Intelligent Network-Based Intrusion Detection System (iNIDS)

P.R. Mahalingam

Department of Computer Science
Rajagiri School of Engineering & Technology, Rajagiri valley, Cochin, India
prmahalingam@gmail.com

Abstract. Networks are regarded as one of the biggest advancements in the field of computer science. But they enable outsiders to "intrude" into our information. Intrusions can be in the form of simple eavesdropping, or gaining access to the host itself. Here, intruders are identified using two main methods – signature analysis and anomaly analysis. The proposed method is such that the signature analysis is strengthened by anomaly analysis, which in turn uses some level of intelligence based on the traffic parameters, obtained and processed using neural networks. The initial intelligence is obtained using the KDDCUP99 dataset, which trains a neural network. The neural network will take care of further detections, and it strengthens itself during the run itself. The result obtained suggests that even with minimal initial intelligence, iNIDS can reach accuracy levels of over 70%, and by increasing the initial set a little more, it reaches accuracy levels exceeding 80%.

Keywords: Intrusion detection, neural networks, intelligence, anomaly analysis, signature analysis, KDDCUP99, JpCap.

1 Introduction

Advancements and increased usage of computer networks paved way for increase in variety and complexity of security threats. The scenario is getting worse in the sense even single firewall strategies are insufficient to counter security threats[1]. Nowadays people are aware of the risks involved in securing a computer network. So a system which is capable of detecting network security threats is developed [2]. Here, an Network based Intrusion Detection System (NIDS) is proposed that uses real time internet traffic for analysis. Also, the system uses Artificial Intelligence for improving the performance and speed of detection.

Real time packets in the network are captured online i.e. from the internet as and when they reach the interface of the network, using suitable Java[3]-based packages. iNIDS is designed to provide the basic detection techniques so as to secure the systems inside a computer network that are directly or indirectly connected to the Internet.

Network intrusion[4] can be defined as any deliberate attempt to enter or gain unauthorized access to a network and thereby break the security of the network and thus gaining access to confidential information present in the computers inside the

N. Meghanathan et al. (Eds.): Advances in Computing & Inform. Technology, AISC 176, pp. 1–9.
springerlink.com

network. An IDS[4] captures and inspects all traffic, regardless of whether it's permitted or not. Based on the contents of the packet, their flow, length etc, at either the IP or application level, an alert is generated.

An *intrusion signature*[4] can be defined as a special TCP state set such as [SYN|RST] in one packet, special bytes in the IP header, or a special byte stream in the payload of a packet that will provide a pattern which can be used for packet analysis for identifying threats.

The primary goals of the whole proposal can be summarized to the following.

1) Detect Network intrusions[8][9]
2) Use of Artificial Intelligence to improve detection[5][6]
3) Use Network traffic for analysis and detection[5][7].

2 NIDS and ANN

Dr. Dorothy E Denning proposed an Intrusion detection system in 1987 which became a benchmark in the research in this area[2]. Many researches have been conducted based on this paper and currently researchers are more interested in developing intrusion detection systems based on Artificial Neural Networks. Artificial Neural Networks possess features like generalization, flexibility etc. Wang Zhenqi and Wang Xinyu proposed a Netflow[1] based intrusion detection system, which can resist network attacks and intrusions. It was found to be cost effective and does not affect the performance of backbone network.[1][13]

Usually, sampled data from Kddcup99 dataset[14], an attack or intrusion database is the standard for evaluating the security detection mechanisms. This dataset is used for signature analysis, for training neural network for anomaly analysis and for testing the IDS itself. The advantage of using Backpropagation algorithm is that it can train (learn) data at a faster rate and it provides efficient generalization and flexibility when compared to other existing Neural Network technologies[13]. But, the performance of a Neural Network depends mainly on the amount of training data given[15][16][17].

