Research Journal of Information Technology

ISSN 1815-7432

Academic Journals Inc.

www.academicjournals.com

OPEN ACCESS Research Journal of Information Technology

ISSN 1815-7432
DOI: 10.3923/rjit.2017.1.6

Review Article
Advances in Intrusion Detection Algorithms for Secure E-business Using Artificial Intelligence

[1]Hiren K. Mewada and [2]Sanjay Patel

[1]Department of Electronics and Communication Engineering, Charotar University of Science and Technology, At Post Changa, 388421 Gujarat, India
[2]30200 Telegraph Road, Suite 300, Bingham Farms MI 48025, United States

Abstract

Development in information technology increased the demand of the electronics-commerce business for faster delivery with reduction in the cost. The major challenging task in the development of E-commerce system is the security at the user-end for web transaction. This study analyzed the security algorithms using intrusion detection. Further adaption in the technologies needs continuous monitoring of intrusion detection and hence, comparative analysis of adaptive artificial intelligent techniques based intrusion detection algorithms presented in this study.

Key words: Artificial intelligence, intrusion detection, E-business, network security, Gaussian mixture model, hidden Markov model, fuzzy system, neural network, genetic algorithm.

Received: October 21, 2016 **Accepted:** November 21, 2016 **Published:** December 15, 2016

Citation: Hiren K. Mewada and Sanjay Patel, 2017. Advances in intrusion detection algorithms for secure E-business using artificial intelligence. Res. J. Inform. Technol., 9: 1-6.

Corresponding Author: Hiren K. Mewada, Department of Electronics and Communication Engineering, Charotar University of Science and Technology, At Post Changa, 388421 Gujarat, India Tel: 02697-265062

Competing Interest: The authors have declared that no competing interest exists.

Data Availability: All relevant data are within the paper and its supporting information files.

INTRODUCTION

In the era of digitalization and modernization, the need of business organization can only be satisfied using electronic commerce (E-commerce). E-commerce increases the speed of delivery by providing the quality goods with reduction in cost. Some time it is also refer as paperless business information exchange. Its multi-functional features like non-cash payment, 24×7 services, better marketing management, pre-sales and after sales services with fast and reliable communication with customer makes it popular amongst the business partner[1]. It is uniform platform for sharing the information, maintain the strategy and to support commercial activities across the world. Apart from these advantages, E-commerce has major two deficiencies in terms of technical and non-technical. The main technical disadvantage is its security due to inefficient algorithm implementation. At user end, it is very difficult to provide security on web transaction. The financial loss associated with the user and proprietor is not bearable. The factor, which is responsible for this loss in E-commerce is downtime associated due to attack by intrusion. Therefore, this study presents an analysis of security algorithm using intrusion detection algorithms.

INTRUSION DETECTION SYSTEM

Intrusion can be detected by analyzing the network or audit trail in the operating system[2]. Several organization have developed and are using Intrusion Detection System (IDS) like multics based detection and alerting of intrusion by national security agency[3], computer watched developed by AT and T[4], netwrok anomaly detection and intrusion reporting by Los-Almos national laboratory[5] etc. The popularity amongst the users in day life demands proactive steps for the risk of intrusion. The possible examples of intrusion include browsing from unsecured systems, assessment of file server program, by providing super user status, installation of snooping program and so on[6]. Unauthorized access can be prevented using firewall and authentication system. However, they fail to monitor network traffic and as most attacks are being traffic it is required to monitor network traffic. It is expected that IDS should be robust to the failure, fault or crash occurs in the system. In addition to that, IDS must be configurable with the change in the network by maintaining the high performance[7].

Host based IDS collects activity information on host system. Agents also refer as sensors are installed in host system to collect the information of events running in the host systems and keep records of these collected data inform of audit trails or log files. A host based IDS depends heavily on Operating System (OS) which in turn require attention of OS developer[8]. Host based system is preferable choice due to its existence in the operating system and the detail information it can provide. The weakness of this system is accumulation of large data collected by the sensors. The finer level data collection makes it complex and hence it is easier for intruders to hide their footprint. Another challenge with the host-based system is that sensors must be compatible with the operating system.

