International Journal of Scientific Research & Engineering Trends
Volume 6, Issue 4, July-Aug-2020, ISSN (Online): 2395-566X

An Automated Intrusion Detection System by Using Artificial Immune System

Assistant Professor Anjul Rai , M. Tech. Scholar Kalyani Takpure
anjulrai@puibm.in
Peoples University Bhopal,MP,India

Abstract:-In recent years, this area has seen significant advances in with the increasing network attacks worldwide, intrusion detection (ID) has become a hot research topic in last decade. Technologies such as neural networks and fuzzy logic have been applied in ID. The results are varied. Intrusion detection accuracy is the main focus for intrusion detection systems (IDS). Most research activities in the area aim to improve the ID accuracy. This paper focuses on an artificial advanced immune system (AIMS) based network intrusion detection scheme is proposed. An optimal feature selection and parameter quantization algorithms are defined. The difficult subject is addressed in the design of the algorithms. The scheme is tested on the widely used KDD CUP 99 dataset. We found good agreement when comparing our method with results from previous scheme outperforms other schemes in detection accuracy. In our experiments, a number of feature sets have been tried and compared. Compromise between complexity and detection accuracy has been discussed in the paper.

Key words- Intrusion Detection, Negative selection, Artificial Immune System, KDD CUP 99

I.INTRODUCTION

With the enormous development of the computer and network technologies, the security of the network information is becoming increasingly important. New access technologies and devices have increased the possibilities of malicious attacks or service abuse by various hackers. The traditional passive defense technologies like encryptions and firewalls can not fully meet the current security requirements. Therefore, the Intrusion Detection Systems (IDSs) which serves as special purpose systems to detect attacks and misuses in the network is needed.

Generally speaking, two approaches, misuse detection and anomaly detection, can be used in computer systems and computer networks. The misuse detection is used to detect the intrusion when the behavior of the system matches with any of the intrusion signatures. And the anomaly detection, also called as outlier detection [3], is used to detect the intrusion when the given data set does not match with the establishedbeen studied and applied in intrusion detection [11] aiming for better performance. Algorithms such as Genetic Algorithm (GA), Artificial Neural Networks (ANN) and Artificial Immune Systems are widely studied. Among them, AIS is a relatively new comer. Further investigation on AIS based network intrusion detection is needed. The concept of AIS was proposed in mid 1980s. Farmer, Packard and Perelson [29], Bersini and Varela's [30] work have started the area. AIS became a subject of its own in mid 90s. It has been defined by Castro and Timmis [5] as: "Adaptive systems inspired by theoretical immunology and observed immune functions, principles and models, which are applied to problem solving." The early work of applying AIS to IDS can be found in [11]. A multilayer AIS based IDS was proposed by Dasgupta [12] in order to provide systematic defense. These AIS based algorithms have achieved good detection results. But their computing complexity is quite high. In IDS, responding time is an important issue. The more complex the system, the more computing time and the longer responding time will be. Large parameter set in IDS can increase the detection accuracy. However, the more parameters used, the more complex the system. The trade off between the complexity and the accuracy is a challenge. Our study on AIS based IDS is to further improve its detection accuracy while keeping a low algorithm complexity.

In this paper, an AIS based intrusion detection system with some efficient feature selection algorithms is presented. The anomaly detection in the system is set up based on AIS negative selection algorithm. The feature selection algorithm is used to reduce the complexity of the system. The artificial immune system and the negative selection algorithm are introduced in Section II. The AIS based IDS is presented in Section III. Our experiment and results are illustrated in Section IV. Section V draws a conclusion and some future works are discussed

II. ARTIFICIAL IMMUNE SYSTEM

The artificial immune system (AIS) is a branch of bioinspired computational intelligence, and it has attacked increasing interest from the researchers after it was first proposed. Three main algorithms: negative selection, clonal selection, immune network theory compose the

most popular theories of the current AIS research. The negative selection, due to its ability to distinguish the difference between self and non-self, fits naturally into the area of intrusion detection [20].

Negative selection, which is proposed by Forrest et al. [4], is inspired from the negative selection process of the adaptive immune system [5]. The lymphocytes undergo the negative selection during the maturation of T cells in the thymus. It is the major algorithm of the artificial immune system. In the case of an anomaly detection domain, the algorithm prepares a set of exemplar pattern detectors trained on normal (nonanomalous) patterns that model and detect unseen or anomalous patterns [6]. The principle of the negative selection is shown in Figure 1.

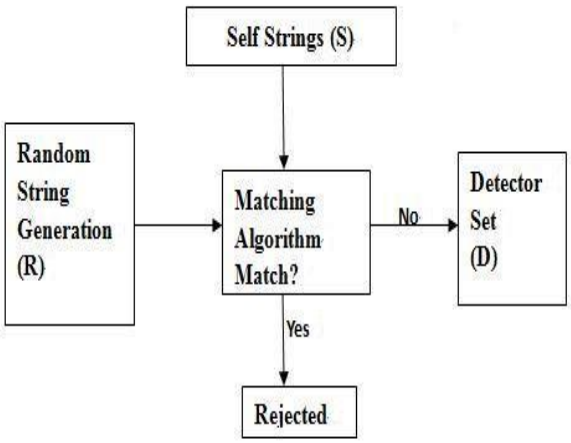

Figure 1 The Principle of Negative Selection.