The strategy[18] dictates that NIDS uses a hybrid detection engine i.e. a combination of Signature detection and Anomaly detection capabilities. "Rule –based" detection technique is used for signature analysis and "Pattern matching" is used for anomaly analysis.

3 Signature Analysis and Anomaly Analysis

Firewalls cannot or do not analyze packets once they are inside the network and it only analyze them while it enters the network.[19] So, if some anomalies or activities happen from inside the network, then firewalls won't respond to those activities. But, NIDS analyze network packets internally as well as while it enters or leaves the network. With the explosive growth of networking and data sharing, NIDS have become the most popular form of Intrusion Detection[19]. A NIDS is capable of detecting network security threats. Many different NIDSs have been developed and each of them has its own advantages and disadvantages.

Signature Analysis: An NIDS use signature based detection, based on known traffic data to analyze network traffic. This type of detection is very fast, and easy to configure[20]. However, an attacker can slightly modify an attack to render it undetectable by a signature based IDS. Still, signature-based detection, although limited in its detection capability, can be very accurate. A common strategy for IDS in detecting intrusions is to memorize signatures of known attacks. These signatures are written based on data collected from known and previous attacks, and this unfortunately ensures that these signatures “will always be a step behind the latest underground exploits” [20][21].

Anomaly Analysis: Anomaly analysis[17][21] is an efficient way to detect intrusions and thus forms a vital part of the next generation Intrusion Detection Systems. The most efficient Anomaly analysis technique is the pattern based anomaly detection. In pattern based anomaly analysis, the Intrusion Detection System is given a pre-defined set of intrusion patterns. Network packets are collected for a specified period of time or till a specified number of packets. These packets are considered as a block for analysis. The predefined patterns are then matched with this packet block and if any patterns match, an alert is given. During anomaly analysis, a normal behavior model is used as the base for analyzing incoming traffic and any deviation or variation from the normal behavior model is considered as an intrusion or threat[22]. But this can produce a rather high degree of false alarms.

The following attacks are stressed upon in iNIDS.

- Denial of Service[23] - UDP Flooding[24], TCP SYN Attack[26], Smurf attacks[28]
- User to Root (U2R)[29] - Type signatures
- Remote to Local (R2L)[8]
- Probing Attack[30] - Portsweep, Satan, Nmap, etc.

Each possesses its own signature, and attack characteristics, which make it easier to detect and handle. They can sometimes be identified directly from the signature, or by using the anomaly detection methods.

4 KDDCUP99 Dataset

The “KDD CUP’99” dataset [14], which derived from the DARPA dataset, was used for the KDD (Knowledge Discovery and Data Mining Tools Conference) Cup 99 Competition. The complete dataset has around 5 million input patterns and each record represents a TCP/IP connection that is composed of 41 features. The dataset used in this study is a smaller subset (10% of the original training set), it has 494 021 instances (patterns) and it was employed as the training set in the original competition. Each record of the KDD Cup 99 dataset captures various features of the connections, as for example, the source and destination bytes of a TCP connection, the number of failed login attempts or the duration of a connection. Complex relationships exist between the features, which are difficult for human experts to discover.

An NIDS must therefore reduce the amount of data to be processed so as to maintain an accurate and real-time detection. Some input data may not be useful to the Network based IDS and thus can be eliminated before processing. In complex classification systems, the features may contain false correlations, which block the process of detecting intrusions/attacks. Furthermore, some features may be redundant since the information they add is contained in other features.[14][31]

KDD Cup 99 dataset feature selection[32] consists of detecting the relevant features and discarding the irrelevant features. Relevant features are features that can be used for analysis with ease and that can deliver relevant information as well as can work without any performance degradation.

KDDCUP99 attributes[33] can be categorized into four. They are: *Intrinsic Attributes*, *Content Attributes*, *Traffic Attributes*, and *Class Attributes*.

