Optimized Intrusion Detection System Using Computational Intelligent Algorithm

P. J. Sajith and G. Nagarajan

Abstract The broad development of the radio frequency identification (RFID) in internet of things (IoT) application has provoked system interruption recognition, which turn into a basic part of intrusion detection system. Because of the open society of the IoT, the security of IoT frameworks and information is dependably in danger. The major objective of this research paper is to design an intrusion detection system framework using Anomaly-Based Detection technique. Optimization of interesting rules from a dense database is determined, using computation intelligent algorithm such as genetic algorithm (GA), genetic programming (GP), and swarm intelligence algorithm.

Keywords Internet of things · RFID · Genetic algorithm · Genetic programming · Swarm intelligence

1 Introduction

An intrusion detection system (IDS) is a software application or hardware appliance that monitors traffic moving on networks and through systems to search for suspicious activity and known threats, sending up alerts when it finds such items. There are two types of IDS system, host-based IDS and network-based IDS. In host-based IDS, a software intelligent agent would monitor the input and output packets from devices. It performs log analysis, file integrity checking, policy monitoring, rootkit detection, real time alerting, and active response. In network-based IDS, sensor will do the monitoring work. The connected network monitors and analyze the network traffics.

P. J. Sajith (✉) · G. Nagarajan
Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Jeppiaar Nagar, Chennai, India
e-mail: pjsajith@gmail.com

G. Nagarajan
e-mail: nagarajanme@yahoo.co.in

P. K. Mallick et al. (eds.), *Advances in Electronics, Communication and Computing*, Lecture Notes in Electrical Engineering 709,
https://doi.org/10.1007/978-981-15-8752-8_64

Similarly, there are two types of IDS techniques, Signature-based IDS and Anomaly-based IDS. In Signature-based IDS, a specific signature pattern is used to analyze the content of each packets in all 7 layers. Whereas, in Anomaly-based IDS, it monitors the network traffic and it compares it against standard baseline for normal use. These classification helps to identify whether it is normal or anomalous network.

2 Background

The Anomaly-Based Detection (ABD) [1] identifies the intrusion detection based on the behavior observation. If there is any change in the normal activity, it will be notify. There are two type of anomaly detection self-learning system and programmed model. Programmed model (ABD), in this model, the system will be trained to detect any abnormal changes. The administrator decide a threshold to flags system if any abnormality was there. Self-learning system (ABD) operated by a set of standard normal operation. This model is structured by observing the network strategies over a set of time. Lu and Traore [2] implemented a genetic programming-based intrusion detection system. They used DARPA dataset. According to them, the FPR is low. Bankovic et al. [3] used KDD99cup dataset. They used principal component analysis-based method to extract data.

3 Proposed System Design

The overall functional diagram of the proposed system is shown in the Fig. 1. The information collected over time regarding the network and the corresponding data are extracted and stored in a relational database after pre-processing. From the database, the required data knowledge are extracted using GNP-based fuzzy rule extraction method [4]. The rules initially defined are updated by computing the support, confidence, and the chi-square attributes. According to this, the datasets are classified.

Using this system, the intruders can be classified accurately using the proposed GNP-based classifier [10]. This classifier [5, 6] used both binary and continuous values for rule extraction. The working principle of the above system is explained below:

The extracted dataset consist of source IP address, destination IP address, and source and destination port number. During pre-processing, the missing elements and redundant data are all eliminated. As shown in Fig. 2.

