ISSN No. 0976-5697

Volume 2, No. 2, Mar-Apr 2011

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Analyzing Soft Computing based Intrusion Detection Systems

Mr. Dharmendra G. Bhatti*
Associate Professor
S. R. Institute of Management and Computer Application,
Bardoli, Gujarat, India
dgbhatti@yahoo.com

Dr. Paresh V. Virparia
Associate Professor
Dept. of Computer Science, S. P. University,
Vallabh Vidyanagar, Gujarat, India
pvvirparia@yahoo.com

Abstract: With the growth of Internet both in its importance and size, network security is also become vital. Intrusion Detection System is one of the important components here. Intrusion Detection Systems based on Soft Computing are currently attracting considerable interest from research community. Characteristics of Soft Computing, such as adaptation, fault tolerance, and error resilience in the face of noisy information, fit the requirements of building a good intrusion detection model. Typically Soft Computing algorithms learn from human knowledge and mimic human skills. Here we have analyzed the application of Soft Computing to the problem of intrusion detection. In this work primarily we have focused on Genetic Algorithm and Neural Network. Such evolutionary techniques are tested effectively using KDD Cup 99 dataset.

Keywords: Network Security, Intrusion Detection System, Soft Computing, Genetic Algorithm, Neural Network

I. INTRODUCTION

Since last few decades computer networks specifically Internet, play an important role in various spheres of human activity. Data and information has become the most precious asset. The world become increasingly dependent on the information since more information is being stored and processed on network-based systems. Due to widespread use of Internet based applications protecting the information to a very high extent becomes crucial. To keep, transmit and automate information processing becomes major problem. Depending on the importance of information different security levels are developed. The security level of processed information rapidly increases from individual to commercial to military. The violation of the information confidentiality, integrity and accessibility may have significant undesirable consequences.

To protect computer systems several traditional approaches and techniques developed over years. Computer networks are typically protected by a number of access restriction techniques like anti-virus, firewall, encryption, secure network protocols, password based authentication, access control list. After applying all these security measures also potential attacker can always find a way to infiltrate into a network. So we need some additional support that would detect different types of security breaches. These support systems are known as intrusion detection systems (IDS) and are placed in Demilitarized Zone and/or inside the protected network, looking for potential threats in network traffic and/or audit data recorded by hosts.

II. INTRUSION DETECTION

Intrusion detection is defined as the process of intelligently monitoring computer system or network, analyzing them for security breaches. The objective of Intrusion Detection System is to protect the availability, confidentiality and integrity of critical networked information systems as per security policy. Intrusion Detection Systems are an important component of defense in depth protecting computer systems and networks from misuse. When IDS is properly deployed it provides alerts indicating that a system is under attack. Intrusion Detection Systems are characterized based on two aspects:

- host based or network based
- anomaly detection or misuse detection

Intrusion detection attempts to detect computer attacks by examining data records observed by processes on the same network. These attacks are typically split into two categories, host-based attacks and network-based attacks. Host-based attack detection routines normally use system call data from an audit process that tracks all system calls made on behalf of each user on a particular machine [22][23]. These audit processes usually run on each monitored machine. Some examples of host-based IDS are: Verisys, OSSEC (a multi-platform open source HIDS), ISS RealSecure Server Sensor (Generic); Tripwire, AIDE (Check host file system), BlackICE, PortSentry (Check host network connections); LogSentry, Swatch (Check host's log files). Network-based attack detection routines typically use network traffic data from a Network sniffer like tcpdump/libpcap. Second aspect is anomaly detection or misuse detection. In a misuse detection based IDS, intrusions are detected by looking for activities that correspond to known signatures of intrusions or vulnerabilities. This is just like anti-virus system. On the other hand, an anomaly based Intrusion Detection System detects intrusions by searching for abnormal network traffic. While anomaly-based IDSs are more capable of identifying new types of attacks than misuse-based system, they also tend to have higher false alarm rates. In this paper we have only considered anomaly-based approach.

We can describe intrusion detection system as a system which processes information coming from the system to be protected. This detector system can also launch investigations to trigger the audit process, such as requesting version numbers for applications. Typically it uses three kinds of information: long-term information related to the technique used to detect intrusions, configuration information about the system's current state, and audit information like the events that are happening to the system. One of the roles of the

detector is to eliminate unneeded information from the audit trail. It then presents either a synthetic view of the actions related to security, taken during normal usage of the system, or a synthetic view of the current security state of the system. Based on these actions or this state, decision is then taken to evaluate the probability of an intrusion or vulnerabilities.

