Information Sciences Letters
An International Journal

http://dx.doi.org/10.18576/isl/1203047

Predication Attacks Based on Intelligent Honeypot Technique

A. Alshahrani

Computer Science Department, Arab Open University, Riyadh, Kingdom Saudi Arabia

Received: 22 Jun. 2022, Revised: 22 Sep. 2022, Accepted: 6 Nov. 2022.
Published online: 1 Mar. 2023.

Abstract: Honeypot combined with machine learning techniques is offered as a model for intrusion detection presented in the current research. Recent years have seen an uptick in the number of security initiatives implemented by every type of business. This requires anticipatory analysis of a potential attack in order to achieve the desired result. Honeypots are one of the instruments used to observe malicious actors in action. A honeypot is a type of network system used to detect intrusions into computer networks by observing and analysing the actions of potential intruders in a controlled, but vulnerable, setting. Improved outcomes in terms of true positives and false positives were also presented thanks to the use of the Decision Tree (DT). Both the overall accuracy in detecting attacks and the false alarm rate are enhanced by the suggested model-based honeypot and machine learning.

Keywords: security; predication; Honeypot; detection.

1 Introduction

These days, every business needs some sort of internet presence to function, which makes internet security important. Despite extensive research efforts, it is currently impossible to ensure the safety of all interconnected systems on the internet. Every day, hackers find novel ways to breach networks' security. Although it is a challenging task, new software is being released to counteract the methods used by attackers to circumvent this security flaw. This is how honeypots have been introduced to the attacker by security analysts.

Organizations often employ "honeypots," or decoy systems, to lure in attackers so that their actions may be recorded and analyzed later [1]. The honeypot is a widely used instrument of intrusion detection and analysis that simulates a legitimate system in order to gather information about an attacker's methods. There are certain advantages to honeypots, such as their minimal resource consumption and ease of installation [2], but there are also some drawbacks, such as the increased danger of takeover due to their susceptibility to detection and fingerprinting and their narrow field of view. As a result, honeypots are typically used in conjunction with NADS [3] to detect network anomalies.

Machine learning's ability to detect malware will be put to the test by the extensive data mining that will be required. When machine learning has little in the way of background knowledge, it has a hard time detecting new forms of malware [4][5]. Honeypot can help update machine learning by prepping data so it can be more accurate. However, the honeypot doesn't have a clear design in its implementation, making it difficult and confusing to use, especially when tailoring it to the needs of developing a machine learning model. As a result, more research into relevant texts is needed to spot shifts and give head in the right way.

In this research, we provide an intelligent honeypot-based prediction model that employs machine learning methods to detect intrusion attempts. Experimental data demonstrate that the DT algorithm can improve accuracy.

The remainder of this paper is organized as follows. Section 2 introduces the related work. Section 3 is the proposed system. Section 4 presents model evaluation metrics. Section 5 is the results and section 6 presents the conclusion.

2 Related Works

Recent studies on honeypots show their increasing prevalence. The authors of [6] examined the use of honeypots in IoT networks and discussed their history, importance, and history in data security. [7] reviews honeypot-based botnet identification, while [5] audits malware recognition using honeypots.

The authors of [8] employed honeypots to get their adversary discovery methodology approved in an IoT enterprise. Midway through 2020, the network defense innovation company Cybereason sent out a network honeypot to learn about the tactics, procedures, and tools used by online criminals. The lessons learned from this approach and the use of honeypots in the security of important framework systems were covered in [9].

*Corresponding author e-mail: A.shahrani@arabou.edu.sa

Using a web-based interface and an RDBMS, the Honey analyzer is utilized in to analyze Honeyed logs [10]. The honeyed tool has been researched in the publication [11]. a technique for producing fake honeypots that mimic network-level services in computer systems. Additionally, the design of honeyed is described, and it is demonstrated how honeyed contributes to system security.

For the purpose of gathering data regarding users connected to the FTP server, an FTP honeypot system was developed [12]. The honeynet system is used [13] to help system administrators find suspicious activity and network intrusions. A honeypot simulation for engaging with attackers is created in [14]. Furthermore, a variety of tools are used to analyze data that was obtained via a honeypot. [15] evaluates a novel threat intelligence method for identifying attack trends.

The authors demonstrate how difficult it is to use general-purpose tools to assess the massive amount of data produced by honeypots. Elasticsearch, an open-source search analytics engine, is employed to find a solution. The authors of [16] used supervised machine learning techniques to find attacks on IoT devices by utilizing honeypots. In addition, honeypots are employed to defend against specific attacks. Attacks using denial of service (DoS) can take down entire networks without spotting security loopholes.

A honeypot concept is suggested for Internet of Things (IoT) devices [17] to reduce DoS attacks. With relation to mobile, spontaneous organizations, Tiruvakadu and Pallapa discuss wormhole attacks and suggest a honeypot-based solution to confirm these attacks [18]. A wormhole attack involves at least two attackers placing themselves in a target organization on purpose and creating a path between them. This shortens the route between those particular hubs, encouraging actual traffic to use the attacker's passage. They argue against the necessity of a system for confirming attacks and employ a honeypot to do so while using a wormhole assault tree.

