ISSN (Print) : 2320 – 3765
ISSN (Online) : 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering
Vol. 2, Issue 2, February 2013

Design and Implementation of H-IDS Using Snort, Feature Extraction, Honey pot and Rank and Reduce Alert

Neha chaudhary[1], Shailendra Mishra[2]

Assistant Professor, Dept. of C.S.E, Greater Noida Institute of Technology, Greater Noida, India [1]

Professor and Head, Dept. of C.S.E, Bipin Tripathi Kumaon Institute of Technology, Dwarahat, India [2]

ABSTRACT: The Internet is being used by increasing number of users day by day. Security is a big issue for all networks in today's enterprise environment. The security of a computer is compromised when an intrusion takes place. Many methods have been developed to secure the network infrastructure and communication over the Internet, among them the use of encryption algorithm, virtual private network and firewall. An intrusion detection system (IDS) is a device or software application that identify the suspicious activity on a target system or network. Many approaches have been used for batter intrusion detection system. There are two techniques of intrusion detection: misuse detection and anomaly detection. Some of the approaches use misuse based and some and some use anomaly based technique. Misuse detection can detect known attacks but the Main problem with misuse based technique is its vulnerability to unknown attacks. Anomaly detection can detect unknown intrusions, But the problem with anomaly based technique is that they give a lot of false alarms that is very difficult to realize. Entropy used in intrusion detection, is one of the anomaly detection technique. In this paper we are designing a new system that uses both technique(misuse and anamoly) with the help of Snot ,Entropy and honeypot . Also another issue of IDS is a lot of fault alarm, has also been addressed by developing alert reduction and ranking system. The results shows our system which is working in real time in efficient manner.

Keywords: Intrusion detection system (IDS), snort, entropy, alartrank, suspectincex (SI)

I. INTRODUCTION

Many users now days has broadband connection and capably to connect directly from their homes. They install a verity of applications that enable them to shop on line, play on line game, book e –tickets, transfer funds from one account to another, and so on. these applications required confidentiality and user must take care not to disclose their IDs and password to prevent unauthorized access. At the same time, networks have to robust enough to protect confidential data. So many cyber crimesof hackers invading the security barriers and stealing critical data so Intrusion detection system comes in to picture. Intrusion detection system is used to detecting unwanted traffic on a network or a device. The unwanted traffic are the traffic that can be harmful with respect to system security. The intrusion detection system is hardware or software application or combination of both that is used to detect the unwanted traffic or intruder activity with the help the expert system. Generally all IDS had a dual approach with a rule-based expert system to detect known types of intrusions plus a statistical anomaly detection component based on profiles of users, host systems, and target systems [1]. The systems based on misuse detection are called Misuse-based systems which can detect known attacks with high success rate , but when faces the unknown attacks it becomes powerless [2]. This is also called signature based detection. An anomaly detection technique identifies the observed activities that deviate significantly from the normal usage as intrusions [3]. Thus anomaly detection can detect unknown intrusions, which cannot be detected by misuse detection. But, the main problem with anomaly detection approaches is false positive alarms. False positive alarms are the activities detected by IDS as attacks but they are not. Whereas false negative alarms are the activities that are attacks but missed by IDS. IDS can also be classified into two categories based on network level and host level. Network based IDS (NIDS) detects activities that are suspicious with respect to whole network. It detects the malicious activities in packets flowing in the network. Host based IDS depend upon the system log files to find intrusions.

Snort [4] is one the most famous open NIDS. It is signature based. It uses alert based system to denote the suspicious activities. Its alert comprises of packet source and destination information along with signature id and timestamp. The main problems with Snort are vulnerability to unknown attacks, huge amount of alerts and no ways to identify the importance of alerts.

ISSN (Print) : 2320 – 3765
ISSN (Online) : 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering
Vol. 2, Issue 2, February 2013

The remaining part of this paper is organized as follows: Section 2 gives the background and the related work. Section 3 explains the system design and implementation. The experimentation results are discussed in section 4 and the paper concludes with challenges and future work in section 5.