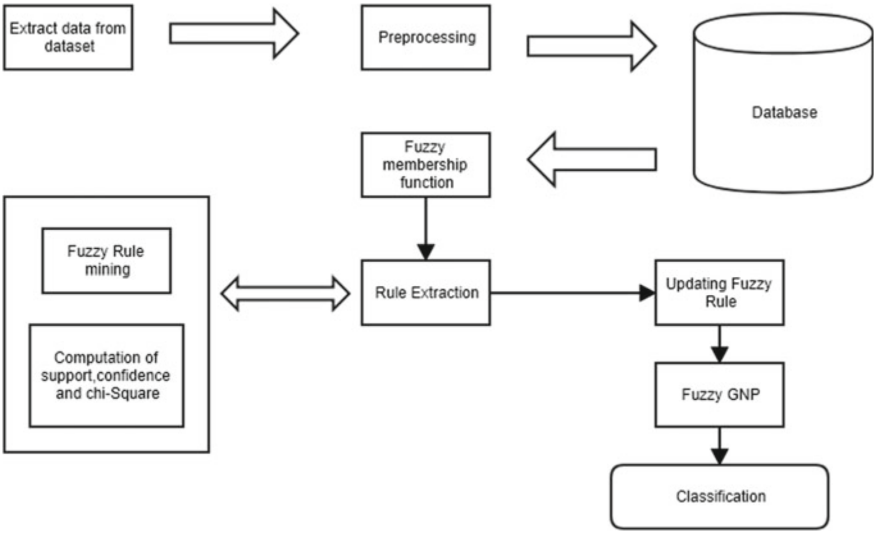
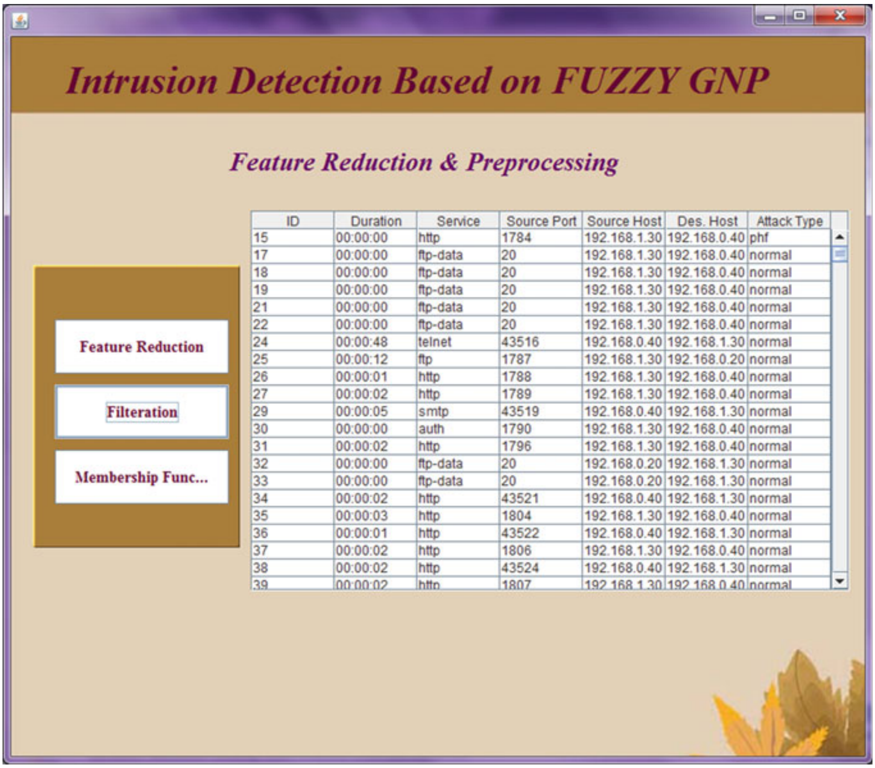


Fig. 1 Overall system design



ID	Duration	Service	Source Port	Source Host	Des. Host	Attack Type
15	00:00:00	http	1784	192.168.1.30	192.168.0.40	phf
17	00:00:00	ftp-data	20	192.168.1.30	192.168.0.40	normal
18	00:00:00	ftp-data	20	192.168.1.30	192.168.0.40	normal
19	00:00:00	ftp-data	20	192.168.1.30	192.168.0.40	normal
21	00:00:00	ftp-data	20	192.168.1.30	192.168.0.40	normal
22	00:00:00	ftp-data	20	192.168.1.30	192.168.0.40	normal
24	00:00:48	telnet	43516	192.168.0.40	192.168.1.30	normal
25	00:00:12	ftp	1787	192.168.1.30	192.168.0.20	normal
26	00:00:01	http	1788	192.168.1.30	192.168.0.40	normal
27	00:00:02	http	1789	192.168.1.30	192.168.0.40	normal
29	00:00:05	smtp	43519	192.168.0.40	192.168.1.30	normal
30	00:00:00	auth	1790	192.168.1.30	192.168.0.40	normal
31	00:00:02	http	1796	192.168.1.30	192.168.0.40	normal
32	00:00:00	ftp-data	20	192.168.0.20	192.168.1.30	normal
33	00:00:00	ftp-data	20	192.168.0.20	192.168.1.30	normal
34	00:00:02	http	43521	192.168.0.40	192.168.1.30	normal
35	00:00:03	http	1804	192.168.1.30	192.168.0.40	normal
36	00:00:01	http	43522	192.168.0.40	192.168.1.30	normal
37	00:00:02	http	1806	192.168.1.30	192.168.0.40	normal
38	00:00:02	http	43524	192.168.0.40	192.168.1.30	normal
39	00:00:02	http	1807	192.168.1.30	192.168.0.40	normal