3 The Proposed System

In this paper, a prediction model is based on an intelligent honeypot technique using machine learning algorithms to identify attacks. The proposed model is a combination of honeypot systems and ML techniques for threat predictions by gathering information and analysis to ensure network security. The major steps of the proposed system are presented in Figure 1.

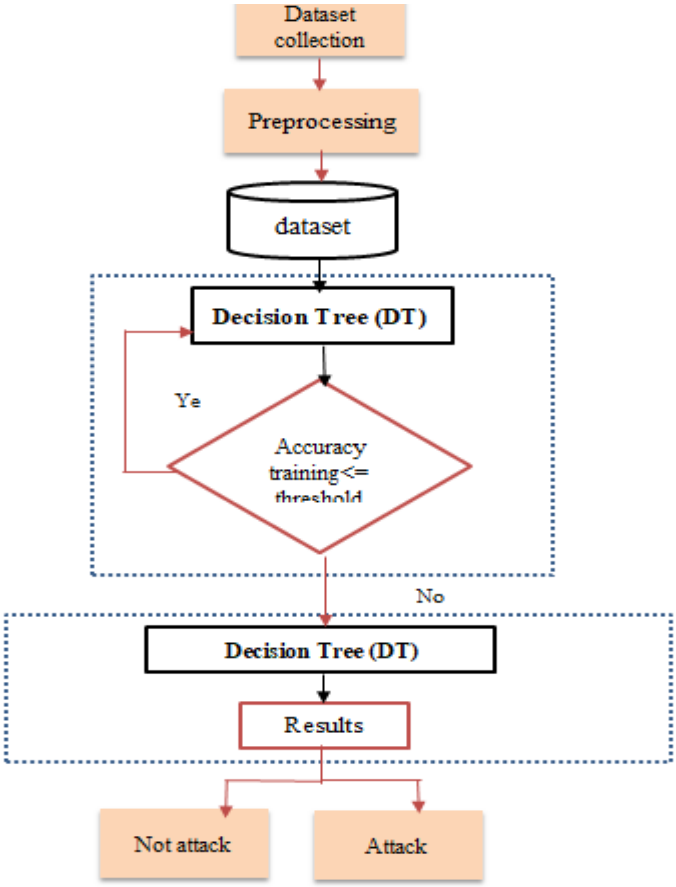

Fig. 1. The architecture of the Proposed System

The suggested system presents pre-processing after dataset gathering. To improve system speed, features are needed to convert some characters and symbols into numbers. The dataset used to generate the suggested system is then fed into the training phase DT. The testing phase is offered to ascertain the rate of false alarms and detection accuracy. 28 features from the gathered dataset in [19] were used in the suggested detection model. To prevent DNN issues, or overfitting of the data set, the acquired data was divided into two subsets: the testing set (40%), and the training set (60%).

4 Model Evaluation

The following performance metrics were used in this study to assess the effectiveness of the suggested system: detection accuracy rate; false alarm rate; and error rate:

$$\text{Accuracy} = \frac{\text{Number of correctly classified patterns}}{\text{Total number of patterns}} \quad (1)$$

$$\text{Error rate} = 1 - Accuracy \quad (2)$$

Additionally, four alarms (a confused matrix) were shown to have a false alarm rate in this study, as shown in Equations 3 and 4:

$$\text{TN}_{\text{Rate(specificity)}} = \frac{\text{TN}}{\text{TN+FP}} \quad (3)$$

$$\text{FN}_{\text{Rate(1-sensitivity)}} = \frac{\text{FN}}{\text{FN+TP}} \quad (4)$$

- True negative (TN) is the number of normal servers determined as normal servers.
- False negative (FN) is the number of honeypot servers determined as normal servers.

5 Results and Discussion

The results of the experiment demonstrate the potential of the DT-based proposed model to improve performance detection. More can be done to improve detection accuracy while decreasing the frequency of false alarms. Online single honeypot and online multi-honeypot technologies are the focus of the experiment in order to increase the predictability and authenticity of the outcome. Performance indicators for the proposed model are shown in Tables 1 and 2.

Table 1: Performance Metrics

Online Single Honeypot System				
Detection rate	Avg. False alarm	Time	Error rate	Overhead
95.52%	4.81%	13 s	1.04%	2584p/s

Table 2: Performance Metrics

Online Multi-Honeypot System				
Detection rate	Avg. False alarm	Time	Error rate	Overhead
90.87%	12.3%	43 s	3.17%	59547p/s

The main functions of the suggested DT-based system are to decrease false alarm rates, increase detection rates, and reduce error rates. According to the findings shown in Tables 1 and 2, there are variations in the ratio of detections to false alarms. By using a DT, you may improve detection rates while reducing the number of false alarms. When contrasting our findings with the earlier investigation in [20]. While the prior study only managed to attain 91% accuracy, our proposed model's total detection accuracy was at 93.19%, enabling an effective detection rate based on DT.