5 JPCAP

Jpcap(Java Packet Capturer)[34] is a Java library for sniffing, capturing and sending network packets, from an available network interface. It also facilitates visualization, creation and analysis of network packets by appropriate coding in Java. The Java language gives it the capability to work in multiple platforms (Operating systems). Jpcap has been tested on Microsoft Windows (98/2000/XP/Vista), Linux (Fedora, Mandriva, Ubuntu), Mac OS X (Darwin), FreeBSD, and Solaris and was found to be working successfully. Jpcap can capture Ethernet, IPv4, IPv6, ARP/RARP, TCP, UDP, and ICMPv4 packets[34]. It is open source, and is licensed under GNU LGPL.

6 Artificial Neural Networks

Artificial Neural Networks, also known as “Artificial Neural Nets”, “neural nets”, or ANN for short, is a computational tool modeled based on the interconnection of the biological neurons in the nervous systems of the human body. ANN can be trained / taught to solve certain type of problems using a training method and data to train. [35] By following this method, Artificial Neural Networks made can be used to perform different tasks depending on the amount training given. A properly trained ANN is capable of generalization, the ability to recognize similarities among many different number and type of input patterns, especially patterns that have been corrupted by noise which has a wide variety of applications.

Artificial Neural Networks have a number of features and characteristics that make them an attractive alternative to traditional problem-solving techniques. We can design a complex network of neurons using multiple layers[35][36], the first of which was called the *Multi-layered Perceptron*. There are a variety of thresholds[37] available for ANN, an example of which is the *sigmoid*.

No matter what organization is used, the ANN has to be trained (or it learns)[38]. It can be *supervised* or *unsupervised*. Here, Backpropagation is preferred due to its speed, which is essential in networked environments. It works on a supervised training model. So, the initial training data, extracted fron KDDCUP99 dataset is important.

7 Implementation

During development, an existing packet capture program based of JpCap was used. The anomaly analysis module was added on top of that by analyzing the parameters of the packets on the fly, and reporting any threats before the attacker has a chance to enter.

The overall system looks like below.

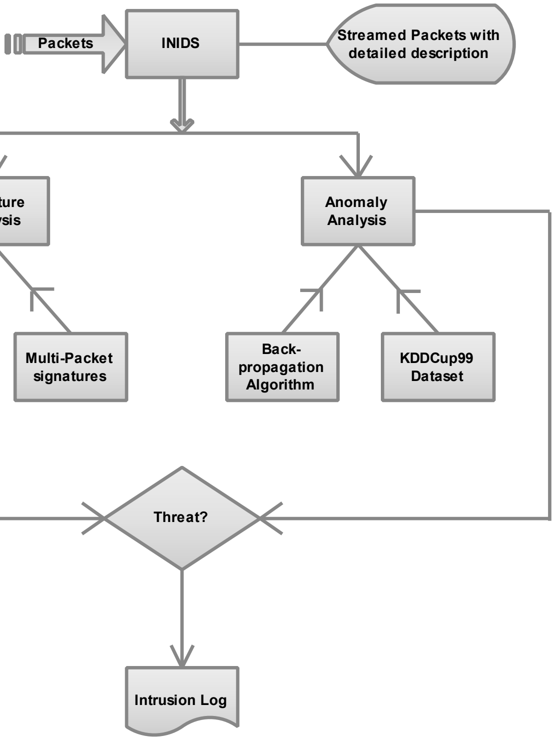

Fig. 1. Implementation

When newer threats are identified by the neural network, it implicitly trains itself to identify that pattern, and it is in turn added to the library. So, further analysis will check for this pattern also. Also, since we use neural networks, it doesn't need an exact match. Any similar pattern of that genre can be caught. So, iNIDS improves on the go, by learning newer ways in which attackers can threaten the integrity and security of the data.

8 Results

The performance was as expected during the design phase. The initial training data was taken such that there was only one sample for each type of intrusion expected.

So, the initial performance was very bad, at about 10%. But eventually, it improved by adding more patterns to the arsenal, and finally, it was able to reach accuracy of more than 70%, in just 21 runs.