Fig. 2 Data pre-processing

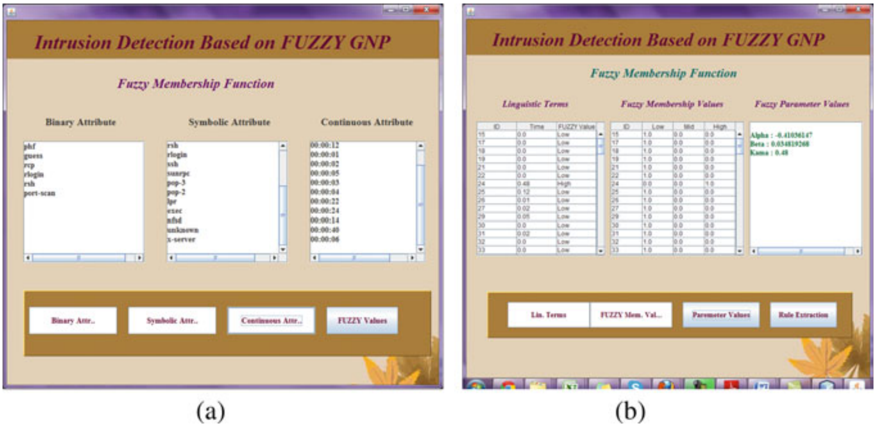


Fig. 3 Fuzzy attribute calculation

For the convenience of fuzzy rule formation, the continuous attributes of the database are linguistically transformed as α, β, and γ to represent low, mild, and high attributes, respectively. To combine the discrete and continuous values in this paper, GNP-based fuzzy rule mining is used. The fuzzy rules are extracted and updated using the confidence and support values [11]. This above process is shown in Fig. 3a, b.

Another important parameter used to update are chi-square value (C). If (X, Y) be the support value of a x_i and y_j. Then, the updated C value for N tuples is calculated as shown in Eq. (1). Where z is the union of (x) and (y). The implementation result shown in Fig. 4.

$$C = \frac{N(z - x.y)^2}{xy(1 - x)(1 - y)} \tag{1}$$

The fitness (f) of the fuzzy rule [7] is determined by the following equation Eq. (2) and shown in Fig. 5. Where d_{r}, d_{ir} are the correctly and incorrectly determined data. T and N are the total number of trained and test data, respectively. The value is scaled between [−1, 1]. If the value is high, then the positive false rate is low and vice versa.

$$f = \frac{d_{\mathrm{r}}}{T} - \frac{d_{\mathrm{ir}}}{N} \tag{2}$$

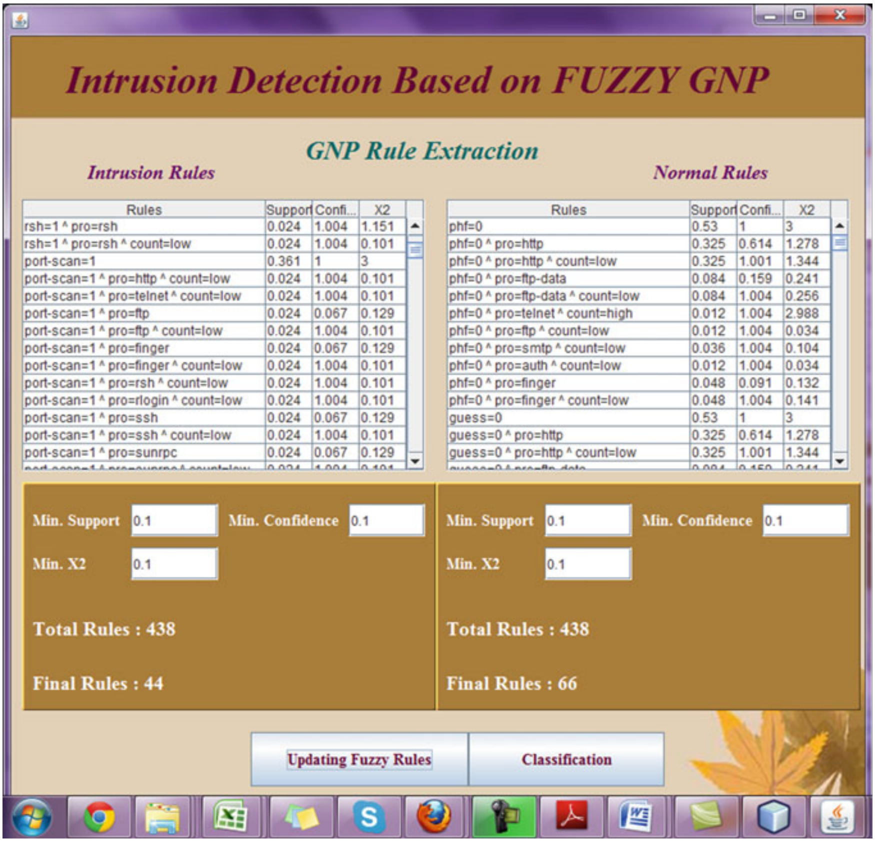


Fig. 4 Chi-square updation

4 Conclusion

In this work, based on fuzzy rule generation a GNP classifier is designed for sub-attribute selection and utilization. This intrusion detection-based classifier [8, 9] is used to detect anomaly in the network. This proposed system extract many effective rules, which can be used for anomaly detection.