Network based system monitor the traffic data over the network. It checks the contents and header information of the packets within the network. Rules are defined to identify the types of attacks. Using these rules, the sensors identify the packets in traffic. The network-based IDS is cheaper than the host-based IDS. It is portable and independent of operating system. It is the best choice for the businesses running on self-develop software. It is compatible with all types of network configuration and adaption to the system resources. The weakness of this system is as follows: (a) Identification of attacks heavily depends on the rules or signatures defined and these rules are designed from the known and previous attacks. (b) Sensors shall inspect all the packets and it is not capable to handle high-speed traffic data i.e., more than gigabyte speeds[9]. (c) The network-based system is not able to monitor the traffic moving over other media i.e., PSTN network. (d) The identification of network attacks may be limited due to encryption and switching technology.

The IDS can also be classified into signature-based and anomaly-based IDS, using the methods used to check its vulnerability. As suggested previously, the rules also refer as signatures are used as features to compare the analyzing packets in the network then it is called signature based IDS. In anomaly-based IDS, the normal tranfic profile is used to detect the intrusion in the system.

Overall network attacks can be categorized as follows[10]:

- **Denial of service:** In this attack, the large computer memories have been occupied for other processes and hence take more time to respond the server
- **Remote to user attack:** In this attack, a packet is transmitted over the network and this packet is not accessible by the users. This packet exposes the vulnerability in the network
- **User to root attack:** In this type of attack, the intruder uses the system as a normal user and and attempts for vulnerability in the network

- **Probing:** In this attack, the intruder studies and determines the weakness of network system and later it probe the vulnerabilities

As intrusion detection is the basic requirement for the E-commerce applications, use of any one system is not preferable. Therefore, hybrid approach of both host and network-based approaches seems to be preferable. The common weakness in both systems is signatures used and this requires large collection of data and knowledge of past attacks. Artificial intelligence plays a vital role to learn about various attacks from limited collection of data. The learning characteristic makes it generalized and ideal candidate to classify the intrusion. The quantitative evaluation of IDS is obtained by measuring the False Positive (FP) and False Negative (FN). The FP means normal event is predicted as intrusive. The FN means identification of event is classified as false event in the system. True Positive (TP) measures the actual positive events. Performance using these parameters can be calculated using dice similarity coefficients defined as:

$$DSC = \frac{TP \times 100\%}{(TP+FN)}$$

This study discusses various AI based system for both host and network based IDS.

ARTIFICIAL INTELLIGENCE BASED INTRUSION DETECTION SYSTEM

Artificial intelligence has been widely used in decision-making system and for the classification in computer vision applications. It is utilized in various image processing applications including pattern recognition, image analysis and image registration. Further, this AI can be classified in supervised, unsupervised and hybrid learning methods. In supervised learning, desired output is used to learn the algorithm for better classification. In unsupervised AI inputs are analyzed without target output[11]. Hybrid approach integrates both supervise and unsupervised learning method to generates specified rules and/or signature. The block diagram for AI based intrusion detection system is shown in Fig. 1.

Maximum likelihood and hidden Markov model based IDS: In this system, hidden Markov model contains numbers of hidden layers and the events are observed at each layer. The probabilities of events are calculated using these event's observation. Initially HMM model is trained using both normal events and intrusion detected events. Later the expectation maximization algorithm is used to estimate the maximum likelihood parameters. Yeung and Ding[11] used to normal data only to train the HMM. They implemented dynamic modeling approach in opposed to static modeling approach with reduction in storage and computational time. Jia and Zhong[12] used HMM model and defined the transition rules of system call. They obtained the short sequence call using viterbi algorithm. To detect the intrusion, the sequence generated by the system call is compared with maximum likelihood sequence. This model offers faster computation speed with better detection.