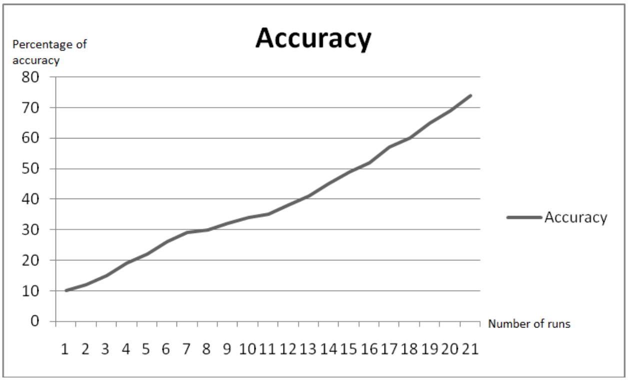

Fig. 2. Accuracy plot with smaller training data, over 21 runs

On a later run, the initial training data itself was improved, with multiple patterns per type of attack. In this case, the initial performance itself went up to 40%, and later, it improved to more than 85% in the same number of runs.

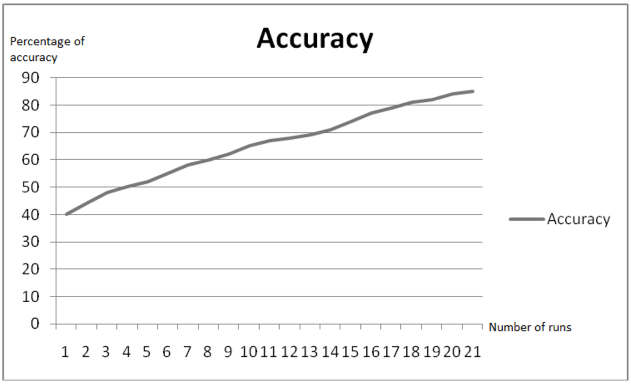

Fig. 3. Accuracy plot with more comprehensive training data, over 21 runs

9 Conclusion

iNIDS initially kicked off as an abstract deign that can improve the current packet capture and analysis routines in intrusion detection, by adding a level of knowledge and intelligence to the analysis. In this case, the threat detection is reinforced by past knowledge of similar attacks, and we are able to achieve levels of accuracy much higher than conventional systems that use the same level of knowledge.

10 Future Work

Even though the research proposes a new technique, it can't be claimed as the best possible technique. The field of network security is subject to continuous research in search for new and better techniques that can ensure better security capabilities. Applying Artificial Intelligence to an NIDS gives it the capability to detect unknown threats which is an important characteristic.

The intelligence thus given to the NIDS can be improved by the use of a better Neural Network architecture as well as a much better training algorithm than Backpropagation which is used in this implementation. A better Neural Network architecture will give better input processing capabilities such as improved speed, pattern matching etc. This will thus enable the ANN to detect patterns much more efficiently with very little data. Better training algorithms can provide reduced training time thereby allowing more training data to be used in very limited time. This will thus improve the training speed and also the training data thereby resulting in an improved performance. So a thorough research is recommended in the field of NIDSs based on Artificial Intelligence.

References

[1] Wang, Z., Wang, X.: NetFlow Based Intrusion Detection System. In: International Conference on MultiMedia and Information Technology, MMIT 2008, December 30-31, pp. 825–828 (2008)

[2] Denning, D.E.: An Intrusion-Detection Model. IEEE Transactions on Software Engineering SE-13(2), 222–232 (1987)

[3] The JavaTM Tutorials, `http://download.oracle.com/javase/tutorial/` (accessed August 20, 2011)

[4] Garuba, M., Liu, C., Fraites, D.: Intrusion Techniques: Comparative Study of Network Intrusion Detection Systems. In: Fifth International Conference on Information Technology: New Generations, ITNG 2008, April 7-9, pp. 592–598 (2008)