As shown in Figure 1, the matching algorithm is the core of the negative selection. The affinity between the Ab (Antibody) and Ag (Antigen) is decided by using the matching algorithm. Several algorithms have been proposed in this area to determine the affinity, like Eucliden algorithm, hamming distance algorithm, r-contiguous bit rule algorithm, etc.

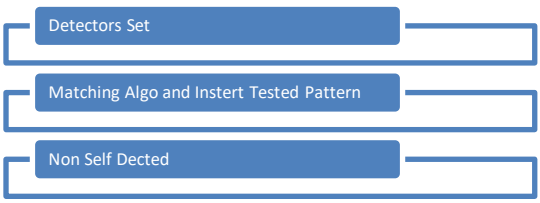

Figure 2 The detection process.

After the detector set is created, it can be used for detection of non-self elements. As shown in Figure 2, for any pattern to be checked, it needs to be compared with all the patterns in the detector set. If it is matched to any pattern in the detector set it will be considered as a non-self element. AIS has been found applications in many areas such as optimization, data analysis, machine learning, pattern recognition, etc and network intrusion detection which is the focus of this paper.

III. PROPOSED SOLUTION

Generally, network intrusion detection is based on the examination of monitored network parameters. Different examine algorithms lead to different IDS. The general AIS based IDS [16]can be divided into two parts, i.e. detector set generation and non-self-detection. To form the detection set, negative selection algorithm is applied. A large set of normal network parameter patterns are required. Initially, the immature detectors (parameter pattern in this case) are randomly generated as shown in figure 1. Then, these immature detectors are compared with the normal network parameter patterns. If a random generated pattern matches a normal pattern, the immature detector will be rejected and deleted. Those which do not match any normal network parameter patterns will be saved as mature detectors. In the live detection stage, a monitored network parameter pattern is compared with detectors in the detector set. If it is matched with any detector, then a network intrusion is detected.

A compact and effective detector set can reduce the algorithm computing complexity. For detectors which do not contribute any detection in a period of time, they should be removed or put to a sleeping state. Therefore, all the mature detectors will have a time_to_live (TTL) parameter. Whenever detection is occurred, all detectors' TTLs will be deducted by one except for the detector which detects the intrusion. ItsTTLwill be reset to the maximum. When a detector reaches its lifetime, ie its TTL becomes zero; this detector will be become inactive

1. Experimental Database

The KDD Cup 99 data set, which is the most widely used data set for network-based intrusion detection, is used in our project. This data set is built based on the data captured in DARPA'98 IDS evaluation program [13].The data set contains 24 training attack types and 14 additional attack types in the test data only. These attacks fall into four main categories:

1.1. Denial of service (DOS): In this type of attack an attacker makes some computing or memory resources too busy or too full to handle legitimate requests, or denies legitimate users access to a machine. Examples are Apache2, Back, Land, Mailbomb, SYN Flood, Ping of death, Process table, Smurf.

1.2. Remote to user (R2L): In this type of attack an attacker who does not have an account on a remote machine sends packets to that machine over a network and exploits some vulnerability to gain local access as a

user of that machine. Examples are Dictionary, Ftp_write, Guest, Imap, Named, Phf, Sendmail, Xlock.

1.3. User to root (U2R): In this type of attacks an attacker starts out with access to a normal user account on the system and is able to exploit system vulnerabilities to gain root access to the system. Examples are Eject, Loadmodule, Ps, Xterm, Perl, Fdformat.

1.4. Probing: In this type of attacks an attacker scans a network of computers to gather information or find known vulnerabilities. An attacker with a map of machines and services that are available on a network can use this information to look for exploits.
Examples are Ip sweep, Mscan, Saint, Satan, Imap.

The data set has 41 attributes for each connection record plus one class label, which points out whether the data is a normal one or an attack one. The 41 parameters are listed in Table 1.

Table 1 KDD CUP 99 parameter.

No.	Feature	No.	Feature
1	duration	2	protocol_type
3	Service	4	flag
5	src_bytes	6	dst_bytes
7	land	8	wrong_fragment
9	Urgent	10	hot
11	num_failed_logins	12	logged_in
13	num_compromised	14	root_shell
15	su_attempted	16	num_root

2. Feature Selection

As shown in Table 1, forty one parameters (attributes) for each pattern (or record) are too many to be used for intrusion detection as some of them may related to others and do not contribute much in the detection. If all are used, the system complexity will be high. A subset of the parameter set can be used to achieve the similar detection result without exhausting the system in computing. Therefore selecting most important and independent parameters to form the subset is a key process in any IDS. There are some studies on the feature selection of KDD data set. Three popular paradigms of the selection are standing out. They are rough set theory (RS), linear genetic programming (LGP), and multivariate adaptive regression splines (MARS).