II. BACKGROUND AND RELATED WORK

Apart from the known attacks, a lot of unknown types of attacks keep happening. A lot of machine intelligence techniques have been used in intrusion detection area to reduce vulnerability to unknown attacks. If we talk about Fuzzy logic, [5] gives the fuzzy expert system based approach in which automated learning of fuzzy rules. The data mining techniques application is to IDS for anomaly detection . The problem of intrusion detection is reducible to classification of traffic to normal and different attacks accurately, which can be achieved by data mining. W. Lee et al. [6], gives the application of different data mining techniques for building data mining models. MADAM ID [7] is a project of Columbia University which showed utilities of data mining techniques for building better IDS. ADAM [8], IDDM [9] and eBayes [10] all use anomaly detection techniques to detect intrusions. ADAM is a network based IDS which uses association rules for formulating normal behavior. MINDS [11] project at University of Minnesota uses data mining techniques to automatically detect attacks by assigning a score value to each connection. Score value represents the degree of deviation of behavior compared to normal behavior. They are successful in detecting numerous novel intrusions that could not be identified by widely used tools like Snort. It also takes into account network features important to intrusion detection. For this it has a feature extraction module. All these data mining techniques mainly depend on the training data i.e. they build model around feature values which can be changed easily during attack. Hence we must take into account overall traffic pattern and have to use statistical techniques. If we talk about the statistical techniques used for building models to detect intrusions, the first model is proposed by Denning [12]. Here Denning told the general model based on anomalies found in the user profiles developed by the statistical algorithms over a period of time. A lot of methods are borrowed from statistical signal theory and pattern recognition theory for detecting anomalies like Principal Component Analysis (PCA), covariance methods, Auto Regressive Moving Average (ARMA) etc. Entropy is another well-known measure for quantifying the information of network traffic and has been extensively studied for anomaly detection and prevention. G. Nychis et al. [13], uses the entropy based technique for anomaly detection in network traffic. M. Celenk et al. [14], proposed a model for using entropy based anomaly detection which does not use long term statistics. But the problem with this model was that it finds anomalies with respect to current data. So if there are more anomalies than normal data then it will give huge number of false alerts. Also it is not tested in real time.

Another design issue in building IDS is the proper representation of alerts to security analyst. As IDS generateshuge amount of alerts, it is very difficult to analyze them. Various alert correlation techniques are proposed. T. Zang et al. [15], provides a summary of all alert fusion techniques. These techniques mainly use aggregation, verification and correlation. Aggregation combines alerts having same attributes like Source IP, Destination IP etc. within a time window. Then correlation is performed on those alerts to find out the attack patterns. Hence these techniques only help in network forensics and aggregation only focus on grouping of alerts for correlation. No focus is upon the representing alerts according to their importance so that we can stop them timely.Hence in this paper we are designing a IDS by combining various approaches i.e. signature based IDS (Snort), entropy based anomaly detection method with feature extraction and alert reduction and ranking system to address the issues of unknown attacks, alert reduction and better alert visualization.