Intrusion detection has become an integral part of the security process with the advent of soft computing for it. Soft Computing is characterized by the use of inexact solutions to computationally-hard tasks such as the solutions of NP-complete problems, for which an exact solution cannot be derived in polynomial time. Many researchers have identified potential and suggested use of soft computing techniques for intrusion detection [2][6][9][12][15][16]. For intrusion detection some soft computing techniques are used extensively: Genetic Algorithm [7][8][13][17] , Neural Network [1][5][14], Fuzzy techniques[20][21], and Support Vector Machine[19]. In this work we have analyzed use of Genetic Algorithm and Neural Network for intrusion detection.

III. ALGORITHM-1 (GENETIC ALGORITHM)

The Genetic Algorithm is based on Darvin's evolutionary theory in order to evolve optimal solutions to a given problem. In nature, organisms fight for survival, and the fittest members tend to have the best chance of living on. These fit spices are then more likely to reproduce and pass on their successful genes to future generations. In this manner, the average fitness of each subsequent generation is normally greater than that of the parent. To avoid local minima problem, mutation is used. Mutation adds randomness to the process by introducing random change to certain members of a population. These mutations can either be beneficial or detrimental.

For finding computational solution also, it works in much the same manner. Individuals in a population are represented by strings also called the genetic codes. These genetic codes represent potential solutions to a given problem, and their effectiveness is represented by a fitness function which takes into account the optimality. The most fit members of the population are the most likely to cross-breed and reproduce off-spring carrying on their attributes. Weaker members are less likely to reproduce, and their features are therefore more likely to obsolete as time goes on. In this manner, subsequent generations tend to be more fit than the previous ones, as the fittest are more likely to survive and pass on their genes. Random mutation introduced to encourage change and prevent the settling into local minima. For Genetic Algorithm based intrusion detection we have used following model:

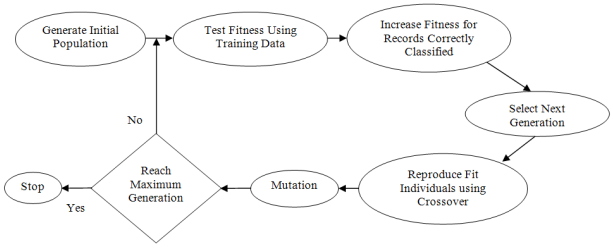

Figure -1 Intrusion Detection using Genetic Algorithm

We have used following fitness function:

$$f(x) = \alpha/A - \beta/B \quad (1)$$

where α is the number of correctly detected attacks, A is the total number of attacks, β is the number of false positives, and B is the total number of normal connections. The fitness function value range was over the closed interval [-1, 1] with -1 being the weakest possible fitness and 1 being the best. A high correct detection rate and a low false positive rate results in a high score on the fitness function. Low detection rate or high false positive rate returns low scores on the fitness function.

IV. ALGORITHM-2 (NEURAL NETWORK)

Neural networks are capable of classification in a wide variety of areas. A neural network is a collection of nodes and weighted connections that are meant to represent biologic neurons. Figure-2 shows the basic structure of a standard neural network.

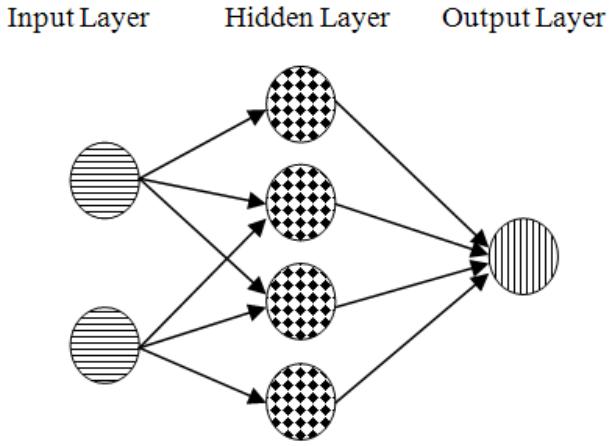

Figure-2 Simple Neural Network

Input is first passed to the input layer, which contains one node for every input into the network. These nodes then pass their input along their connections to the hidden layer. While neural networks generally have only one input and output layer, there can be any number of hidden layers. The hidden layer accepts input from each of its connections to the input layer, and then applies the following operation:

$$y_j = \mathrm{K}\left(\sum_{i=0}^{n} x_i w_{ji}\right) \quad (2)$$

where y_j is the output from the j^{th} hidden node, n is the total number of input nodes, x_i is the i^{th} input, and w_{ij} is the weight on the connection between the i^{th} input and the j^{th} hidden node. K(a) is a predefined function which transforms the weighted sum into a given range. The sigmoidal function is used to transform value into a real number between 0 and 1.