IV. RESULT AND DISCUSSION

The raw dataset which we used to generate detectors contains about five million connection records, 700 million bytes. Meanwhile, the testing data we choose contains 300,000 records, and about 45 million bytes.
Three feature selection algorithms have been tested in our experiment. The results are shown in Table 2. TP (true positive) represents that an abnormal pattern is successful detected. FN (false negative) means an abnormal pattern is falsely recognized as a normal pattern. FP (false positive) means that the normal data is mistakenly detected as an abnormal pattern (i.e. an attack) and this is a false alarm

V. CONCLUSINO AND FUTRUE WORK

This paper investigate several question related to an artificial immune system-based intrusion detection system is presented. Negative selection is the algorithm used in the AIS. A number of feature selection algorithms have been tried and compared in experiments using the KDD Cup 99 dataset. The parameter quantization proposed is aiming to reduce the complexity and maintain the detection performance. The system has shown excellent detection accuracy. The experiments shows MARS feature selection algorithm has the best detection accuracy. In the future work, an adaptive mechanism will be introduced to AIS, so that the detector set will be adaptively updated so that the system can adapt to changes in the network situation. Also more dataset will be tried for the performance testing and verification. Other future works include searching and exploring new feature selection algorithms and computing optimization.

REFERENCE

[1] Rebecca Copeland, Converging NGN Wireline and Mobile 3G Network with IMS, Taylor & Francis Group, U.S.A, 2009

[2] Michael T.Hunter, Russell J.Clark, Frank S. Park, “Security Issues with the IP Multimedia Subsystem (IMS): A White Paper”,

[3] Hans-Peter Kriegel, Peer Kröger, Arthur Zimek (2009). "Outlier Detection Techniques (Tutorial)". 13th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2009) (Bangkok,Thailand). http://www.dbs.ifi.lmu.de/Publikationen/Papers/tutorial_slides.pdf. Retrieved 2010-06-05.

[4] Steven A. Hofmeyr and S. Forrest,”Architecture for an ArtificialImmune System”, Evolutionary Computation Journal, pp. 443-473, 2000.

[5] eandro N. de Castro and Jonathan Timmis, “Artificial Immune Systems: A New Computational Intelligence Approach”, Springer, 2002.

[6] Forrest, S.; Perelson, A.S.; Allen, L.; Cherukuri, R. (1994). "Self-nonself discrimination in a computer" (PDF). Proceedings of the 1994 IEEE Symposium on Research in Security and Privacy. Los Alamitos, CA. pp. 202–212. http://www.cs.unm.edu/~immsec/publications/virus.pdf.

[7] 3GPP Technical Specification of Security: http://www.3gpp.org/ftp/Specs/html-info
[8] 3GPP TS 33.203: Third Generation Partnership Project; Technical Specification Group Services and System Aspects; 3G securtiy; Access security for IP-based services (Release 7).
[9] 3GPP TS 33.210: Third Generation Partnership Project; Technical Specification Group Services and System Aspects; 3G securtiy; Network Domain Security; IP network layer security (Release 7).
[10] Chi-Yuan Chen, Tin-Yu Wu, Yueh-Min Huang, Han-Chieh Chao, "An Efficient end-to-end security mechanism for IP multimedia subsystem", Computer Communications archive. Volume 31, Issue 18, December 2008, page: 68-81.
[11] S. A. Hofmeyr and S. Forrest, "Immunity by design: An artificial immunb system," in Proceedings of the Genetic and Evolutionary Computation Conference. San Mateo, CA: Morgan Kaufmann, July 1999, pp.1289–1296.
[12] D. Dasgupta. Immunity-based intrusion detection systems: A general framework. presented at 22nd Nat. Information Systems Security Conf.. [Online]. Available: http://csrc.nist.gov/nissc/1999/proceedings/papers/p11.pdf
[13] R.P.Lippmann, D.J.Fried, I.Graf, J.W.Haines, K.R.Kendall, D.McClung, D.Weber, S. E. Webster, D. Wyschogrod, R. K. Cunningham, and M. A. Zissman, "Evaluating intrusion detection systems: The 1998 darpa offline intrusion detection evaluation," discex, vol. 02, p. 1012, 2000.
[14] Anazida Zainal, MohdAizainiMaarof and Siti Mariyam Shamsuddin, "Feature Selection Using Rough Set in Intrusion Detection", TENCON 2006. 2006 IEEE Region 10 Conference, 14-17 Nov. 2006, pp.1-4
[15] GurselSerpen and MaheshkumarSabhnani, "Measuring similarity in feature space of knowledge entailed by two separate rule sets", Knowledge-Based Systems, Volume 19, Issue 1, March 2006, pp. 67-76.
[16] Junyuan Shen and Jidong Wang, "Network intrusion detection by artificial immune system," IECON 2011 - 37th Annual Conference of the IEEE Industrial Electronics Society, Melbourne, VIC, 2011, pp. 4716-4720, doi: 10.1109/IECON.2011.6119993.