is reviewed.

II. KDD Cup DARPA 99 Dataset

The analysis phase of this survey employs of KDD cup 99 dataset. The KDD Cup 99 is the accumulation tcpdump network traffic segment of DARPA volumes 4GB. The packet dissemination of the dataset contains 41 features and 24 types of attacks. The attacks are classified into 4 types.
Denial of Service (DoS) is an attack event in which the perpetrators or intruders disrupts the legitimate user from accessing the system or network resources temporarily. The DoS is achieved by making the memory of the resources busy through traffic flooding. The traffic flooding is an act of generating huge mass of well-planned requisition with the objective of prohibition the service and there by degrading the system performance.
Probing refers to the acquisition of vulnerable information about the objective network from the external network. On learning the susceptibility the intruders plot the attack plan to exploit the weakness. The attackers in general surfs for the host with open port by sending wisely designed packet to all destination port numbers, once identified the stealthy action starts.
User to Root Attack (U2R) is a class of exploit in which the attacker starts out with access to a normal user account on the system (perhaps gained by sniffing passwords, a dictionary attack, or social engineering) and is able to exploit some vulnerability to gain root access to the system.
Remote to Local Attack (R2L) occurs when an attacker has the ability to send packets to a machine over a network but does not have an account on that machine exploit some vulnerability to gain local access as a user of that machine.

Table 1: Redundancy Table Training Dataset

Class of Record	Native Record	Distinct Record	Redundant Record	Redundancy Rate
Normal	11358	8745	2613	23.01
Attack	45430	3694	41736	91.87
Total	56788	12439	44349	78.10

Table 2: Redundancy Table Test Dataset

Class of Record	Native Record	Distinct Record	Redundant Record	Redundancy Rate
Normal	7061	5648	1413	20.01
Attack	29223	2981	26242	89.8
Total	36284	8629	27655	76.22

Data Acquisition

The preliminary initiative in detection of in detection of intrusion involves online monitoring of system resources for any anomalous behavior. The customary method for misuse or attack detection involves inspecting the log files, event statistics user connectivity etc. But these techniques are infeasible for complex network monitoring so packet sniffing technique is used. The packet sniffing is an act of capturing the data stream packets that are advancing over the network. The captured packet is intercepted to analyze the network activities for detection of any intrusion, troubleshooting or forensics etc.

Several packet sniffing tools like tcpdump, wireshark, ngrep, snort and snoop can be used for capturing the packets. The packet loss is one of the noticeable issues in the packet sniffer. The packet lost must be prevented for improving the reliability of the attack detection , so more than one tool are configured for traffic capturing. Compare the packet captures of each tool and find for any missing packet of either of the tool.

Data Preprocessing

Data preprocessing is a technique used in transforming the inconsistent raw fact to complete understandable information. This is prerequisite for eliminating the redundant data, unconcerned features etc to avoid confusion and yield of inaccurate knowledge. The most common task in data preprocessing is the feature removal of outlier.

Feature Selection

Feature selection or variable selection or attribute selection involves extraction of attributes the data that are most relevant to the predictive modelling problem we are working on. Feature selection is the process of selecting a subset of relevant features. There are three general classes of feature selection algorithms: filter methods, wrapper methods and embedded methods. Filtering involves selection methods that apply statistical measure to assign a scoring to each feature. The features are ranked by the score and either selected to be kept or removed from the dataset. The methods are often univariate and consider the feature independently, or with regard to the dependent variable. The filtering methods are used for feature selection are, Wrapper Methods, Embedded Methods, Regularization methods

Outlier Detection

An outlier is an observation which deviates so much from the other observations and arouses suspicions that it was generated by different mechanism.

K-means clustering: Clustering analysis algorithm that groups objects based on their feature values into K disjoint clusters. Objects that are classified into the same cluster have similar feature values. K is a positive integer number specifying the number of clusters, and has to be given in advance.

Genetic Algorithm: This algorithms uses adaptive heuristics and robust in nature can be applied in problems of any domain with slight context based modification.Theyare simple and optimal nature

Naive Bayes Algorithm: Naive Bayes algorithm is based on the model called probability based model.The Bayes rule provides a way to calculate the probability of a hypothesis based on its prior probability.

Experimental Evaluation

The comparison of J48, Random forest, Random tree using KDD Cup 99 DARPA is conducted. The experiment is conducted by evaluating the detection rate and false attack detection rate.