V. IMPLEMENTATION AND RESULTS

The KDD Cup 99 dataset is popularly used as a benchmark dataset in several different research works [24]. The KDD Cup 99 intrusion detection Data Set, which are based on DARPA 98 Data Set, provides labeled data for researchers working in the field of intrusion detection and is the only labeled dataset publicly available. This Data Set is a benchmark Data Set for comparing performance of various IDSs [2][3][5][6][7]. For analyzing GA and NN algorithms, we have also used this

benchmark Data Set. Following is the Class Distributions of 10% KDD99 Data Set:

Table-1 Class Distributions of 10% KDD99	
Class	Number of Connections
Normal	97277
DoS	391458
U2R	52
R2L	1126
Probe	4107
Total	494021

IDS have three common issues: speed, accuracy and adaptability. In our experiment we have focused on detection rate which is computed as the ratio between the number of correctly detected attacks and the total number of attacks.

Table-2 Experimental Results		
	GA Detection Rate	NN Detection Rate
Normal	96.22 %	96.82 %
DoS	97.46 %	97.24 %
U2R	48.54 %	42.54 %
R2L	39.85 %	44.16 %
Probe	97.57 %	97.23 %

These experimental results indicate that both soft computing techniques performed well for specific attack class. Specifically both are performing poor for U2R and R2L attack classes. Compare to statistical approach soft computing techniques are capable of identifying unknown attacks. Genetic Algorithm has good learning capability while Neural Network is a good classifier. Comparison of different approaches is given below:

Table-3 Comparison of Different Algorithms		
	Advantages	Disadvantages
Statistical approach	Higher detection accuracy	Not applicable to NP-complete (non-deterministic polynomial-time complete)
Genetic Algorithm	Good learning ability	Early constringency and parameter selection
Neural Network	Good classifier	Poor knowledge representation and explanation

VI. CONCLUSION

Since many years Soft Computing Techniques are used to solve many computational problems. In the field of Intrusion Detection also researchers have identified potential of Soft Computing Techniques. We have analyzed usage of soft computing techniques in Intrusion Detection Systems. Our experiment has shown that Genetic Algorithm and Neural Network are promising Soft Computing Techniques for Intrusion Detection. We have also observed that both techniques are good in specific attack types. Here we suggest that hybrid soft computing techniques should be experimented for efficient Intrusion Detection.

VII. REFERENCES

[1] Syed Muhammad Aqil Burney, M. Sadiq Ali Khan, Dr.Tahseen A. Jilani, "Feature Deduction and Ensemble Design of Parallel Neural Networks for Intrusion Detection System", IJCSNS International Journal of Computer Science and Network Security, VOL.10 No.10, October 2010

[2] S Selvakani Kandeeban, and Rengan S Rajesh, "Integrated Intrusion Detection System Using Soft Computing", International Journal of Network Security, Vol.10, No.2, PP.87-92, Mar. 2010

[3] Shilpa lakhina, Sini Joseph, and Bhupendra verma, "Feature Reduction using Principal Component Analysis for Effective Anomaly–Based Intrusion Detection on NSL-KDD", International Journal of Engineering Science and Technology Vol. 2(6), 2010, 1790-1799

[4] Kusum Bharti, Shweta Jain, Sanyam Shukla, "Fuzzy K-mean Clustering Via J48 For Intrusiion Detection System", International Journal of Computer Science and Information Technologies, Vol. 1 (4) , 2010, 315-318

[5] Ian Stewart, "A Modified Genetic Algorithm and Switch-Based Neural Network Model Applied to Misuse-Based Intrusion Detection", Master of Science Thesis, Queen's University, Kingston, Ontario, Canada, February 2009

[6] A.Chandrasekar, V. Vasudevan, and P. Yogesh, "Evolutionary Approach for Network Anomaly Detection Using Effective Classification", IJCSNS International Journal of Computer Science and Network Security, VOL.9 No.1, January 2009