Gaussian mixture model based IDS: It is statistical parameter based learning model. In this model, the traffic are grouped according to its statistical parameters. For grouping k-means clustering is preferred. Then GMM is used to train the algorithm. Fawwaze[13] initially formed a group of traffic data in the training phase using unsupervised learning method.

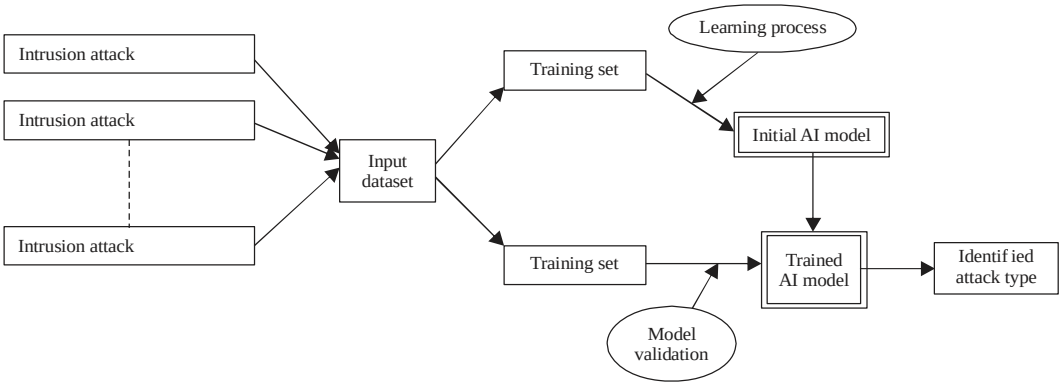

Fig. 1: AI based intrusion detection system

To cluster the traffic, Gaussian based finite mixture model has been used which offers fast computation and no intrusive data for training the model. Apostriori (MAP) algorithm is used to detect and classify the intrusion. This model is validated using NS2 and found reduction in false alarm rates. Bahrololum and Khaleghi[14] proposed hierarchical GMM approach. They evaluated their model using KDD99 datasets. This model has less-missing rate in comparison with bianary tree classifier, SOM and ART model.

Fuzzy based intrusion detection system: In fuzzy based system, rules are defined based on the reasons about the quantity like how many IP addresses approached the system in last few seconds. If rate of change of IP address is high then it will be considered as detection of intrusion. Sekeh and Maarof[15] integrated data mining with fuzzy model with reduction in size of datasets and less computation time. Khurana *et al.*[16] reviewed the various fuzzy based model for intrusion detection system and concluded that fuzzy based system requires few datasets for training the model with better performance.

Genetic algorithm for intrusion detection system: Genetic algorithm selects the sample points from the search space and recombines them to form new sample points. Features extracted from the sample data are assigned different weight-age and k-nearest neighbor classifiers are used as a fitness function to evaluate the performance of new weighted feature sets. Many literatures proposed the use of GA for intrusion detection[17]. The comparative analysis of GA with SVM classifier, neural network has been established using KDD99 datasets[18]. Benaicha *et al.*[19] presented improvement in initial population and selection operator. They verified the attacks in the audit file in reasonable time.

Neural network based intrusion detection system: In comparison with fuzzy, a neural network generates results that are more accurate. However, training and optimization steps in neural network increase its complexity then fuzzy based system. Thus for better accuracy neural network is preferred over the fuzzy[20]. Srinivasu and Avadhani[21] concluded that training of neural network using GA provides better performance over SVD based training algorithm.

Support Vector Machine (SVM) based intrusion detection: The SVM offers scalable implementation of IDS. The major limitation of SVM is its larger training time. Traditional