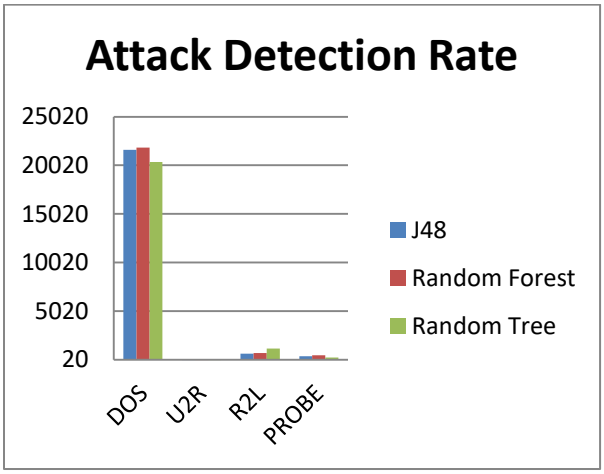
Table: III Instance Table

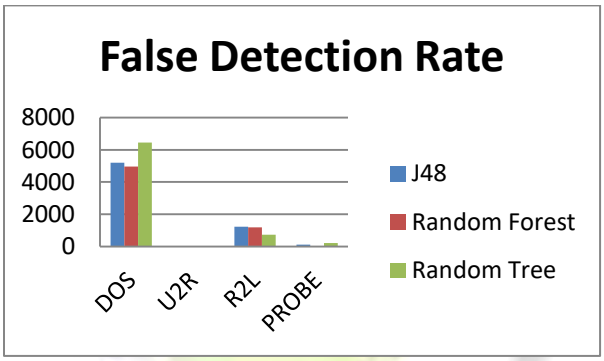
Attack class	Training Data Set	Test Data Set
DoS	43736	26814
U2R	12	29
R2L	136	1889
Probe	455	485
Normal	12449	7067
Total	56788	36284

Table: IV Categorization Table

Classifier	DoS		U2R	
	CORRECT	FALSE	CORRECT	FALSE
J48	21612	5202	8	21
Random Forest	21856	4958	3	26
Random Tree	20368	6446	14	15

Classifiers	R2L		PROBE	
	CORRECT	FALSE	CORRECT	FALSE
J48	649	1240	364	121
Random Forest	705	1184	468	17
Random Tree	1161	728	260	225





III. Conclusion

This paper, reviews various intrusion detection technique and estimates that the best among the machine learning techniques are J48, RF, RT algorithms. The experimental result reveals RF and RT produce optimal detection rate and false attack detection rate. The J48 produces fair results when comparing with the other two. Huge challenges are involved in detecting the intrusion in real time in environment like cloud based infrastructure. The indexing of known attack pattern in the profile database is a tedious task. The future work involves the devising intrusion detection model that implements the optimal indexing feature.

References

[1] KalpanaJaswal, Pravween Kumar and Seema Rawat, “Design and Development of a Prototype Application for Intrusion Detection using Data Mining,” in UP, 978-1-4673-7321-2/15/$31.00 IEEE, 2015.

[2] B. Raju and B. Srinivas, “Network Instruction Detection System Using KMP Pattern Matching Algorithm,” in Warangal, India, IJCST vol. 3, pp. 33-36, January 2012.

[3] Zhou Chunyue, Liu Yun and Zhang Hongke, “A Pattern Matching Based Network Intrusion Detection System,” in Beijing, China, 1-4244-0342-1/06/$20.00 IEEE, 2006.

[4] L. Vokorokos and A. Balaz, “Host – Based Intrusion Detection System,” in Kosice, 978-1-4244-7652-7/10/$26.00 ©2010 IEEE, 2010.

[5] Alaoui- AdibSaad, Chougdakli Khalid and Jedra Mohamed, "Network Intrusion Detection System Based on Direct LDA," in Rabat, 978-1-4673-9669-1/15/$31.00 IEEE, 2015.

[6] Anuradha and Anita Singhrova, "A Host Based Intrusion Detection System for DDOS Attack in WLAN," in Murthal Sonepat, India, 978-1-4577-1386-611$26.00 IEEE, 201I.

[7] F. Lydia Catherine, Ravi Pathak and V. Vaidehi, "Efficient Host Based Intrusion Detection System Using Partial Decision Tree and Correlation Feature Selection Algorithm," in Chennai, India, 978-1-4799-4989-2/14/$31.00 IEEE, 2014.

[8] MahbodTavallaee, EbrahimBagheri, Wei Lu and Ali A. Ghorbani, "A Detailed Analysis of the KDD CUP 99 Data Set," 978-1-4244-3764-1/09/$25.00 IEEE, 2009.

[9] Robert Moskovitch, Shay Pluderman, Ido Gus, Dima Stopel, Clint Feher, Yisrael Parmet, Yuval Shahar and Yuval Elovici, "Host Based Intrusion Detection Using Machine Learning" in Israel, 1-4244-1330-3/07/$25.00 IEEE, 2007.

[10] Ed' Wilson Tavares Ferreira, Ailton Akira Shinoda, Ruy De Oliveira, Valtemir Emerencio Nasciment and Nelcileno Virgilio De Souza Araujo, "A Methodology for building a Dataset to Assess Intrusion Detection Systems in Wireless Networks," in Brazil, E-ISSN: 2224 - 2864, vol. 14, pp. 113-119, 2015.

[11] Firkhan Ali Bin Hamid Ali and Yee Yong Len, "Development of Host Based Intrusion Detection System for Log Files," in Langkawi, Malaysia, 978-1-4577-1549-5/11/$26.00 IEEE, 2011.

[12] Lata and KashyapIndu, "Novel Algorithm for Intrusion Detection System," in Haryana, India, IJARCCE, vol. 2, Issue 5, May 2013.

[13] S. Sobinsoniya and S. Maria Celestin Vigila, "Intrusion Detection System: Classification and Techniques," in Tamilnadu, India, 978-1-5090-1277-0/16/$31.00 IEEE, 2016.

[14] Yanjie Zhao, "Network Intrusion Detection System Model Based on Data Mining," in Weifang. China, 978-1-5090-2239-7/16/$31.00 IEEE, 2016.