[7] Zorana Bankovic,"A Genetic Algorithm-based Solution for Intrusion Detection", Journal of Information Assurance and Security, (2009) 192-199

[8] Maath. K. Al-anni and V. Sundarajan, "Detecting a denial of service using artificial intelligent tools, genetic algorithm", Indian Journal of Science and Technology Vol.2 No 2 (Feb. 2009)

[9] P. Kiran Sree, "Exploring a Novel Approach for providing Software Security Using Soft Computing Systems", International Journal of Security and its Applications Vol. 2, No. 2, April, 2008

[10] Michael Wilkison, "How to Evaluate Network Intrusion Detection Systems?", http://www.sans.org/resources/idfaq/index.php, Retrieved December 2010

[11] Marcus J. Ranum, "Experience Benchmarking Intrusion Detection Systems", December 2001. Retrieved December, 2010 from http://www.snort.org/docs/Benchmarking-IDS-NFR.pdf

[12] Ajith Abraham, Crina Grosan, and Carlos Martin-Vide, "Evolutionary Design of Intrusion Detection Programs", International Journal of Network Security, Vol.4, No.3, PP.328–339, Mar. 2007

[13] Ian Stewart, "A Modified Genetic Algorithm and Switch-Based Neural Network Model Applied to Misuse-Based Intrusion Detection", MS Thesis, Queen's University, Kingston, Ontario, Canada, February 2009

[14] S. SELVAKANI and R.S.RAJESH, "Escalate Intrusion Detection using GA - NN", Int. J. Open Problems Compt. Math., Vol. 2, No. 2, June 2009

[15] Mahmoud Jazzar and Aman Jantan, "A Novel Soft Computing Inference Engine Model for Intrusion Detection", IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.4, April 2008

[16] Jonatan Gomez Perdomo, "SOFT COMPUTING TECHNIQUES FOR INTRUSION DETECTION", A Dissertation Presented for the Doctor of Philosophy Degree The University of Memphis, 2004

[17] Zorana Bankovic, José M. Moya, Álvaro Araujo, Slobodan Bojanic and Octavio Nieto-Taladriz, "A Genetic Algorithm-based Solution for Intrusion Detection", Journal of Information Assurance and Security 4 (2009) 192-199

[18] Sung-Hae Jun and Kyung-Whan Oh, "An Evolutionary Statistical Learning Theory", International Journal of Computational Intelligence 3;3 © www.waset.org Summer 2007

[19] Sung-Hae Jun and Kyung-Whan Oh, "An Evolutionary Support Vector Machine for Intrusion Detection", Asian Journal of Information Technology 5(7): 778-783, 2006

[20] Jonatan Gomez and Dipankar Dasgupta, "Evolving Fuzzy Classifiers for Intrusion Detection", Proceedings of the 2002 IEEE Workshop on Information Assurance, United States Military Academy, West Point, NY June 2001

[21] Mohammad Saniee Abadeh, Jafar Habibi, and Emad Soroush, "Induction of Fuzzy Classification Systems via Evolutionary Aco-Based Algorithms", IJSSST, Vol. 9, No. 3, September 2008

[22] Mohd Nizam Omar and Rahmat Budiarto, "Intelligent Host-based Stepping Stone Detection Approach", Proceedings of the World Congress on Engineering and Computer Science 2009 Vol II WCECS 2009, San Francisco, USA

[23] Wang DongLiang, Wang Hongxin, "The Application and Research of IDS model Based on Multi-technique Fusion", Proceedings of the 2009 International Symposium on Web Information Systems and Applications (WISA'09) Nanchang, P. R. China, May 22-24, 2009, pp. 148-151

[24] H. Güneş Kayacık, A. Nur Zincir-Heywood, Malcolm I. Heywood, "Selecting Features for Intrusion Detection: A Feature Relevance Analysis on KDD 99 Intrusion Detection Datasets", the NIMS Laboratory, http://www.cs.dal.ca/projectx, 2006

[25] Jamil Farshchi, "Statistical-Based Intrusion Detection" http://www.securityfocus.com/print/infocus/1686, Retrieved December 2010

[26] Christos Dimitrakakis, Aikaterini Mitrokotsa, "Statistical decision making for authentication and intrusion detection", IAS technical report IAS-UVA-09-03, http://www.science.uva.nl/research/isla/MetisReports.php, Retrieved December 2010