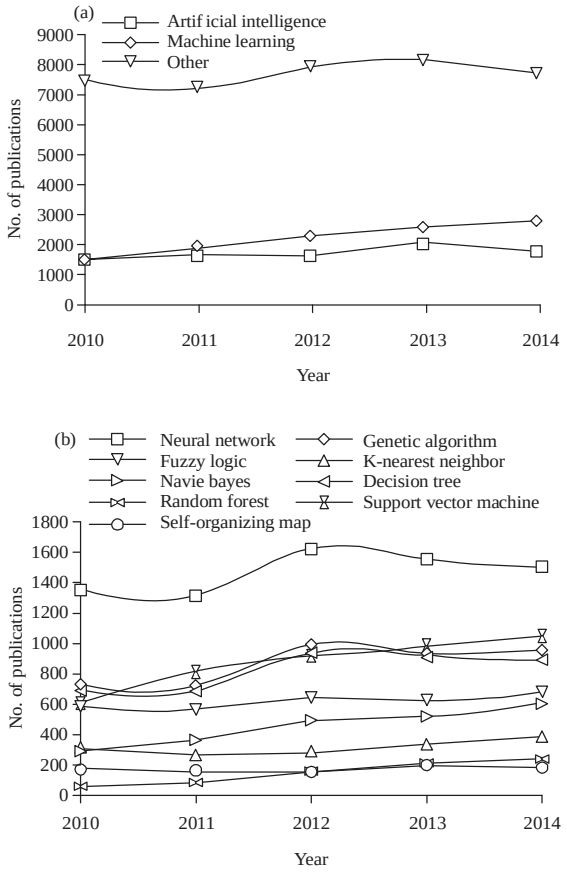

Fig. 2(a-b): (a) AI and machine learning intrusion detection popularity in comparison with other methods and (b) Popularity analysis of various AI methods

clustering algorithm used in SVM was replaced with self-organizing tree that grows dynamically in the system[22]. It improved the training phase significantly. Heba *et al.*[23] proposed the selection of feature sets using principle components and these feature sets are used in SVM based IDS. Thus reduction in number of feature sets reduced the training phase of SVM.

Quantitative analysis: It can be quantified by the number of publication accepted over the years. Stamper[24] presented publication for the artificial intelligence and machine learning for the period of 2010-14. Figure 2a presents that AI and machine learning are most outstanding methods for intrusion detection. Maximum likelihood based IDS has been focused more in comparison to other AI techniques. Figure 2b presents the various AI techniques for 2010-14 duration. Neural network is most accepted algorithm amongst the genetic algorithm and SVM methods.

CONCLUSION

The use of E-commerce system is increasing day by day. With digitization in system, security becomes major apprehension today. This demand continuous refinement and evolution in the intrusion detection systems. Due to properties of artificial intelligence like its adaptively, self-learning and independent of platform make it suitable choice for the intrusion detection system. Further integration of above discussed machine learning algorithm can overcome the limitation offered by individual machine learning algorithms.

SIGNIFICANCE STATEMENTS

This study discusses the network security issues faced in the development of E-commerce system. Major contributions of this study are:

- Network intrusion detection is classified and various network attacks are summarized
- Role of Artificial Intelligence (AI) in intrusion detection method is explained
- Use of AI methods in comparison with other intrusion detection methods is summarized