III. SYSTEM DESIGN AND IMPLEMENTATION

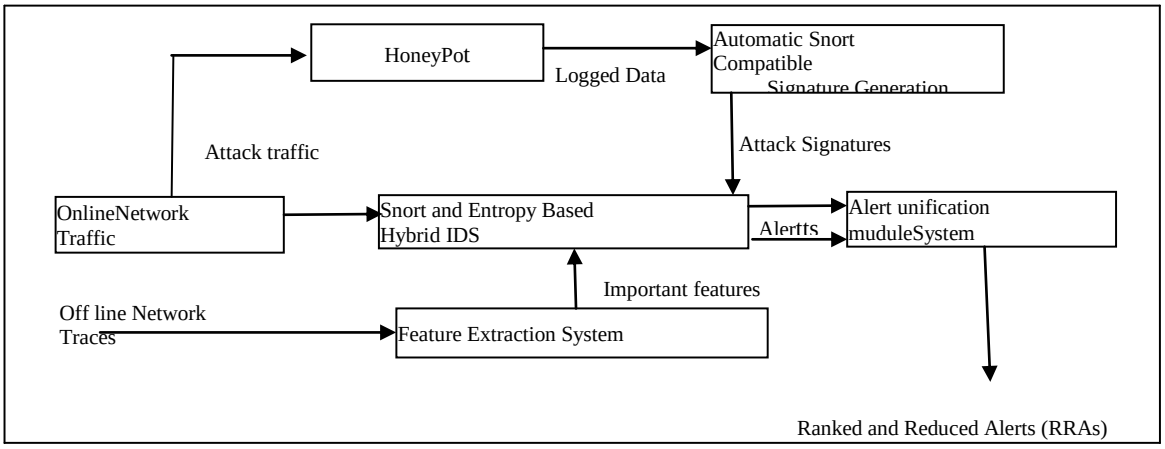


Fig. 1 System architecture

ISSN (Print) : 2320 – 3765
ISSN (Online) : 2278 – 8875

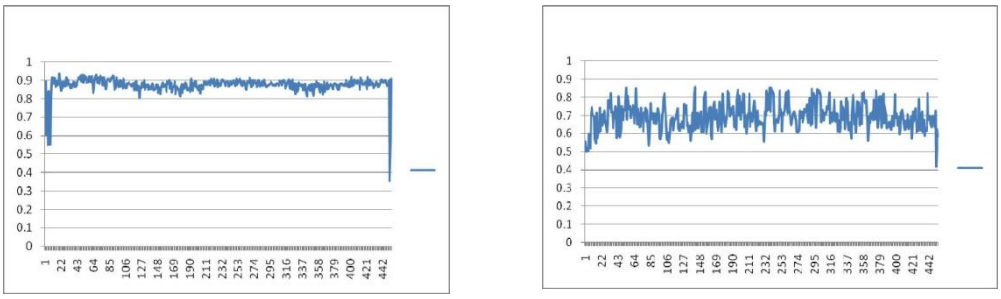
International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering
Vol. 2, Issue 2, February 2013

Fig 1 shows the model of IDS which we are proposing which uses the technique of feature extraction [16] and entropy based anomaly based technique [13]. Both these modules are integrated with Snort to make Snort more robust. Feature extraction module gives the important features which are important to detect intrusion. Kayasik [15], uses the KDD data set for feature selection but results are quite biased due to various limitations of KDD dataset. Hence NSL-KDD dataset is used for off line study of network traffic.

All the features which are relevant are inserted into attack-feature database. This database contains features that are most relevant to an attack. The main utility of this database is that entropy is calculated for these features. Now there is entropy based module for detecting anomalies in the network traffic. Important features are extracted from online traffic with the help of Information Extraction Module. Features are taken from the attack feature database. Features must have discrete values to calculate entropy. Here we are using mainly four features Src_ ip, Src_port, Dst_ip, and Dst_port. Hence all real time data which is coming is converted into records having 5 attributes including timestamp. For real time data, Snort is run in packet sniffing mode. For converting in this format, python script is implemented. Mean and standard deviation is used for entropy based detection. This can be done in two ways. In first method, historical data is used for finding these values. But this method is biased towards historical data. In second method, we can do entropy detection in real time data itself [14]. This method creates problem when huge part in real time data is itself anomaly. In this system, we use both techniques. We are using entropy based technique because attacks like port scanning, DDoS change the pattern of network traffic. They make network traffic more uniform or more random. Hence entropy based technique is used. As signature based systems use particular signatures for attack detection, hence these are ineffective when encountered with new patterns of attacks but with the help of anomaly based module in our system we can eliminate this problem. An intrusion detection system cannot be totally relied upon anomaly based module due to its limitation. Hence we are using Snort as its signature based system. A parallel Snort system is put with anomaly based module. As entropy based module can only detect attacks that change the network traffic pattern. Hence to detect other attacks we have to use a signature based system. For increasing the performance of system, a robust alert and logging system is integrated with Snort. With better alert and logging system we can have improved alert visualization and management. Once alerts are generated by both systems, they must be integrated.