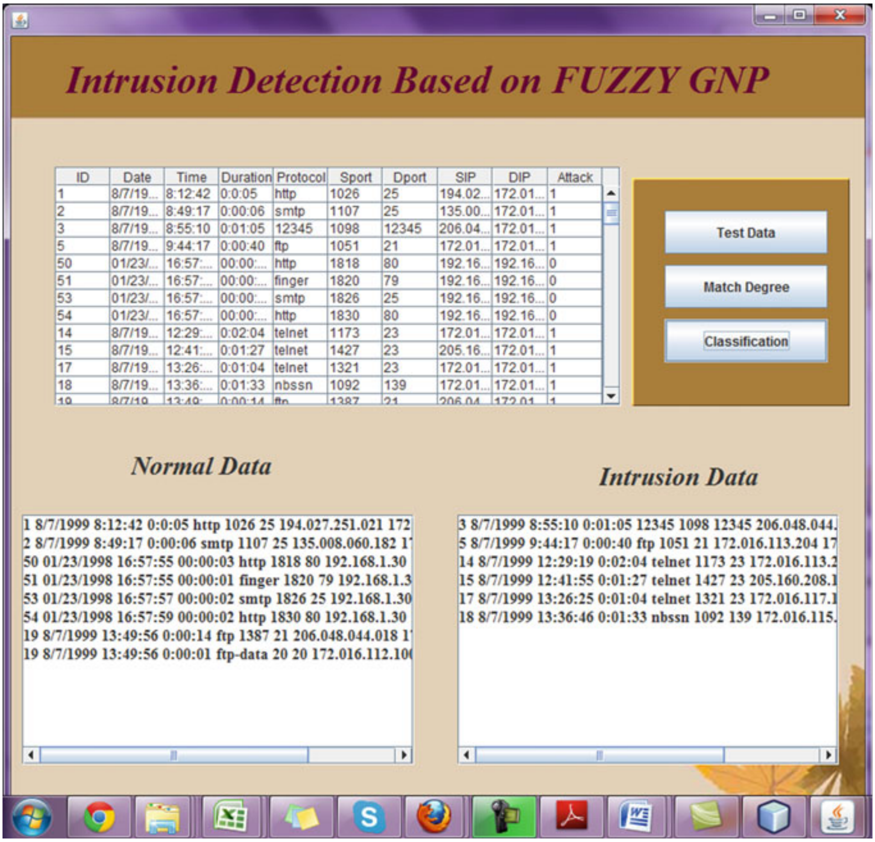


Fig. 5 GNP-based fuzzy rule implementation

References

1. Elhag, S., Fernández, A., Altalhi, A., Alshomrani, S., & Herrera, F. (2019). A multi-objective evolutionary fuzzy system to obtain a broad and accurate set of solutions in intrusion detection systems. *Soft Computing, 23*(4), 1321–1336.
2. Lu, W., & Traore, I. (2004). Detecting new forms of network intrusion using genetic programming. *Computational Intelligence, 20*(3), 475–494.
3. Banković, Z., Stepanović, D., Bojanić, S., & Nieto-Taladriz, O. (2007). Improving network security using genetic algorithm approach. *Computers & Electrical Engineering*, *33*(5-6), 438–451.
4. Elhag, S., Fernández, A., Alshomrani, S., & Herrera, F. (2019). Evolutionary fuzzy systems: A case study for intrusion detection systems. In *Evolutionary and Swarm Intelligence Algorithms* (pp. 169–190). Springer, Cham.
5. Minu, R. I., & Thyagharajan, K. K. (2011). Scrutinizing the video and video retrieval concept. *International Journal of Soft Computing & Engineering, 1*(5), 270–275.
6. Thyagharajan, K. K., & Minu, R. I. (2013). Prevalent color extraction and indexing. *International Journal of Engineering and Technology, 5*(6), 4841–4849.

7. Rajalakshmi, T., & Minu, R. I. (2014, February). Improving relevance feedback for content based medical image retrieval. In *International Conference on Information Communication and Embedded Systems (ICICES2014)* (pp. 1–5). IEEE.
8. Thamilarasu, G., & Sridhar, R. (2008, November). Intrusion detection in RFID systems. In *MILCOM 2008–2008 IEEE Military Communications Conference* (pp. 1–7). IEEE.
9. Yang, H., Guo, J., & Deng, F. (2011). Collaborative RFID intrusion detection with an artificial immune system. *Journal of Intelligent Information Systems, 36*(1), 1–26.
10. Ezhilarasi, R., & Minu, R. I. (2012). Automatic emotion recognition and classification. *Procedia Engineering, 38,* 21–26.
11. Madhu, K., & Minu, R. I. (2013, February). Image segmentation using improved JSEG. In *2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering* (pp. 37–42). IEEE.