[5] Shun, J., Malki, H.A.: Network Intrusion Detection System Using Neural Networks. In: Fourth International Conference on Natural Computation, ICNC 2008, October 18-20, vol. 5, pp. 242–246 (2008)

[6] Wang, Y., Huang, G.X., Peng, D.G.: Model of Network Intrusion Detection System based on BP Algorithm. In: 2006 1st IEEE Conference on Industrial Electronics and Applications, May 24-26, pp. 1–4 (2006)

[7] Liu, B., Lin, C., Ruan, D., Peng, X.: Netfiow Based Flow Analysis and Monitor. In: International Conference on Communication Technology, ICCT 2006, November 27-30, pp. 1–4 (2006)

[8] Ahmad, I., Abdullah, A.B., Alghamdi, A.S.: Remote to Local attack detection using supervised neural network. In: 2010 International Conference for Internet Technology and Secured Transactions (ICITST), November 8-11, pp. 1–6 (2010)

[9] Stolfo, S.J., Fan, W., Lee, W., Prodromidis, A., Chan, P.K.: Cost-based modeling for fraud and intrusion detection: results from the JAM project. In: Proceedings of DARPA Information Survivability Conference and Exposition, DISCEX 2000, vol. 2, pp. 130–144 (2000)

[10] Zihao, S., Hui, W.: Network Data Packet Capture and Protocol Analysis on Jpcap-Based. In: 2009 International Conference on Information Management, Innovation Management and Industrial Engineering, December 26-27, vol. 3, pp. 329–332 (2009)

[11] Al-Shaer, E.: Managing firewall and network-edge security policies. In: IEEE/IFIP Network Operations and Management Symposium, NOMS 2004, April 23-23, vol. 1, p. 926 (2004)

[12] Yang, Y., Mi, J.: Design and implementation of distributed intrusion detection system based on honeypot. In: 2010 2nd International Conference on Computer Engineering and Technology (ICCET), April 16-18, vol. 6, pp. V6-260–V6-263 (2010)

[13] Ahmad, I., Ansari, M.A., Mohsin, S.: Performance Comparison between Backpropagation Algorithms Applied to Intrusion Detection in Computer Network Systems. In: 9th WSEAS International Conference on Neural Networks, May 2-4, pp. 47–52 (2008)

[14] KDD Cup 1999 Data (1999), http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html (accessed August 13, 2011)

[15] Lee, S.M., Kim, D.S., Park, J.S.: A Hybrid Approach for Real-Time Network Intrusion Detection Systems. In: 2007 International Conference on Computational Intelligence and Security, December 15-19, pp. 712–715 (2007)

[16] Abdel-Azim, M., Abdel-Fatah, A.I., Awad, M.: Performance analysis of artificial neural network intrusion detection systems. In: International Conference on Electrical and Electronics Engineering, ELECO 2009, November 5-8, pp. II-385–II-389 (2009)

[17] Yu, X.: A new model of intelligent hybrid network intrusion detection system. In: 2010 International Conference on Bioinformatics and Biomedical Technology (ICBBT), April 16-18, pp. 386–389 (2010)

[18] Successful software development - Google Books, http://books.google.co.uk/books?id=lrix5MNRiu4C&pg=PA184&dq=software+development+life+cycle+preliminary+design+detailed+design&hl=en&ei=0h1SToP2GdKwhAf9rNHaBg&sa=X&oi=book_result&ct=result&resnum=1&ved=0CDAQ6AEwAA#v=onepage&q=software%20development%20life%20cycle%20preliminary%20design%20detailed%20design&f=false (accessed August 22, 2011)

[19] Network-Based IDS (NIDS) overview | IDStutorial, http://idstutorial.com/network-based-ids.php (accessed August 13, 2011)

[20] Wu, T.M.: Intrusion Detection Systems, September 25 (2009), http://iac.dtic.mil/iatac/download/intrusion_detection.pdf (accessed August 12, 2011)