6 Conclusion

In this paper, a model is proposed for attack detection and classification by combining machine learning and honeypot. Moreover, the DT algorithm is exploited to present more accurate results in terms of decreasing of detection rate with a low false alarm rate. The proposed system overcomes all the challenges presented by the previous research such as the high rate of false alarms and low detection rate. With the dataset collected, this proposed model presented efficient results obtained for a high rate of detection accuracy and decreasing false alarm rate.

Acknowledgment:

The author would like to thank Arab Open University, Saudi Arabia for supporting this study.

Conflict of interest

The authors declare that there is no conflict regarding the publication of this paper.

References

[1] I. Barak. Critical infrastructure under attack: lessons from a honeypot. *Netw. Secur.*, **9**, 16– 17 (2020).

[2] T. Campbell, R. M., Padayachee, K. and Masombuka. A survey of honeypot research: Trends and opportunities. *10th Int. Conf. internet Technol. Secur. Trans. (ICITST)', IEEE*, 208–212 (2015).

[3] I. Mokube and M. Adams. Honeypots: Concepts, Approaches, and Challenges. in *Proceedings of the 45th Annual Southeast Regional Conference*, 321–326 (2007).

[4] R. Baykara, M. and Das. A novel honeypot-based security approach for real-time intrusion detection and prevention systems. *J. Inf. Secur. Appl.*, **41**, 103–116 (2018).

[5] B. Matin, I. M. M. and Rahardjo. The Use of Honeypot in Machine Learning Based on Malware Detection: A Review. *'2020 8th Int. Conf. Cyber IT Serv. Manag.* (2020).

[6] S. E. W. Y. Roh and G. Heo. A Survey on Data Collection for Machine Learning: A Big Data - AI Integration Perspective. *IEEE Trans. Knowl. Data Eng*, **4347**, 1–10 (2019).

[7] F. Z. Razali, M. F., Muruti, G., Razali, M. N., Jamil, N. and Mansor. IoT honeypot: A review from researcher's perspective. *'IEEE Conf. Appl. Inf. Netw. Secur. AINS*, 93–98 (2018).

[8] N. Z. Seungjin, L., Abdullah, A. and Jhanjhi. A review on honeypot-based botnet detection models for smart factory. *Int. J. Adv. Comput. Sci. Appl.*, **11**(6), 418–435 (2020).

[9] F. Zhang, J., Bhuiyan, M. Z. A., Yang, X., Wang, T., Xu, X., Hayajneh, T. and Khan. AntiConcealer: Reliable Detection of Adversary Concealed Behaviors in EdgeAI Assisted IoT. *IEEE Internet Things J.*, 1–11 (2021).

[10] U. Thakar. HoneyAnalyzer – Analysis and Extraction of Intrusion Detection Patterns & Signatures Using Honeypot. *The Second International Conference on Innovations in Information Technology (IIT'05)* (2017).

[11] N. Provos. A Virtual Honeypot Framework. *Proceedings of the 13th USENIX Security Symposium* (2004).

[12] I. V. V. A. Perevozchikov, T. A. Shaymardanov and Chugunkov. New techniques of malware detection using FTP Honeypot systems. *Proc. 2017 IEEE Russ. Sect. Young Res. Electr. Electron. Eng. Conf. ElConRus*, 204–207 (2017).

[13] I. Mahmood. Computer Science & Systems Biology The Use of Honeynets to Detect Exploited Systems Across the Wireless Networks. *Journal of Computer Science & Systems Biology*, **11(3)**, 219–223 (2018).

[14] N. Bhagat and B. Arora. Intrusion detection using honeypots. *PDGC 2018 - 2018 5th Int. Conf. Parallel, Distrib. Grid Comput.*, 412–417 (2018).

[15] L. Almohannadi, H., Awan, I., Al Hamar, J., Cullen, A., Disso, J. P. and Armitage. Cyber threat intelligence from honeypot data using elasticsearch. *'Proceedings - Int. Conf. Adv. Inf. Netw. Appl. AINA*, 900–906 (2018).

[16] C. Shrivastava, R. K., Bashir, B. and Hota. Attack Detection and Forensics Using Honeypot in IoT Environment. *G. Fahrnberger, S. Gopinathan L. Parida, eds, 'Distributed Comput. Internet Technol. Springer Int. Publ.*, 402–409 (2019).

[17] D. J. N. M. Anirudh and S. A. Thileeban. Use of honeypots for mitigating DoS attacks targeted on IoT networks. *Proc. 2017 Int. Conf. Comput. Commun. Signal Process. (ICCCSP), Chennai, India*, 1–4 (2017).

[18] V. Tiruvakadu, D. S. K. and Pallapa. Confirmation of wormhole attack in MANETs using honeypot. *Comput. Secur.*, **76**, 32–49 (2018).

[19] CIRA-CIC-DoHBrw-2020. [Online]. Available: https://www.unb.ca/cic/datasets/dohbrw-2020.html. [Accessed: 22-Nov-2021].

[20] J. L. W. Wang, Y. Shang and Y. He, Y. Li. An explainable sentiment prediction model based on the portraits of users sharing representative opinions in social sensors. *Inf. Sci. (Ny).*, **511** (2020).