REFERENCES

1. Anonymous, 2016. E-commerce-overview. https://www.tutorialspoint.com/e_commerce/ecommerce_overview.htm
2. Boyce, C.A.P. and A.N. Zincir-Heywood, 2016. A comparison of four intrusion detection systems for secure E-business. https://www.cs.dal.ca/~zincir/bildiri/icecr03-cn.pdf
3. Sebring, M.M., E. Shellhouse, M.E. Hanna and R.A. Whitehurst, 1988. Expert systems in intrusion detection: A case study. Proceedings of the 11th National Computer Security Conference, October 17-20, 1988, Baltimore, MD., pp: 74-81.
4. Dowell, C. and P. Ramsted, 1990. The computer watch: Data reduction tool. Proceedings of the 13th National Computer Security Conference, October 1-4, 1990, Washington, DC., pp: 99-108.
5. Hochberg, J., K. Jackson, C. Stallings, J.F. McClary, D. DuBois and J. Ford, 1993. NADIR: An automated system for detecting network intrusion and misuse. Comput. Secur., 12: 235-248.
6. Puketza, N.J., K. Zhang, M. Chung, B. Mukherjee and R.A. Olsson, 1996. A methodology for testing intrusion detection systems. IEEE Trans. Software Eng., 22: 719-729.
7. Pappas, N., 2008. Network IDS deployment strategy. SANS Institute, USA. https://www.sans.org/reading-room/whitepapers/detection/network-ids-ips-deployment-strategies-2143.
8. Anonymous, 2005. Host vs Intrusion detection system. SANS Institute, USA. https://www.giac.org/paper/gsec/1377/host-vs-network-based-intrusion-detection-systems/102574.
9. Shipley, G., 1999. Intrusion detection, take two. J. Network Comput., 10: 44-48.
10. Hoque, M.S., M. Mukit, M. Bikas and A. Naser, 2012. An implementation of intrusion detection system using genetic algorithm. Int. J. Network Secur. Applic., 4: 109-120.
11. Yeung, D.Y. and Y. Ding, 2003. Host-based intrusion detection using dynamic and static behavioral models. Pattern Recognit., 36: 229-243.
12. Jia, C. and A. Zhong, 2016. An intrusion detection model based on the maximum likelihood short system call sequence. Proceedings of the International Conference on Intelligent Computing, August 16-19, 2006, Kuniming, China, pp: 709-714.
13. Fawwaze, E.M., 2008. Intrusion detection for mobile *ad hoc* network. Ph.D. Thesis, Cairo University, Egypt.
14. Bahrololum, M. and M. Khaleghi, 2008. Anomaly intrusion detection system using Hierarchical Gaussian mixture model. Int. J. Comput. Sci. Network Secur., 8: 264-271.
15. Sekeh, M.A. and M.A.B. Maarof, 2009. Fuzzy intrusion detection system via data mining technique with sequences of system calls. Proceedings of the 5th International Conference on Information Assurance and Security, Volume 1, August 18-20, 2009, China, pp: 154-157.
16. Khurana, K., P.S. Sajja and Z. Bhatt, 2016. Fuzzy based research techniques for intrusion detection and analysis: A survey. Int. Res. J. Eng. Technol., 3: 1223-1227.
17. Pawar, S.N., 2013. Intrusion detection in computer network using genetic algorithm approach: A survey. Int. J. Adv. Eng. Technol., 6: 730-736.
18. Dastanpour, A., S. Ibrahim, R. Mashinchi and A. Selamat, 2014. Comparison of genetic algorithm optimization on artificial neural network and support vector machine in intrusion detection system. Proceedings of the IEEE Conference on Open Systems, October 26-28, 2014, Malaysia, pp: 72-77.
19. Benaicha, S.E., L. Saoudi, S.E.B. Guermeche and O. Lounis, 2014. Intrusion detection system using genetic algorithm. Proceedings of the Science and Information Conference, August 27-29, 2014, London, pp: 564-568.
20. Xie, T.T., H. Yu and B.M. Wilamowski, 2012. Comparison of Fuzzy and Neural Systems for Implementation of Nonlinear Control Surfaces. In: Human-Computer Systems Interaction: Backgrounds and Applications, Hippe, Z.S., J.L. Kulikowski and T. Mroczek (Eds.). Springer, New York, pp: 313-324.

21. Srinivasu, P. and P.S. Avadhani, 2012. Genetic algorithm based weight extraction algorithm for artificial neural network classifier in intrusion detection. Procedia Eng., 38: 144-153.
22. Khan, L., M. Awad and B. Thuraisingham, 2007. A new intrusion detection system using support vector machines and hierarchical clustering. VLDB J., 16: 507-521.
23. Heba, F.E., A. Darwish, A.E. Hassanien and A. Abraham, 2010. Principle components analysis and support vector machine based intrusion detection system. Proceedings of the 10th International Conference on Intelligent Systems Design and Applications, November 29-December 1, 2010, Cairo, Egypt, pp: 363-367.
24. Stamper, M., 2016. Artificial intelligence in network intrusion detection. Information Systems Security Bureau, Croatia. https://www.fer.unizg.hr/_download/repository/KDI,_Miroslav_Stampar.pdf