[21] SANS institute, Host- vs. Network-Based Intrusion Detection Systems (2005), http://www.giac.org, http://www.giac.org/paper/gsec/1377/host-vs-network-based-intrusion-detection-systems/102574 (accessed August 12, 2011)

[22] Zhang, W., Yang, Q., Geng, Y.: A Survey of Anomaly Detection Methods in Networks. In: International Symposium on Computer Network and Multimedia Technology, CNMT 2009, January 18-20, pp. 1–3 (2009)

[23] Gill, K., Yang, S.-H.: A scheme for preventing denial of service attacks on wireless sensor networks. In: 35th Annual Conference of IEEE Industrial Electronics, IECON 2009, November 3-5, pp. 2603–2609 (2009)

[24] Chang, R.K.C.: Defending against flooding-based distributed denialof-service attacks: a tutorial. IEEE Communications Magazine 40(10), 42–51 (2002)

[25] IntelliGuard I.T. - Eliminate DDoS and Flash crowd problems, http://www.intelliguardit.net/library_attackscenarios.html (accessed August 19, 2011)

[26] Wang, H., Zhang, D., Shin, K.G.: Detecting SYN flooding attacks. In: Proceedings of the Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies, p.1530 (June 2002)

[27] ASA/PIX 7.x and Later: Mitigating the Network Attacks - Cisco Systems, http://www.cisco.com/en/US/products/ps6120/products_tech_note09186a00809763ea.shtml (accessed August 19, 2011)

[28] Mirkovic, J., Reiher, P.: A taxonomy of DDoS attack and DDoS defense mechanisms. ACM SIGCOMM Computer Communication Review, 39–53 (2004)

[29] Ahmad, I., Abdullah, A.B., Alghamdi, A.S.: Applying neural network to U2R attacks. In: 2010 IEEE Symposium on Industrial Electronics & Applications (ISIEA), October 3-5, pp. 295–299 (2010)

[30] Ahmad, I., Abdullah, A.B., Alghamdi, A.S.: Application of artificial neural network in detection of probing attacks. In: IEEE Symposium on Industrial Electronics & Applications, ISIEA 2009, October 4-6, vol. 2, pp. 557–562 (2009)

[31] MIT Lincoln Laboratory: Communication Systems and Cyber Security: Cyber Systems and Technology: DARPA Intrusion Detection Evaluation, http://www.ll.mit.edu/mission/communications/ist/corpora/ideval/data/index.html (accessed August 14, 2011)

[32] Bolon-Canedo, V., Sanchez-Maroo, N., Alonso-Betanzos, A.: A combination of discretization and filter methods for improving classification performance in KDD Cup 99 dataset. In: International Joint Conference on Neural Networks, IJCNN 2009, June 14-19, pp. 359–366 (2009)

[33] Index of /acwaldap/gureKddcup, http://www.sc.ehu.es/acwaldap/gureKddcup/README.pdf (accessed August 14, 2011)

[34] Jpcap - a Java library for capturing and sending network packets, http://netresearch.ics.uci.edu/kfujii/Jpcap/doc/index.html (accessed August 14, 2011)

[35] Artificial Neural Networks/Neural Network Basics - Wikibooks, open books for an open world, http://en.wikibooks.org/wiki/Artificial_Neural_Networks/Neural_Network_Basics (accessed August 14, 2011)

[36] Advances in Data Analytical Techniques, http://www.iasri.res.in/ebook/EBADAT/5-Modeling%20and%20Forecasting%20Techniques%20in%20Agriculture/5-ANN_GKJHA_2007.pdf (accessed August 14, 2011)

[37] Neural Networks: Tutorials: Paras Chopra, http://paraschopra.com/tutorials/nn/index.php (accessed August 14, 2011)

[38] Basic Concepts for Neural Networks, http://www.cheshireeng.com/Neuralyst/nnbg.html (accessed August 14, 2011)