Here main power of our system lies. As snort generates huge number of alerts which cause problems to security analyst to analyze the results. Alert unification module reduces the number of alerts generated by Snort and also classifies the alerts into some categories which help the analyst to better analyze the result. Anomaly based module gives the time stamps between which anomalies are detected. Now the suspect index is calculated for different features within that time stamps. Suspect index is the measure of how a particular feature is contributing to that anomaly. Once the suspect indexes of features are calculated, they are passed to alert unification module. Alert and logging system of Snort gives the full information corresponding to each alert to unification module. Each alert comprises of Signature-id, Src_ip, Src_port, Dst_ip, Dst_port, timestamp and other information. Here we use some threshold for suspect index. All alerts given by Snort within same timestamp as the anomaly based module and particular feature having valid suspect index (suspect index > threshold) are unified. We have used classifiers for alerts. Alerts can be of three types. Alerts detected by Snort and Entropy based module, alerts detected by Snort only and alerts detected by entropy based module only. Final alerts are shown to analyst according to their importance i.e. Snort and Anomaly based then Snort based, and then Entropy based. Entropy based can be moved up if deviation in network traffic pattern is very high. System is implemented in Ubuntu-9.04. Snort 2.8.6 is used as signature based system. Barnyard and Acid-base are used as alert and logging system. Feature selection algorithm is implemented in java. NSL-KDD data is used as off line traffic. Entropy based Anomaly detection module is implemented in Java.

IV. EXPERIMENATL RESULT

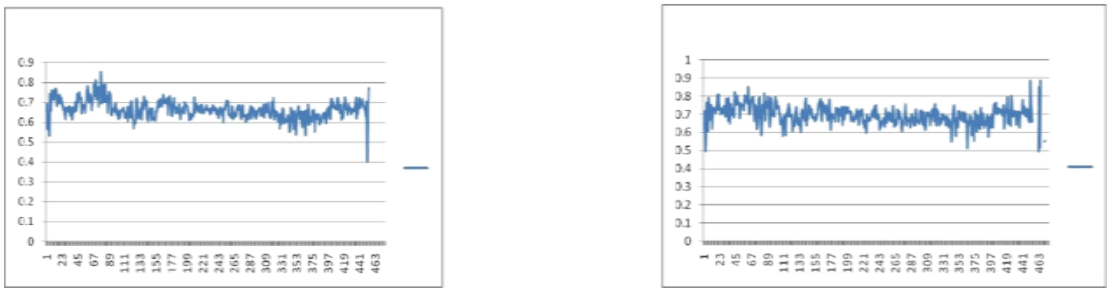


I. II.

Fig. 2 Graphs representing relation between entropy (Vertical Axis) and time stamp (horizontal axis) for different features. (I) Source IP, (II) Source Port

ISSN (Print) : 2320 – 3765
ISSN (Online) : 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering
Vol. 2, Issue 2, February 2013



I. II.

Fig. 3 Graphs representing relation between entropy (Vertical Axis) and time stamp (horizontal axis) for different features. (I) Destination IP, (II) Destination Port

First we have to prove the validity of entropy based module. For this, network traffic of Institute Hostel is used. Traffic of Institute is also used as the historical data. Fig. 2 and fig. 3 gives the entropy change in different features of the network traffic of a particular day. Graph is between entropy and a fixed time window (e.g. 60 seconds) which is represented by numbers. As we have seen in the figure, Source IP is most uniform and entropy is very high most of the time due to randomness of data. But here there are major drops in the entropy values which show the drastic change of traffic patterns due to port scanning attacks being done on different systems. Other features are also showing the deviation in entropy of that feature value but not very uniform. This is due to property of port scanning attack. So we first train our entropy model on this data and find the threshold values for anomalies. As it is the most difficult thing to decide the threshold factor so we have must have data for deciding this. Threshold factor means deciding mean and standard deviation value with the help of historical data. After finding threshold values, we will check each observation with respect to these values.

TABLE I
RELATION BETWEEN THRESHOLD AND ANOMALOUS POINT

Feature	No. of Anomalous Point Detected when t =2	No. of Anomalous Point Detected when t=3
Src_IP	96	9
Dst_IP	94	1
Source Port	96	4
Destination Port	71	6

Table 1 shows the drastic impact of threshold (t) values used for anomalies detection. Number of anomalous points drastically increases when t is decreased to 2 from 3. This causes a lot of points to be declared as anomalous points. If we see closely on this data we can find out that we can find anomalies without need of historical data as if we take a second time window parameter. So we can find out anomalies in that time window. In this way, both techniques of entropy based anomaly detection which are described earlier are used.

Table 2 shows the results when suspect index is calculated for different feature values within a single time window (60 seconds).

TABLE II
FEATURE VALUES AND THEIR SUSPECT INDEX

Feature	Suspect Index
172.17.12.231	0.85
172.17.14.25	0.31
172.17.12.236	0.29
172.17.12.49	0.27
…….	…….

As at five points anomaly is detected, we calculate which feature value is contributing most in anomaly. Here we are calculating suspect index for Src_ip but it can be used for other features also. As results show that for each anomaly point, it is showing all suspect indexes. IP address 172.17.12.231 has highest suspect index within a particular time

ISSN (Print) : 2320 – 3765
ISSN (Online) : 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering
Vol. 2, Issue 2, February 2013

window as attack is done from this IP address. All other IP addresses have very low suspect indexes. Hence we have incorporated a threshold value for this also. Here it is taken as .50

TABLE III
COMPARISON BETWEEN SNORT AND PROPOSED SYSTEM

Time-Stamp	No. of Snort Alert	No. of Unified Alert	Alert Rank
11:49:16-11:50:16	0	0	3
13:51:16-13:52:16	0	0	3
19:18:16-19:19:16	5	1	1
19:19:16-19:20:16	18	1	1
….....	….....	…..	….....

Table 3 shows the number of alerts generated by Snort. Now all the Snort alert in the same time window in which anomaly is detected and having same value as feature value for which anomaly is detected is grouped into one because Snort generates huge number of alerts between same pair of systems during attack and some of them can be false. Now we have verified with the anomaly based module that it is attack hence classified as class1 attack. Similarly attacks which are not verified by Anomaly based module are given the class2 rank and similarly class3 rank is given to those that are not verified by Snort. Class1 is the highest rank alerts. This type of ranking system of alerts gives the analyst to better analyze the result and makes him to take corrective action timely.

V. CONCLUSION

In this paper, we have designed and implemented real time Intrusion detection system with the help of integration of Snort (Signature based system) and Entropy based (Anomaly based) system. We have validated the anomaly based module. We have developed an alert reduction and ranking system of alerts. Experimental results show that our system is real time and also is better than Snort. In future, we will try to extend this system to find anomalies for diverse type of attacks by incorporating the anomaly based routines which are not totally based upon traffic patterns. Also we will try to include the dynamic addition of rules in Snort in real time.

REFERENCES

[1] Wikipedia Article of IDS. Available online at http://en.wikipedia.org/ids.html.
[2] M. Yang; D. Chen; X. Zhang, "Anomaly Detection based on Contiguous Expert Voting Algorithms", Apperceining Computing and Intelligence Analysis, ICACIA , pages 158-161, 2009..
[3] E. Eskin, A. Arnold, M. Preau, L.Portnoy, and S. Stolfo, "A Geometric Framework for Unsupervised Anomaly Detection: Detecting Intrusions in UnlabeledData " Applications of Data Mining in Computer Society, Kluwer Academic Publishers, 2002.
[4] Rafeeq Ur Rehamn, "Intrusion Detection Systems with Snort, Advanced Intrusion Techniques using Snort, PHP, MySQL, Apache and ACID", Pearson Education ISBN 0-13-140733-3, 2003 ..
[5] W. Yunwu, "Using Fuzzy Expert System Based on Genetic Algorithms for Intrusion Detection System", Information Technology and Applications, IFITA, pages 221-224, 2009.
[6] W. Lee; S. J. Stolf; K. W. Mok, "A Data Mining Framework for Building Intrusion Detection Model", Security and Privacy, Proceedings of the 1999 IEEE Symposium, pages 120-132, 1999.
[7] W. Lee; S. J Stolfo, "Data mining approaches for intrusion detection". Proc. Of Seventh USENIX Security Symposium. San Antonio, TX, 1998.
[8] D. Barbara; J. Couto; S. Jajodia; N. Wu , "Adam: Dtecting Intrusions by Data Mining" Proc. Of 2nd Annual IEEE Information Assurance Workshop. West Point, NY, 2001
[9] T. Abraham; "IDDM: Intrusion Detection Using Data Mining Techniques." Technical report DSTO-GD-0286, DSTO Electronics and Surveillance Research Laboratory, 2001
[10] A. Valde; K. Skinner , "Adaptive, model based monitoring for cyber attack detection", Recent advances on Intrusion Detection, France, Springer Verlag, pp 80-93, 2000
[11] L. Eetoz; E. Eilertson; A. Lazarevic; P. Tan; P. Dokes; V. Kumar; J. Srivastava, "Detection of Novel Attacks using Data Mining". Proc. IEEE Workshop on Data Mining and Computer Security, 2003
[12] D. E. Denning, "An Intrusion-Detection Model" , IEEE Transaction on Software Engineering, Vol. SE-13 No.2, 1987.
[13] G. Nychis, V. Sekar, D.G. Andersenm, H. Kim and H. Zhang, "An empirical evaluation of entropy-based traffic anomaly detection", Proceedings of 8th ACM SIGCOMM conference on Internet Measurement, pp 151-156, 2008.
[14] M. Celenk; T. Colony; J. Willies; J. Graham; , "Predictive Network Anomaly Detection and Visualization," *Information Forensics andSecurity, IEEE Transactions on* , vol.5, no.2, pp.288-299, June 2010.
[15] T. Zang; X. Yun; Y. Zhang; , "A Survey of Alert Fusion Techniques for Security Incident," *Web-Age Information Management, 2008.WAIM '08. The Ninth International Conference on* , vol., no.,pp.475-481, 20-22 July 2008
[16] Kayasik; G. Heywood; Z. Heywood, " Selecting Features for Intrusion Detection", Proceedings of the IEEE ISI, pp. 796-800, 2005

ISSN (Print) : 2320 – 3765
ISSN (Online) : 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering
Vol. 2, Issue 2, February 2013

BIOGRAPHY

Neha Chaudhary received the B.Tech. degreein computer science from the College of Technology, Pantnagar, India and She is currently working toward the M.Tech. degree in the Department ofComputerScience, Uttarakhand technical University . She is an Assistant Professor, Department ofComputer Science and Engg., Greater noida Institute of Technology, Mahamaya technical University.Her research area include information systems security and privacy.

Dr .Shailendra Mishra received Ph.D degree in CSE & Master of Engineering Degree (M E) in Computer Science & Engineering (Specialization: Software Engineering) from MNREC Allahabad(Now MotiLal Nehru National Institute of Technology (MNNIT)) India, Presently he is Professor &Head , Department of Computer Science &Engineering , KEC Dwarahat India.
His recent research has been in the field of Mobile Computing & Communication, Advance Network Architecture and Software Engineering. He has also been conducting research on Communication System & Computer Networks with Performance evaluation and design of Multiple Access Protocol for Mobile Communication Network.
He handled many research projects during the last 5 years; Power control and recourse management for WCDMA System funded and sponsored by UCOST Dehradun Uttrakhand, Code and Time complexity for WCDMA System, OCQPSK spreading techniques for third generation communication system, "IT mediated education and dissemination of health information via Training & e-Learning Platform" sponsored and funded by Oil Natural Gas Commission (ONGC), New Delhi, India (November 2006),"IT based Training and E-Learning Platform", sponsored and funded by UCOST, Department of Science and Technology,Govt. of Uttarakhand, India (December 2006) etc. He received Young Scientist Award in the Yr 2006 and 2008 from DST UCOST Govt.ofUttrakhand.
He had supervised 6 Ph.D and currently guiding five research scholars. He had authored four books in the area of Computer Network and Security and published and presented 60 research papers in international journals and international conferences and wrote more than 10 articles on various topics in national magazines. He is recipient of Young Scientist Award in IIIrdUttrakhand State Science Congress &Ist Uttaranchal State Science Congress organized by Uttaranchal Council for Science &Technology, Department of Science &Technology, Govt. of Uttarkhand, India (10,11Nov 2008 &11 Nov 2006).He is Member of Institution of Engineers India (IEI) and ISTE.